TET1 attenuates human pluripotency-related heterogeneities and enhances differentiation capabilities

Makoto Motono[1,2,4,*], Keiko Hiraki-Kamon[1,4], Masayoshi Kamon[1], Hidenori Kiyosawa[3], Yoichi Kondo[2], and Hidemasa Kato[1,5,*]

[1]Department of Developmental Biology and Functional Genomics, Ehime University Graduate School of Medicine, 454, Shitsukawa, Toon, Ehime 791-0295, Japan

[2]Department of Anatomy and Cell Biology, Osaka Medical and Pharmaceutical University, 2-7, Daigakumachi, Takatsuki, Osaka 569-8686, Japan

[3]Department of Life Science, Chiba Institute of Technology, 2-17-1, Tsudanuma, Narashino, Chiba 275-0016, Japan

[4]These authors contributed equally.

[5]Lead Contact

[*]Correspondence: makoto.motono@ompu.ac.jp (M.M.); hidemasa.kato.zz@ehime-u.ac.jp (H. Kato)

Abstract

Human induced pluripotent stem cells (hiPSCs) offer transformative potential for developmental research and therapeutic discovery; however, their variable characteristics are a significant barrier to their broader application. We reprogrammed hiPSCs by adding TET1, a DNA-demethylating dioxygenase, to produce TET1-iPSCs (T-iPSCs) with enhanced epithelialization and differentiation capabilities. A default differentiation survey—designed and implemented in our hiPSC-production pipeline—revealed considerable variation in differentiation capacity across qualitatively pluripotent iPSC clones. By comparing high-differentiation T-iPSCs and conventional iPSCs (C-iPSCs) at the single-cell level, we identified a subpopulation of C-iPSCs in the G_0/G_1 phase (*CDKN2A*-positive) exhibiting low *TET1* expression and an elevated extraembryonic mesenchymal gene signature. The extraembryonic mesenchymal gene signature is common to iPSCs and quantitatively correlates to default differentiation results. Deviation of C-iPSCs from the cell-cycling embryonic epithelial gene profile was accompanied by DNA hypermethylation concomitant with *TET1* insufficiency and the accompanying derepression of poised and Polycomb-repressed enhancers, leading to the derepression of senescence-associated and extraembryonic genes. TET1-facilitated reprogramming can ameliorate these deviations. This study provides a standardized hiPSC-production pipeline and discovers an aspect of TET1 for establishing human pluripotency by remedying the inherent heterogeneities resulting from C-iPSC reprogramming.

38 **Introduction**

39 Human pluripotent stem cells (hPSCs) exhibit variable capabilities for differentiation, which is
40 more pronounced in human induced PSCs (hiPSCs) (Koyanagi-Aoi et al. 2013; Ohnuki et al.
41 2014) than in human embryonic stem cells (hESCs) (Bock et al. 2011). Efforts to optimize
42 human reprogramming have focused on enhancing the efficiency of iPSC colony derivation
43 (Esteban et al. 2010; Wang et al. 2017); however, strategies to improve differentiation capacities
44 are limited. In mice, Tet1, a DNA-demethylating dioxygenase, increased the differentiation
45 potential of iPSCs by demethylating enhancers (Bartoccetti et al. 2020). Mesenchymal-to-
46 epithelial transition is an essential step of reprogramming (Schiebinger et al. 2019), in which
47 Tet1 plays a crucial role (Hu et al. 2014). Compared to that in mice, in humans, the onset of
48 *TET1* expression during reprogramming is more delayed, which hinders mesenchymal-to-
49 epithelial transition and leads to epigenetic aberrations (Cacchiarelli et al. 2015; Bartoccetti et
50 al. 2020). Therefore, the Yamanaka factors alone may be insufficient to demethylate the genome
51 of human mesenchymal cells and attain genuine pluripotency.

52 We propose a novel hiPSC-production pipeline that adds TET1 to the Yamanaka factors,
53 yielding TET1-iPSCs (T-iPSCs). This pipeline integrates differentiation tests of individual
54 clones, selecting hiPSCs with enhanced differentiation capabilities. Assessment of
55 differentiation capabilities and the corresponding systematic transcriptomics revealed inter- and
56 intra-clonal cell heterogeneities incidental to the differentiation defects of conventional hiPSCs
57 (C-iPSCs). Analyses of the DNA methylome and histone modification support the hypothesis
58 that the dysregulation of this gene is inherent to conventional reprogramming from somatic
59 cells.

60

61 **Results**

62 *TET1-aided human cell reprogramming: T-iPSCs*

63 During hiPSC reprogramming, the establishment of pluripotent DNA methylation patterns is
64 delayed compared with that in mice (Cacchiarelli et al. 2015). By reanalyzing single-cell
65 transcriptome data from human reprogramming (Liu et al. 2020), we found that *TET1*
66 expression was selectively pronounced during a successful reprogramming, followed by
67 successive *DNMT3A* and *DNMT3B* upregulation (Supplemental Fig. S1A,B). Given the known
68 pleiotropic roles of TET1 in reprogramming and pluripotency (Verma et al. 2018; Bartoccetti
69 et al. 2020; Cheng et al. 2022), we generated hiPSCs using episomal vectors carrying
70 conventional reprogramming factors alone (C-iPSCs) (Okita et al. 2011) and with *TET1*
71 addition (T-iPSCs) (Fig. 1A). We cultured hiPSCs in an epiblast stem cell medium (Brons et al.
72 2007). This enabled single-cell passage and improved morphological characterization (Fig. 1B).
73 During colony formation, most emerging T-iPSCs organized into a simple columnar epithelium.
74 These epithelialized colonies were rare for C-iPSCs, which bore mesenchymal morphologies
75 (Fig. 1C,D). Some clones reverted to mesenchymal morphologies upon expansion (Fig. 1E). As
76 the emerging epiblasts manifested as epithelialized structures *in vivo* and the epithelialization
77 of mesenchymal cells during reprogramming results from TET activities, we aborted
78 mesenchymal clones for both C-iPSCs and T-iPSCs (Fig. 1B,E). Finally, we established 25 C-
79 iPSC and 145 T-iPSC clones (from 140 and 420 colonies, respectively) (Fig. 1F). A vector-free
80 threshold of 0.03 was employed to exclude clones with episomal vector genome integration
81 (Supplemental Fig. S1C), yielding 46 vector-free clones (7 C-iPSCs and 39 T-iPSCs) from four
82 independent parallel inductions.

83 To characterize the hiPSC clones at the earliest point, we quantitatively analyzed the
84 endogenous expression of "pluripotency-associated" factors in all clones. We observed no
85 noticeable difference between C-iPSCs and T-iPSCs in expression levels, marker-positive ratios,
86 or sex variance for *NANOG*, *POU5F1,* and *SOX2,* indicating the overall uniformity of the
87 hiPSC clones obtained, and all clones conformed to the qualitative criteria for pluripotency (Fig.

88 1G; Supplemental Figs. S1D, S2A). By contrast, *KLF4* markedly varied across clones
89 (Supplemental Fig. S2A). Despite the addition of TET1 during reprogramming, T-iPSC clones
90 did not show higher endogenous *TET1* levels than C-iPSC clones (Fig. 1G).
91
92 *Interclonal heterogeneity in the differentiation* cap*abilities of hiPSCs*
93 To set a tentative threshold for hiPSC differentiation capability, we chose the default model of
94 neural induction (Fig. 2A). According to this evolutionarily conserved model, dissociated early
95 embryonic cells default to neural induction when incubated in a developmentally neutral culture
96 medium devoid of signaling molecules (Munoz-Sanjuan and Brivanlou 2002; Patani et al. 2009).
97 We selected well-expandable C-iPSCs (1C6) and T-iPSCs (1T35 and 1T47) and pilot-
98 differentiated them. For 1T35 and 1T47, neural differentiation peaked on day 12 of default
99 differentiation, and the expression of PAX6 and SOX1 was maintained until day 18 (Fig. 2B;
100 Supplemental Fig. S2B) (Chambers et al. 2009; Zhang et al. 2010), but 1C6 cells failed to
101 express neural markers and regained pluripotency markers, indicating a defect in pluripotency
102 exit (Supplemental Fig. S2B). We noticed substantial cavities in 1C6-derived spheres, whereas
103 1T47 produced solid spheres (Fig. 2B). Thus, hiPSCs exhibited considerable differentiation
104 variegation across clones. However, some hiPSC clones can exhibit quasi-total differentiation
105 toward the neural lineage comparable to hESCs (Patani et al. 2009).
106 To select high-quality hiPSC clones, we investigated each clone using default
107 differentiation. We evaluated expression on day 8 when neural differentiation efficiency
108 diverged (Supplemental Fig. S2B). In addition, we analyzed *HES5* expression (Tchieu et al.
109 2017). Substantial clones exhibited low differentiation potential, but T-iPSCs contained high-
110 potential outliers (Fig. 2C). *NANOG* (Loh and Lim 2011) and *TBXT* (Sasaki et al. 2016)
111 correlated negatively with neural markers and were used to disqualify clones (Supplemental
112 Fig. S2C,D). We selected clones that outperformed 1C6 in all tests (Fig. 2D) as "default-elite"

113 clones (karyotypes shown in Supplemental Fig. S3).

114 A retrospective analysis confirmed that the copy number of residual vectors below the set
115 standard of a vector-free threshold (0.03 vector/genome) did not correlate with neural marker
116 expression following default differentiation (Supplemental Fig. S2E). Based on these criteria,
117 only 0.00005% of fibroblasts used for C-iPSC induction acquired the default differentiation
118 capability, a standard that is >100 times more stringent than another report (Fig. 2E) (Okita et
119 al. 2011). TET1-assisted reprogramming increased efficiency to 0.0002%—a fourfold increase
120 in the production of hiPSC clones with enhanced differentiation potential.

121

122 *Transcriptome features affecting differentiation efficiency*

123 To link undifferentiated gene expression profiles to differentiation potential, we measured gene
124 expression before (day 0) and after (day 8) default differentiation. We found no gene positively
125 correlated with due differentiation capability (*HES5*/*SOX1*), including classical pluripotency
126 genes (Supplemental Fig. S2C). Conversely, *KLF4* and *MYCN* correlated negatively with neural
127 markers (Supplemental Fig. S2C). Because KLF4 is a factor in differentiation resistance (Kim
128 et al. 2011; Lister et al. 2011; Ohnuki et al. 2014), we investigated KLF4-related factors
129 (DNMT3L and NLRP7) (Cacchiarelli et al. 2015) and HAND1 (Yang et al. 2021). We found
130 that these factors (collectively HDKM genes) negatively correlated with neural markers and
131 *HAND1* positively correlated with aberrant markers (*BMP4* and *NANOG*) (Supplemental Fig.
132 S2C). Clone-by-clone inspection of individual clones confirmed the unrelatedness of classical
133 pluripotency factors to differentiation capability (Fig. 2F). By contrast, clones expressing low
134 *HDKM* in the undifferentiated state tended to express higher levels of the three neural markers
135 (Fig. 2G).

136

137 *Neuronal differentiation beyond default neurogenesis*

138 To evaluate differentiation potential beyond the early neural stages assessed by default
139 differentiation, we selected the precursors of the ventral midbrain (VM). We modified a
140 published protocol (Steinbeck et al. 2015) and achieved high efficiency in inducing FOXA2
141 and LMX1A—VM precursor markers—using the default-elite clones (Andersson et al. 2006).
142 Differentiation efficiency was the highest for 1T47 (Fig. 3A,B; Supplemental Fig. S4A–C).
143 However, some cells lacked LMX1A, NKX6-1, or PAX6, which are markers of VM precursors
144 (Duan et al. 2013). The proportion of these "undefined" cells, presumably induced to cells
145 outside the VM, was significantly higher for 6C6 than for 1T47 (Fig. 3C; Supplemental Table
146 S1). Furthermore, the neuronal dendrites derived from 1T47 were longer than those from 6C6
147 (Fig. 3D,E).
148
149 *Interclonal heterogeneity of hiPSC clones*
150 We tentatively designated 1T47, with elevated differentiation potentials, as the reference clone.
151 We aimed to identify the distinguishing features unique to 1T47 (T-iPSC^{1T47}) compared with
152 other default-elite clones. We selected 6C6 (C-iPSC^{6C6}) as our best C-iPSC clones (6C4 was
153 aneuploid; Supplemental Fig. S3). Bulk RNA-sequencing (RNA-seq) analysis identified
154 differentially expressed genes (DEGs) between these clones (Fig. 4A). Gene ontology (GO)
155 analysis indicated few GO terms associated with DEGs expressed at higher levels in T-iPSC^{1T47}
156 (T-iPSC^{1T47}-high-DEGs) (Fig. 4B). C-iPSC^{6C6} expressed genes of mesenchymal traits (*TWIST1*,
157 *MYL7, ACTC1, ANXA1*, *VIM*, and *BMP4*) (Fig. 4B,C) (Pham et al. 2022) and genes related to
158 aging, such as cell cycle inhibitors (*CDKN1A/2A/2B*) and senescence-associated secretory
159 phenotype (SASP: *GDF15, IL6, IL1A,* and *IGFBP7*) (Fig. 4C) (Kumari and Jat 2021).
160 Next, we evaluated the versatility of mesenchymal-related and aging-associated gene
161 expression across hiPSC clones. Reanalysis of array data from 49 hiPSC clones revealed
162 significant variation in mesenchymal-related and aging-associated gene expression

163 (Supplemental Fig. S5A,B,D) (Koyanagi-Aoi et al. 2013). Differentiation-defective markers
164 (Ohnuki et al. 2014) were examples of substantial interclonal expression heterogeneities
165 (Supplemental Fig. S5C,D). These results suggest that interclonal heterogeneity is common
166 across hiPSC clones and not unique to our selected reference C-iPSC^{6C6} clone.
167
168 *Mesenchymal gene noise in C-iPSCs is related to extraembryonic lineages*
169 The mesenchymal propensity was surprising because hiPSCs are supposedly equivalent to
170 postimplantation epiblasts—an epithelialized tissue (Nakamura et al. 2016). Because obtaining
171 human embryonic gene profiles from implantation to gastrulation is difficult, we used single-
172 cell RNA-seq (scRNA-seq) data of peri-gastruloids (Liu et al. 2023) to determine lineages that
173 express mesenchymal genes during human development. Peri-gastruloids encompass
174 developmental trajectories from the epiblast to the definitive mesoderm/endoderm and
175 extraembryonic lineages (Supplemental Fig. S6A–D). We integrated peri-gastruloid scRNA-
176 seq data with data obtained from T-iPSC^{1T47}, C-iPSC^{6C6}, and a standard hESC H9 (Messmer et
177 al. 2019). Although the mesenchymal markers (*ACTC1*, *BMP4*, *HAND1*, *MYL7*, *TWIST1*, and
178 *VIM*) are widely recognized as mesodermal markers, cells expressing these genes in the peri-
179 gastruloids did not coexpress the definitive mesodermal markers, *MESP1/2* and *MIXL1*
180 (Supplemental Fig. S6C,D) (Liu et al. 2023). Notably, this population exclusively expressed
181 *ANXA1*, a human extraembryonic mesoderm marker (Supplemental Fig. S6D) (Pham et al.
182 2022). T-iPSC^{1T47} and hESC H9 cosegregated close to the epiblast clusters, whereas C-iPSC^{6C6}
183 additionally scattered within the extraembryonic mesoderm (Supplemental Fig. S6A,E). These
184 data indicate that a subpopulation of C-iPSC^{6C6} is akin to the extraembryonic mesoderm rather
185 than the definitive mesoderm (Fig. 4C; Supplemental Fig. S6C,D). To consolidate the
186 developmental contexts of hPSCs, we referred to cultured monkey blastocysts up to the
187 gastrulation stage (Supplemental Fig. S7A) (Yang et al. 2021). Consistently, C-iPSC^{6C6}-high-

188 DEGs were expressed in the amnion or extraembryonic mesoderm and low in the definitive
189 mesodermal lineages (Fig. 4C; Supplemental Fig. S7B). These results indicate that C-iPSC^{6C6}
190 expresses extraembryonic mesodermal/mesenchymal rather than embryonic mesodermal genes.
191

192 *Intraclonal heterogeneities in C-iPSCs*

193 scRNA-seq analyses revealed a proportion of C-iPSC^{6C6} cells with halted cell cycles (Fig. 4D).
194 We defined cells with higher *CDKN2A* expression and lower *MKI67* expression as being in the
195 G_0 phase (Supplemental Fig. S8A). Based on this, 7% of C-iPSC^{6C6} and no T-iPSC^{1T47} were in
196 the G_0 phase (Fig. 4E). *TET1*, *DNMT3B*, and *SOX2* expression was lower in the G_0 phase (Fig.
197 4D,F,G; Supplemental Fig. S8B–D). Notably, C-iPSC^{6C6} cells in the G_0 phase highly expressed
198 extraembryonic genes (Fig. 4D,F,G; Supplemental Fig. S8B–D). Given the lower expression of
199 *TET1* in cells expressing cell cycle inhibitors and extraembryonic markers, we reanalyzed the
200 RNA-seq data for *TET1/2/3*-triple-knockout hESCs (TET-TKO) (Charlton et al. 2020). TET-
201 TKO exhibited increased expression of cell cycle inhibitors, extraembryonic markers, and
202 differentiation-defective markers, as in C-iPSC^{6C6} (Fig. 4H; Supplemental Fig. S8E).
203 Collectively, a subpopulation of C-iPSCs in the G_0 phase had low *TET1* expression, exhibiting
204 mesenchymal/extraembryonic gene noise (Fig. 4I).

205 We evaluated whether the observed differentiation superiority of T-iPSC^{1T47} (Fig. 3) is a
206 trade-off for their inability to differentiate into extraembryonic lineages. We proactively
207 differentiated T-iPSC^{1T47} by adding BMP4, which induces extraembryonic cells (Chhabra et al.
208 2019; Pham et al. 2022). *BMP4*, a marker of extraembryonic cells, was self-induced at a higher
209 level in T-iPSC^{1T47} than in C-iPSC^{6C6} (Supplemental Fig. S9A). BMP4 induced *TWIST1* and
210 *ANXA1* expression in T-iPSC^{1T47}, indicating its potential to differentiate into extraembryonic
211 lineages (Supplemental Fig. S9A). T-iPSC^{1T47} exhibited higher differentiation efficiency into
212 somatic mesendodermal lineages when C-iPSC^{6C6} retained undifferentiated pluripotency

213 markers (Supplemental Fig. S9B,C). These results indicate that T-iPSCs are not biased toward
214 the neural lineage and can differentiate into the three germ layers and extraembryonic tissues—
215 validating human pluripotency.
216
217 *TET1 insufficiency leads to hypermethylation of poised extraembryonic gene enhancers*
218 To determine the effect of TET1 insufficiency, we compared DNA methylomes of wild-type
219 (WT) hESCs, TET-TKO (Charlton et al. 2020), fibroblasts, T-iPSC^{1T47}, and C-iPSC^{6C6} using
220 whole genome bisulfite sequencing (WGBS). We examined differentially methylated regions
221 (DMRs) and identified 4769 C-iPSC^{6C6}- and 1244 T-iPSC^{1T47}-hypermethylated DMRs (hyper-
222 DMRs). DMRs were observed around the C-iPSC^{6C6}-high-DEG loci of extraembryonic
223 markers, aging-related markers, and *KLF4* (Fig. 5A; Supplemental Fig. S10A). Although C-
224 iPSC^{6C6}-hyper-DMRs were enriched in both promoters and gene bodies, T-iPSC^{1T47}-hyper-
225 DMRs were enriched in promoters (Fig. 5B). C-iPSC^{6C6}-hyper-DMRs were associated with
226 genes involved in developmental morphogenesis and epithelial-mesenchymal plasticity
227 (Supplemental Fig. S10B) (Haerinck et al. 2023). Although DNA hypermethylation silenced
228 gene expression in T-iPSC^{1T47}, it facilitated gene expression in C-iPSC^{6C6} (Fig. 5C;
229 Supplemental Table S2) (Jones 1999). This paradox was resolved by analyzing the Polycomb
230 repression mark, H3K27me3. Both C-iPSC^{6C6}- and T-iPSC^{1T47}-hyper-DMRs attracted
231 H3K4me1, H3K4me3, H3K27ac, and H3K27me3 in hPSCs (Supplemental Fig. S10C,D).
232 Especially, C-iPSC^{6C6}-hyper-DMRs exhibited significant enrichment of H3K4me1 and
233 H3K27me3 in hESCs (Fig. 5D). The regions co-occupied with a putative enhancer mark
234 (H3K4me1) and the Polycomb repression mark (H3K27me3) are categorized as "poised
235 enhancers" (Rada-Iglesias et al. 2011). We found gene body-embedded poised enhancers
236 intersecting C-iPSC^{6C6}-hyper-DMRs in hPSCs (Fig. 5A; Supplemental Fig. S10A). In a
237 reference hESC clone, H3K27me3 was significantly enriched in the C-iPSC^{6C6}-hyper-DMRs

238 associated with the C-iPSC^{6C6}-high-DEGs (Fig. 5E; Supplemental Fig. S10E). Notably, C-
239 iPSC^{6C6}-hyper-DMRs associated with C-iPSC^{6C6}-high-DEGs were significantly enriched in
240 putative enhancer regions (Fig. 5F; Supplemental Fig. S10F). Hypermethylation of poised
241 enhancers could cause target gene derepression because DNA methylation repels Polycomb
242 repressive complex 2 (PRC2) (Blackledge et al. 2015). These results suggest that TET1
243 insufficiency leads to the hypermethylation of poised enhancers and the paradoxical
244 derepression of the related genes. Of the 1232 C-iPSC^{6C6}-high-DEGs, 350 were upregulated in
245 PRC2 component EZH2-knockout hESCs (EZH2-KO) (Fig. 5G,H) (Collinson et al. 2016). We
246 identified 127 genes commonly overexpressed in C-iPSC^{6C6}, EZH2-KO, and TET-TKO (Fig.
247 5H; Supplemental Table S3). GO analysis of these common genes showed enrichment of
248 epithelial-mesenchymal transition-related terms (Fig. 5I). These findings suggest that the
249 hypermethylation of poised enhancers in C-iPSC^{6C6} resulting from TET1 insufficiency may be
250 the cause of the derepression of cell cycle inhibitors and extraembryonic mesenchymal genes
251 by evading Polycomb-mediated repression and the ensuing morphological changes.
252
253 *Aberrant methylation is acquired during iPSC reprogramming*
254 Most C-iPSC^{6C6}-hyper-DMRs within gene bodies were hypomethylated in fibroblasts (Fig. 5A),
255 suggesting that these regions acquired methylation during reprogramming (Ma et al. 2014;
256 Roost et al. 2017; Buckberry et al. 2023). We reanalyzed methylome data from conventionally
257 reprogrammed iPSCs (Cacchiarelli et al. 2015) and found 1732 autosomal enhancers that were
258 poised in hPSCs and acquired *de novo* DNA methylation during reprogramming (Fig. 5J).
259 Approximately 70% of the enhancers were within gene bodies, and *de novo* methylated poised
260 enhancers were similarly distributed (Supplemental Fig. S10G). These results indicated that
261 many poised enhancers in hiPSCs acquired *de novo* methylation in the gene body as the sole
262 consequence of reprogramming. We assessed their expression during reprogramming because

263 gene-body methylation primarily correlates with gene expression in the presence of DNMT3B
264 (Jeziorska et al. 2017). As expected, these extraembryonic genes were transiently expressed
265 during reprogramming (Supplemental Fig. S10H) (Liu et al. 2020). The DNA methylation
266 pattern in hPSCs is constantly regulated by counteracting methylating (DNMT3A/B) and
267 demethylating (TET1) activities (Charlton et al. 2020). This balance may be disrupted in a
268 subset of cells with insufficient TET1 and DNMT3B activity during reprogramming
269 (Supplemental Fig. S1A,B) and in the G_0 phase in hiPSCs (Fig. 4F).
270

271 **Discussion**

272 Our default differentiation survey showed that the expression of individual pluripotency-
273 associated markers not only did not positively correlate to differentiation results but also
274 negatively correlated to neural induction (*KLF4*) in some instances. The "precarious balance
275 model" argues that pluripotency factors act as lineage specifiers, and the balance of these factors
276 is important for maintaining the undifferentiated pluripotent state (Loh and Lim 2011).
277 Therefore, the individual expression of pluripotency markers should be interpreted cautiously
278 and not be used alone to validate the pluripotency of the relevant PSC clone. Notably, no
279 markers showing a positive correlation with differentiation capabilities were identified. This
280 observation supports the idea that variations in the quality of pluripotent stem cells are reflected
281 in the aberrant upregulation of lineage-specific, typically silenced genes. Such derepression is
282 likely driven similarly to the epigenetic drift proposed in the Information Theory of Aging (Lu
283 et al. 2023).

284 Default differentiation is not merely a single lineage differentiation test for neural induction.
285 Human default differentiation may discern the propensity between somatic/embryonic versus
286 extraembryonic differentiation. In addition to its allowance for quantitative measurements, the
287 default test has practical advantages. Compared to *in vivo* teratoma formation, it is a short-term

288 test, making it less labor-intensive, and the defaultness helps to provide robustness and
289 reproducibility, making it suitable for a wide range of laboratory settings.

290 The expression of extraembryonic markers in hPSCs evokes a different scenario to the
291 primed pluripotency proposed for mouse PSCs, where cells are halted before differentiation of
292 the germ lineage (Smith 2017). Considering the wide occurrence of *CDKN2A* and
293 extraembryonic gene expression across hiPSCs, a subset of cells residing in the G_0 phase might
294 be common. We found that cells in the G_0/G_1 phase had lower *TET1* expression (Fig. 4F), which
295 may have led to the hypermethylation of the poised enhancers (Fig. 5D–F). In turn, the
296 hypermethylation of nearby poised enhancers may have caused the derepression of
297 extraembryonic genes (Blackledge et al. 2015). Consistently, in extraembryonic tissue, hyper-
298 DMRs frequently overlapped with PRC2-regulated genes in mouse pluripotent cells (Smith
299 2017) and EZH2-KO exhibited extraembryonic gene dysregulation (Collinson et al. 2016).
300 *BMP4* expression is commonly found across hiPSCs (Buckberry et al. 2023) and may assist in
301 an autocrine/paracrine manner to aggravate aberrant extraembryonic gene expression. This
302 highlights the importance of standardization when optimizing target cell derivation in
303 regenerative medicine because the outcome would be affected by the selection of appropriate
304 hiPSC clones and the combined differentiation protocol, which should be established using a
305 standardized PSC clone. This may explain the current scarcity of universally applicable
306 differentiation protocols.

307 The bystander effect of secreted proteins, such as BMP4, from cell cycle-arrested cells is
308 strongly reminiscent of cellular senescence. Indeed, G_0 cells exhibited a senescence-associated
309 secretory phenotype (SASP) (Fig. 4C). The appearance of G_0/senescent cells was unexpected
310 in stem cells. Considering that aging-related hypermethylated gene loci are Polycomb domains
311 (Lu et al. 2023; Moqri et al. 2024), a wide survey of the "ages" of various hPSCs would be
312 informative.

313 This also raises the possibility that these epigenetic controls are reciprocally linked to cell
314 cycle regulation. TET1 and PRC2 establish bivalent chromatin, a cardinal feature of
315 pluripotency (Verma et al. 2018). Cells that fail to achieve adequate pluripotency or bivalency
316 during development may be "outcompeted" through CDKN2A expression, leading to premature
317 senescence and preventing their contribution to further development. This developmental
318 safeguard mechanism may cause suboptimal hPSCs to exhibit heterogeneity.

319 This study has some limitations. We could not decipher how the ectopic TET1—during
320 reprogramming—directed homogeneous endogenous *TET1* expression or mechanistically
321 conferred differentiation advantages to T-iPSCs partly as complete reprogramming is a rare
322 event (one T-iPSC^{1T47} clone/11 default-elite clones/46 integration-free clones/398 epithelialized
323 colonies). As the episomal vectors used in this study exhibit only transient expression
324 (approximately 7 days) and were integrated in half of our established clones, introducing TET1
325 as synthetic RNA multiple times during reprogramming may enhance the efficiency of hiPSC
326 clones with attenuated heterogeneities. Currently, we are working on obtaining female T-iPSCs,
327 which may exhibit equivalent capabilities to the male 1T47 clone, as this would allow further
328 investigations related to female cell-specific events (oogenesis, X-chromosome inactivation).
329 In this study, we proposed the derepression of PRC2 target genes as a mechanism underlying
330 the inconsistent differentiation potential of iPSCs. While our sample size is limited, we hope
331 our findings motivate a consortium effort to validate whether the differentiation potential of
332 iPSCs globally correlates with the DNA methylome of poised enhancers.

333 In summary, we have developed a new hiPSC induction system coupled with default
334 differentiation screening, showing the quasi-total induction of VM progenitors from T-iPSCs.
335 Our results strongly suggest that the currently proposed method for creating hiPSCs induces
336 heterogeneities due to TET1 insufficiency during and after reprogramming. Developing

strategies to avoid these heterogeneities will be critical for advancing the practical application of hiPSCs, and T-iPSCs may be a promising avenue for accomplishing this.

Competing interest statement

K.H-K. and H. Kato are inventors of patent no. WO2014-069479, which covers the procedure for T-iPSC production. The other authors declare that they have no competing interests.

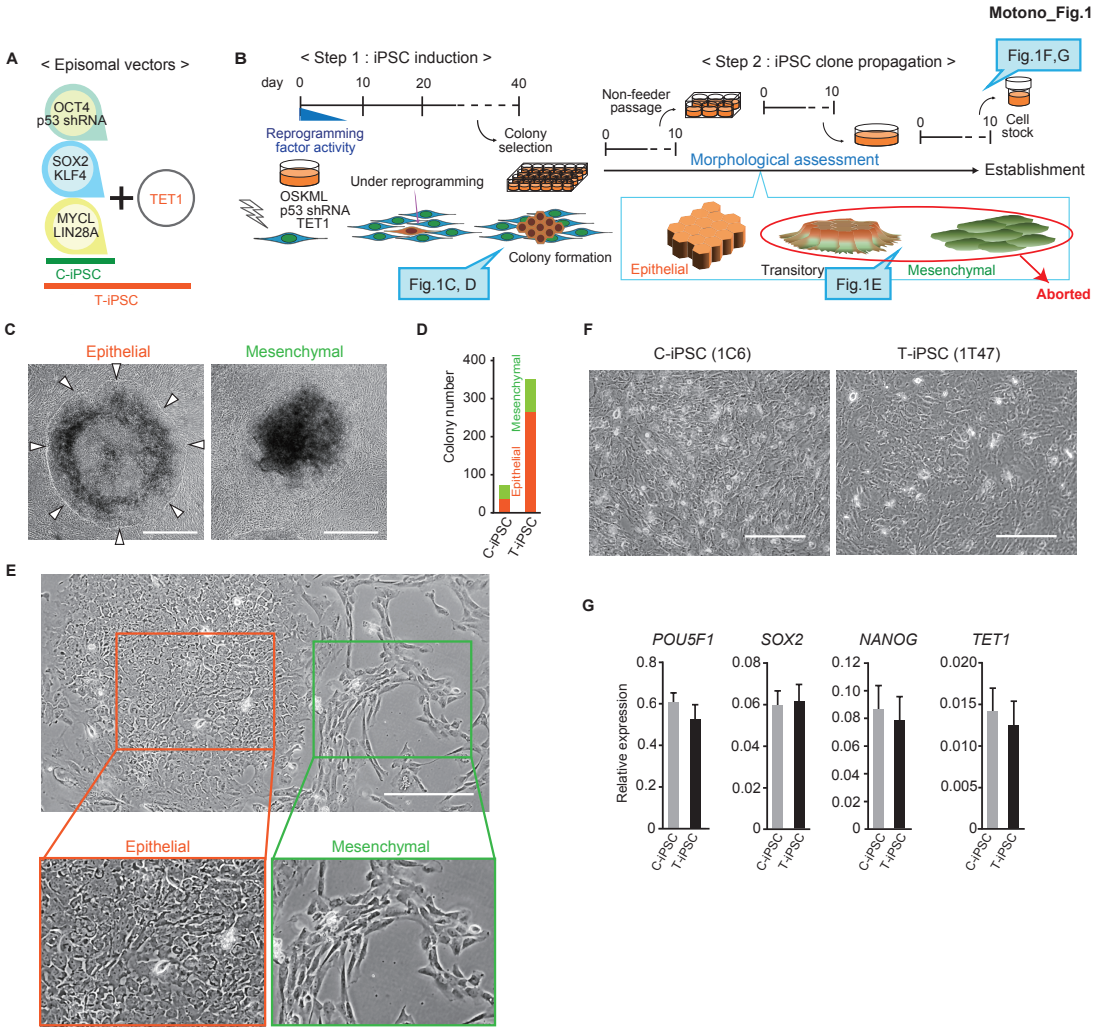
Acknowledgments

We thank Y. Moriyama, Y. Tokuzawa, S. Funayama, R. Hamada, and H. Kawakami for earlier contributions and technical assistance. This study was supported by Zeon Corporation, SATAKE MultiMix Corporation, Yokogawa Electric Corporation, SyntheticGestalt, and AMED under Grant Number JP20bk0104090. We thank Perfect C. (Edanz) and Laura J (Enago) for editing.

Author contributions

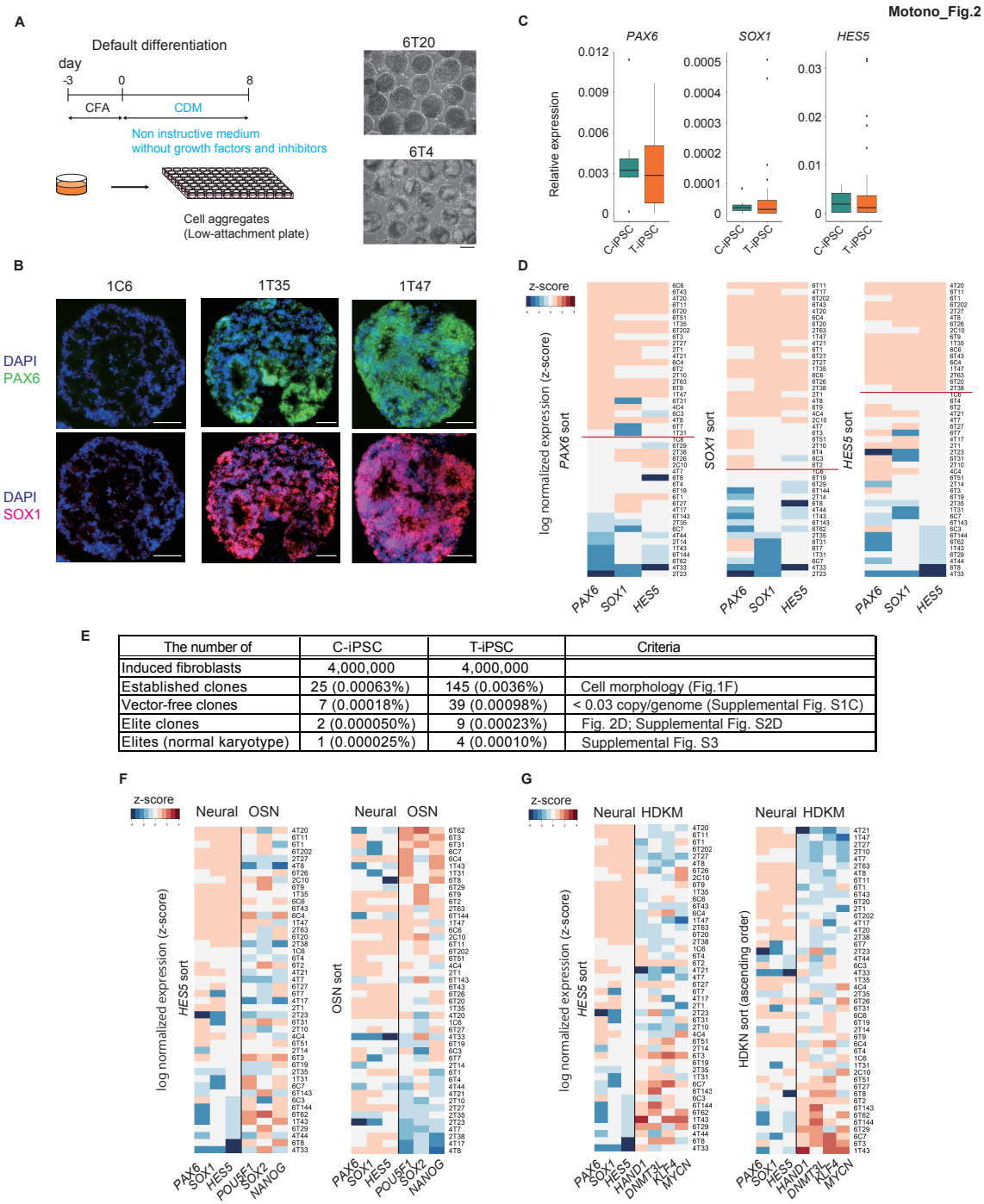
H. Kato conceived the project. M. M., K. H-K., and H. Kato wrote the original draft. M. M., K. H-K., M. K., and H. Kato performed the culture and differentiation experiments and analyzed the data. K. H-K. and M. K. analyzed the sequencing data and performed the bioinformatic analyses under the supervision of H. Kato. M. M., K. H-K., M. K., H. Kiyosawa, Y. K., and H. Kato reviewed, edited, and approved the final version of the manuscript.

358 **Figures and legends**



359

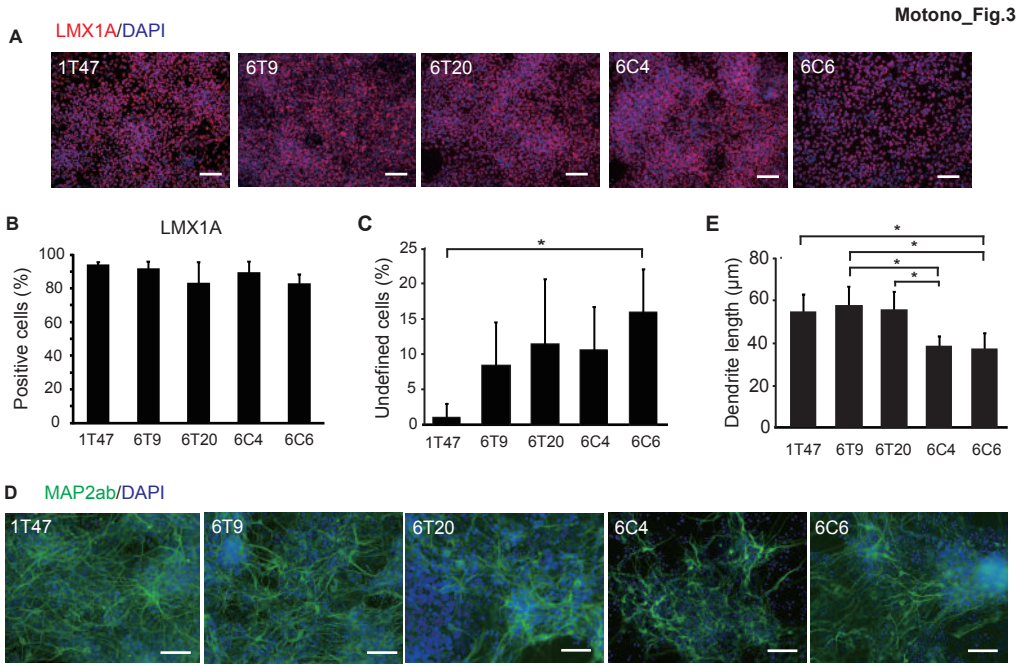
360 **Figure 1.** hiPSC-production pipelines. *(A)* Episomal vector combinations for producing C-
361 iPSCs and T-iPSCs. OCT4, an alias for POU5F1. *(B)* Timeline and procedure for hiPSC
362 production. OSKML: OCT4, SOX2, KLF4, MYCL, and LIN28A. *(C)* Characteristic
363 morphologies of emerging hiPSC colonies in Step 1. Arrowheads indicate epithelia-like colony
364 edges. Scale bar = 50 µm. *(D)* Colony numbers with epithelial or mesenchymal morphologies
365 (n = 3). *(E)* A subconfluent culture during the expansion of a colony in Step 2. Scale bar = 100
366 µm. *(F)* Established C-iPSC and T-iPSC clones. Scale bar = 50 µm. *(G)* Expression of
367 pluripotency markers and *TET1*. Data are presented as the mean ± standard deviation (C-iPSC,
368 n = 7; T-iPSC, n = 39).



The number of	C-iPSC	T-iPSC	Criteria
Induced fibroblasts	4,000,000	4,000,000	
Established clones	25 (0.00063%)	145 (0.0036%)	Cell morphology (Fig.1F)
Vector-free clones	7 (0.00018%)	39 (0.00098%)	< 0.03 copy/genome (Supplemental Fig. S1C)
Elite clones	2 (0.000050%)	9 (0.00023%)	Fig. 2D; Supplemental Fig. S2D
Elites (normal karyotype)	1 (0.000025%)	4 (0.00010%)	Supplemental Fig. S3

Figure 2. Interclonal differences in differentiation capability between hiPSC clones. *(A)* Schematic representation of the default differentiation timeline and procedure. Clones with high-differentiation capability formed solid spheres of uniform size, whereas clones with low differentiation capability had sac-like structures. Scale bar = 100 μm. *(B)* Representative immunostaining images with DAPI (blue), PAX6 (green), and SOX1 (red) on day 18 of default

375 differentiation. Scale bars = 100 μm. *(C)* Box plots showing the relative expression levels of
376 neural markers on day 8 (C-iPSC, n = 7; T-iPSC, n = 39). *(D)* Heatmap of the relative expression
377 levels of neural markers on day 8. Expression levels were normalized to *GAPDH*, log-
378 transformed, and z-score scaled. Lines indicate tentative thresholds between default-elite and
379 nonelite clones. The rows are sorted by the expression levels of each neural marker. *(E)*
380 Summary of the number of cells (ratio) meeting our tentative criteria in each iPSC production
381 process. *(F)* Heatmap showing relative expression of neural and pluripotency-related markers
382 on day 8. The rows are sorted by the expression levels of *HES5* (left) and the expression levels
383 of pluripotency-related markers (right). O, *OCT4*; S, *SOX2*; and N, *NANOG*. *(G)* Heatmap
384 showing relative expression of neural markers and genes negatively correlated with neural
385 markers on day 8. The rows are sorted by the expression levels of *HES5* (left) and by the
386 expression levels of genes negatively correlated with neural markers (right). H, *HAND1*; D,
387 *DNMT3L*; K, *KLF4*; M, *MYCN*.
388



389

390 **Figure 3.** Rationale-based induction of ventral midbrain (VM) progenitors. *(A)*
391 Immunostaining of the hiPSC-derived VM progenitor marker LMX1A (red) with DAPI (blue)
392 on day 19 for default-elite clones. Scale bars = 50 μm. *(B)* Ratio of LMX1A in each clone. *(C)*
393 Ratio of undefined cells (related immunostainings in Supplemental Fig. S4C). *(D)*
394 Immunostaining images with DAPI (blue) and MAP2ab (green) on day 29. Scale bars = 50 μm.
395 *(E)* Lengths of MAP2ab-positive dendrites. Values are presented as mean ± standard deviation
396 (n = 3–4). *$P < 0.05$; One-way analysis of variance and Tukey's multiple-comparison tests.

397

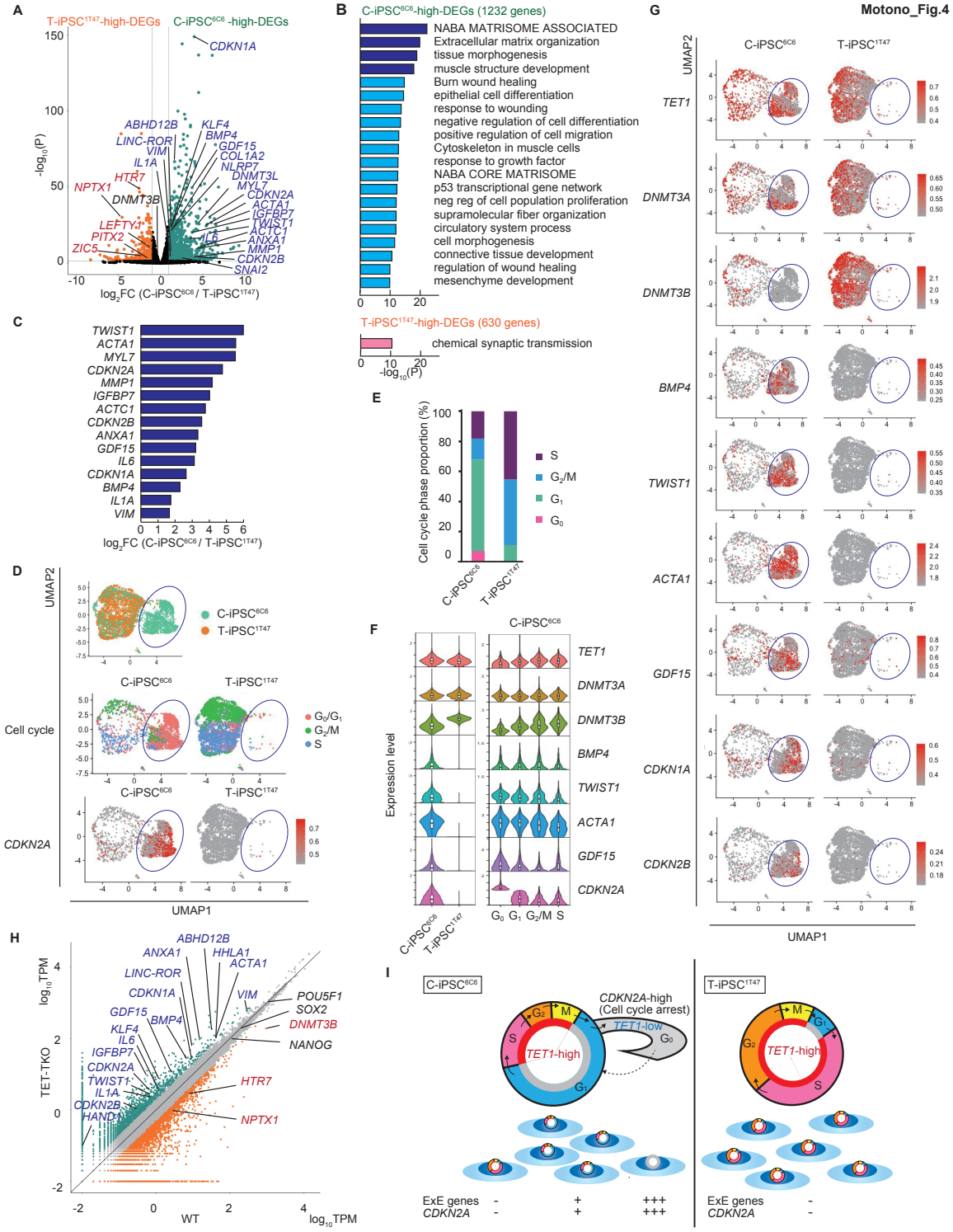


Figure 4. Characterization of interclonal and intraclonal heterogeneities in C-iPSCs. *(A)* DEGs between C-iPSC^{6C6} and T-iPSC^{1T47}. Green and orange dots indicate genes with a $\log_2$(fold change (FC)) of >1.0 (adjusted *P* value < 0.05). *(B)* GO terms of the DEG sets between C-iPSC^{6C6} and T-iPSC^{1T47}. GO terms with $-\log_{10}$(P) >10 are listed. *(C)* Bar plots of relative gene

403 expression of mesenchyme-associated and aging-associated C-iPSC^{6C6}-high-DEGs. *(D)*
404 Uniform Manifold Approximation and Projection (UMAP) plot showing C-iPSC^{6C6} and T-
405 iPSC^{1T47} scRNA-seq data, showing the distribution (upper panel) and cell cycle phase
406 distribution (middle panel) as well as *CDKN2A* gene expression (lower panel). Blue circles
407 indicate the C-iPSC^{6C6}-specific population. *(E)* Proportions of cell cycle phases in C-iPSC^{6C6}
408 and T-iPSC^{1T47}. *(F)* Violin plots showing the expression levels of the regulatory genes for DNA
409 methylation, extraembryonic-associated genes, and aging-associated genes. Gene expression
410 levels in C-iPSC^{6C6} and T-iPSC^{1T47} (left) and at each cell cycle phase in C-iPSC^{6C6} (right). *(G)*
411 Single-cell expression of DNA methylation-, extraembryonic- and aging-associated genes in
412 C-iPSC^{6C6} (left) and T-iPSC^{1T47} (right). Blue circles indicate the C-iPSC^{6C6}-specific population.
413 *(H)* Comparison of gene expression between wild-type hESC HUES8 (WT) and TET1/2/3
414 triple-knockout cells (TET-TKO). Gene expression was plotted as $\log_{10}$(TPM) in WT and TET-
415 TKO. TPM: Transcripts per million. *(I)* The cell cycle affects *TET1* and extraembryonic (ExE)
416 gene expression. C-iPSC^{6C6} exhibits a pronounced cell cycle arrest in the G_0/G_1 phase, whereas
417 T-iPSC^{1T47} shuttles between S and G_2/M. G_0/G_1 cells express lower levels of *TET1* (*TET1*-
418 low/*CDKN2A*-high) and are the source of ExE gene expression.
419

Figure 5. Hypermethylation of poised enhancers on extraembryonic gene loci. *(A)* WGBS (5-mC) and ChIP-seq data of extraembryonic genes. DMRs between C-iPSC^{6C6} and T-iPSC^{1T47} are indicated by red rectangles. WT, hESC HUES8; TET-TKO, TET1/2/3-triple-knockout HUES8. *(B)* Bar plots showing the distribution of hyper-DMRs in the genomes of C-iPSC^{6C6}

425 and T-iPSC^{1T47}. PLS; promoter-like signature. *(C)* The proportion of high- and low-DEGs
426 associated with DMRs between C-iPSC^{6C6} and T-iPSC^{1T47}. (P < 1e-100; chi square test). *(D)*
427 Proportions of hyper-DMRs overlapping with histone modifications in hESC H1. *(E)* Heatmap
428 of ChIP-seq signals of histone modifications (hESC H1) in C-iPSC^{6C6}-hyper-DMRs associated
429 with the DEGs between C-iPSC^{6C6} and T-iPSC^{1T47}. The rows are sorted in the order of
430 H3K27me3-densities. *(F)* Bar plots showing the number of C-iPSC^{6C6}-hyper-DMRs associated
431 with C-iPSC^{6C6}-high-DEGs on *cis*-regulatory elements. *(G)* Bar plots of selected common gene
432 expression levels to C-iPSC^{6C6}-high-DEGs and EZH2-KO-high-DEGs. WT, hESC H9; EZH2-
433 KO, EZH2-knockout H9. *(H)* Common high-DEGs across C-iPSC^{6C6}, TET-TKO, and EZH2-
434 KO. *(I)* GO terms of common high-DEGs across C-iPSC^{6C6}, TET-TKO, and EZH2-KO. *(J)*
435 Heatmap of the DNA methylation profile at poised enhancers during human reprogramming
436 (Cacchiarelli et al. 2015). In the comparison between hiF-T and hiPSC-T, hiPSC-hyper-DMRs
437 and hiPSC-hypo-DMRs are indicated in pink and blue, respectively. hiF-T, TERT-
438 immortalized fibroblast; hiPSC-T, reprogrammed hiPSC. ***P < 1e-100, **P < 1e-20, and *P
439 <0.01; chi square test.
440
441
442

443 **Materials and methods**

444 *hiPSC induction and culture*

445 Human fibroblasts (from male neonatal foreskin, 106-05N; female fetal fibroblasts, 106-05F)
446 were purchased from Cell Applications Inc., and cultured in Dulbecco's modified Eagle's
447 medium (DMEM; Nacalai Tesque) supplemented with 10% fetal bovine serum (FBS; Hyclone),
448 100 μg/mL L-ascorbic acid 2-phosphate trisodium salt (L-AA; FUJIFILM Wako), and 1×
449 penicillin–streptomycin (P/S; Gibco). hiPSCs were induced using episomal vectors
450 (Supplemental Table S4) (Okita et al. 2011). To establish T-iPSCs, a TET1 episomal vector
451 (comprising complementary DNA encoding the human *TET1* open-reading frame, cloned into
452 the pCXLE vector) was used with the three episomal vectors upon reprogramming. One μg of
453 each vector was transfected into approximately 1×10^6 low passage fibroblasts using Human
454 Dermal Fibroblast Nucleofection Kit and Nucleofector Transfection 2b Device (Amaxa
455 Biosystems [Lonza]), according to the manufacturer's instructions, and the cells were plated in
456 a human fibroblast medium. The next day, the culture medium was replaced with Essential 8
457 Medium (E8M; Gibco) supplemented with 100 nM hydrocortisone (Sigma-Aldrich) and 1 ×
458 P/S. E8M, which contains TGF-β, was selected because of its potentiating effect of activating
459 SMAD2/3 during reprogramming (Ruetz et al. 2017). The medium was changed every other
460 day. E8M without hydrocortisone was used from day 10 of iPSC induction. iPSC colonies were
461 picked up, triturated into E8M supplemented with 10 mM Y-27632 (FUJIFILM Wako), and
462 plated onto STO feeder cells (RIKEN BRC, RCB0536; RRID: CVCL_3420) treated with
463 mitomycin C (FUJIFILM Wako). The culture medium was replaced the following day with
464 KnockOut DMEM/F12 (Thermo Fisher Scientific) supplemented with 20% KnockOut Serum
465 Replacement (KSR), 2 mM GlutaMAX^{TM}-I (Glu), 1 × MEM Non-Essential Amino Acids
466 Solution (NEAA) (Gibco), 0.1 mM 2-mercaptoethanol (2-ME), 5 mM NaOH (FUJIFILM
467 Wako), 5 ng/mL fibroblast growth factor 2 (FGF2; Peprotech), and 1 × P/S. After the colonies

468 had been picked, the medium was changed daily. Grown iPSC colonies were dissociated with
469 0.01% trypsin/0.1 mM EDTA buffer (trypsin/EDTA; Nacalai Tesque) and plated onto 5 μg/mL
470 vitronectin (Invitrogen)-coated 6-well plates in KnockOut DMEM/F12 supplemented with 20%
471 KSR, 2 mM Glu, 1 × NEAA, 0.1 mM 2-ME, 15 ng/mL FGF2, 10 ng/mL Activin A (R&D
472 Systems), 5 mM NaOH, and 1 × P/S (KFA medium) supplemented with 10 mM Y-27632. The
473 culture medium was replaced daily with KFA medium without Y-27632. KFA medium was
474 routinely used throughout this study to expand and maintain hiPSCs. Clones that reached >90%
475 confluency in a 100-mm dish (Becton Dickinson) without a mesenchymal morphology were
476 considered "established." One-half of the cells were cryopreserved in a cell stock solution (10%
477 dimethyl sulfoxide [Nacalai Tesque] and 90% FBS). Genomic DNA (gDNA) was obtained from
478 the remaining cells to analyze the genomic integration of the plasmid vectors.
479
480 *Setting a threshold for the genomic integration of episomal vectors*
481 gDNA was extracted using DNeasy Blood & Tissue Kit (QIAGEN). Conventional PCR was
482 performed using Ex Taq DNA polymerase (TaKaRa). Quantitative PCR (qPCR) was performed
483 using Thunderbird SYBR qPCR Mix (Toyobo) and CFX Connect Real-Time System (Bio-Rad).
484 The cycle threshold (Ct) values of the episomal vector were normalized against gDNA
485 measured using the CFTR primer set (Supplemental Table S4). The pEP4-SF1 and pEP4-SR1
486 primers were used to detect episomal vector backbones for conventional PCR (International
487 Stem Cell et al. 2007). For qPCR, the pEP4-SR1 primer was replaced with the pEP4-SR3 primer,
488 which was designed in-house. The template for conventional PCR and qPCR was 50 and 20 ng
489 gDNA, respectively. gDNA extracted from male human dermal fibroblasts was used as a
490 negative control.
491
492 *Differentiation of hiPSCs by default differentiation*

493 Pilot experiments using hiPSCs and attempts to derive neural cells from monkey ESCs (Ono et
494 al. 2014) indicated that the KSR-containing culture media are incompatible with default
495 differentiation. Therefore, we gradually changed KSR medium to a chemically defined medium
496 (CDM) before differentiation to facilitate the efficient derivation of neural cells (Ono et al.
497 2014). KFA-cultured hiPSCs were dissociated with trypsin/EDTA, and 1 × 10^6 cells were
498 replated onto a vitronectin-coated 100-mm dish in CDM supplemented with 10 ng/mL FGF2,
499 2 ng/mL Activin A, 3% KSR, and 10 μM Y-27632 (CFA). CDM is composed of Iscove's
500 modified Dulbecco's medium/F-12 (Thermo Fisher Scientific), 2 mM Glu, 7 μg/mL Insulin
501 (FUJIFILM Wako), 15 μg/mL transferrin, 450 μM monothioglycerol, 5 mg/mL bovine serum
502 albumin fraction V, 80 nM sodium selenite (Sigma-Aldrich), 200 μM L-AA, 1 × lipid
503 concentrate (Gibco), 1 × NEAA, and 1 × P/S. One-half of the media was carefully replaced
504 with CFA without Y-27632 daily. On day 3, the cells were dissociated with trypsin/EDTA and
505 plated at 500 cells per well in a 96-well low-binding Nunclon Sphera Plate (Thermo Fisher
506 Scientific) in 150 μL of CDM supplemented with 10 mM Y-27632. The day on which the cells
507 were seeded was designated differentiation day 0. Two days after seeding, 100 μL of the
508 medium was removed to avoid aspiration of the spheres, and 150 μL of fresh CDM without Y-
509 27632 was added. On differentiation day 5, 100 μL of the medium was removed, and 100 μL
510 of fresh CDM was added. The medium was changed every 3 days using the same procedure.
511 The induced aggregates were collected and analyzed on differentiation days 8 and 18.
512
513 *Differentiation of ventral midbrain neural progenitor cells*
514 A step-by-step protocol can be found in Supplemental materials and methods. We optimized
515 the timing of adding the reagents mainly involved in WNT signaling because WNT activity
516 initially inhibits neurogenesis but is imperative for coordinating the anterior–posterior axis of
517 the midbrain. Briefly, 100 nM LDN193189 was added from days 0 to 11, 10 μM SB431542

518 from days 0 to 4, and 10 nM WNT-C59 (C59) from days 0 to 2. Ventralization was initiated by
519 adding 100 ng/mL recombinant sonic hedgehog (C24II) (R&D Systems) and 2 μM
520 purmorphamine (FUJIFILM Wako) from days 3 to 8. Furthermore, 3 μM CHIR99021 (Sigma-
521 Aldrich) was added from days 3 to 12. From differentiation day 13, neural maturation was
522 initiated by adding 10 ng/mL brain-derived neurotrophic factor, 10 ng/mL glial-derived
523 neurotrophic factor (FUJIFILM Wako), 1 ng/mL TGF-β3 (R&D Systems), 200 μM L-AA, 200
524 μM *N*6,2-*O*-dibutyryladenosine 3, 5-cyclic monophosphate (Sigma-Aldrich), and 5 μM DAPT
525 (Tocris Bioscience). The induced cells were dissociated with Accumax (Innovative Cell
526 Technologies, Inc.) on day 16. Then, 4×10^5 cells/cm^2 were seeded onto culture slides (BD
527 Biosciences) coated with 15 μg/mL poly-L-ornithine, 2 μg/mL fibronectin (Sigma-Aldrich),
528 and 0.5 μg/cm^2 laminin. The medium was changed every day from days 0 to 15, and one-half
529 was changed every 2–3 days from day 16 onward.
530

531 *Immunohistochemical analysis*

532 Default differentiated spheres were washed with phosphate-buffered saline (PBS; TaKaRa),
533 fixed in 4% paraformaldehyde (PFA; FUJIFILM Wako), embedded in OCT compound,
534 cryosectioned at 10 μm using a cryostat (Leica Biosystems), and collected onto glass slides.
535 Undifferentiated hiPSCs and ventral midbrain precursors were cultured on glass chamber slides.
536 The slides were fixed in 4% PFA and blocked in 5% (v/v) normal donkey serum
537 (Millipore)/0.3% (v/v) Triton X-100 (Nacalai Tesque)/PBS and then incubated overnight with
538 primary antibodies (Supplemental Table S5). The slides were washed and incubated with
539 secondary antibodies and DAPI (Invitrogen) (Supplemental Table S5). Finally, the slides were

540 washed with 0.05% Tween-20/PBS and mounted with an anti-photobleaching mounting
541 medium. The stained cells were imaged using a Biorevo BZ-9000 fluorescence microscope
542 (Keyence).
543
544 *Reverse transcription-qPCR analysis*
545 Total RNA was purified from the samples using RNeasy Plus Mini Kit (QIAGEN). More than
546 50 ng of total RNA was reverse transcribed using ReverTra Ace and oligo-dT primers (Toyobo).
547 qPCR analysis was performed using Thunderbird SYBR qPCR Mix and CFX Connect Real-
548 Time System. Data were assessed using the delta-Ct method and normalized to *GAPDH*
549 expression. Error bars represent the standard deviation of biological replicates. Primer
550 sequences used for qPCR are listed in Supplemental Table S4. Statistical significance was
551 determined using Student's unpaired *t*-test or one-way analysis of variance and Tukey's
552 multiple-comparison tests using R software. $P < 0.05$ indicated statistically significant
553 differences.
554
555 *Evaluating standards for default-elite clones*
556 Considering the conducted correlation analyses of differentiation markers and differentiation
557 efficiencies between 1C6, 1T35, and 1T47, we set our default-elite thresholds to *HES5*/*GAPDH*
558 >0.002 and *TBXT*/*GAPDH* <0.0001 detected by qPCR.
559
560 *RNA-seq processing*
561 RNA-seq data analyses were performed for T-iPSC^{1T47} and C-iPSC^{6C6} cultured under an
562 undifferentiated culture condition (KFA medium). Total RNA was purified using RNeasy Plus
563 Mini Kit, according to the manufacturer's instructions. Then, library preparation and

564 sequencing were performed by Macrogen Japan Corp. Briefly, RNA-seq libraries were prepared
565 using TruSeq Stranded Total RNA LT Sample Prep Kit with Ribo-Zero Human/Mouse/Rat
566 (Illumina). By harvesting cells at different passage numbers, we generated biological triplicate
567 data for T-iPSC^{1T47} and C-iPSC^{6C6}. RNA-seq libraries were sequenced on an Illumina NovaSeq
568 6000, generating 100-base paired-end reads.
569
570 *RNA-seq data analysis*
571 The software and resource data are listed in Supplemental Table S6. RNA-seq reads were
572 aligned to the human reference genome (GRCh38/hg38) using STAR software with default
573 parameters. We calculated transcripts per million mapped reads using RSEM with the default
574 parameters and considered their expression levels. The DEGs between T-iPSC^{1T47} and C-
575 iPSC^{6C6} and between EZH2-KO and WT hESCs were calculated using the R package DESeq2
576 with the cutoffs of absolute $\log_2$(FC) >1 and adjusted *P* value of <0.05. The DEGs between
577 TET-TKO and WT hESCs were calculated using a cutoff of absolute $\log_2$(FC) >1. We only
578 considered autosomal chromosomes for analyzing RNA-seq data.
579
580 *Isolation and RNA-seq of single cells*
581 scRNA-seq was performed for T-iPSC^{1T47} and C-iPSC^{6C6} cultured in an undifferentiated culture
582 condition (KFA medium). Azenta Life Sciences performed library preparation and sequencing.
583 Briefly, scRNA-seq was performed using the 10 × Chromium platform and Chromium Single-
584 Cell 3′ Reagent Kits (v3 Chemistry), according to the manufacturer's protocol (10 × Genomics).
585 Libraries were sequenced on DNBSEQ-G400 (MGI) with paired-end sequencing.
586
587 *scRNA-seq data analysis*
588 Alignment of reads to the human reference genome (GRCh38/hg38) and preparation of the raw

589 gene expression matrix were conducted using the Cell Ranger pipeline. scRNA-seq integration
590 analysis was performed using the Seurat package. C-iPSC^{6C6} cells and T-iPSC^{1T47} cells were
591 excluded if they contained a portion of mitochondrial reads >20% or had <200 genes detected.
592 The resulting matrix contained 2119 C-iPSC^{6C6} and 3249 T-iPSC^{1T47}. Unique molecular
593 identifier counts were normalized, followed by variable gene identification, anchor
594 identification, data integration, and data scaling with the default settings of the Seurat package.
595 Cell cycle phases (G_0/G_1, S, and G_2/M) were determined using Seurat's default settings. Cells
596 in the G_0 phase were selected from the G_0/G_1 phase subset based on *CDKN2A* expression (>1).
597 For Uniform Manifold Approximation and Projection (UMAP), we used the first 15 principal
598 component analysis dimensions with the default settings using the RunUMAP function. Gene
599 expression profiles were visualized on a UMAP plot using the FeaturePlot function (min.cutoff
600 = q50, max.cutoff = q80). For clustering, we used a resolution of 0.5 with default settings.
601 To reanalyze scRNA-seq data from peri-gastruloids, we integrated scRNA-seq data from T-
602 iPSC^{1T47}, C-iPSC^{6C6}, and hESC H9. Cells were excluded if they had <200 genes detected. We
603 randomly selected 200 cells each from the T-iPSC^{1T47}, C-iPSC^{6C6}, and primed hESC H9 datasets.
604 The resulting matrix contained 200 cells (C-iPSC^{6C6}), 200 cells (T-iPSC^{1T47}), 200 cells (hESC
605 H9), and 12,008 cells (peri-gastruloids). Unique molecular identifier counts were normalized,
606 followed by variable gene identification with default settings of the Seurat package. For
607 integration with peri-gastruloids, merged hiPSCs, and hESCs in the primed culture condition,
608 we identified anchors using the default settings, except for anchor.features = 1500 and k.filter
609 = 10. Integration and data scaling were performed with the default settings. For *t*-distributed
610 stochastic neighbor embedding, we used the first 14 principal component analysis dimensions
611 with the default settings using the RunTSNE function. For clustering, we used a resolution of
612 0.7 with default settings using the FindClusters function and identified tissue-specific markers
613 expressed in the clusters using the FindAllMarkers function with default settings. The scRNA-

614 seq data of the human reprogramming process (Liu et al. 2020) was reanalyzed under the
615 conditions described in the paper using Seurat. Cultured monkey blastocysts (Yang et al. 2021)
616 were reanalyzed using interactive online tools.
617
618 *WGBS processing*
619 gDNA was extracted using DNeasy Blood & Tissue Kit, according to the manufacturer's
620 instructions. Library preparation and sequencing were performed by Macrogen Japan Corp.
621 Briefly, bisulfite conversion was performed using EZ DNA Methylation-Gold Kit (ZYMO
622 RESEARCH), according to the manufacturer's instructions. The eluted DNA was processed
623 using Accel-NGS Methyl-Seq DNA Library Kit (IDT), following the manufacturer's
624 recommendations. The libraries were sequenced on Illumina NovaSeq 6000, generating 150-
625 base paired-end reads. The WGBS samples had a ≥99.5% sodium bisulfite conversion rate.
626
627 *WGBS data analysis*
628 Raw sequencing reads were trimmed using Trim Galore with the recommended conditions for
629 Accel-NGS Methyl-Seq DNA Library Kit and aligned to the human reference genome
630 (GRCh38/hg38) using Bismark with default parameters. Bismark eliminated duplicated reads,
631 and methylation calls were made using the Bismark methylation extractor with a minimum
632 sequencing depth = 4 at the CpG site. CpG coverage against GRCh38 was 88% for T-iPSC^{1T47}
633 and 87% for C-iPSC^{6C6}. DMRs between T-iPSC^{1T47} and C-iPSC^{6C6} were identified using the R
634 package DSS (≥10% difference of CpG methylation, p.threshold = 0.02, minCG = 15, and
635 minlen = 300). The averaged methylation levels of CpGs in each enhancer were calculated
636 using bedtools. Hypermethylated poised enhancers in C-iPSC^{6C6} compared with fibroblasts
637 were identified using the following criteria: ≥10% higher CpG methylation in C-iPSC^{6C6} and
638 <5% higher CpG methylation in T-iPSC^{1T47}, both compared with fibroblasts. We only

639 considered autosomal chromosomes for DNA methylation analysis.

640

641 *Characterization of the DMRs*

642 BED files for the promoter-like signature and enhancer-like signature were downloaded from
643 SCREEN (ENCODE3) (Consortium et al. 2020). Promoter-associated DMRs and enhancer-
644 associated DMRs were identified by intersecting with promoter-like signature and enhancer-
645 like signature using bedtools, respectively. DMRs within the gene bodies were identified by
646 intersecting nonpromoter-associated DMRs with GENCODE V45 genes. DMRs overlapping
647 the histone modifications in hESC H1 were identified using bedtools and deeptools.

648

649 *GO analysis of DEGs and DMR neighboring genes*

650 DEGs were analyzed using Metascape for GO analysis. GREAT was used to find the genes near
651 DMRs and to perform GO analysis. GREAT was run under the "single nearest gene" association
652 strategy. The DMRs that associated with DEGs were selected based on the genes nearest to the
653 DMRs, and DEGs were identified between C-iPSC^{6C6} and T-iPSC^{1T47}.

654

655 *Statistical analysis of the characteristics of DMR*

656 We used shuffled control bed files as a control for statistical analysis. We shuffled the location
657 of the DMRs on the same chromosome using the "shuffle" tool from bedtools and recalculated
658 the overlaps to promoter-like signature, enhancer-like signature, and gene bodies. Statistical
659 significance was determined by the chi-square test. We used these bed files to write heatmaps
660 of histone modification for comparison.

661

662 *Data availability*

663 All data supporting the conclusions are present in the paper or Supplemental Material. The

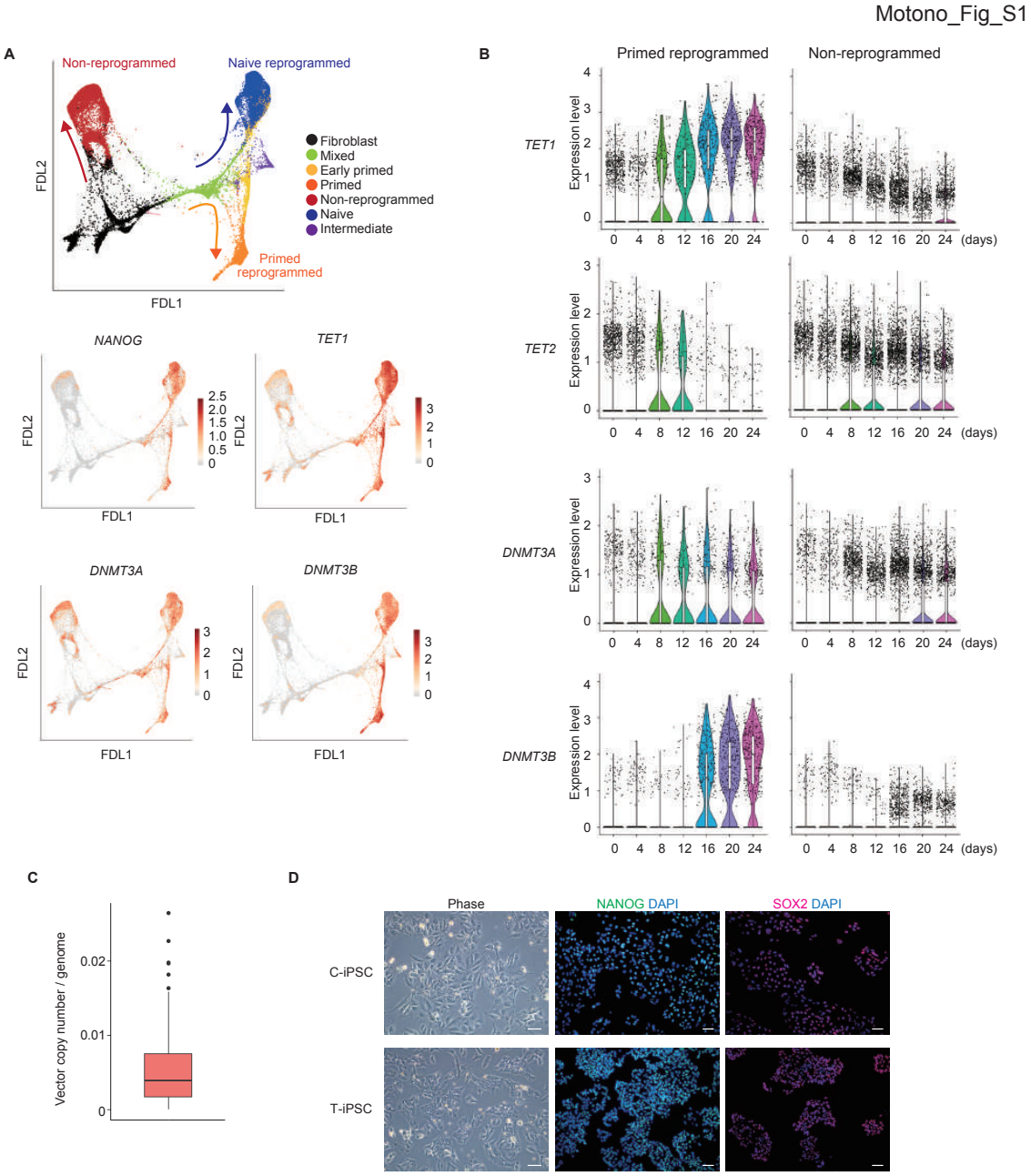
664 sequencing data generated in this study are available through the DNA Data Bank of Japan,

665 DRA013769 (scRNA-seq), DRA014828 (bulk RNA-seq), and DRA017148 (WGBS).

666

667 **Supplemental Figures**

668



669

670 **Supplemental Figure S1**, related to Figure 1. Gene expression profiles of *NANOG*, *TETs*, and
671 *DNMT3*s during naïve and primed reprogramming from human fibroblasts to iPSCs. *(A)* Force-
672 directed layout (FDL) visualization of the fibroblast, naïve, and primed scRNA-seq libraries
673 (reanalyzed from Liu et al. 2020). Arrows in the upper panel show the cell trajectories to non-
674 reprogrammed (red), naïve (blue), and primed (orange) pluripotent cell states. *(B)* Violin plots

675 of *TET1*, *TET2*, *DNMT3A*, and *DNMT3B* expression during reprogramming. Expression data
676 for primed reprogrammed (left) and non-reprogrammed (right) cell trajectories were scored
677 separately. *DNMT3B*, a CpG *de novo* methylase, was induced under pluripotency acquisition
678 (compared with the *DNMT3A* expression profile). *(C)* Box plot showing copy numbers of
679 integrated reprogramming episomal vectors in vector-free clones determined by conventional
680 PCR. Reanalysis using qPCR showed that conventional PCR only detected clones with >0.16
681 plasmid copies/genome as positive. The mean copy number of conventional PCR-deduced
682 "vector-free" clones was 0.0062 ± 0.0067 per genome (mean ± standard deviation, n = 46). (*D)*
683 Phase contrast images and representative immunostaining data with NANOG (green), SOX2
684 (red), and DAPI (blue) in established C-iPSCs and T-iPSCs. Scale bars = 50 μm.
685

Supplemental Figure S2, related to Figure 2. Marker profiling of each C-iPSC and T-iPSC before and after default differentiation. *(A)* Relative expression of pluripotency-associated markers on day 0 of default differentiation. *(B)* Changes in gene expression during default differentiation. The *x* axis indicates the number of days since the start of default differentiation.

691 Day 8 of default differentiation (blue dotted lines) was the earliest reliable time to assess the
692 differentiation capabilities of various clones (n = 1). *(C)* Correlation matrix for marker
693 expression levels and default differentiation behaviors of all vector-free hiPSC clones. Colored
694 circles indicate significant correlations ($P < 0.05$) between the two markers. Red and blue
695 indicate positive and negative correlation, respectively. Color intensity and circle size indicate
696 correlation strength (R). Blue dotted lines indicate the boundary between day 0 and day 8
697 expression values. *(D)* Relative expression of neural markers, *TBXT* and *NANOG*, on day 8 of
698 default differentiation in all established clones. Each dot represents the expression level of the
699 corresponding clone. Horizontal lines in each panel indicate the tentative threshold values for
700 default differentiation capability. *(E)* Correlation between episomal vector copy number and
701 neural markers *NANOG* or *TBXT* (n = 46). A residual episomal copy number of <0.03 per
702 genome did not influence differentiation output.
703

704

Motono_Fig_S3

Cell line	Results of Karyotyping
1T35	46, XY [18] ; 47, XY, +X [2]
1T47	46, XY [20]
2T27	46, XY [5] ; 47, XY, +20 [8] ; 47, XY, +X [7]
2T63	47, XY, +X [9] ; 47, XY, +mar [7] ; 47, XY, +20 [4]
4T20	46, XY [7] ; 47, XY, +20 [13]
6T9	46, XX [20]
6T11	46, XX [6] ; 47, XX, +13 [14]
6T20	46, XX [20]
6T43	46, XX [20]
6C4	46, XX [19] ; 46, XX, -18, +mar [1]
6C6	46, XX [20]

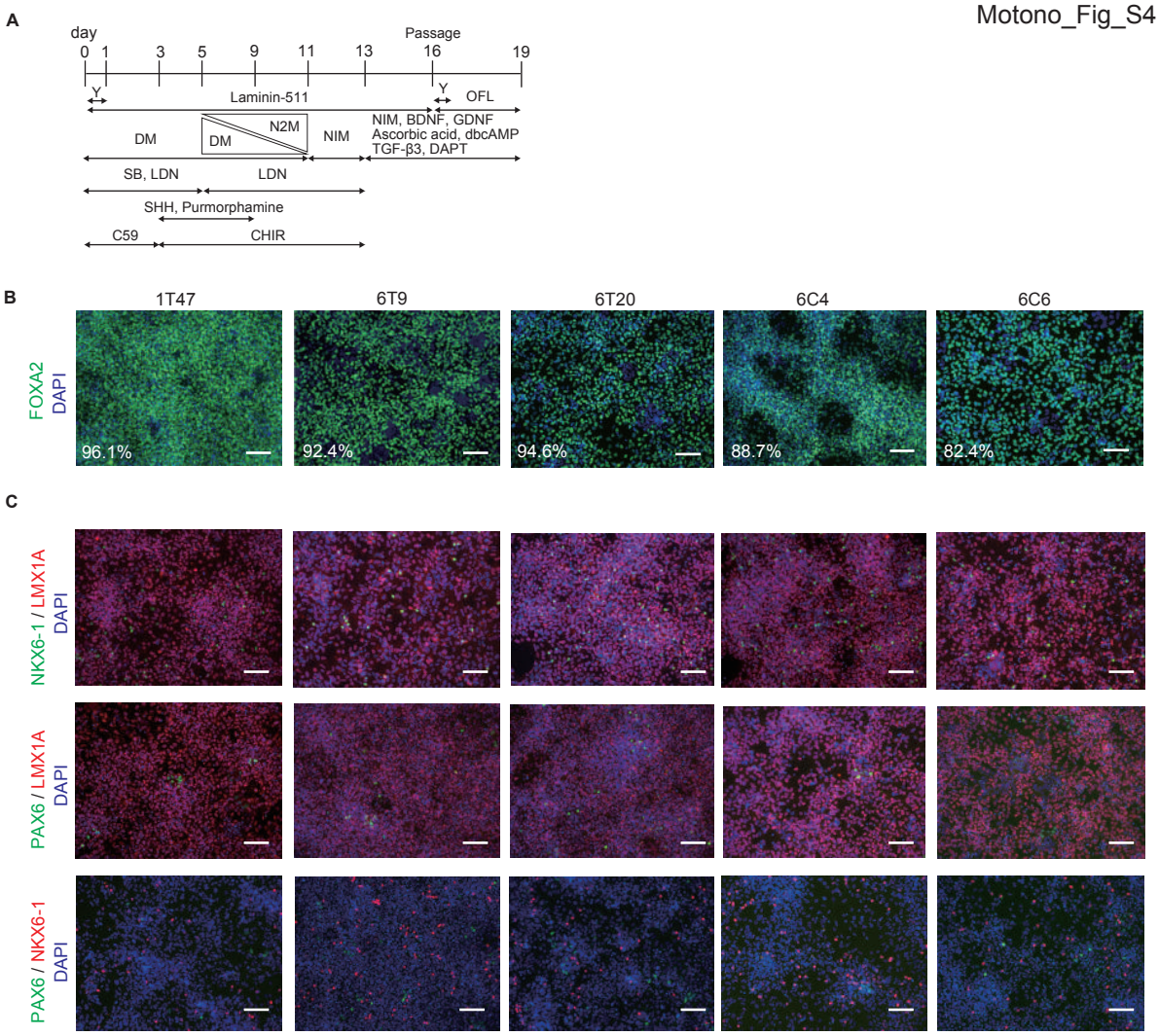
705

706 **Supplemental Figure S3**, related to Figure 2. Karyotyping of default-elite clones. Numbers in

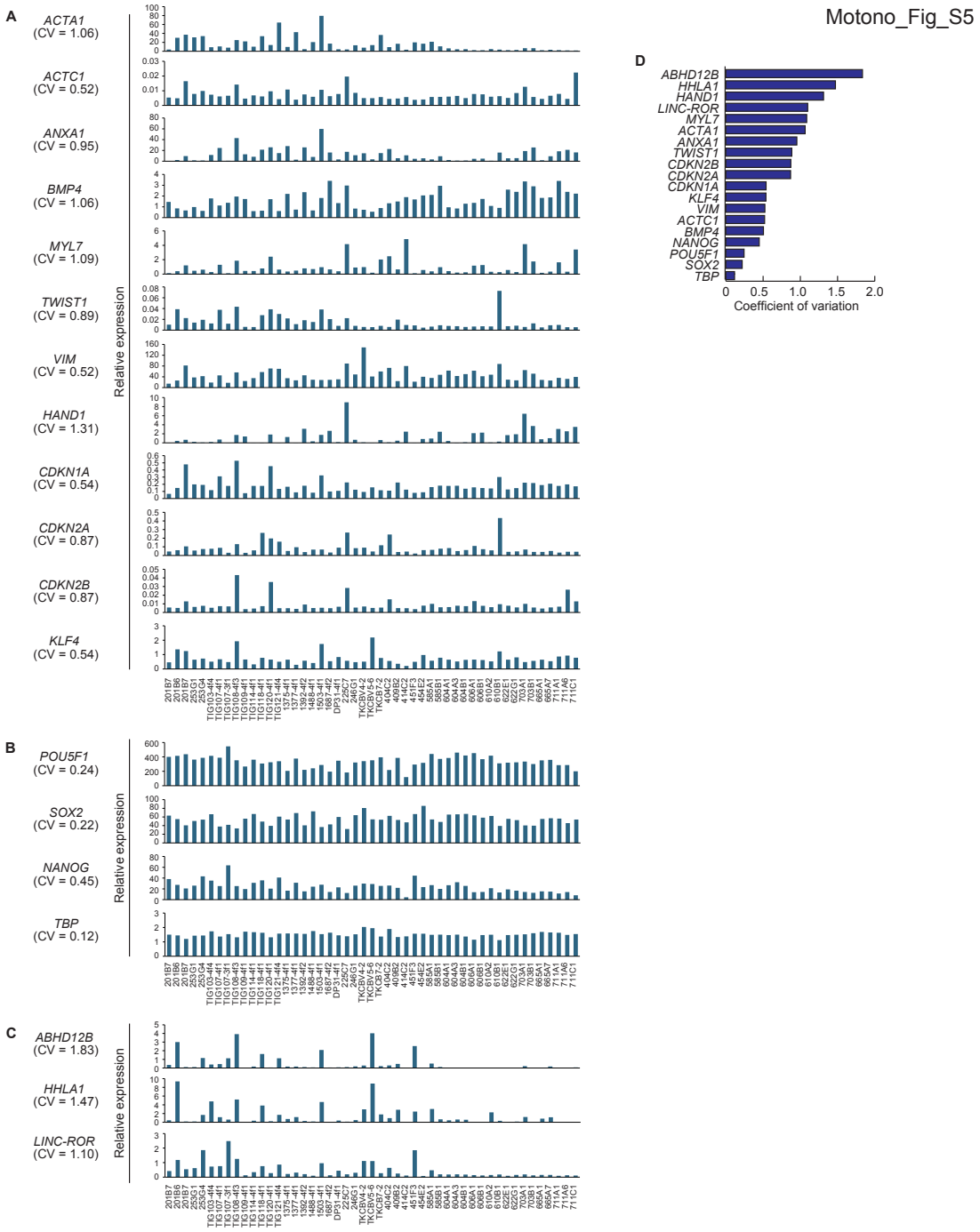
707 square brackets represent the number of observed karyotypes in the 20 nuclei counted. +mar,

708 marker chromosome of unknown origin.

709



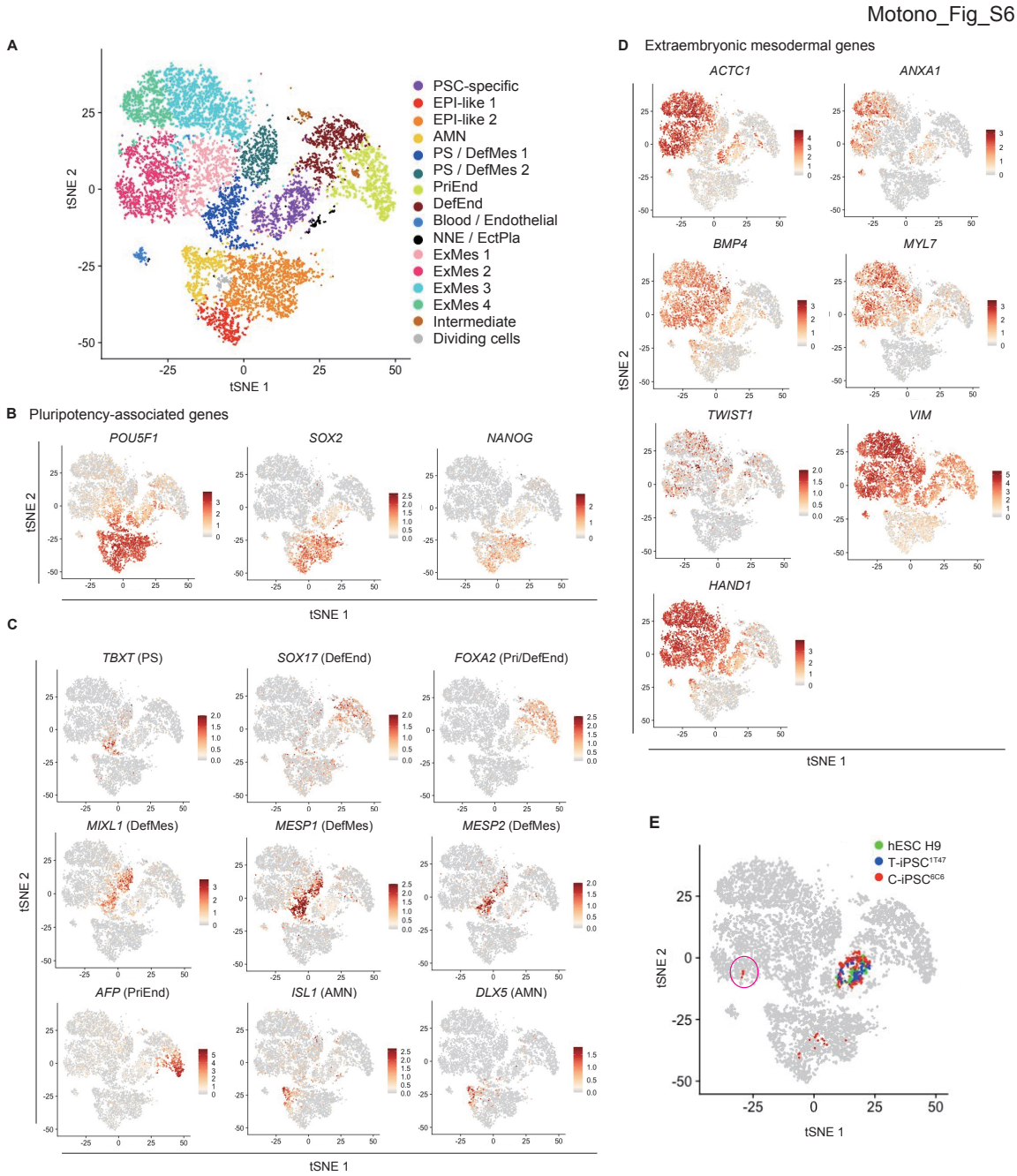
Supplemental Figure S4, related to Figure 3. Differentiation outputs of default-elite clones for VM progenitors. *(A)* Differentiation protocol for the induction of VM progenitors. Triple inhibition of the TGFβ-, BMP-, and WNT-signaling pathways was executed with SB, LDN, and C59, respectively. Y, Y-27632; SHH, sonic hedgehog; OFL, ornithine, fibronectin, and laminin-511. KSR-medium (DM) was gradually changed to N2-medium (N2M) from days 5 to 11, and neural induction medium (NIM) was changed from day 11. *(B)* Immunostaining images of the VM-region marker, FOXA2, in default-elite clones. Numbers indicate the percentage of positive cells. Blue, DAPI; green, FOXA2. Scale bars, 50 μm. *(C)* Immunostaining images of other VM-region markers in the default-elite clones. Blue, DAPI; green, NKX6-1 or PAX6; red, LMX1A or NKX6-1. Scale bars = 50 μm.



721

722 **Supplemental Figure S5**, related to Figure 4. Interclonal heterogeneity of extraembryonic
723 genes and cell cycle regulators in 49 hiPSCs (Koyanagi-Aoi et al. 2013). *(A)* Bar plots of the
724 relative expression levels of extraembryonic genes and cell cycle regulators. *(B)* Bar plots of
725 the relative expression levels of pluripotency-associated genes and the housekeeping gene *TBP*.
726 *(C)* Bar plots of the relative expression levels of previously assigned differentiation-defective

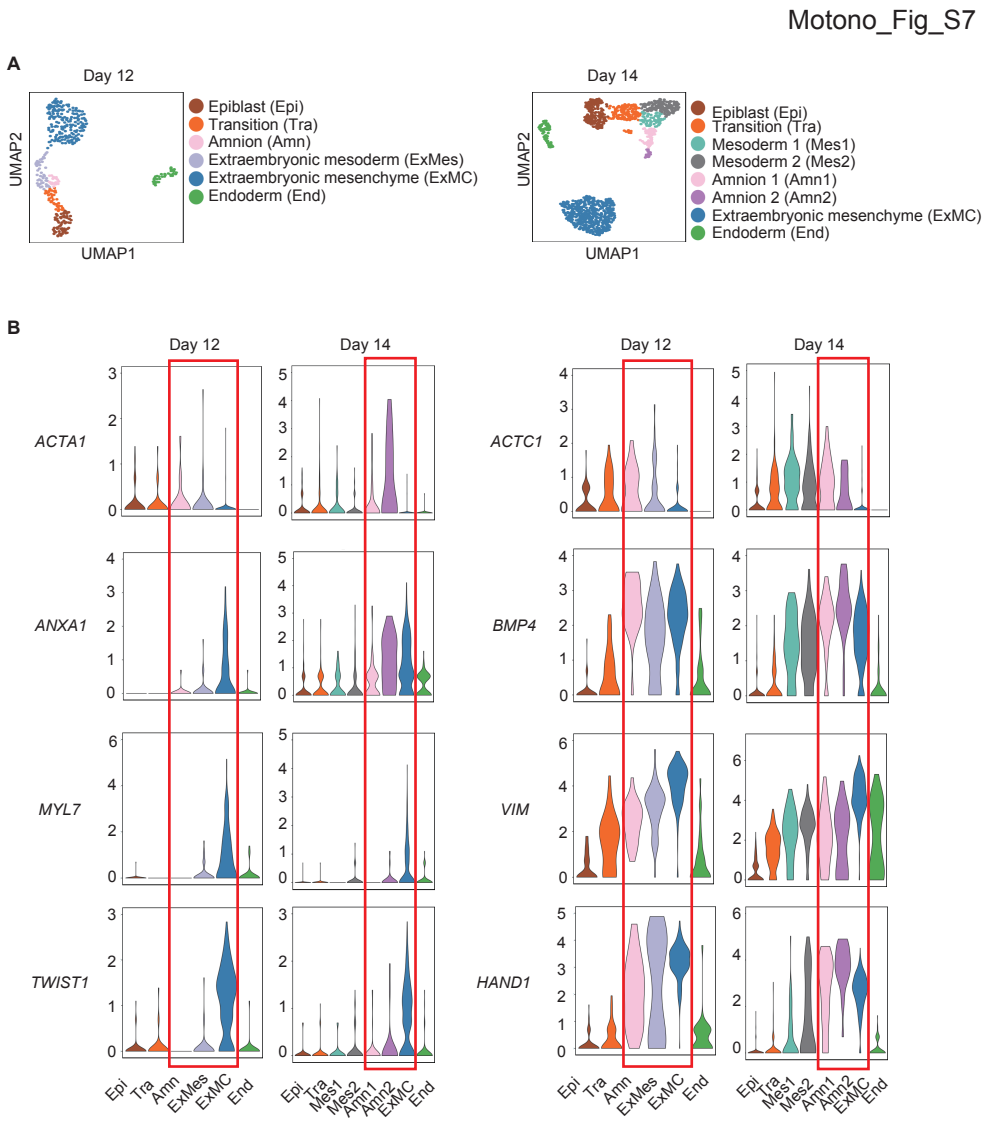
727 markers. *(D)* Bar plots of CV of gene expression for extraembryonic genes and cell cycle
728 regulators in 49 hiPSCs. CV; coefficient of variation.
729



730

731 **Supplemental Figure S6**, related to Figure 4. Characterization of mesenchymal genes
732 associated with extraembryonic lineages by the reanalysis of human peri-gastruloids. *(A)*
733 Single-cell transcriptomic integration of human peri-gastruloids (Liu et al. 2023), T-iPSC^{1T47},
734 C-iPSC^{6C6}, and hESC H9 (Messmer et al. 2019). tSNE plot of human peri-gastruloids integrated
735 with T-iPSC^{1T47}, C-iPSC^{6C6}, and hESC H9. EPI, epiblast; AMN, amnion; PS, primitive streak;
736 DefMes, definitive mesoderm; PriEnd, primitive endoderm; DefEnd, definitive endoderm;

737 NNE, non-neural ectoderm; EctPla, ectoderm placode; ExMes, extraembryonic mesoderm. *(B)*
738 tSNE plot showing pluripotency-associated gene expression. *(C)* tSNE plot showing the
739 expression of *TBXT* (PS), *MESP1/2* and *MIXL1* (DefMes), *AFP* (PriEnd), *SOX17* and *FOXA2*
740 (DefEnd), and *ISL1* and *DLX5* (AMN). *(D)* tSNE plot showing mesodermal and
741 extraembryonic gene expression. *(E)* tSNE plot showing the locations of T-iPSC^{1T47}, C-iPSC^{6C6},
742 and hESC H9.
743



744

745 **Supplemental Figure S7**, related to Figure 4. C-iPSC^{6C6}-high-DEGs are expressed in

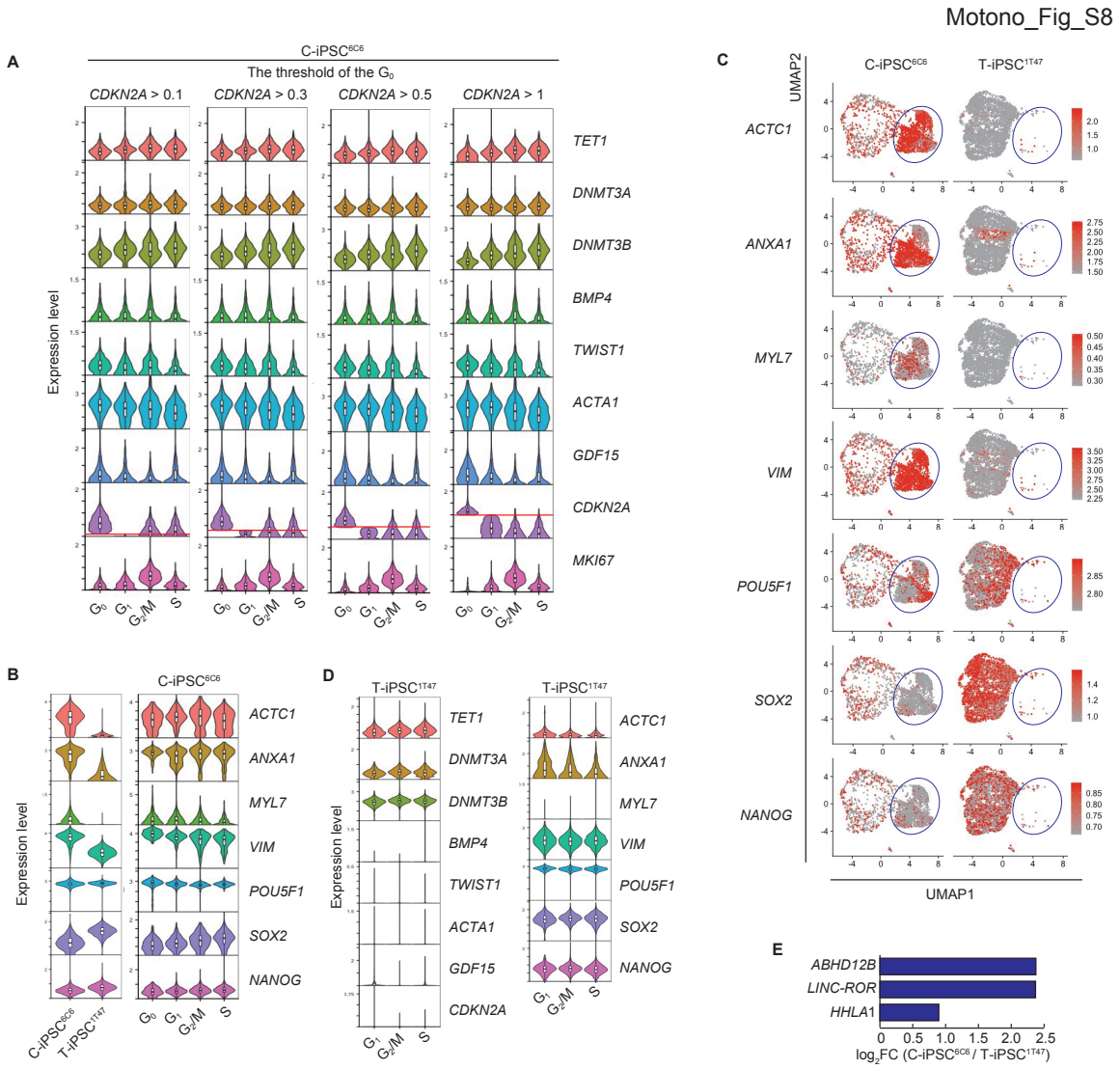
746 extraembryonic lineages during early monkey development (Yang et al. 2021). *(A)* UMAP plot

747 of cells from *in vitro* cultured monkey embryos on days 12 and 14. Day-12 monkey embryos

748 are phenotypically comparable to hPSCs. *(B)* Violin plot showing relative expression of

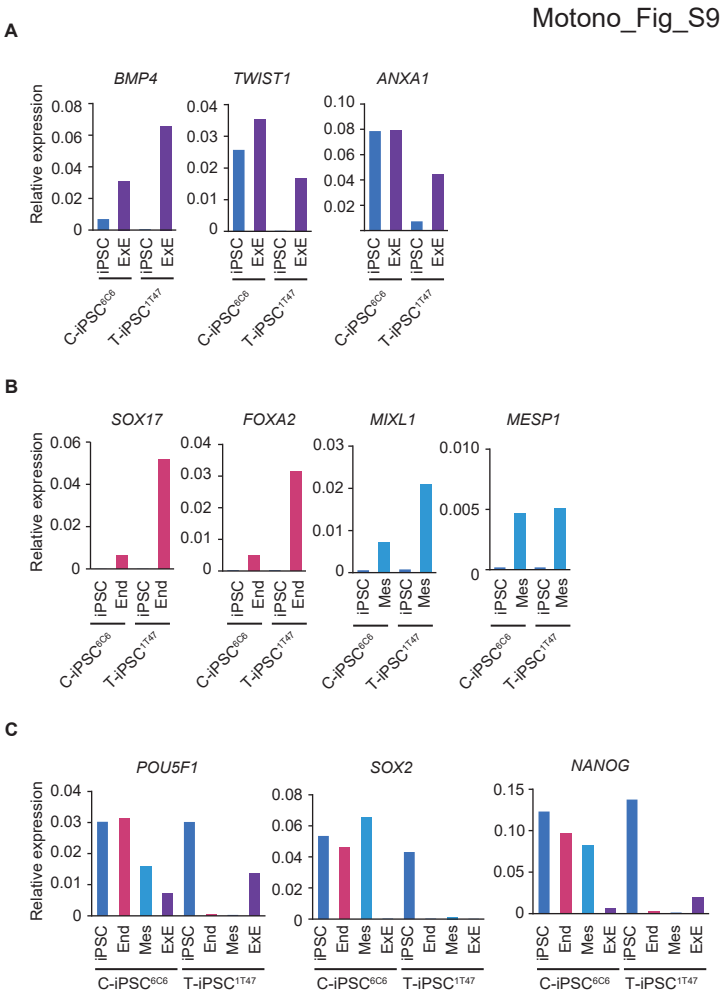
749 extraembryonic-associated genes. Red rectangles surround extraembryonic lineage plots.

750



Supplemental Figure S8, related to Figure 4. G_0 phase annotation and characteristics of C-iPSC^{6C6} and T-iPSC^{1T47} in each cell cycle. *(A)* Violin plots showing the distribution of gene expression levels of regulatory genes for DNA methylation, extraembryonic, aging-, and cell cycle-associated genes in each cell cycle phase in C-iPSC^{6C6}. Each panel represents differences in the distribution of gene expression for each factor when different expression levels of *CDKN2A* define the G_0 phase. We defined *CDKN2A* >1 as the G_0 phase. *(B)* Violin plots showing the distribution of expression of extraembryonic and pluripotency-associated genes. The left panel shows the distribution of gene expression in C-iPSC^{6C6} and T-iPSC^{1T47}. Right

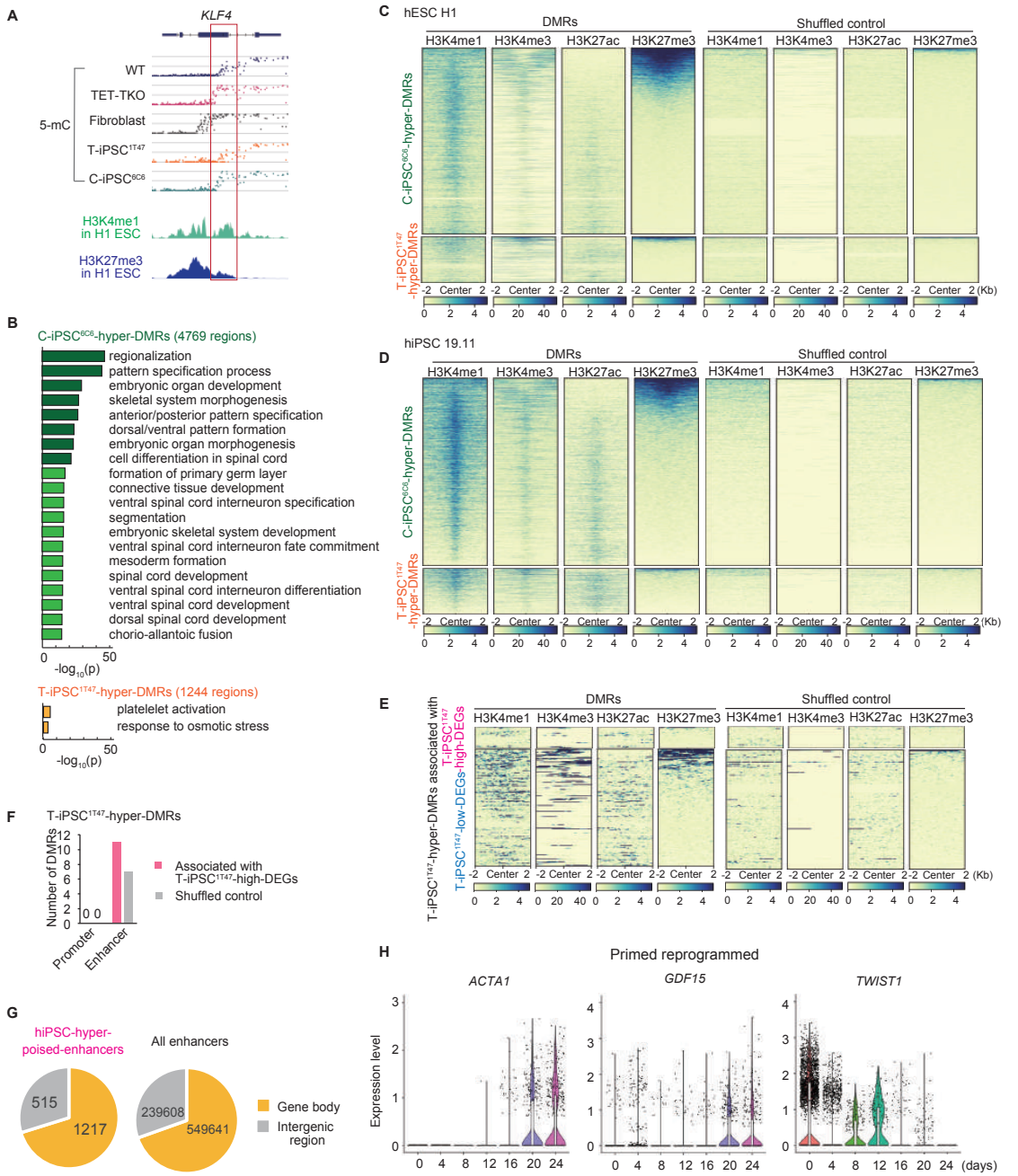
761 panel shows gene expression in each cell cycle phase in C-iPSC^{6C6}. *(C)* UMAP plot showing
762 expression levels of extraembryonic- and pluripotency-associated genes in C-iPSC^{6C6} and T-
763 iPSC^{1T47}. Blue circles indicate the C-iPSC^{6C6}-specific population. *(D)* Violin plots showing the
764 distribution of regulatory genes for DNA methylation, extraembryonic, aging-, and
765 pluripotency-associated genes in each cell cycle phase in T-iPSC^{1T47}. *(E)* Bar plots of the gene
766 expression levels of previously assigned differentiation-defective markers in C-iPSC^{6C6}
767 compared with T-iPSC^{1T47}.
768



769

770 **Supplemental Figure S9**, related to Figure 4. Directed differentiation into the extraembryonic
771 and mesendodermal lineages in C-iPSC^{6C6} and T-iPSC^{1T47}. *(A)* Relative expression levels of
772 extraembryonic-associated genes in C-iPSC^{6C6} and T-iPSC^{1T47} before and after differentiation
773 into the extraembryonic lineage on day 4 (n = 1). iPSC, before differentiation; ExE, after
774 differentiation into the extraembryonic lineage. *(B)* Relative expression levels of marker genes
775 after day 4 of directed differentiation into the endoderm (End) and mesoderm (Mes) (n = 1).
776 *(C)* Relative expression levels of pluripotency-associated genes before and after differentiation
777 into various lineages (n = 1). Pluripotency-associated markers were retained after differentiation
778 in C-iPSC^{6C6}, indicating differentiation defects.

779



780

781 **Supplemental Figure S10**, related to Figure 5. Characterization of DMRs between C-iPSC^{6C6}

782 and T-iPSC^{1T47} and the extraembryonic-associated gene expression dynamics during

783 reprogramming. *(A)* Genome browser tracks for WGBS data (upper panel, 5-mC) and ChIP-

784 seq data (lower panel) of *KLF4*. Red rectangle shows the DMRs between C-iPSC^{6C6} and T-

785 iPSC^{1T47}. WT, hESC HUES8; TET-TKO, TET1/2/3-triple-knockout HUES8. *(B)* Functional

786 enrichment of genes associated with DMRs between C-iPSC^{6C6} and T-iPSC^{1T47} using GREAT.
787 *(C)* Heatmap of ChIP-seq signals of histone modifications (hESC H1) in C-iPSC^{6C6}-hyper-
788 DMRs and T-iPSC^{1T47}-hyper-DMRs (left panel). Right panel shows histone modifications in
789 the shuffled control regions. The rows are sorted in order of H3K27me3-densities. *(D)* Heatmap
790 of ChIP-seq signals of histone modifications (hiPSC 19.11). *(E)* Heatmap of ChIP-seq signals
791 of histone modifications (hESC H1) in T-iPSC^{1T47}-hyper-DMRs associated with the DEGs
792 between C-iPSC^{6C6} and T-iPSC^{1T47}. *(F)* Bar plots showing the number of T-iPSC^{1T47}-hyper-
793 DMRs associated with T-iPSC^{1T47}-high-DEGs overlapped in promoters and enhancers. *(G)* The
794 proportion of *de novo* methylated poised enhancers (hiPSC-hyper-poised-enhancers) in gene
795 bodies and intergenic regions. All enhancers described are those on autosomes. *(H)* Violin plots
796 of *ACTA1*, *GDF15*, and *TWIST1* expression during primed reprogramming (Liu et al. 2020).
797
798
799 **Supplemental information**
800
801 **This PDF file includes**
802 Supplemental Figures S1–S10.
803 **Other Supplemental Materials for this manuscript include the following:**
804 Supplemental methods and materials.
805 Supplemental Tables S1–S6.
806

807 **References**

808 Andersson E, Tryggvason U, Deng Q, Friling S, Alekseenko Z, Robert B, Perlmann T, Ericson
809 J. 2006. Identification of intrinsic determinants of midbrain dopamine neurons. *Cell*
810 **124**: 393-405. doi:10.1016/j.cell.2005.10.037.

811 Bartoccetti M, van der Veer BK, Luo X, Khoueiry R, She P, Bajaj M, Xu J, Janiszewski A,
812 Thienpont B, Pasque V et al. 2020. Regulatory dynamics of Tet1 and Oct4 resolve stages
813 of global DNA demethylation and transcriptomic changes in reprogramming. *Cell Rep*
814 **30**: 2150-2169 e2159. doi:10.1016/j.celrep.2020.01.065.

815 Blackledge NP, Rose NR, Klose RJ. 2015. Targeting Polycomb systems to regulate gene
816 expression: modifications to a complex story. *Nat Rev Mol Cell Biol* **16**: 643-649.
817 doi:10.1038/nrm4067.

818 Bock C, Kiskinis E, Verstappen G, Gu H, Boulting G, Smith ZD, Ziller M, Croft GF, Amoroso
819 MW, Oakley DH et al. 2011. Reference maps of human ES and iPS cell variation enable
820 high-throughput characterization of pluripotent cell lines. *Cell* **144**: 439-452.
821 doi:10.1016/j.cell.2010.12.032.

822 Brons IG, Smithers LE, Trotter MW, Rugg-Gunn P, Sun B, Chuva de Sousa Lopes SM, Howlett
823 SK, Clarkson A, Ahrlund-Richter L, Pedersen RA et al. 2007. Derivation of pluripotent
824 epiblast stem cells from mammalian embryos. *Nature* **448**: 191-195.
825 doi:10.1038/nature05950.

826 Buckberry S, Liu X, Poppe D, Tan JP, Sun G, Chen J, Nguyen TV, de Mendoza A, Pflueger J,
827 Frazer T et al. 2023. Transient naive reprogramming corrects hiPS cells functionally and
828 epigenetically. *Nature* **620**: 863-872. doi:10.1038/s41586-023-06424-7.
829 Cacchiarelli D, Trapnell C, Ziller MJ, Soumillon M, Cesana M, Karnik R, Donaghey J, Smith
830 ZD, Ratanasirintrawoot S, Zhang X et al. 2015. Integrative analyses of human
831 reprogramming reveal dynamic nature of induced pluripotency. *Cell* **162**: 412-424.
832 doi:10.1016/j.cell.2015.06.016.
833 Chambers SM, Fasano CA, Papapetrou EP, Tomishima M, Sadelain M, Studer L. 2009. Highly
834 efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD
835 signaling. *Nat Biotechnol* **27**: 275-280. doi:10.1038/nbt.1529.
836 Charlton J, Jung EJ, Mattei AL, Bailly N, Liao J, Martin EJ, Giesselmann P, Brandl B,
837 Stamenova EK, Muller FJ et al. 2020. TETs compete with DNMT3 activity in
838 pluripotent cells at thousands of methylated somatic enhancers. *Nat Genet* **52**: 819-827.
839 doi:10.1038/s41588-020-0639-9.
840 Cheng S, Mittnenzweig M, Mayshar Y, Lifshitz A, Dunjic M, Rais Y, Ben-Yair R, Gehrs S,
841 Chomsky E, Mukamel Z et al. 2022. The intrinsic and extrinsic effects of TET proteins
842 during gastrulation. *Cell* **185**: 3169-3185 e3120. doi:10.1016/j.cell.2022.06.049.
843 Chhabra S, Liu L, Goh R, Kong X, Warmflash A. 2019. Dissecting the dynamics of signaling
844 events in the BMP, WNT, and NODAL cascade during self-organized fate patterning in

845 human gastruloids. *PLoS Biol* **17**: e3000498. doi:10.1371/journal.pbio.3000498.

846 Collinson A, Collier AJ, Morgan NP, Sienerth AR, Chandra T, Andrews S, Rugg-Gunn PJ. 2016.
847 Deletion of the Polycomb-group protein EZH2 leads to compromised self-renewal and
848 differentiation defects in human embryonic stem cells. *Cell Rep* **17**: 2700-2714.
849 doi:10.1016/j.celrep.2016.11.032.

850 Consortium EP, Moore JE, Purcaro MJ, Pratt HE, Epstein CB, Shoresh N, Adrian J, Kawli T,
851 Davis CA, Dobin A et al. 2020. Expanded encyclopaedias of DNA elements in the
852 human and mouse genomes. *Nature* **583**: 699-710. doi:10.1038/s41586-020-2493-4.

853 Duan D, Fu Y, Paxinos G, Watson C. 2013. Spatiotemporal expression patterns of Pax6 in the
854 brain of embryonic, newborn, and adult mice. *Brain Struct Funct* **218**: 353-372.
855 doi:10.1007/s00429-012-0397-2.

856 Esteban MA, Wang T, Qin B, Yang J, Qin D, Cai J, Li W, Weng Z, Chen J, Ni S et al. 2010.
857 Vitamin C enhances the generation of mouse and human induced pluripotent stem cells.
858 *Cell Stem Cell* **6**: 71-79. doi:10.1016/j.stem.2009.12.001.

859 Haerinck J, Goossens S, Berx G. 2023. The epithelial-mesenchymal plasticity landscape:
860 principles of design and mechanisms of regulation. *Nat Rev Genet* **24**: 590-609.
861 doi:10.1038/s41576-023-00601-0.

862 Hu X, Zhang L, Mao SQ, Li Z, Chen J, Zhang RR, Wu HP, Gao J, Guo F, Liu W et al. 2014.
863 Tet and TDG mediate DNA demethylation essential for mesenchymal-to-epithelial

transition in somatic cell reprogramming. *Cell Stem Cell* **14**: 512-522. doi:10.1016/j.stem.2014.01.001.

International Stem Cell I, Adewumi O, Aflatoonian B, Ahrlund-Richter L, Amit M, Andrews PW, Beighton G, Bello PA, Benvenisty N, Berry LS et al. 2007. Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. *Nat Biotechnol* **25**: 803-816. doi:10.1038/nbt1318.

Jones PA. 1999. The DNA methylation paradox. *Trends Genet* **15**: 34-37. doi:10.1016/S0168-9525(98)01636-9.

Kim H, Lee G, Ganat Y, Papapetrou EP, Lipchina I, Socci ND, Sadelain M, Studer L. 2011. miR-371-3 expression predicts neural differentiation propensity in human pluripotent stem cells. *Cell Stem Cell* **8**: 695-706. doi:10.1016/j.stem.2011.04.002.

Koyanagi-Aoi M, Ohnuki M, Takahashi K, Okita K, Noma H, Sawamura Y, Teramoto I, Narita M, Sato Y, Ichisaka T et al. 2013. Differentiation-defective phenotypes revealed by large-scale analyses of human pluripotent stem cells. *Proc Natl Acad Sci U S A* **110**: 20569-20574. doi:10.1073/pnas.1319061110.

Kumari R, Jat P. 2021. Mechanisms of cellular senescence: cell cycle arrest and senescence associated secretory phenotype. *Front Cell Dev Biol* **9**: 645593. doi:10.3389/fcell.2021.645593.

Lister R, Pelizzola M, Kida YS, Hawkins RD, Nery JR, Hon G, Antosiewicz-Bourget J,

883 O'Malley R, Castanon R, Klugman S et al. 2011. Hotspots of aberrant epigenomic
884 reprogramming in human induced pluripotent stem cells. *Nature* **471**: 68-73.
885 doi:10.1038/nature09798.
886 Liu L, Oura S, Markham Z, Hamilton JN, Skory RM, Li L, Sakurai M, Wang L, Pinzon-Arteaga
887 CA, Plachta N et al. 2023. Modeling post-implantation stages of human development
888 into early organogenesis with stem-cell-derived peri-gastruloids. *Cell* **186**: 3776-3792
889 e3716. doi:10.1016/j.cell.2023.07.018.
890 Liu X, Ouyang JF, Rossello FJ, Tan JP, Davidson KC, Valdes DS, Schroder J, Sun YBY, Chen
891 J, Knaupp AS et al. 2020. Reprogramming roadmap reveals route to human induced
892 trophoblast stem cells. *Nature* **586**: 101-107. doi:10.1038/s41586-020-2734-6.
893 Loh KM, Lim B. 2011. A precarious balance: pluripotency factors as lineage specifiers. *Cell*
894 *Stem Cell* **8**: 363-369. doi:10.1016/j.stem.2011.03.013.
895 Lu AT, Fei Z, Haghani A, Robeck TR, Zoller JA, Li CZ, Lowe R, Yan Q, Zhang J, Vu H et al.
896 2023. Universal DNA methylation age across mammalian tissues. *Nat Aging* **3**: 1144-
897 1166. doi:10.1038/s43587-023-00462-6.
898 Lu YR, Tian X, Sinclair DA. 2023. The Information Theory of Aging. *Nat Aging* **3**: 1486-1499.
899 doi:10.1038/s43587-023-00527-6
900 Ma H, Morey R, O'Neil RC, He Y, Daughtry B, Schultz MD, Hariharan M, Nery JR, Castanon
901 R, Sabatini K et al. 2014. Abnormalities in human pluripotent cells due to

902 reprogramming mechanisms. *Nature* **511**: 177-183. doi:10.1038/nature13551.

903 Messmer T, von Meyenn F, Savino A, Santos F, Mohammed H, Lun ATL, Marioni JC, Reik W.
904 2019. Transcriptional heterogeneity in naive and primed human pluripotent stem cells
905 at single-cell resolution. *Cell Rep* **26**: 815-824 e814. doi:10.1016/j.celrep.2018.12.099.

906 Moqri M, Cipriano A, Simpson DJ, Rasouli S, Murty T, de Jong TA, Nachun D, de Sena
907 Brandine G, Ying K, Tarkhov A et al. 2024. PRC2-AgeIndex as a universal biomarker
908 of aging and rejuvenation. *Nat Commun* **15**: 5956. doi:10.1038/s41467-024-50098-2.

909 Munoz-Sanjuan I, Brivanlou AH. 2002. Neural induction, the default model and embryonic
910 stem cells. *Nat Rev Neurosci* **3**: 271-280. doi:10.1038/nrn786.

911 Nakamura T, Okamoto I, Sasaki K, Yabuta Y, Iwatani C, Tsuchiya H, Seita Y, Nakamura S,
912 Yamamoto T, Saitou M. 2016. A developmental coordinate of pluripotency among mice,
913 monkeys and humans. *Nature* **537**: 57-62. doi:10.1038/nature19096.

914 Ohnuki M, Tanabe K, Sutou K, Teramoto I, Sawamura Y, Narita M, Nakamura M, Tokunaga Y,
915 Nakamura M, Watanabe A et al. 2014. Dynamic regulation of human endogenous
916 retroviruses mediates factor-induced reprogramming and differentiation potential. *Proc*
917 *Natl Acad Sci U S A* **111**: 12426-12431. doi:10.1073/pnas.1413299111.

918 Okita K, Matsumura Y, Sato Y, Okada A, Morizane A, Okamoto S, Hong H, Nakagawa M,
919 Tanabe K, Tezuka K et al. 2011. A more efficient method to generate integration-free
920 human iPS cells. *Nat Methods* **8**: 409-412. doi:10.1038/nmeth.1591.

921 Ono T, Suzuki Y, Kato Y, Fujita R, Araki T, Yamashita T, Kato H, Torii R, Sato N. 2014. A
922 single-cell and feeder-free culture system for monkey embryonic stem cells. *PLoS One*
923 **9**: e88346. doi:10.1371/journal.pone.0088346.
924 Patani R, Compston A, Puddifoot CA, Wyllie DJ, Hardingham GE, Allen ND, Chandran S.
925 2009. Activin/Nodal inhibition alone accelerates highly efficient neural conversion from
926 human embryonic stem cells and imposes a caudal positional identity. *PLoS One* **4**:
927 e7327. doi:10.1371/journal.pone.0007327.
928 Pham TXA, Panda A, Kagawa H, To SK, Ertekin C, Georgolopoulos G, van Knippenberg S,
929 Allsop RN, Bruneau A, Chui JS et al. 2022. Modeling human extraembryonic mesoderm
930 cells using naive pluripotent stem cells. *Cell Stem Cell* **29**: 1346-1365 e1310.
931 doi:10.1016/j.stem.2022.08.001.
932 Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. 2011. A unique
933 chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**:
934 279-283. doi:10.1038/nature09692.
935 Roost MS, Slieker RC, Bialecka M, van Iperen L, Gomes Fernandes MM, He N, Suchiman
936 HED, Szuhai K, Carlotti F, de Koning EJP et al. 2017. DNA methylation and
937 transcriptional trajectories during human development and reprogramming of isogenic
938 pluripotent stem cells. *Nat Commun* **8**: 908. doi:10.1038/s41467-017-01077-3.
939 Ruetz T, Pfisterer U, Di Stefano B, Ashmore J, Beniazza M, Tian TV, Kaemena DF, Tosti L,

940 Tan W, Manning JR et al. 2017. Constitutively active SMAD2/3 are broad-scope
941 potentiators of transcription-factor-mediated cellular reprogramming. *Cell Stem Cell* **21**:
942 791-805 e799. doi:10.1016/j.stem.2017.10.013.
943 Sasaki K, Nakamura T, Okamoto I, Yabuta Y, Iwatani C, Tsuchiya H, Seita Y, Nakamura S,
944 Shiraki N, Takakuwa T et al. 2016. The germ cell fate of cynomolgus monkeys is
945 specified in the nascent amnion. *Dev Cell* **39**: 169-185.
946 doi:10.1016/j.devcel.2016.09.007.
947 Schiebinger G, Shu J, Tabaka M, Cleary B, Subramanian V, Solomon A, Gould J, Liu S, Lin S,
948 Berube P et al. 2019. Optimal-transport analysis of single-cell gene expression identifies
949 developmental trajectories in reprogramming. *Cell* **176**: 928-943 e922.
950 doi:10.1016/j.cell.2019.01.006.
951 Smith A. 2017. Formative pluripotency: the executive phase in a developmental continuum.
952 *Development* **144**: 365-373. doi:10.1242/dev.142679.
953 Steinbeck JA, Choi SJ, Mrejeru A, Ganat Y, Deisseroth K, Sulzer D, Mosharov EV, Studer L.
954 2015. Optogenetics enables functional analysis of human embryonic stem cell-derived
955 grafts in a Parkinson's disease model. *Nat Biotechnol* **33**: 204-209.
956 doi:10.1038/nbt.3124.
957 Tchieu J, Zimmer B, Fattahi F, Amin S, Zeltner N, Chen S, Studer L. 2017. A modular platform
958 for differentiation of human PSCs into all major ectodermal lineages. *Cell Stem Cell* **21**:

959 399-410.e397. doi:10.1016/j.stem.2017.08.015.

960 Verma N, Pan H, Dore LC, Shukla A, Li QV, Pelham-Webb B, Teijeiro V, Gonzalez F, Krivtsov
961 A, Chang CJ et al. 2018. TET proteins safeguard bivalent promoters from de novo
962 methylation in human embryonic stem cells. *Nat Genet* **50**: 83-95. doi:10.1038/s41588-
963 017-0002-y.

964 Wang L, Huang D, Huang C, Yin Y, Vali K, Zhang M, Tang Y. 2017. Enhanced human somatic
965 cell reprogramming efficiency by fusion of the MYC transactivation domain and OCT4.
966 *Stem Cell Res* **25**: 88-97. doi:10.1016/j.scr.2017.10.014.

967 Yang R, Goedel A, Kang Y, Si C, Chu C, Zheng Y, Chen Z, Gruber PJ, Xiao Y, Zhou C et al.
968 2021. Amnion signals are essential for mesoderm formation in primates. *Nat Commun*
969 **12**: 5126. doi:10.1038/s41467-021-25186-2.

970 Zhang X, Huang CT, Chen J, Pankratz MT, Xi J, Li J, Yang Y, Lavaute TM, Li XJ, Ayala M et
971 al. 2010. Pax6 is a human neuroectoderm cell fate determinant. *Cell Stem Cell* **7**: 90-
972 100. doi:10.1016/j.stem.2010.04.017.