Oct4 redox sensitivity potentiates reprogramming and differentiation

Zuolian Shen[1,2,¶], Yifan Wu[1,2,¶], Asit Manna[1,2], Kimberley J. Evason[1,2], Mahesh B. Chandrasekharan[2,3] and Dean Tantin[1,2,*]

[1]Department of Pathology,

[2]Huntsman Cancer Institute,

[3]Department of Radiation Oncology, University of Utah School of Medicine, Salt Lake City, UT 84112, USA

[*]Correspondence: dean.tantin@path.utah.edu.

[¶]Co-first authors

Characters in title: 68

Word count (excluding STAR Methods, supplemental item legends, and references): 6,287

16 **SUMMARY**

17 **The transcription factor Oct4/Pou5f1 is a component of the regulatory circuitry governing**
18 **pluripotency and is widely used to generate induced pluripotent stem cells (iPSCs) from somatic**
19 **cells. Here we use domain swapping and mutagenesis to compare Oct4's reprogramming activity**
20 **with the reprogramming-incompetent paralog Oct1/Pou2f1, identifying a redox-sensitive DNA**
21 **binding domain cysteine residue (Cys48) as a key determinant of reprogramming and**
22 **differentiation. In combination with the Oct4 N-terminus, mutating Oct1's serine at this position to**
23 **cysteine is sufficient to confer reprogramming activity. Conversely, Oct4^{C48S} reduces**
24 **reprogramming potential. Oct4^{C48S} desensitizes the protein to oxidative inhibition and prevents**
25 **oxidation-mediated protein ubiquitylation. *Pou5f1^{C48S}* point mutation has little effect on**
26 **undifferentiated embryonic stem cells (ESCs), but upon retinoic acid (RA) treatment causes**
27 **retention of Oct4 expression and deregulated differentiation. *Pou5f1^{C48S}* ESCs also contribute**
28 **poorly to adult somatic tissues and form less differentiated teratomas. Collectively, these findings**
29 **uncover a new Oct4 redox mechanism involved in both entry into and exit from pluripotency.**

30

31 **Keywords** Oct4 (Pou5f1); Oct1 (Pou2f1); induced pluripotent stem cells (iPSCs); oxidative stress;
32 ubiquitylation

33 **INTRODUCTION**

34 The early mammalian embryo contains undifferentiated, pluripotent cells capable of generating all
35 embryonic tissue lineages. Cultured embryonic stem cells (ESCs) have similar capabilities.[1] The
36 transcription factor Oct4 governs the establishment of undifferentiated inner cell mass and epiblast cells,
37 as well germline cells and ESCs.[2] Both ESC differentiation and blastocyst implantation are accompanied
38 by Oct4 loss.[3]. Together with transcription factors such as Nanog and Sox2, Oct4 forms a network of
39 pluripotency transcriptional regulators that reinforce their own and each other's expression.[4]

40 Oct4 is widely used to generate induced pluripotent stem cells (iPSCs) from somatic cells.[5,6]
41 Reprogramming methods that omit Oct4 nevertheless rely on activation of the endogenous *Pou5f1* (*Oct4*)
42 gene.[7,8] Reprogramming strategies that increase efficiency[9-13], or improve the quality of iPSC clones[14-17]
43 reveal trade-offs: efficiently generating iPSC that are of non-optimal quality, or inefficiently generating
44 iPSCs with improved developmental potential.

45 Pluripotent cells co-express Oct4 together with the widely expressed prototypic POU transcription
46 factor Oct1/Pou2f1.[18-20] The two proteins share similar DNA binding domains (DBDs) and consensus
47 DNA binding sequences.[21] The two proteins also associate with many of the same target genes.[22,23]
48 Nevertheless, Oct1 and other mammalian Oct4 paralogs do not efficiently reprogram.[5,24] Oct1 in particular
49 is poor at reprogramming relative to most other Oct transcription factors.[25] In contrast, other
50 reprogramming factors such as Sox2, Klf4 and c-Myc can be replaced with close paralogs.[26,27]
51 Structure/function studies comparing Oct4 with its paralogs have identified reprogramming
52 determinants.[28-30] An example is three residues present in the Oct4 DBD but absent in Oct6/Pou3f1 needed
53 to form Sox2 heterodimers. Introducing these residues, plus the Oct4 N- and C-termini, results in efficient
54 reprogramming.[29]. However, several Oct4 paralogs including Oct1 contain these residues but are
55 nevertheless reprogramming-incompetent, indicating the presence of additional reprogramming
56 determinants present in Oct4 (and Oct6) but absent in Oct1.

57 Recent studies show that omission of Oct4 from reprogramming cocktails is inefficient, but results
58 in qualitatively superior iPSC clones with enhanced developmental potential.[17,31] The counterproductive
59 activities of Oct4 are caused by transient promotion gene expression programs that deviate cells from
60 reprogramming trajectories.[17] A logical extension of these findings is that transient down-regulation of
61 counterproductive Oct4 activities may enable reprogramming.

62 Here, we identify an Oct4 DBD cysteine residue (Cys48) as a central reprogramming determinant.
63 A serine residue is present in Oct1 at this position. In the presence of the Oct4 N-terminus, changing
64 Oct1's serine to cysteine confers robust reprogramming activity. Conversely, mutating Oct4's cysteine to
65 serine reduces reprogramming efficiency by ~60%. Cys48 sensitizes DNA binding activity and protein
66 ubiquitylation to oxidative stress. CRISPR-edited Oct4^{C48S} embryonic stem cell (ESC) lines are
67 phenotypically normal until differentiated, following which they abnormally retain Oct4 expression,
68 proliferate less and undergo apoptosis at higher rates compared to parental controls. These cells also form
69 less differentiated teratomas and contribute poorly to the development of viable adult mice. Cumulatively,
70 our findings identify a redox-sensitive Oct4 residue that restrains protein function, but nevertheless
71 functions positively in reprogramming and normal differentiation.

72 **RESULTS**

73

74 **Oct4 vs. Oct1 reprogramming potential concentrates in a DBD cysteine**

75 The POU-specific and POU-homeodomain (POU_S and POU_H) DNA binding subdomains of mouse Oct1
76 and Oct4 share 71% and 59% identity, respectively. In contrast, the N- and C-termini, and the linker
77 between the subdomains, are divergent (Figure 1A and not shown). The human proteins are structured
78 similarly. The amino acid numbering of the DBDs can be standardized by referring to the first amino acid
79 of POU_S as residue 1 (Figure 1A). To identify Oct4 determinants important for reprogramming, we
80 generated Oct4/Oct1 chimeric constructs containing the distinct domains (Figure 1B). We included C-
81 terminal FLAG-tags on all proteins to track expression. Unlike an N-terminal FLAG tag which was
82 deleterious (Figure S1), constructs with engineered C-terminal tags reprogram with similar efficiency
83 (Figure 1C). Prior work comparing Oct4 to Oct6 also identified a lysine residue within Oct4 POU_S (K40,
84 Figure 1A) as important for reprogramming.[30] We therefore mutated $Oct1^{N40}$ to lysine in some constructs
85 (Figure 1D-F).

86 To measure reprogramming efficiency, we transduced mouse embryonic fibroblasts (MEFs)
87 expressing GFP under the control of endogenous *Pou5f1*[32] with a lentivirus expressing the four Yamanaka
88 factors[33] (Figure 1B). Substituting Oct1 for Oct4 eliminated reprogramming potential (Figure 1D,E),
89 setting a baseline for domain-swaps and mutations. We then replaced the Oct1 N- and C-termini with that
90 of Oct4, and mutated $Oct1^{N40}$ to the lysine present in Oct4. Only fusion of the Oct4 N-terminus generated
91 detectable colony-forming potential, at ~5% the level of WT Oct4 (Figure 1D,E). These findings are
92 consistent with work establishing the importance of the Oct4 N-terminus.[34] Similar data were obtained at
93 multiple timepoints (not shown), indicating that this result does not simply reflect kinetic differences in
94 reprogramming. Anti-FLAG immunoblotting showed similar expression (Figure 1F). Viral titers and
95 expression of Klf4 from the mSTEMCCA vector were also similar (not shown). Because a construct with
96 both the Oct4 N- and C-termini displayed only ~5% efficiency (Figure 1D,E), we conclude that

97 determinants within the Oct4 DNA-binding domain confer the bulk of reprogramming potential vis-à-vis
98 Oct1.

99 In the context of the Oct4 N-terminus ("4N"), we mutagenized DBD amino acid residues that
100 differ between Oct1 and Oct4, focusing on post-translationally modifiable residues: Lys, Ser, Thr and Cys.
101 This analysis identified 14 different amino acids, not including Lys40 (Figure 1A). Oct6 is capable of
102 reprogramming when modified to enable Sox2 dimerization.[35] Filtering these residues using conservation
103 between Oct4 and Oct6 pinpointed six possible residues. Replacing one of these, Oct1 Ser48, with the
104 Oct4 cysteine residue found at this position increased activity >7-fold over the "4N" construct containing
105 the Oct4 N-terminus alone (Figure 2A,B, "1-4All"). iPSC colonies formed under these conditions showed
106 a pluripotent morphology and expressed GFP from the endogenous *Pou5f1* locus at levels comparable to
107 Oct4 (Figure S2). The remaining activity was concentrated in the linker domain, as a construct containing
108 the Oct4 N-terminus, Cys48 and the Oct4 linker (4N-S48C-4L) had 100% activity (Figure 2A,B). The
109 Oct4 linker has been shown to regulate reprogramming via selective interactions with different
110 cofactors.[28,36] The differences were maintained over time, with the 1-4All construct consistently
111 generating ~60% activity over multiple days, indicating that the differences were qualitative and not a
112 function of differing rates (Figure S3).

113 An Oct4^{C48S} construct reduced activity by 60% (Figure 2A,B), indicating that over half of Oct4's
114 differential reprogramming activity vis-à-vis Oct1 is lost by mutating this residue. The reciprocal Oct1^{S48C}
115 mutation resulted in no reprogramming (Figure 2A,B, Oct1-S48C), indicating that Cys48 is not sufficient,
116 and highlighting its combined importance with the Oct4 N-terminus. The proteins were equivalently
117 expressed (Figure 2C). These results indicate that in the context of the Oct4 N-terminus, Cys48 confers
118 robust reprogramming activity. In an Oct1 DBD:DNA structure[37], Ser48 hydrogen bonds with the DNA
119 backbone (Figure 2D). In a recent Oct4:Sox2:nucleosome structure[38], the Cys48 sulfhydryl also allows
120 for hydrogen bonding. Predictably, Cys48 mutation to glycine greatly diminishes Oct4 DNA binding in
121 vitro.[39] Oct6 also contains a cysteine at this position (Figure 2E). These structural studies and our findings

122 indicate that Ser48 and Cys48 have minimal differences in DNA binding, but that Cys48 nevertheless
123 plays key roles in reprogramming.
124
125 **Oct4 Cys48 mediates degradation in response to oxidative stress**
126 Oct4 is hypersensitive to oxidative stress[39-41] in a manner reversed by thioredoxin[41], implicating cysteine
127 thiols in redox sensitivity. We reproduced this result (Figure S4), which suggested that Cys48 may
128 potentiate reprogramming through a mechanism involving redox sensitivity. We used lysates from 293T
129 cells transfected with WT or C48S Oct4 to measured DNA binding using electrophoretic mobility shift
130 assays (EMSA) with DNA containing a canonical binding element.[22] To measure redox sensitivity, we
131 used methods similar to Guo et al.[41] in which Oct4 is pre-treated with the oxidizing agent diamide prior
132 to addition of the DNA. 293T cells express endogenous Oct1 as an internal control (Figure 3A, lanes 2-
133 3). Oct4 bound to octamer DNA specifically (Figure 3A, lanes 7-8) and was hypersensitive to diamide
134 treatment (lanes 9-11). Oct4^{C48S} bound DNA equivalently (lane 13) and was comparatively resistant to
135 oxidation (lanes 14-16). In contrast, little sensitivity was observed with Oct1. Quantification from three
136 replicates is shown in Figure 3B. The recombinant Oct4 contained a C-terminal FLAG tag, allowing for
137 purification (Figure S5). Similar relative resistance to oxidative inhibition of DNA binding was observed
138 using purified C48S recombinant Oct4 (Figure 3C). Quantification from three replicates is shown in
139 Figure 3D. These results indicate that Oct4 Cys48 mediates oxidative inhibition of DNA binding.
140 Oct4 oxidation has been linked to its degradation.[39] We noted in over-exposed immunoblots the
141 presence of multiple slow-migrating high molecular weight forms of wild-type Oct4 which were strongly
142 attenuated with C48S mutation (Figure 3E, lanes 2-3). Hydrogen peroxide treatment augmented the
143 banding pattern using wild-type but not mutant protein (lanes 4-5). The bands were also increased using
144 MG-132, consistent with ubiquitylated species (lanes 6-7). H_2O_2 and MG-132 co-treatment further
145 increased the pattern (lanes 8-9). None of these treatments affected short-term cell viability (Figure S6).
146 We purified Oct4 expressed from 293T cells using a C-terminal streptavidin tag, and subjected the

147 isolated proteins to immunoblotting using antibodies specific to different ubiquitin species. Antibodies
148 recognizing mono- or poly-ubiquitylated proteins revealed strong diminution with Oct4^{C48S} as expected
149 (Figure 3F, FK2). Strong diminution was observed with K48 linkage-specific polyubiquitin antibodies
150 (Figure 3F), while K63 linkage-specific antibodies showed modest differences (Figure S7). Similar results
151 were observed using endogenous Oct4 in wild-type ESCs (Figure 3G, lane 2) that were augmented by
152 H_2O_2 and MG-132 (lanes 4 and 6). We conclude that Cys48 promotes Oct4 ubiquitylation and degradation
153 in response to oxidative stress, largely via catalysis of K48-linked polyubiquitin chains.
154

155 **Oct4^{C48S} ESCs differentiate abnormally**

156 We generated ESCs with a single-base missense mutation in the endogenous *Oct4* (*Pou5f1*) locus that
157 converts Cys48 to Ser. Two silent point mutations were also engineered to create an *Xba*I site to monitor
158 targeting and facilitate genotyping. Three low-passage parent ESC lines of different strain backgrounds
159 were chosen for targeting. Multiple independent clones were generated for each line. Homozygous
160 mutation was confirmed by HRMA and resequencing (not shown and Figure 4A). Tested clones were
161 karyotypically normal, with the proportion of diploid cells ranging between 71 and 93%. Undifferentiated
162 mutant ESC clones were morphologically indistinguishable from their parent clones and expressed Oct4,
163 Sox2 and Nanog at similar levels (Figure 4B,C).

164 To test if C48S mutation protects endogenous Oct4 against oxidative stress, we studied Oct4
165 ubiquitination in parent and mutant ESCs. Oct4 was immunoprecipitated from ESC lysates, and the
166 resulting material was immunoblotted using pan-Ub (FK2) and Oct4 antibodies. Ubiquitylation of
167 endogenous Oct4 in ESCs was clearly visualized and increased with H_2O_2 exposure (Figure 4D, lanes 1-
168 4). Ubiquitylation was attenuated in mutant cells (lanes 5-6), and unaffected by oxidative stress (lanes 7-
169 8). Oct4 immunoblotting confirmed that the protein was immunoprecipitated, with decreased parent Oct4
170 recovered with oxidative stress and more Oct4 recovered in mutant cells (Figure 4D, right side).

171 Cells were then differentiated in vitro using retinoic acid (RA) to assess Oct4 expression,

172 morphology and cell counts. Oct4 protein expression was abnormally retained early during differentiation
173 (Figure 4E, lanes 3 and 7). Quantification from multiple experiments revealed that the effect was transient,
174 with a >2.8-fold increase of Oct4 at day 2 that was lost by day 3 (Figure 4F). These results were
175 recapitulated using a second clone from the same parental ESC line (Figure S8A,B), as well as a clone
176 from a different 129×C57BL/6 ESC line, G4-N, which differentiated more slowly and showed greater
177 differences (Figure S8C).

178 We performed transcriptional profiling of parent and mutant cells at RA differentiation day 2, as
179 well as undifferentiated controls. Three replicates per condition were used, which showed tight
180 concordance (Figure S9A). 20-31 million sequence reads were generated for each sample, ~80% of which
181 aligned to the *mm39* reference genome. We identified 421 genes with significantly elevated gene
182 expression in differentiating mutant relative to parent cells (p<0.001, >2.5-fold increased, Figure 4H,
183 Table S1). Intersecting these genes with the set of genes normally decreased upon differentiation of parent
184 cells at this timepoint (1536 genes) showed that 40% of the 421 genes with retained expression in mutant
185 cells (169 genes) are normally down-regulated with differentiation. These included *Oct4* itself (*Pou5f1*),
186 as well as *Sox2*, *Nanog* and *Klf4* (Figure 4H). Comparing these genes with established Oct4 ChIP-seq[42]
187 identified 114 of 169 genes (67%) as direct Oct4 targets (Table S1). The complete set of 169 genes were
188 enriched for gene ontology (GO) terms associated with pluripotency such as "response to leukemia
189 inhibitory factor" and "stem cell population maintenance" (Figure 4G). Similar results were obtained
190 using a hierarchical clustering approach. We utilized the 100 top differentially expressed genes upon
191 differentiation for both genotypes (150 total genes). Although the differentiating parental and mutant
192 replicates clustered together, there was a cohort of pluripotency associated genes (including *Pou5f1*) that
193 were less down-regulated in mutant cells (Figure S9B). A second cohort of differentiation-specific genes
194 showed poorer expression in mutant cells upon differentiation. These included *Krt18*, *Mest*/*Peg1* and
195 *Dag1* (Figure S9B). Comparing the complete set of 150 genes to Oct4 ChIP-seq[42] identified 117 (78%) as
196 direct Oct4 targets (Table S1). These results indicate that Oct4^{C48S} ESCs inappropriately delay loss of

197 pluripotency gene expression signatures during differentiation.

198 The duration of retained Oct4 expression during differentiation can affect lineage specification.[43-
199 45] To determine the effects of Oct4^{C48S} mutation on long-term differentiation, we visualized cells treated
200 with RA over 12 days microscopically. These experiments revealed a significant depletion of cells later
201 during the second half differentiation (Figure 5A,B, D6-D12). The decreased cell numbers observed
202 during differentiation of C48S-edited cells with RA could be due to decreased proliferation and/or cell
203 death. We studied proliferation using cell trace violet (CTV) and apoptosis using TUNEL staining.
204 Adherent, viable differentiating C48S cells showed decreased proliferation at days 3 and 4 (Figure 5C)
205 and increased apoptosis at day 2 (Figure 5D). These results define a functional role for Oct4 Cys48 in
206 ESC differentiation.

207

208 **Oct4^{C48S} ESCs form less differentiated teratomas and contribute poorly to embryonic development**

209 We used mutant and control ESC lines to form teratomas in mice, which were monitored for two weeks
210 before tumor excision. Tumors were fixed, H&E stained, and pathologically scored in a blinded fashion.
211 Parent tumors contained areas of maturing structures such as squamous tissue, respiratory epithelium (cilia
212 and/or goblet cells), maturing neural components, gastric foveolar epithelium and fat. In contrast, tumors
213 formed with Oct4^{C48S} ESCs injected into the contralateral flanks of the same mice were more
214 homogeneous with few maturing components, instead entirely harboring immature neuroepithelial
215 elements, immature cartilage, embryonal carcinoma and yolk sac (Figure 6A,B). Additional images are
216 shown in Figure S10. At later time points, larger mutant teratomas began to accumulate mature structures,
217 but less robustly than parent controls (not shown).

218 One hallmark of pluripotency is the ability to contribute to cells and tissues of the embryo.[46] We
219 conducted chimerism assays to assess the ability of Oct4^{C48S} ESCs to contribute to development. We used
220 eight lines of mutant ESCs, derived from three parent lines of different strain backgrounds to minimize
221 background and clonal effects. Compared to the parent cells, which contributed to offspring at a rate of

222 10-20% (Figure 6C), none of the derived Oct4^{C48S} ESCs were able to contribute to adult mice (Figure 6C).

223 The results indicates that Oct4^{C48S} ESCs contribute poorly to embryonic development. Collectively, these

224 results uncover a critical role for the redox-sensitive Oct4 Cys48 in reprogramming and development.

225 **DISCUSSION**

226 The mechanisms by which Oct4 is regulated, particularly at a post-transcriptional level, are poorly defined.
227 Here, we pinpoint Oct4 redox sensitivity as an important determinant of both reprogramming and
228 differentiation. Pluripotent cells co-express Oct4 and Oct1.[47] Despite having similar DBDs and
229 recognizing the same 8-bp consensus octamer sequence[21], Oct1 and other Oct4 paralogs do not function
230 in reprogramming.[5] Structure/function studies comparing Oct4 with these paralogs have identified
231 different Oct4 reprogramming determinants.[28-30,34] We identify Cys48 of the Oct4 DBD as a critical
232 reprogramming determinant that sensitizes the protein to oxidative inhibition and degradation. In Oct1,
233 this residue is a serine. Both residues hydrogen bond with the DNA phosphate backbone.

234 Transcription factors with both positive and negative roles in reprogramming include Oct4 and
235 Smad3.[17,31,48] Recent studies show that removal of Oct4 from reprogramming cocktails is inefficient but
236 results in qualitatively superior iPSC clones.[17,31] Consequently, the few iPSC colonies derived in the
237 absence of Oct4 have enhanced developmental potential. For example, iPSCs reprogrammed without Oct4
238 are better able to generate viable mice in a single generation using tetraploid complementation. The
239 counter-productive activity of Oct4 is due to transient promotion of off-target, ectoderm-specific gene
240 expression.[17] These findings suggest that transient down-regulation of counterproductive Oct4 activities
241 may be an important step for efficient reprogramming. Our work identifies Oct4 redox sensitivity as
242 fulfilling such a role. The findings may provide a route to generating superior iPSC reprogramming
243 outcomes with higher efficiency.

244 Oct4 is hypersensitive to oxidative stress[39-41] in a manner reversed by thioredoxin.[41] Relative to
245 other Oct proteins, Oct4 has unusually high cysteine content in the DBD, with two cysteines in each
246 subdomain, including cysteine 142 (Cys50 of the POU-homeodomain), which lies in the recognition helix
247 of POU_H. Cys142 is the distinguishing feature of POU-homeodomains relative to other members of the
248 homeodomain family.[21] Cys48 is also located within a DNA recognition helix, of the POU_S domain. The
249 hydrogen bond contact between Cys48 and the DNA predicts that cysteine oxidation will block DNA

250 binding. Consistently, Oct4^{C48S} mutation maintains DNA binding and confers resistance to oxidation-
251 induced DNA binding inhibition in vitro while reducing reprogramming efficiency by more than half.
252 Cysteine redox sensitivity is known to be enhanced by adjacent basic amino acid residues[49], and Cys48 is
253 followed by a conserved Arg. Oct1 S48C mutation is sufficient to confer reprogramming activity, but only
254 in the context of the Oct4 N-terminus. The Oct4 N-terminus is known to mediate liquid-liquid phase
255 transitions to control target gene transcription.[34]

256 Prior studies have also identified a link between Oct4 oxidation and degradation.[39] In addition to
257 mediating sensitivity to oxidative stress, Cys48 strongly promotes protein ubiquitylation, in particular
258 following exposure to oxidative stress or MG-132. Oct4 is known to be ubiquitylated via K63-linked
259 chains by the Stub1 E3 Ub ligase[50], however we found that Oct4 marked with K48-specific linkages were
260 more strongly affected with Oct4^{C48S} mutation, suggesting a distinct pathway.

261 In ESCs, endogenous Oct4^{C48S} mutation does not affect the undifferentiated state, but results in
262 abnormalities upon RA-mediated differentiation, including retention of Oct4 expression, deregulated gene
263 expression, decreased proliferation and increased apoptosis. Prominent roles for Oct4 in cell survival and
264 proliferation during post-implantation embryonic development have been documented.[45] Consistent with
265 their pluripotency, when injected into adult mice, ESCS are normally able to form teratomas containing
266 differentiating structures representing all three germ layers. When injected into blastocysts, they can
267 contribute to adult mouse tissues.[46] We found that both of these features become defective with Cys48
268 mutation. Cumulatively, these results indicate that Cys48 is as important for differentiation and exit of
269 pluripotency as it is for reprogramming and its establishment, and suggest a physiological role for Oct4
270 redox regulation through Cys48 during peri-implantation development.

271 **EXPERIMENTAL PROCEDURES**

272

273 **Key Resources table**

274 Key Resources and their catalog numbers are provided in a separate table.

275

276 **Resource availability, lead contact, materials availability and data and code availability**

277 Further information and requests for reagents should be directed to and will be fulfilled by the lead contact,
278 Dean Tantin, PhD. (dean.tantin@path.utah.edu). All the materials generated in this study are accessible
279 upon request. The high-throughput data have been uploaded to the gene expression omnibus (GEO) server
280 under series record GSE241332. The data are accessible for download through private token
281 gtyfyseqljufjwz while in private status. This manuscript does not report original code.

282

283 **Cell culture**

284 ESCs/iPSCs were cultured as described previously[51,52] with "2iV" supplementation: the MEK inhibitor
285 PD03259010 (1 μM, LC laboratories) and the glycogen synthase kinase-3β (GSK3β) inhibitor
286 CHIR99021 (3 μM, LC Laboratories). Culture medium was also supplemented with 50 μg/mL ascorbic
287 acid (Sigma). H_2O_2 (Sigma) was supplied at 1 mM for 2 hr and MG-132 (Sigma) at 10 μM for 2 hr. CTV
288 proliferation assays were conducted using a kit (CTV Proliferation Kit for Flow Cytometry,
289 ThermoFisher). TUNEL assays were conducted using a kit (Apo-Direct, Becton-Dickinson).

290

291 **FLAG-tagging lentiviral constructs for iPSC generation**

292 The mSTEMCCA lentiviral vector expressing Oct4, Klf4, Sox2 and c-Myc was used for these studies.[53]
293 The Oct4 cDNA cassette was replaced with c-terminally FLAG-tagged full-length mouse Oct4, mouse
294 Oct1 or chimeric constructs. FLAG-tagged Oct4 was generated by overlap PCR using primers containing
295 the DYKDDDDK FLAG peptide sequence: (Fwd: 5'-

296 GATTACAAGGATGACGACGATAAGGGAAGTGGCGTGAAACAGA; Rev: 5'-
297 CTTATCGTCGTCATCCTTGTAATCGTTTGAATGCATGGGAGAG). The first fragment was
298 amplified using a Fwd primer containing a *Not*I restriction enzyme site (5'-
299 AATGAAAAAAGCGGCCGCCATGGCTGGACACCT) and overlapping Rev primer containing the
300 FLAG sequence. A second fragment was amplified using the Rev primer containing an *Nde*I restriction
301 site (5'-GGGAATTCCATATGTGTGGCCATATTATCATCGTGT) and overlapping Fwd primer. The
302 first and second fragments were mixed as template and amplified Fed and Rev primers containing *NotI*
303 and *Nde*I. After purification and digestion with *Not*I and *Nde*I , the fragment was ligated with the
304 mSTEMCCA vector digested by *Not*I and *Nde*I. Similar strategies were applied to generate other
305 constructs. For point mutation, a site-directed mutagenesis kit (New England Biolabs Q5) was used. When
306 making fusions between the different linker domains and POU_H, we maintained the "RKRKR" sequence
307 at the N-terminus of Oct4 POU_H, which has been shown to be important for reprogramming.[29]
308
309 **Lentiviral IRES-Oct1**
310 The *Oct1* (*Pou2f1*) cDNA and IRES (internal ribosomal entry site) elements were amplified and cloned
311 together by overlap PCR. In the first PCR, primers to the 5' end of *Oct1* containing a *Not*I restriction site
312 and to the 3' end of *Oct1* that contained a 5' extension of IRES-complementary DNA were used. The
313 sequences were: Oct1-NotI-For: 5'-AATGAAAAAAGCGGCCGCCATGAATAATCCATCAGAAAC-
314 3'; Oct1-Rev-IRES: 5'-TTAGGGGGGGGGGAGGGATCTTCACTGTGCCTTGGAG-3'. In the second
315 PCR, an IRES sequence was amplified using primers to the 5' end of the IRES containing a 5' extension
316 of DNA complementary of the *Oct1* 3' end, and primers to the 3' end of the IRES containing an *Nde*I
317 restriction site. The sequences were: IRES-overlap-For: 5'
318 AGATCCCTCCCCCCCCCCTAACGTTACTGGCCGAA-3'; IRES-Rev-NdeI: 5'-
319 GGGAATTCCATATGTGTGGCCATATTATCATCGTGT-3'. The third PCR used as a template the
320 PCR products from the first two rounds, and the Oct1-NotI-For and IRES-Rev-NdeI primers. This process

321 generated a DNA fragment containing an *Oct1* cDNA fused to an IRES at the 3' end, along with a *Not*I
322 site at the 5' terminus and an *Nde*I site at the 3' terminus. The fragment was cloned into the optimized,
323 self-inactivating, nonreplicative pHAGE lentiviral vector using the *Not*I and *Nde*I restriction sites. To
324 insert a *Puro* cassette after the IRES, the cDNA was amplified using primers containing 5' *Nde*I and 3'
325 *Cla*I restriction sites. The sequences were Puro-NdeI-For: 5'-
326 GGAATTCCATATGATGACCGAGTACAAGCCCACGGT-3'; Puro-ClaI-Rev: 5'
327 GGTTTATCGATTCAGGCACCGGGCTTGC-3'. Because the IRES apparently attenuated expression of
328 the *Puro* resistance cassette in this vector, puromycin selection was performed at 0.75 μg/mL. To generate
329 an empty vector control, the vector was cut with *Nde*I and *Not*I, filled in with Klenow fragment, and re-
330 ligated.
331
332 **iPSC generation**
333 293T cells were transfected with mSTEMCCA encoding FLAG-tagged Oct4, Oct1 and different domain-
334 swaps and mutants as well as the packaging plasmids (pMDLg/pRRE and pRSV-Rev) and envelope
335 plasmid (VSV-G) using polyethylenimine. Lentiviral supernatants were filtered and used to transduce
336 passage 3 primary MEFs expressing EGFP from the endogenous *Pou5f1* locus (Jackson Labs strain
337 #008214). Cells were cultured in ESC medium (above). GFP^+ iPSC colony emergence was quantified
338 after 12 days using an Olympus IX-71 epifluorescence microscope.
339
340 **Immunoblotting**
341 Antibodies for immunoblotting were: Oct4, α-tubulin and β-actin, Santa Cruz; Oct1, Bethyl; Nanog and
342 Sox2, GeneTex; GAPDH, EMD Millipore; FK2 (pan-Ub), Enzo Life Sciences; Ub-K48, Cell Signaling.
343
344 **Electrophoretic mobility shift assay**

345 EMSA was performed using published methods.[54,55] Diamide (Sigma) was added at final concentration of
346 up to 300 μM on ice during reaction assembly. 5'Cy5-labeled double-stranded DNA probes were as
347 follows: wild-type octamer, 5'-TGTCGAATGCAAATCACT-3'; mutant octamer, 5'-
348 TGTCGAATGCAAGCCACT-3'. Reactions were incubated for 30 min at room temperature prior to gel
349 loading. For experiments involving reversal of oxidative inhibition, thioredoxin (Sigma) was supplied at
350 a final concentration of 0 to 60 μM after 15 min incubation with diamide, and both thioredoxin-treated
351 and control samples were incubated for a further 15 min at 37°C prior to gel loading. DNA was visualized
352 using a Molecular Dynamics Typhoon system.
353
354 **Recombinant Oct4 protein**
355 Wild-type and C48S Oct4 constructs were generated bearing dual C-terminal FLAG and Twin-Strep tags.
356 A g-block (IDT) fragment coding for the Twin-Strep and FLAG epitope tags was inserted between the
357 *Nco*I and *Kpn*I sites in the transient expression vector pACE-MAM2 (Geneva Biotech). The resultant
358 construct was then digested with *Nco*I and *Pvu*II to insert the g-block fragment coding for either wild-type
359 or C48S murine Oct4. After Sanger sequencing to confirm veracity, the plasmids were transfected into
360 Expi293F cells (ThermoFisher) using ExpiFectamine 293 Tranfection Kit (ThermoFisher) following the
361 manufacturer instructions. After 2 days, transfected cells were harvested and washed with ice-cold PBS.
362 Cells were resuspended in Buffer A (20 mM HEPES pH 8.0, 1.5 mM $MgCl_2$, 10 mM KCl, 0.25% NP-40,
363 0.5 mM DTT, 2x protease inhibitor cocktail/Roche), incubated for 10 min on ice, and homogenized using
364 a Dounce homogenizer (Wheaton). The homogenate was centrifuged at 4°C for 5 min at 3,800×*g*). Pelleted
365 nuclei were resuspended in Buffer C (20 mM HEPES pH 8.0, 25% glycerol, 1.5 mM $MgCl_2$, 420 mM
366 KCl, 0.25% NP-40, 0.2 mM EDTA, 0.5 mM DTT, 2x protease inhibitor cocktail) and Dounce
367 homogenized again. Nuclei were extracted for 30 min at 4°C with end-to-end rotation, then centrifuged at
368 4°C for 30 min at 20,000×*g*. The nuclear extract was then subjected to immunoprecipitation with anti-

369 FLAG (M2)-agarose beads (Sigma). Bead-bound proteins were eluted with 200 ng/μL 3×FLAG peptide
370 (Sigma) by incubating at 4°C for 30 min. The protein was pure and monodisperse as verified using
371 Coomassie blue staining and size exclusion chromatography.
372

373 **Generation of Oct4^{C48S} mouse ESCs**

374 Mutant ESCs with a single-base missense mutation that converts Cys48 to Ser were generated using
375 templated CRISPR. A modified synthetic sgRNA (Synthego) was designed to target Cas9 and cutting near
376 the target nucleotide, 5' GGCCTCGAAGCGA<u>C</u>AGATGG 3', where the underlined C indicates the target
377 nucleotide. The sgRNA was complexed with Cas9 protein (IDT) to form a ribonucleoprotein (RNP)
378 complex that was electroporated with program CG-104 into 6×10^4 mouse ESCs using a LONZA 4D-
379 Nucleofector X-unit. The RNP was co-electroporated with a corresponding single-stranded
380 oligodeoxynucleotide (ssODN) with 5' and 3' phosphorothioate modifications, 5'-
381 ctcactctgtttgatcggcctttcagGAAAGGTGTTCAGCCAGACCA**CCATCT<u>C</u>T<u>A</u>G<u>A</u>TTCGAGGCC**TTGC
382 AGCTCAGCCTTAAGAACATGTGTAAGCTG-3' where lower-case denotes intronic sequence and the
383 sgRNA site is in bold. The ssODN contained the target mutation (underlined), silent mutations in the seed
384 region (double underlined) to block CRISPR re-cutting after incorporation of the donor and to generate a
385 unique *Xba*I restriction enzyme site for screening purposes. Incorporation of the ssODN sequence was
386 identified via PCR and restriction enzyme digestion, and confirmed by sequencing of the 709 bp region
387 flanking the mutation. Mutant clones were generated from a pure C57BL/6 ESC line, a 129/BL6 hybrid
388 line expressing constitutive GFP and a reporter line used to de-differentiate ESCs into a 2C-like state.
389 Clones were tested by karyotype and any clones with <70% normal diploid content were excluded.
390

391 **RNA-seq**

392 ESCs were first depleted by feeder cells by incubation on plates coated with gelatin. Parental and derived
393 Oct4^{C48S} ESC lines were differentiated in media lacking LIF and containing 1.5mg/L RA for 2 days.

394 Undifferentiated and RA-differentiated cells were collected, and RNA was prepared using a kit (RNeasy,
395 Qiagen). RNA poly-A selected for mRNA using NEBNext rRNA Depletion Kit v2 (Human, Mouse, Rat)
396 (E7400). Bulk RNA was processed and sequenced as reported previously.[23] Differential genes were
397 identified using 2.5-fold change and $p<0.001$ as cutoffs. GO terms were identified using PANTHER.[56]
398

399 **Teratoma formation**

400 Feeder-depleted parental and derived Oct4^{C48S} ESC lines were injected subcutaneously into contralateral
401 flanks of 6-8 week-old female NOD.*Rag1*$^{-/-}$*Il2rg*$^{-/-}$ (NRG, Jackson labs) mice to generate teratomas. 1×10^6
402 cells were implanted. After two weeks, mice were humanely euthanized, and tumors were excised. After
403 fixation, embedding and sectioning, slides were stained with H&E and evaluated in a blinded fashion.
404

405 **Chimeric mouse assays**

406 Targeted or parental ESCs were collected and resuspended in injection medium consisting of DMEM with
407 HEPES (ThermoFisher) supplemented with 10% FBS (Peak Serum) and 0.1 mM 2-mercaptoethanol
408 (Sigma). C57B6 clones were injected into albino B6 blastocysts, while the G4N and C4 clones (B6:129
409 and 129 in origin, respectively) were injected into C57B/6 blastocysts. E3.5 blastocysts were used. 10-25
410 ESCs were injected into each blastocyst using micromanipulators (Eppendorf) and a Leica inverted
411 microscope. After culturing injected blastocysts in KSOM (CytoSpring), approximately 12 to 18
412 blastocysts were surgically transferred into the uterine horns of day 2.5 pseudopregnant mice or the
413 oviduct of day 0.5 pseudopregnant mice. Mice were scored based on coat and eye color by members of
414 the transgenic core facility at weaning.

415 **ACKNOWLEDGEMENTS**

416 We thank S. Buckley for critical reading of the manuscript. We thank C. Davey and the University of Utah
417 Health Sciences Center Mouse Transgenic Core and Mutation Generation Core Facilities for assistance
418 with generating CRISPR-edited ESC lines, chimerism assays and karyotyping. We thank B. Dalley and
419 the High-Throughput Genomics core facility, and Q. Li and the Bioinformatics Analysis core facility for
420 assistance with RNA-seq. We thank D. Lum and the university Preclinical Research Resource core facility
421 for assistance with teratoma formation. We thank B. Weaver and B. Cairns for the gift of the C4 ESC line.
422 This work was supported by NIH/NIGMS grant (R01GM122778) to DT.

423

424 **AUTHOR CONTRIBUTIONS**

425 D.T. conceived the study. Z.S., Y.W. and D.T. designed experiments. Z.S., Y.W. and M.B.C. generated
426 new reagents for the study. Z.S., Y.W., A.M. and K.J.E. performed experiments and/or analyzed data.
427 K.J.E. performed blinded pathological analysis of teratoma samples. generated figures. All authors
428 contributed to the writing of the manuscript and read and approved the final version of the manuscript
429 prior to submission for publication.

430 **FIGURE LEGENDS**

431

432 **Figure 1. Oct4 reprogramming activity vis-à-vis Oct1 maps to the DBD**

433 (A) Mouse Oct1 vs Oct4 DBD alignment. POU_S and POU_H refer to the two DNA binding sub-domains.
434 Asparagine 40 (blue) was mutated to lysine in some constructs. The key Oct1 serine 48 residue (cysteine
435 in Oct4) is boxed in red.

436 (B) Schematic of the reprogramming vector with C-terminal Oct4 FLAG tag. The different domains of
437 Oct4 and Oct1 are color coded.

438 (C) Reprogramming activity of C-terminally tagged Oct4 relative to the untagged control. Results show
439 an average of N=3 independent experiments. Error bars show ±standard deviation. Anti-FLAG and anti-
440 Oct4 immunoblots using tagged and untagged constructs transiently transfected into 293T cells are shown
441 at right.

442 (D) iPSC formation assay using target Oct4-GFP MEFs. Reprogramming was assessed using the number
443 of GFP-expressing iPSC colonies. Oct4 was used as a positive control and Oct1 as a negative control. All
444 constructs bore C-terminal FLAG-tags. N=3 experiments were performed. Error bars show ±standard
445 deviation.

446 (E) Schematic and summary of size and activity of the constructs used in D. The activity of Oct4 in D was
447 set to 100%. C-terminal FLAG-tags are not shown for simplicity.

448 (F) Anti-FLAG immunoblot showing expression of the different constructs when transiently transfected
449 into 293T cells transduced with the reprogramming vectors. Histone H3 is shown as a loading control.

450

451 **Figure 2. Activity of Oct4 cysteine 48 in reprogramming**

452 (A) iPSC formation assay using Oct1-Oct4 lentiviral constructs that manipulate cysteine 4. Oct4 was used
453 as a positive control and Oct1 as a negative control. All constructs bore C-terminal FLAG-tags. N=3
454 experiments. Error bars show ±standard deviation.

455 (B) Schematic and summary of size and activity of the constructs used in A. The activity of Oct4 in A was
456 set to 100%.
457 (C) Anti-FLAG immunoblot showing expression of the different constructs when transiently transfected
458 into 293T cells transduced with the reprogramming vectors. Histone H3 is shown as a loading control.
459 (D) Detail from an Oct1:octamer DNA co-crystal structure (PDB ID: 1OCT)[37] showing Ser48 making
460 hydrogen bond contacts with the DNA backbone.
461 (E) Paralog comparison of mouse POU protein amino acid residues corresponding to α-helix 3 of the
462 POU_S sub-domain.
463
464 **Figure 3. Oct4 cysteine 48 confers DNA sensitivity to oxidative stress and oxidative stress-induced**
465 **ubiquitylation**
466 (A) EMSA using canonical octamer DNA and wild-type of C48S full-length Oct4. 293T cells were
467 transiently transfected with plasmids encoding C-terminal FLAG-tagged Oct4 constructs. Protein lysates
468 were made and immediately used for EMSA with increasing amounts of diamide. The larger redox-
469 resistant Oct1 protein that binds wild-type (WT) but not mutant (Mut) probe is highlighted with an arrow.
470 Asterisk shows a redox sensitive band with superior binding to mutant rather than octamer DNA.
471 (B) Quantification of Oct4 band intensity relative to no diamide, which was set to 100%. N=3 independent
472 experiments. Error bars show ±standard deviation.
473 (C) Oct4 containing a C-terminal streptavidin tag was purified from 293T cell lysates using anti-FLAG
474 beads. Isolated protein was used in assays similar to A.
475 (D) Quantification of C and two similar experiments. Error bars show ±standard deviation.
476 (E) Highly exposed Oct4 immunoblot using lysates prepared from 293T cells transfected with constructs
477 expressing C-terminally FLAG/Twin-Strep-tagged Oct4. Subsets of cells were pre-treated with H_2O_2 (1
478 mM for 2 hr), MG-132 (10 μM for 2 hr) or both. All samples contained MG-132 in the lysis buffer to
479 prevent degradation. Histone H3 is shown as a loading control.

480 (F) Purified wild-type and C48S Oct4 was immunoblotted using antibodies against specific ubiquitin
481 linkages.
482 (G) Oct4 and control immunoprecipitates from wild-type ESCs were immunoblotted using FK2
483 antibodies. Cells were treated with 10 μM MG-132 for 2 hr or 2.5 mM H_2O_2 for 2 hr.
484
485 **Figure 4. Oct4^{C48S} ESCs are phenotypically normal but retain Oct4 expression when differentiated**
486 (A) *Top:* Schematic for CRISPR-mediated *Pou5f1* (*Oct4*) mutation. A single nucleotide change (arrow)
487 converts the wild-type cysteine to the serine present in Oct1. Asterisks depict two silent point mutations
488 present in the guide RNA to monitor mutagenesis. *Bottom:* DNA sequencing trace of homozygous mutant
489 clone.
490 (B) An example C57BL/6 ESC clone (#46) was immunoblotted for Oct4, Sox2, Nanog and Histone H3
491 as a loading control.
492 (C) Example images of pluripotent colonies from a parental ESC line and derived Oct4^{C48S} ESC line (#46).
493 Images were collected at 10× magnification. Scale bar = 0.2 mm.
494 (D) Pan-Ub and control Oct4 immunoblots of Oct4 immunoprecipitates from parent and Oct4^{C48S} (clone
495 #2) ESC lines. Cells were grown on gelatin without feeders. Normal culture medium was replaced with
496 medium lacking supplemental glutamine and β-mercaptoethanol 2 hr prior to the experiment. Where
497 indicated, cells were treated with 2.5 mM H_2O_2 for a further 2 hr. Oct4 was immunoprecipitated using a
498 rabbit polyclonal antibody (Abcam) and immunoblotted using monoclonal mouse antibodies against pan-
499 Ub and Oct4.
500 (E) Parent or derived Oct4^{C48S} ESCs were differentiated using RA. Lysates were collected at the indicated
501 times and probed for Oct4 by immunoblotting. β-actin was used as a loading control.
502 (F) Average quantified Oct4/β-actin levels from N=3 immunoblots conducted with mutant ESC clone
503 #46. Error bars depict ±standard deviation.

504 (G) mRNA from parental and #46 mutant cells differentiated for 2 days with RA, and undifferentiated
505 controls, were transcriptionally profiled using RNA-seq. Significantly increased genes in differentiating
506 C48S relative to parent cells (421 genes) were intersected with the set of genes down regulated in parental
507 cells upon differentiation (1536 genes), to yield 169 genes.
508 (H) PANTHER GO terms associated with 169 genes from G identified using the "biological process"
509 ontology.
510
511 **Figure 5. Oct4^{C48S} ESCs differentiate abnormally**
512 (A) Phase-contrast images of parent and derived Oct4^{C48S} ESCs (clone #46) differentiated using RA for
513 12 days. Images were collected at the indicated times at 10× magnification. Scale bar = 0.2 mm.
514 (B) Differentiating cells were trypsinized and counted using trypan blue exclusion using a hemacytometer.
515 Cell counts (N=3) were averaged and plotted. Error bars depict ±standard deviation.
516 (C) Proliferation was measured in parent and C48S ESCs differentiating with RA at days 2, 3 and 4. Cells
517 were loaded with cell trace violet (CTV) at day 2.
518 (D) TUNEL assay showing Oct4^{C48S} ESCs differentiated using retinoic acid. Cells were differentiated for
519 2 days and incubated with FITC-labeled dUTP for 1 hour at 37°C prior to flow cytometry.
520
521 **Figure 6. Oct4^{C48S} ESCs form teratomas lacking maturing structures and contribute poorly to adult**
522 **chimeras**
523 (A) 1×10^6 parental or Oct4^{C48S} C57BL/6:129 ESCs (clone G4-N) were injected into contralateral flanks
524 or NRG mice. After two weeks, tumors were harvested, fixed, sectioned, stained with H&E, deidentified
525 and scored pathologically. The percent of cells within the tumor containing undifferentiated components
526 was averaged and plotted. N=5 independent tumors for each condition combined from two separate
527 experiments. For each tumor, values represent the average from a pair of slides. Two nonadjacent sections

528 from each tumor were used. Error bars depict ±standard deviation. *P*-value was generated using an
529 unpaired one-tailed T-test.
530 (B) Representative images are shown. Mature (differentiated) components, e.g. airway epithelium in
531 teratomas derived from parental ESCs are highlighted by yellow arrows. Asterisk indicates area of
532 necrosis.
533 (C) N=3 different parent (left side) or CRISPR-mutated Oct4^{C48S} ESCs (right side) were injected into
534 recipient blastocysts and implanted in pseudo-pregnant animals. For each line, the proportion of chimeric
535 offspring at weaning is shown as a pie chart. All karyotypes were confirmed to be similar to parent lines
536 (between 72% and 92% normal).
537
538 **SUPPLEMENTAL ITEM LEGENDS**
539
540 **Figure S1. Effect of N-terminal FLAG-tag on iPSC formation**
541 (A) Similar to Figure 1C, except cells were transduced with lentiviral vectors expressing untagged and N-
542 terminal rather than C-terminal FLAG-tagged Oct4.
543 (B) Anti-FLAG and Oct4 immunoblots showing expression of tagged and untagged Oct4. α-tubulin is
544 shown as a loading control.
545
546 **Figure S2. iPSC colony morphology and GFP intensity using different chimeric and mutant**
547 **reprogramming constructs**
548 Phase-contrast and epifluorescence images are shown of example iPSC colonies generated using the
549 constructs shown in Figure 2. Scale bar=200 μm
550
551 **Figure S3. GFP^{+} iPSC numbers over time using different Oct1-Oct4 chimeric constructs**

552 Similar to Figure 2A except multiple time points are shown. Wild-type Oct4 is shown as a positive control.
553 Results show an average of experiments performed in biological triplicate. Error bars show ±standard
554 deviation.
555
556 **Figure S4. Oxidative inhibition of Oct4 DNA binding can be reversed with the addition of**
557 **thioredoxin**
558 EMSA is shown similar to Figure 3C, except that after 15 min incubation with diamide, thioredoxin was
559 added at the indicated concentrations. Reactions were incubated for a further 15 min at 37°C prior to
560 native electrophoresis.
561
562 **Figure S5. Purification of recombinant wild-type and C48S mutant Oct4 from expi-293T cells**
563 Oct4 was FLAG-tagged at the C-terminus. Protein was purified using beads covalently coupled to anti-
564 FLAG antibodies. Lysates were prepared 48 hr post-transfection.
565
566 **Figure S6. H_2O_2 and MG-132 treatment of primary MEFs largely maintains viability**
567 Phase-contrast images are shown of Oct4-GFP MEFs treated with the same concentrations and timepoints
568 of H_2O_2 and MG-132 used in Figure 3E. Scale bar=200 μm.
569
570 **Figure S7. H_2O_2 and MG-132 treatment of primary MEFs largely maintains viability**
571 Similar to Figure 3F except additional antibodies were used for immunoblotting: an antibody detecting
572 K63-linked Ub chains (Cell Signaling) and a pan-Oct4 antibody as a positive control (Enzo Life Sciences).
573 Oct4 was Twin-Strep- and FLAG-tagged at the C-terminus.
574
575 **Figure S8. The effects of Oct4 C48S mutation are not parent line- or clone-specific**
576 (A) Similar to Figure 4B except a different clone from the same parent line (#2) was tested.

577 (B) Similar to Figure 4E except a different clone from the same parent line (#2) was tested.

578 (C) Similar to Figure 4E except a different clone (#18) from a different parent line (G4-N) was tested.

579

580 **Figure S9. RNA-seq analysis of differentiating parental and Oct4^{C48S} ESCs**

581 (A) Similarity between bulk RNA-seq replicates and conditions. Clustering of the RNA-seq replicates

582 using Euclidean distance is shown.

583 (B) The union of the top 100 most differentially expressed genes upon differentiation for both genotypes

584 (150 genes total) was subjected to hierarchical clustering and displayed as a heat map. Example genes are

585 shown at right.

586

587 **Figure S10. Additional tumor images similar to Fig.6D**

588 Asterisk represents necrotic regions.

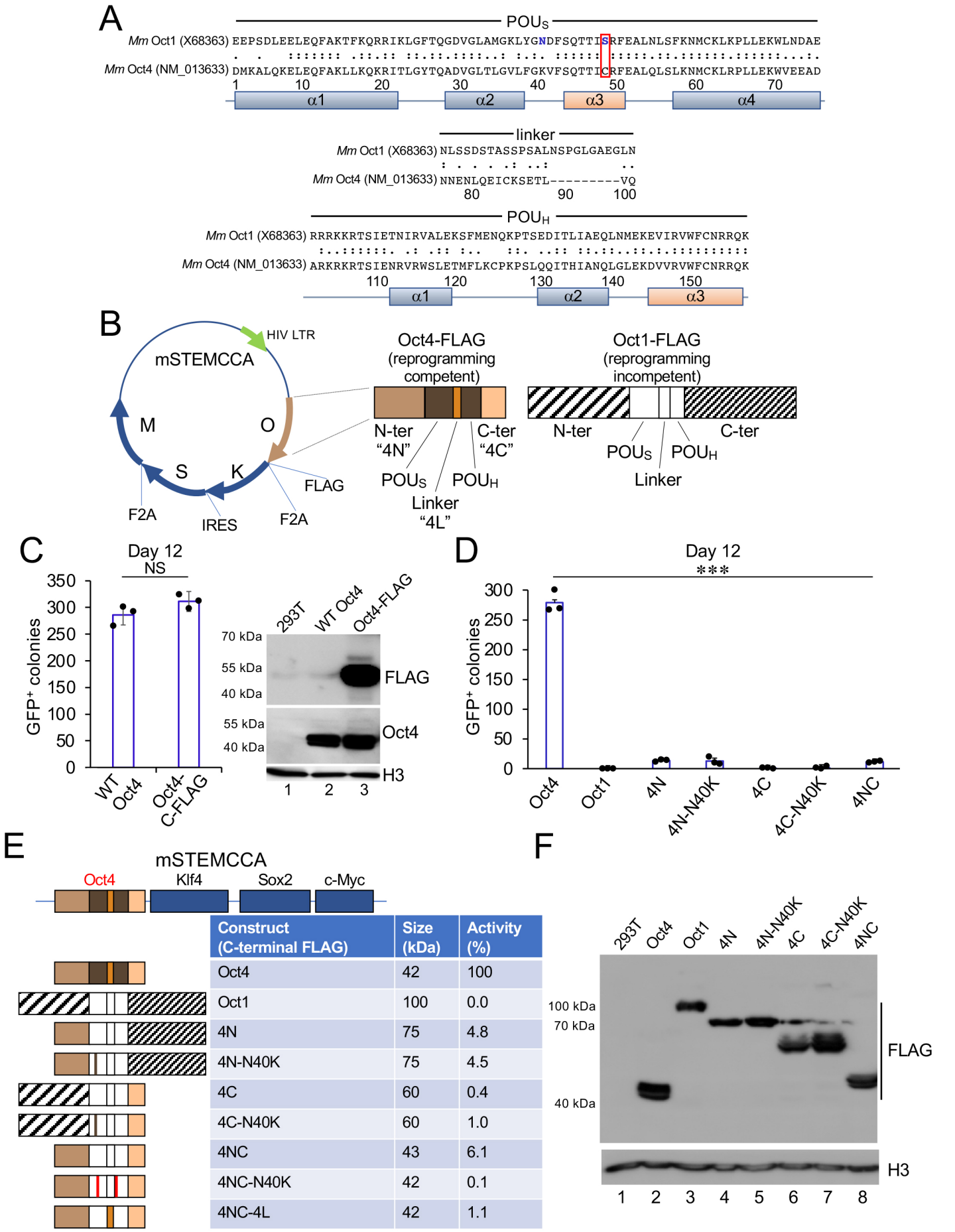
589 **REFERENCES**

590 1. Abranches, E., Silva, M., Pradier, L., Schulz, H., Hummel, O., Henrique, D., and Bekman, E.
591 (2009). Neural differentiation of embryonic stem cells in vitro: a road map to neurogenesis in the
592 embryo. PLoS ONE *4*, e6286. 10.1371/journal.pone.0006286 [doi].
593 2. Morey, L., Santanach, A., and Di Croce, L. (2015). Pluripotency and Epigenetic Factors in Mouse
594 Embryonic Stem Cell Fate Regulation. Mol Cell Biol *35*, 2716-2728.
595 3. Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer,
596 H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends
597 on the POU transcription factor Oct4. Cell *95*, 379-391.
598 4. Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G.,
599 Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry
600 in human embryonic stem cells. Cell *122*, 947-956.
601 5. Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse
602 embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676. DOI
603 10.1016/j.cell.2006.07.024.
604 6. Li, Y., Zhang, Q., Yin, X., Yang, W., Du, Y., Hou, P., Ge, J., Liu, C., Zhang, W., Zhang, X., et al.
605 (2011). Generation of iPSCs from mouse fibroblasts with a single gene, Oct4, and small molecules.
606 Cell Res *21*, 196-204. cr2010142 [pii] 10.1038/cr.2010.142 [doi].
607 7. Gao, Y., Chen, J., Li, K., Wu, T., Huang, B., Liu, W., Kou, X., Zhang, Y., Huang, H., Jiang, Y.,
608 et al. (2013). Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of
609 DNA methylation and hydroxymethylation in reprogramming. Cell Stem Cell *12*, 453-469.
610 8. Heng, J.C., Feng, B., Han, J., Jiang, J., Kraus, P., Ng, J.H., Orlov, Y.L., Huss, M., Yang, L., Lufkin,
611 T., et al. (2010). The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine
612 somatic cells to pluripotent cells. Cell Stem Cell *6*, 167-174. 10.1016/j.stem.2009.12.009.
613 9. Esteban, M.A., Wang, T., Qin, B., Yang, J., Qin, D., Cai, J., Li, W., Weng, Z., Chen, J., Ni, S., et
614 al. (2010). Vitamin C enhances the generation of mouse and human induced pluripotent stem cells.
615 Cell Stem Cell *6*, 71-79. S1934-5909(09)00624-9 [pii] 10.1016/j.stem.2009.12.001 [doi].
616 10. Rais, Y., Zviran, A., Geula, S., Gafni, O., Chomsky, E., Viukov, S., Mansour, A.A., Caspi, I.,
617 Krupalnik, V., Zerbib, M., et al. (2013). Deterministic direct reprogramming of somatic cells to
618 pluripotency. Nature *502*, 65-70. 10.1038/nature12587.
619 11. Costa, Y., Ding, J., Theunissen, T.W., Faiola, F., Hore, T.A., Shliaha, P.V., Fidalgo, M., Saunders,
620 A., Lawrence, M., Dietmann, S., et al. (2013). NANOG-dependent function of TET1 and TET2 in
621 establishment of pluripotency. Nature *495*, 370-374. nature11925 [pii] 10.1038/nature11925 [doi].
622 12. Vierbuchen, T., and Wernig, M. (2012). Molecular roadblocks for cellular reprogramming. Mol
623 Cell *47*, 827-838. S1097-2765(12)00784-8 [pii] 10.1016/j.molcel.2012.09.008 [doi].
624 13. Soufi, A., Donahue, G., and Zaret, K.S. (2012). Facilitators and impediments of the pluripotency
625 reprogramming factors' initial engagement with the genome. Cell *151*, 994-1004.
626 14. Buganim, Y., Markoulaki, S., van Wietmarschen, N., Hoke, H., Wu, T., Ganz, K., Akhtar-Zaidi,
627 B., He, Y., Abraham, B.J., Porubsky, D., et al. (2014). The developmental potential of iPSCs is
628 greatly influenced by reprogramming factor selection. Cell Stem Cell *15*, 295-309.
629 10.1016/j.stem.2014.07.003.
630 15. Chen, J., Gao, Y., Huang, H., Xu, K., Chen, X., Jiang, Y., Li, H., Gao, S., Tao, Y., Wang, H., et
631 al. (2015). The combination of Tet1 with Oct4 generates high-quality mouse-induced pluripotent
632 stem cells. Stem Cells *33*, 686-698. 10.1002/stem.1879.
633 16. Yuan, X., Wan, H., Zhao, X., Zhu, S., Zhou, Q., and Ding, S. (2011). Combined Chemical
634 Treatment Enables Oct4-Induced Reprogramming from Mouse Embryonic Fibroblasts. Stem Cells
635 *29*, 549-553. 10.1002/stem.594 [doi].

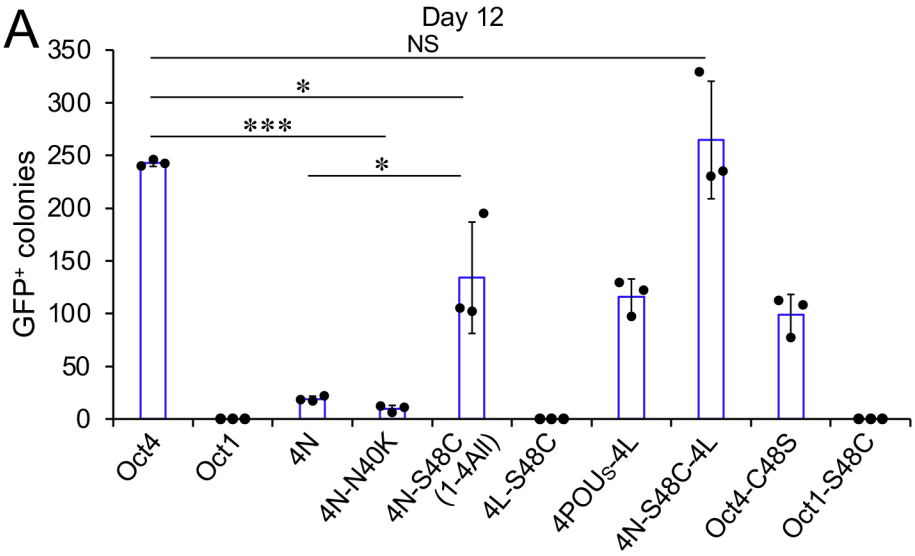
636 17. Velychko, S., Adachi, K., Kim, K.P., Hou, Y., MacCarthy, C.M., Wu, G., and Scholer, H.R.
637 (2019). Excluding Oct4 from Yamanaka Cocktail Unleashes the Developmental Potential of
638 iPSCs. Cell Stem Cell *25*, 737-753. 10.1016/j.stem.2019.10.002.
639 18. Okamoto, K., Okazawa, H., Okuda, A., Sakai, M., Muramatsu, M., and Hamada, H. (1990). A
640 novel octamer binding transcription factor is differentially expressed in mouse embryonic cells.
641 Cell *60*, 461-472.
642 19. Rosner, M.H., Vigano, M.A., Ozato, K., Timmons, P.M., Poirier, F., Rigby, P.W., and Staudt,
643 L.M. (1990). A POU-domain transcription factor in early stem cells and germ cells of the
644 mammalian embryo. Nature *345*, 686-692. 10.1038/345686a0 [doi].
645 20. Suzuki, N., Rohdewohld, H., Neuman, T., Gruss, P., and Scholer, H.R. (1990). Oct-6: a POU
646 transcription factor expressed in embryonal stem cells and in the developing brain. EMBO J *9*,
647 3723-3732.
648 21. Tantin, D. (2013). Oct transcription factors in development and stem cells: insights and
649 mechanisms. Development *140*, 2857-2866. 10.1242/dev.095927.
650 22. Ferraris, L., Stewart, A.P., Kang, J., DeSimone, A.M., Gemberling, M., Tantin, D., and
651 Fairbrother, W.G. (2011). Combinatorial binding of transcription factors in the pluripotency
652 regions of the genome. Genome Res. *21*, 1055-1064.
653 23. Perovanovic, J., Wu, Y., Abewe, H., Shen, Z., Hughes, E.P., Gertz, J., Chandrasekharan, M.B.,
654 and Tantin, D. (2023). Oct1 cooperates with the Smad family of transcription factors to promote
655 mesodermal lineage specification. Sci. Signal. *16*, eadd5750. 10.1126/scisignal.add5750.
656 24. Feng, B., Jiang, J., Kraus, P., Ng, J.H., Heng, J.C., Chan, Y.S., Yaw, L.P., Zhang, W., Loh, Y.H.,
657 Han, J., et al. (2009). Reprogramming of fibroblasts into induced pluripotent stem cells with
658 orphan nuclear receptor Esrrb. Nat Cell Biol *11*, 197-203.
659 25. Kim, K.P., Choi, J., Yoon, J., Bruder, J.M., Shin, B., Kim, J., Arauzo-Bravo, M.J., Han, D., Wu,
660 G., Han, D.W., et al. (2021). Permissive epigenomes endow reprogramming competence to
661 transcriptional regulators. Nat Chem Biol *17*, 47-56. 10.1038/s41589-020-0618-6.
662 26. Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K.,
663 Mochiduki, Y., Takizawa, N., and Yamanaka, S. (2008). Generation of induced pluripotent stem
664 cells without Myc from mouse and human fibroblasts. Nat Biotechnol *26*, 101-106.
665 27. Shi, Y., Desponts, C., Do, J.T., Hahm, H.S., Scholer, H.R., and Ding, S. (2008). Induction of
666 pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule
667 compounds. Cell Stem Cell *3*, 568-574.
668 28. Esch, D., Vahokoski, J., Groves, M.R., Pogenberg, V., Cojocaru, V., Vom Bruch, H., Han, D.,
669 Drexler, H.C., Arauzo-Bravo, M.J., Ng, C.K., et al. (2013). A unique Oct4 interface is crucial for
670 reprogramming to pluripotency. Nat Cell Biol *15*, 295-301. ncb2680 [pii] 10.1038/ncb2680 [doi].
671 29. Jerabek, S., Ng, C.K., Wu, G., Arauzo-Bravo, M.J., Kim, K.P., Esch, D., Malik, V., Chen, Y.,
672 Velychko, S., MacCarthy, C.M., et al. (2017). Changing POU dimerization preferences converts
673 Oct6 into a pluripotency inducer. EMBO Rep *18*, 319-333.
674 30. Jin, W., Wang, L., Zhu, F., Tan, W., Lin, W., Chen, D., Sun, Q., and Xia, Z. (2016). Critical POU
675 domain residues confer Oct4 uniqueness in somatic cell reprogramming. Sci Rep *6*, 20818.
676 10.1038/srep20818.
677 31. An, Z., Liu, P., Zheng, J., Si, C., Li, T., Chen, Y., Ma, T., Zhang, M.Q., Zhou, Q., and Ding, S.
678 (2019). Sox2 and Klf4 as the Functional Core in Pluripotency Induction Without Exogenous Oct4.
679 Cell Rep *29*, 1986-2000. 10.1016/j.celrep.2019.10.026.
680 32. Lengner, C.J., Camargo, F.D., Hochedlinger, K., Welstead, G.G., Zaidi, S., Gokhale, S., Scholer,
681 H.R., Tomilin, A., and Jaenisch, R. (2007). Oct4 expression is not required for mouse somatic
682 stem cell self-renewal. Cell Stem Cell *1*, 403-415. 10.1016/j.stem.2007.07.020 [doi].

683 33. Sommer, C.A., Stadtfeld, M., Murphy, G.J., Hochedlinger, K., Kotton, D.N., and Mostoslavsky,
684 G. (2009). Induced pluripotent stem cell generation using a single lentiviral stem cell cassette.
685 Stem Cells *27*, 543-549.
686 34. Boija, A., Klein, I.A., Sabari, B.R., Dall'Agnese, A., Coffey, E.L., Zamudio, A.V., Li, C.H.,
687 Shrinivas, K., Manteiga, J.C., Hannett, N.M., et al. (2018). Transcription Factors Activate Genes
688 through the Phase-Separation Capacity of Their Activation Domains. Cell *175*, 1842-1855 e1816.
689 35. Kim, K.P., Wu, Y., Yoon, J., Adachi, K., Wu, G., Velychko, S., MacCarthy, C.M., Shin, B., Ropke,
690 A., Arauzo-Bravo, M.J., et al. (2020). Reprogramming competence of OCT factors is determined
691 by transactivation domains. Sci Adv *6*, eaaz7364. 10.1126/sciadv.aaz7364.
692 36. Han, D., Wu, G., Chen, R., Drexler, H.C.A., MacCarthy, C.M., Kim, K.P., Adachi, K., Gerovska,
693 D., Mavrommatis, L., Bedzhov, I., et al. (2022). A balanced Oct4 interactome is crucial for
694 maintaining pluripotency. Sci Adv *8*, eabe4375. 10.1126/sciadv.abe4375.
695 37. Klemm, J.D., Rould, M.A., Aurora, R., Herr, W., and Pabo, C.O. (1994). Crystal structure of the
696 Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding
697 modules. Cell *77*, 21-32.
698 38. Michael, A.K., Grand, R.S., Isbel, L., Cavadini, S., Kozicka, Z., Kempf, G., Bunker, R.D., Schenk,
699 A.D., Graff-Meyer, A., Pathare, G.R., et al. (2020). Mechanisms of OCT4-SOX2 motif readout on
700 nucleosomes. Science *368*, 1460-1465. 10.1126/science.abb0074.
701 39. Marsboom, G., Zhang, G.F., Pohl-Avila, N., Zhang, Y., Yuan, Y., Kang, H., Hao, B.,
702 Brunengraber, H., Malik, A.B., and Rehman, J. (2016). Glutamine Metabolism Regulates the
703 Pluripotency Transcription Factor Oct4. Cell Rep *16*, 323-332. 10.1016/j.celrep.2016.05.089.
704 40. Lickteig, K., Lamb, K., Brigman, K., and Rizzino, A. (1996). Effects of oxidation and reduction
705 on the binding of transcription factors to cis-regulatory elements located in the FGF-4 gene. Mol
706 Reprod Dev *44*, 146-152.
707 41. Guo, Y., Einhorn, L., Kelley, M., Hirota, K., Yodoi, J., Reinbold, R., Scholer, H., Ramsey, H., and
708 Hromas, R. (2004). Redox regulation of the embryonic stem cell transcription factor Oct-4 by
709 thioredoxin. Stem Cells *22*, 259-264.
710 42. King, H.W., and Klose, R.J. (2017). The pioneer factor OCT4 requires the chromatin remodeller
711 BRG1 to support gene regulatory element function in mouse embryonic stem cells. Elife *6*.
712 10.7554/eLife.22631.
713 43. Aksoy, I., Jauch, R., Chen, J., Dyla, M., Divakar, U., Bogu, G.K., Teo, R., Leng Ng, C.K., Herath,
714 W., Lili, S., et al. (2013). Oct4 switches partnering from Sox2 to Sox17 to reinterpret the enhancer
715 code and specify endoderm. EMBO J *32*, 938-953. emboj201331 [pii] 10.1038/emboj.2013.31
716 [doi].
717 44. Wang, Z., Oron, E., Nelson, B., Razis, S., and Ivanova, N. (2012). Distinct lineage specification
718 roles for NANOG, OCT4, and SOX2 in human embryonic stem cells. Cell Stem Cell *10*, 440-454.
719 10.1016/j.stem.2012.02.016.
720 45. DeVeale, B., Brokhman, I., Mohseni, P., Babak, T., Yoon, C., Lin, A., Onishi, K., Tomilin, A.,
721 Pevny, L., Zandstra, P.W., et al. (2013). Oct4 is required ~E7.5 for proliferation in the primitive
722 streak. PLoS Genet *9*, e1003957. 10.1371/journal.pgen.1003957.
723 46. De Los Angeles, A., Ferrari, F., Xi, R., Fujiwara, Y., Benvenisty, N., Deng, H., Hochedlinger, K.,
724 Jaenisch, R., Lee, S., Leitch, H.G., et al. (2015). Hallmarks of pluripotency. Nature *525*, 469-478.
725 47. Shen, Z., Kang, J., Shakya, A., Tabaka, M., Jarboe, E.A., Regev, A., and Tantin, D. (2017).
726 Enforcement of developmental lineage specificity by transcription factor Oct1. ELife *6*, e20937.
727 10.7554/eLife.20937.
728 48. Ruetz, T., Pfisterer, U., Di Stefano, B., Ashmore, J., Beniazza, M., Tian, T.V., Kaemena, D.F.,
729 Tosti, L., Tan, W., Manning, J.R., et al. (2017). Constitutively Active SMAD2/3 Are Broad-Scope
730 Potentiators of Transcription-Factor-Mediated Cellular Reprogramming. Cell Stem Cell *21*, 791-
731 805 e799. 10.1016/j.stem.2017.10.013.

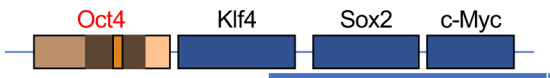
732 49. Britto, P.J., Knipling, L., and Wolff, J. (2002). The local electrostatic environment determines
733 cysteine reactivity of tubulin. J Biol Chem *277*, 29018-29027. 10.1074/jbc.M204263200.
734 50. Mamun, M.M.A., Khan, M.R., Zhu, Y., Zhang, Y., Zhou, S., Xu, R., Bukhari, I., Thorne, R.F., Li,
735 J., Zhang, X.D., et al. (2022). Stub1 maintains proteostasis of master transcription factors in
736 embryonic stem cells. Cell Rep *39*, 110919. 10.1016/j.celrep.2022.110919.
737 51. Shen, Z., Formosa, T., and Tantin, D. (2018). FACT Inhibition Blocks Induction But Not
738 Maintenance of Pluripotency. Stem Cells Dev *27*, 1693-1701. 10.1089/scd.2018.0150.
739 52. Shakya, A., Callister, C., Goren, A., Yosef, N., Garg, N., Khoddami, V., Nix, D., Regev, A., and
740 Tantin, D. (2015). Pluripotency transcription factor Oct4 mediates stepwise nucleosome
741 demethylation and depletion. Mol Cell Biol *35*, 1014-1025. 10.1128/MCB.01105-14.
742 53. Somers, A., Jean, J.C., Sommer, C.A., Omari, A., Ford, C.C., Mills, J.A., Ying, L., Sommer, A.G.,
743 Jean, J.M., Smith, B.W., et al. (2010). Generation of transgene-free lung disease-specific human
744 induced pluripotent stem cells using a single excisable lentiviral stem cell cassette. Stem Cells *28*,
745 1728-1740. 10.1002/stem.495.
746 54. Kang, J., Gemberling, M., Nakamura, M., Whitby, F.G., Handa, H., Fairbrother, W.G., and Tantin,
747 D. (2009). A general mechanism for transcription regulation by Oct1 and Oct4 in response to
748 genotoxic and oxidative stress. Genes Dev *23*, 208-222. 10.1101/gad.1750709.
749 55. Tantin, D., Gemberling, M., Callister, C., and Fairbrother, W. (2008). High-throughput
750 Biochemical Analysis of in-vivo Location Data Reveals Novel Distinct Classes of
751 POU5F1(Oct4)/DNA complexes. Genome Res *18*, 631-639.
752 56. Mi, H., Muruganujan, A., Casagrande, J.T., and Thomas, P.D. (2013). Large-scale gene function
753 analysis with the PANTHER classification system. Nat Protoc *8*, 1551-1566.
754 10.1038/nprot.2013.092.



Construct (C-terminal FLAG)	Size (kDa)	Activity (%)
Oct4	42	100
Oct1	100	0.0
4N	75	4.8
4N-N40K	75	4.5
4C	60	0.4
4C-N40K	60	1.0
4NC	43	6.1
4NC-N40K	42	0.1
4NC-4L	42	1.1

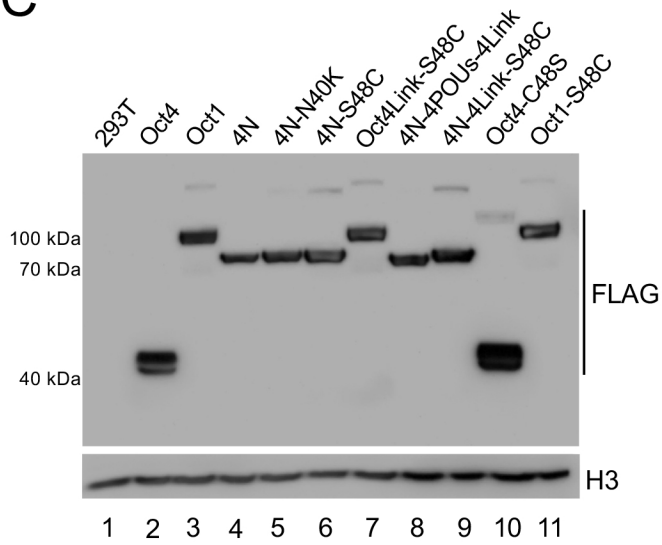


B

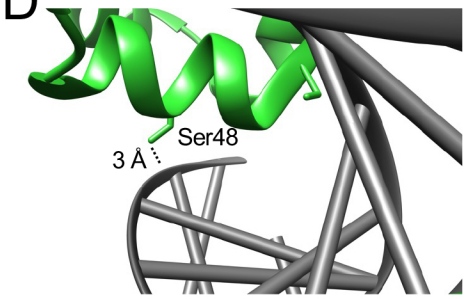


Construct (C-terminal FLAG)	Size (kDa)	Activity (%)
Oct4	42	100
Oct1	100	0.0
4N	75	7.8
4N-N40K	75	4.0
4N-S48C “1-4All”	75	55.2
4L-S48C	98	0
4N/POUs/L	72	47.8
4N/L-S48C	73	109
Oct4-C48S	42	40.8
Oct1-S48C	100	0

C



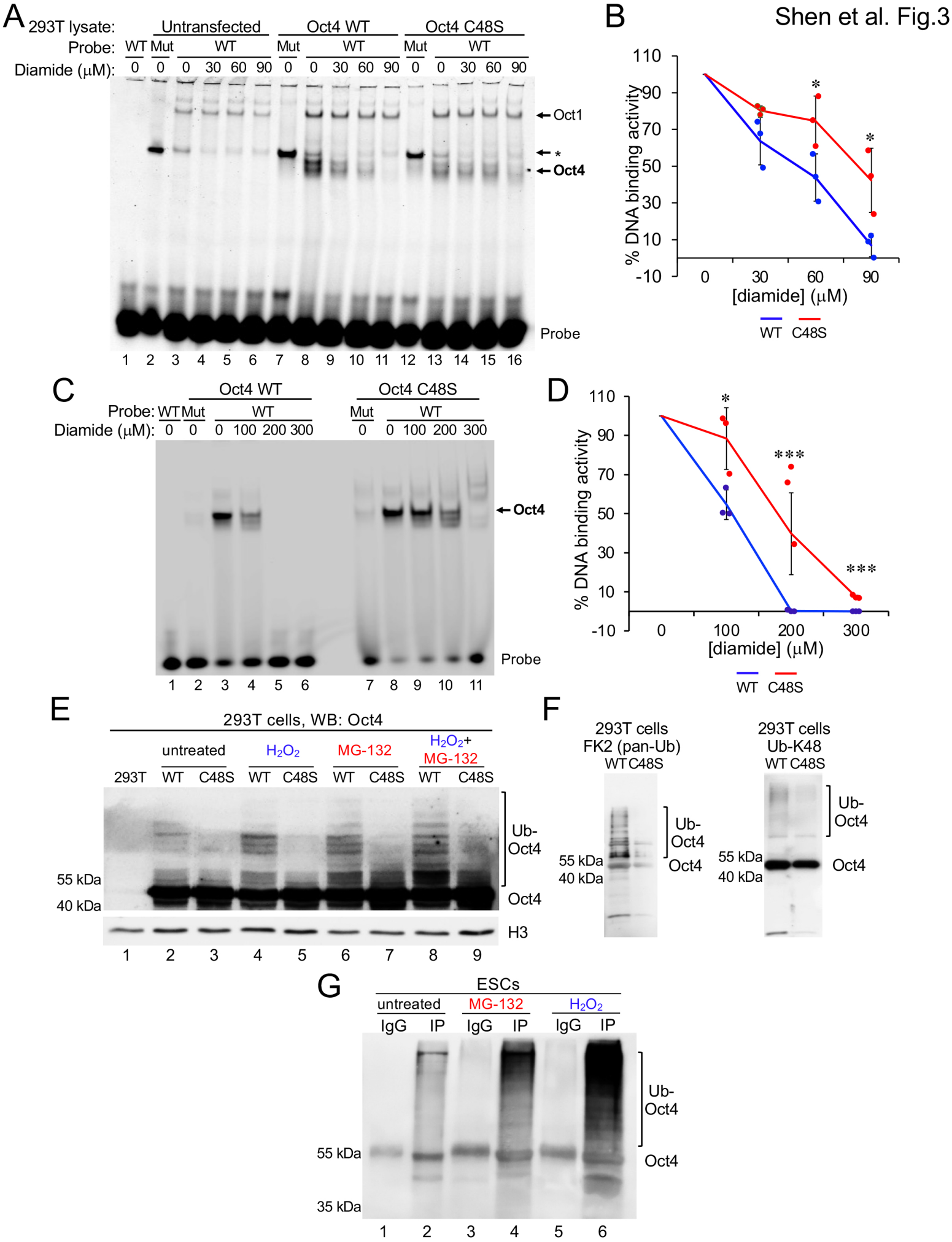
D

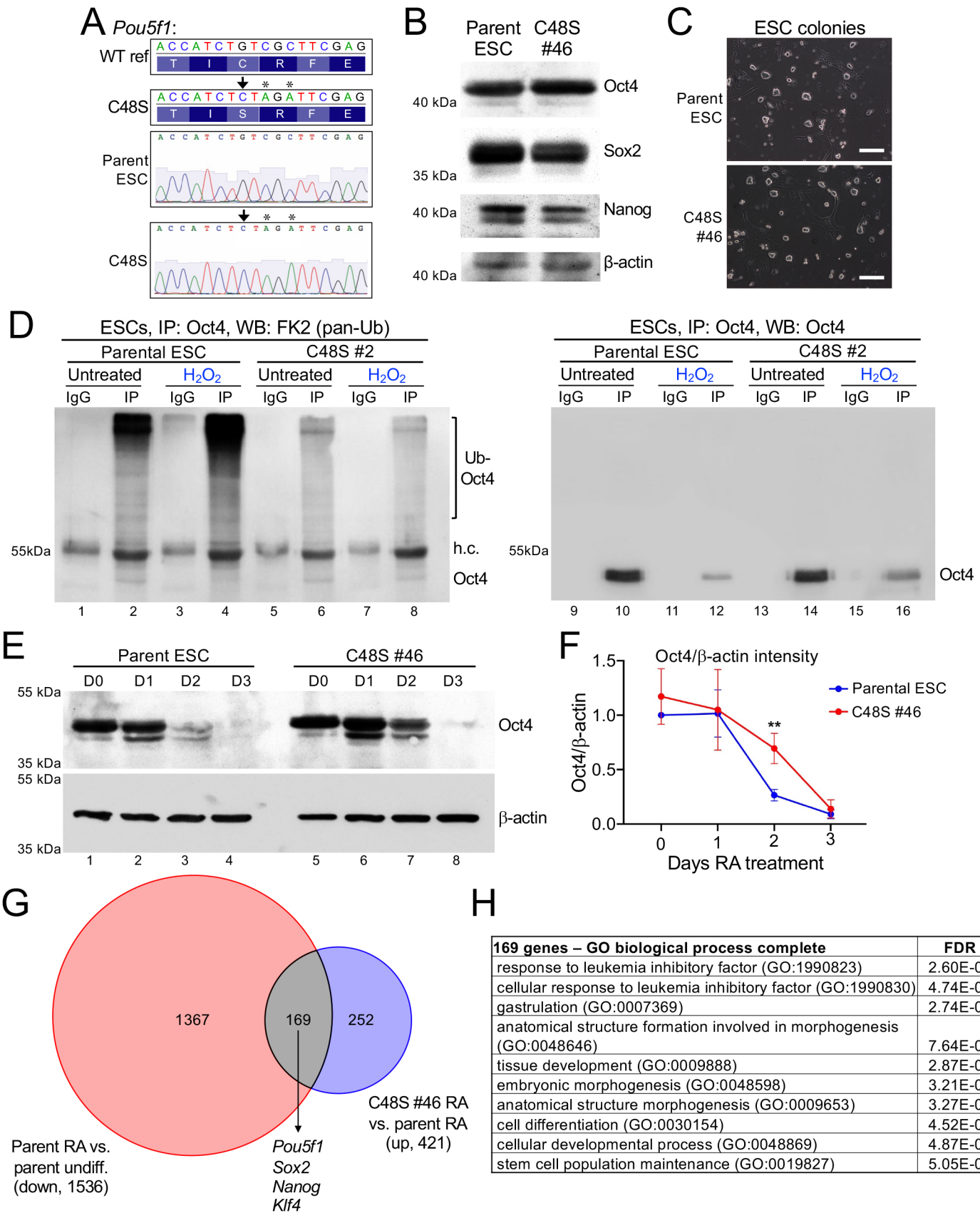


E

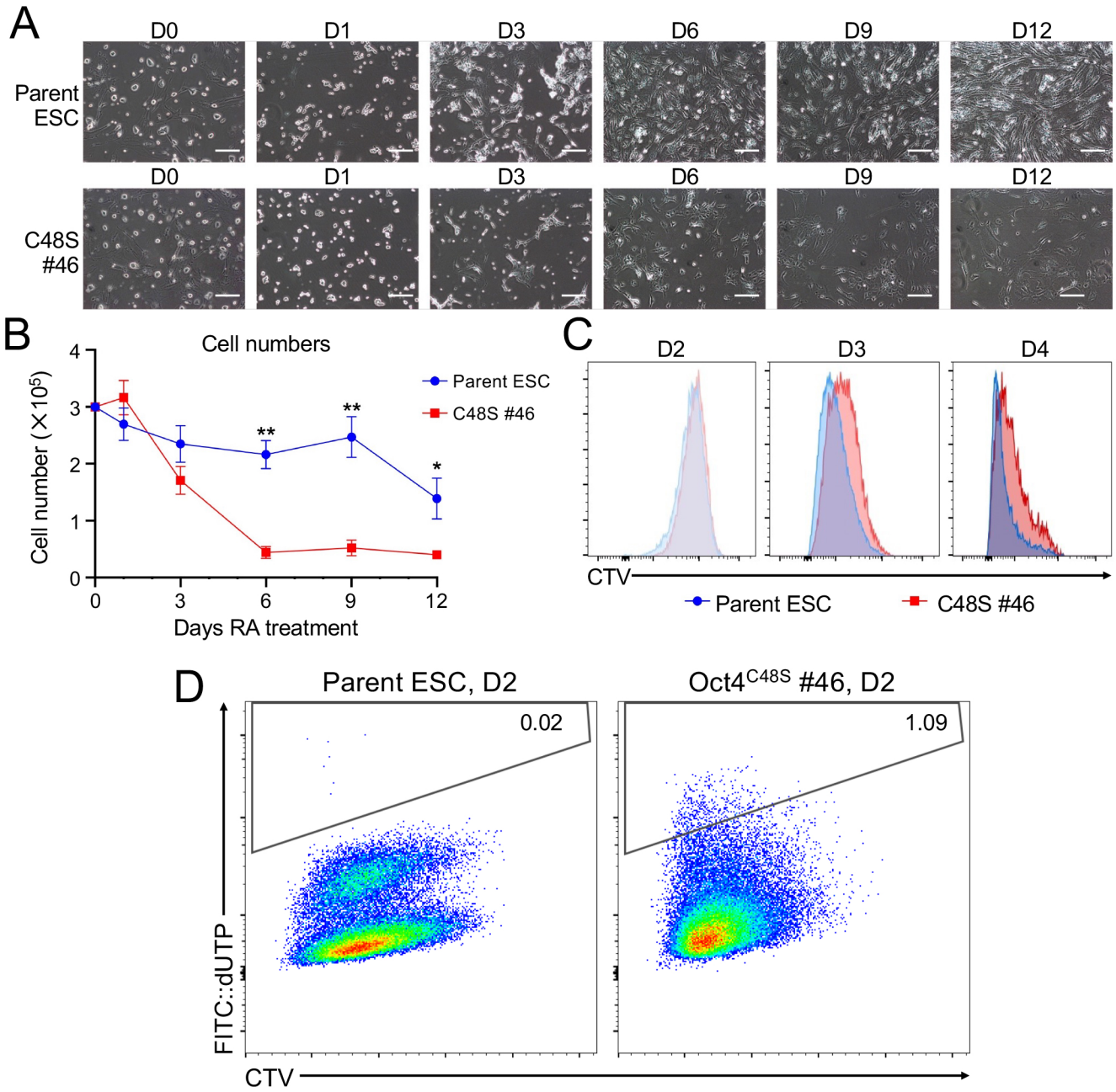
FSQTTISRFE *Mm* Oct1/Pou2f1 (NP_035267)
FSQTTICRFE *Mm* Oct4/Pou5f1 (NP_038661)
FSQTTISRFE *Mm* Oct2/Pou2f2 (Q00196-1)
FSQTTICRFE *Mm* Oct6/Pou3f1 (P21952)
FSQTTISRFE *Mm* Oct11/Pou2f3 (XP_011240722)
FSQTTICRFE *Mm* Brn2/Pou3f2 (XP_021017060)
FSQTTICRFE *Mm* Pit1/Pou1f1 (Q00286)

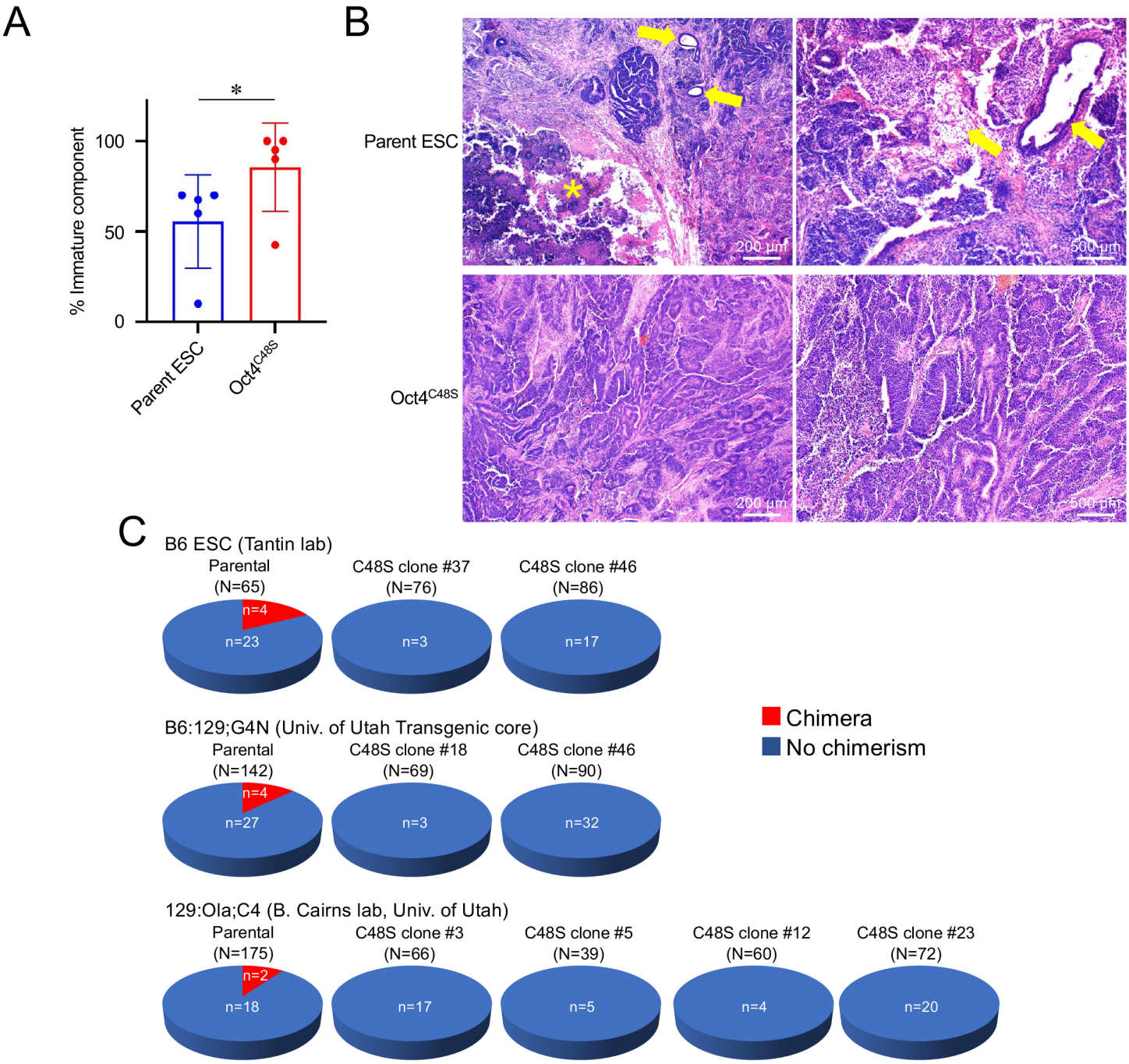
Shen et al. Fig.3





169 genes – GO biological process complete	FDR
response to leukemia inhibitory factor (GO:1990823)	2.60E-08
cellular response to leukemia inhibitory factor (GO:1990830)	4.74E-08
gastrulation (GO:0007369)	2.74E-06
anatomical structure formation involved in morphogenesis (GO:0048646)	7.64E-06
tissue development (GO:0009888)	2.87E-05
embryonic morphogenesis (GO:0048598)	3.21E-05
anatomical structure morphogenesis (GO:0009653)	3.27E-05
cell differentiation (GO:0030154)	4.52E-05
cellular developmental process (GO:0048869)	4.87E-05
stem cell population maintenance (GO:0019827)	5.05E-05


A
D0
D1
D3
D6
D9
D12
Parent ESC
C48S #46
B
Cell numbers
Cell number (×10^5)
Days RA treatment
Parent ESC
C48S #46
**
**
*
C
D2
D3
D4
CTV
Parent ESC
C48S #46
D
Parent ESC, D2
Oct4^C48S #46, D2
0.02
1.09
FITC::dUTP
CTV


A
*
% Immature component
100
50
0
Parent ESC
Oct4C48S
B
Parent ESC
Oct4C48S
200 µm
500 µm
200 µm
500 µm
C
B6 ESC (Tantin lab)
Parental (N=65)
C48S clone #37 (N=76)
C48S clone #46 (N=86)
n=4
n=23
n=3
n=17
B6:129;G4N (Univ. of Utah Transgenic core)
Parental (N=142)
C48S clone #18 (N=69)
C48S clone #46 (N=90)
n=4
n=27
n=3
n=32
Chimera
No chimerism
129:Ola;C4 (B. Cairns lab, Univ. of Utah)
Parental (N=175)
C48S clone #3 (N=66)
C48S clone #5 (N=39)
C48S clone #12 (N=60)
C48S clone #23 (N=72)
n=2
n=18
n=17
n=5
n=4
n=20