Oct4 redox sensitivity potentiates reprogramming and differentiation

Zuolian Shen[1,2,¶], Yifan Wu[1,2,¶], Asit Manna[1,2], Chongil Yi[2,3], Bradley R. Cairns[2,3,4], Kimberley J. Evason[1,2], Mahesh B. Chandrasekharan[2,5] and Dean Tantin[1,2,*]

[1]Department of Pathology,

[2]Huntsman Cancer Institute,

[3]Department of Oncological Sciences,

[4]Howard Hughes Medical Institute,

[5]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

Running Title: Oct4 redox regulation

Words in abstract: 158

Manuscript length (including spaces): 69,013

Abstract

The transcription factor Oct4/Pou5f1 is a component of the regulatory circuitry governing pluripotency and is widely used to induce pluripotency from somatic cells. Here we use domain swapping and mutagenesis to study Oct4's reprogramming ability, identifying a redox-sensitive DNA binding domain cysteine residue (Cys48) as a key determinant of reprogramming and differentiation. Oct4 Cys48 sensitizes the protein to oxidative inhibition of DNA binding activity and promotes oxidation-mediated protein ubiquitylation. *Pou5f1*^{C48S} point mutation has little effect on undifferentiated embryonic stem cells (ESCs), but upon retinoic acid (RA) treatment causes retention of Oct4 expression, deregulated gene expression and aberrant differentiation. *Pou5f1*^{C48S} ESCs also form less differentiated teratomas and contribute poorly to adult somatic tissues. Finally, we describe *Pou5f1*^{C48S} (*Janky*) mice, which in the homozygous condition are severely developmentally restricted after E4.5. Rare animals bypassing this restriction appear normal at birth but are sterile. Collectively, these findings uncover a novel Oct4 redox mechanism involved in both entry into and exit from pluripotency.

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

36 **Introduction**

37 The early mammalian embryo contains undifferentiated, pluripotent cells capable of generating all
38 embryonic tissue lineages. Cultured embryonic stem cells (ESCs) have similar capabilities (Evans and
39 Kaufman 1981). The transcription factor Oct4 governs the establishment of undifferentiated inner cell
40 mass and epiblast cells, as well germline cells and ESCs (Scholer et al. 1990b; Morey et al. 2015). Both
41 ESC differentiation and blastocyst implantation are followed by Oct4 loss (Nichols et al. 1998). Together
42 with factors such as Nanog and Sox2, Oct4 forms a network of pluripotency regulators that enforce their
43 own and each other's expression (Boyer et al. 2005).

44 Oct4 is widely used to generate induced pluripotent stem cells (iPSCs) from somatic cells
45 (Takahashi and Yamanaka 2006). Reprogramming methods that omit Oct4 nevertheless rely on activation
46 of the endogenous *Pou5f1* (*Oct4*) gene (Heng et al. 2010; Gao et al. 2013). Reprogramming strategies that
47 increase efficiency (Esteban et al. 2010; Soufi et al. 2012; Vierbuchen and Wernig 2012; Costa et al. 2013;
48 Rais et al. 2013), or improve the quality of iPSC clones (Yuan et al. 2011; Buganim et al. 2014; Chen et
49 al. 2015; Velychko et al. 2019) reveal trade-offs: efficiently generating iPSC that are of non-optimal
50 quality, or inefficiently generating iPSCs with improved developmental potential.

51 Pluripotent cells co-express Oct4 together with the widely expressed prototypic POU transcription
52 factor Oct1/Pou2f1 (Okamoto et al. 1990; Rosner et al. 1990; Suzuki et al. 1990). The two proteins share
53 similar DNA binding domains (DBDs) and consensus DNA binding sequences (Tantin 2013). They also
54 associate with many of the same target genes (Ferraris et al. 2011; Perovanovic et al. 2023). Nevertheless,
55 Oct1 and other mammalian Oct4 paralogs do not efficiently reprogram (Takahashi and Yamanaka 2006;
56 Feng et al. 2009). In contrast, other reprogramming factors such as Sox2, Klf4 and c-Myc can be replaced
57 with close paralogs (Nakagawa et al. 2008; Shi et al. 2008). Structure/function studies comparing Oct4
58 with its paralogs have identified reprogramming determinants (Esch et al. 2013; Jin et al. 2016; Jerabek
59 et al. 2017). An example is three amino acid residues in the Oct4 DNA binding domain (DBD) but absent
60 in Oct6/Pou3f1 needed to form Sox2 heterodimers. Introducing these residues, plus the Oct4 N- and C-

61 termini, into Oct6 results in efficient reprogramming (Jerabek et al. 2017). Interestingly several Oct4
62 paralogs including Oct1 contain these residues but are still reprogramming-incompetent, indicating the
63 presence of additional reprogramming determinants present in Oct4 (and Oct6) but absent in Oct1.

64 Recent studies show that omission of Oct4 from reprogramming cocktails, despite being inefficient,
65 results in qualitatively superior iPSC clones with enhanced developmental potential (An et al. 2019;
66 Velychko et al. 2019). The counterproductive activities of Oct4 are caused by transient promotion of gene
67 expression programs that deviate cells from correct reprogramming trajectories (Velychko et al. 2019). A
68 logical extension of these findings is that transient mitigation of counterproductive Oct4 activities may
69 enable reprogramming.

70 Here, we identify an Oct4 DBD cysteine residue (Cys48) as a central reprogramming determinant.
71 A serine residue is present in Oct1 at this position. In the presence of the Oct4 N-terminus, mutating Oct1's
72 serine to cysteine confers robust reprogramming activity. Conversely, mutating Oct4's cysteine to serine
73 reduces reprogramming efficiency by ~60%. Cys48 sensitizes Oct4 DNA binding activity to oxidative
74 stress and potentiates protein ubiquitylation. CRISPR-edited Oct4^{C48S} embryonic stem cell (ESC) lines
75 are phenotypically normal until differentiated, following which they abnormally retain Oct4 protein
76 expression and pluripotency gene expression signatures. Differentiating Oct4^{C48S} cells proliferate less and
77 undergo apoptosis at higher rates compared to parental controls. Mutant ESCs also form less differentiated
78 teratomas and contribute poorly to the development of viable adult mice. Oct4^{C48S} mutant mice (dubbed
79 *Janky* mice, *Jky*) appear phenotypically normal and fertile in the heterozygous condition, while
80 homozygotes are severely developmentally restricted after the blastocyst stage. Cumulatively, our findings
81 identify an Oct4 redox sensitivity mechanism that restrains protein function, but nevertheless functions
82 positively in reprogramming and normal differentiation.

83 **Results**

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

85 The two DNA binding subdomains of mouse Oct1 and Oct4, termed the POU-specific and POU-
86 homeodomain (POU_S and POU_H), share 71% and 59% identity, respectively (Fig. 1A). In contrast, the N-
87 and C-termini, and the linker between the subdomains, are divergent. The human proteins are structured
88 similarly. The high similarity allows the amino acid numbering of the DBDs to be standardized by
89 referring to the first amino acid of POU_S as residue 1 (Fig. 1A). To measure reprogramming efficiency,
90 we transduced mouse embryonic fibroblasts (MEFs) expressing GFP under the control of endogenous
91 *Pou5f1* (Lengner et al. 2007) with lentiviruses expressing the four Yamanaka factors (Sommer et al. 2009)
92 (Fig. 1B). To identify Oct4 determinants important for reprogramming, we replaced Oct4 within this
93 construct with Oct1 or chimeric proteins containing the distinct DBDs and N- and C-termini (Fig. 1B).
94 We included C-terminal FLAG-tags on all proteins to track expression. Unlike an N-terminal FLAG tag
95 which was deleterious to Oct4 (Supplemental Fig. S1), a C-terminal FLAG tag reprograms with similar
96 efficiency (Fig. 1C). Prior work comparing Oct4 to Oct6 identified Oct4 K40 (Fig. 1A) as important for
97 reprogramming (Jin et al. 2016). We therefore also mutated $Oct1^{N40}$ within POU_S to lysine in some
98 constructs (N40K, Fig. 1B). Substituting Oct1 for Oct4 eliminated reprogramming potential (Fig. 1D,E),
99 setting a baseline for domain-swaps and mutations. All of the chimeric constructs showed poor activity,
100 with constructs containing the Oct4 N-terminus generating the most colonies (~5% the level of WT Oct4,
101 Fig. 1D,E). These findings are consistent with work establishing the importance of the Oct4 N-terminus
102 (Boija et al. 2018). The raw data for Fig. 1D are shown in Supplemental Table S1. Similar data were
103 obtained at multiple timepoints (Supplemental Fig. S2A), indicating that the results reflect qualitative
104 rather than kinetic differences. Anti-FLAG immunoblotting showed similar expression (Fig. 1F). Viral
105 titers from six tested vectors from this experiment were also similar (Supplemental Fig. S2B). Because a
106 construct with both the Oct4 N- and C-termini ("4N/C") displayed only ~5% efficiency (Fig. 1D,E), we

107 conclude that determinants within the Oct4 DBD confer the bulk of reprogramming potential vis-à-vis
108 Oct1.

109 In the context of the Oct4 N-terminus ("4N"), we mutagenized DBD amino acid residues that
110 differ between Oct1 and Oct4, focusing on post-translationally modifiable residues: Lys, Ser, Thr and Cys.
111 This analysis identified 14 different amino acids, not including Lys40 (Fig. 1A). Oct6 is capable of
112 reprogramming when modified to enable Sox2 dimerization (Kim et al. 2020). Filtering these residues
113 using conservation between Oct4 and Oct6 pinpointed six possible residues. Replacing one of these, Oct1
114 Ser48, with the Oct4 cysteine residue found at this position increased activity >7-fold over the "4N"
115 construct containing the Oct4 N-terminus alone (Fig. 2A,B, "1-4All"). The raw data for Fig. 2A are shown
116 in Supplemental Table S2. iPSC colonies formed under these conditions showed a pluripotent morphology
117 and expressed GFP from the endogenous *Pou5f1* locus at levels comparable to Oct4 (Supplemental Fig.
118 S3). The remaining activity was concentrated in the linker domain, as a construct containing the Oct4 N-
119 terminus, Cys48 and the Oct4 linker (4N/L-S48C) had 100% activity (Fig. 2A,B). The Oct4 linker has
120 been shown to regulate reprogramming via selective interactions with different cofactors (Esch et al. 2013;
121 Han et al. 2022). The differences were maintained over time, with the 1-4All construct consistently
122 generating ~60% activity over multiple days, indicating that the differences were not kinetic in nature
123 (Supplemental Fig. S4A). Similar results were obtained using normal primary MEFs and assessing
124 alkaline phosphatase-positive rather than GFP-positive iPSCs (Supplemental Fig. S4B-C).

125 An Oct4^{C48S} construct reduced activity by 60% (Fig. 2A,B), indicating that over half of Oct4's
126 differential reprogramming activity vis-à-vis Oct1 operates through this residue. The reciprocal Oct1^{S48C}
127 mutation did not reprogram (Fig. 2A,B, Oct1-S48C), indicating that Cys48 is not sufficient, and
128 highlighting its combined importance with the Oct4 N-terminus for reprogramming. The proteins were
129 equivalently expressed (Fig. 2C). These results indicate that in the context of the Oct4 N-terminus, Cys48
130 confers robust reprogramming activity. In an Oct1 DBD:DNA structure (Klemm et al. 1994), Ser48
131 hydrogen bonds with the DNA backbone (Fig. 2D). In an Oct4:Sox2:nucleosome structure (Michael et al.

132 2020), the Cys48 sulfhydryl also allows for hydrogen bonding. Predictably, Cys48 mutation to glycine
133 greatly diminishes Oct4 DNA binding in vitro (Marsboom et al. 2016). Oct6 also contains a cysteine at
134 this position, as does human Oct4 (Fig. 2E). The *Danio rerio* Oct1/4 paralog Pou2 also contains a cysteine
135 at this position (Fig.2E), suggesting that cysteine may be the ancestral residue. These structural studies
136 and our findings indicate that Ser48 and Cys48 have minimal differences in DNA binding, but that Cys48
137 nevertheless plays key roles in reprogramming.
138
139 *Oct4 Cys48 mediates degradation in response to oxidative stress*
140 Oct4 is hypersensitive to oxidative stress (Lickteig et al. 1996; Guo et al. 2004; Marsboom et al. 2016) in
141 a manner reversed by thioredoxin (Guo et al. 2004), implicating cysteine thiols in redox sensitivity. We
142 reproduced this result using electrophoretic mobility shift assays (EMSA) and DNA containing a
143 canonical binding element (Supplemental Fig. S5), suggesting that Cys48 may potentiate reprogramming
144 through a redox mechanism. We used lysates from 293T cells transfected with WT or C48S Oct4 to
145 measure DNA binding. To measure redox sensitivity, we pre-treated Oct4 with the oxidizing agent
146 diamide prior to addition of DNA. 293T cells express endogenous Oct1 as an internal control (Fig. 3A,
147 lanes 2-3). Oct4 bound to octamer DNA specifically (Fig. 3A, lanes 7-8) and was hypersensitive to
148 diamide treatment (lanes 9-11). Oct4^{C48S} bound DNA equivalently (lane 13) but was comparatively
149 resistant to oxidation (lanes 14-16). In contrast, little sensitivity was observed with Oct1. Quantification
150 from three replicates is shown in Fig. 3B. The recombinant Oct4 contained a C-terminal FLAG tag,
151 allowing for purification (Supplemental Fig. S6). Similar relative resistance to oxidative inhibition of
152 DNA binding was observed using purified recombinant C48S Oct4 (Fig. 3C). Quantification from three
153 replicates is shown in Fig. 3D. These results indicate that Oct4 Cys48 mediates oxidative inhibition of
154 DNA binding.
155 Oct4 oxidation has been linked to its degradation (Marsboom et al. 2016). We noted in over-
156 exposed immunoblots the presence of multiple slow-migrating high molecular weight forms of wild type

157 Oct4 which were attenuated with C48S mutation (Fig. 3E, lanes 2-3). Hydrogen peroxide treatment
158 augmented the banding pattern using wild type but not mutant protein (lanes 4-5). The bands were also
159 increased using MG-132, consistent with ubiquitylated species (lanes 6-7). H_2O_2 and MG-132 co-
160 treatment further increased the pattern (lanes 8-9). None of these treatments affected short-term cell
161 viability (Supplemental Fig. S7).

162 We immunoblotted purified Oct4 with antibodies specific to different ubiquitin species.
163 Antibodies recognizing both mono- and poly-ubiquitylated proteins revealed strong diminution with
164 $Oct4^{C48S}$ as expected (Fig. 3F, FK2). Diminution was observed with K48 linkage-specific polyubiquitin
165 antibodies (Fig. 3F), while K63 linkage-specific antibodies showed modest differences (Supplemental
166 Fig. S8). Similar ubiquitylated species were observed using endogenous Oct4 in wild type ESCs (Fig. 3G,
167 lane 2) that were augmented by H_2O_2 and MG-132 (lanes 4 and 6). We conclude that Cys48 promotes
168 Oct4 ubiquitylation and degradation in response to oxidative stress, at least in part via catalysis of K48-
169 linked polyubiquitin chains.

170

171 *$Oct4^{C48S}$ ESCs differentiate abnormally*

172 We generated ESCs with a single-base missense mutation in the endogenous *Oct4* (*Pou5f1*) locus that
173 converts Cys48 to Ser. Two silent point mutations were also engineered to create an *Xba*I site to monitor
174 targeting and facilitate genotyping. Three low-passage parent ESC lines of different strain backgrounds
175 were chosen for targeting. Multiple independent clones were generated for each line. Homozygous
176 mutation was confirmed by high-resolution melting analysis and resequencing (Supplemental Fig. S9A
177 and Fig. 4A). Tested clones were karyotypically normal and similar to parent ESCs, with the proportion
178 of diploid cells ranging between 71 and 93%. Undifferentiated mutant ESC clones expressed Oct4, Sox2
179 and Nanog at similar levels and were morphologically indistinguishable from their parent clones (Fig.4B-
180 D, Supplemental Fig. S9B-D).

181 To test if C48S mutation protects endogenous Oct4 against oxidative stress, we studied Oct4

182 ubiquitination in parent and mutant ESCs. Oct4 was immunoprecipitated from ESC lysates, and the
183 resulting material was immunoblotted using pan-Ub (FK2) and Oct4 antibodies. Ubiquitylation of
184 endogenous Oct4 in ESCs was clearly visualized and increased with H_2O_2 exposure (Fig. 4E, lanes 1-4),
185 while Oct4^{C48S} ubiquitylation was attenuated (lanes 5-6) and unaffected by oxidative stress (lanes 7-8).
186 Oct4 immunoblotting confirmed that the protein was immunoprecipitated, with decreased Oct4 recovered
187 with oxidative stress in parental cells and more Oct4 recovered in mutant cells (Fig. 4E, right side).

188 Cells were then differentiated in vitro using retinoic acid (RA) to assess Oct4 expression,
189 morphology and cell counts. Oct4 protein expression was abnormally retained early during differentiation
190 (Fig. 4F, lanes 3 and 7). Quantification from multiple experiments revealed that the effect was transient,
191 with a >2.8-fold increase of Oct4 at day 2 that normalized by day 3 (Fig. 4G). Similar results were obtained
192 using the same clone compared to parent cells electroporated without CRISPR RNP (Supplemental Fig.
193 S9E). These results were recapitulated using a second clone from the same parental ESC line
194 (Supplemental Fig. S9F), as well as a clone from a different 129×C57BL/6 ESC line, which differentiated
195 more slowly and showed greater differences (Supplemental Fig. S9G).

196 We performed transcriptional profiling of parent and mutant cells at RA differentiation day 2, as
197 well as undifferentiated controls. Three replicates per condition were used, which showed tight
198 concordance (Supplemental Fig. S10A). 20-31 million sequence reads were generated for each sample,
199 ~80% of which aligned to the *mm39* reference genome. We identified 421 genes with significantly
200 elevated gene expression in differentiating mutant relative to parent cells (p<0.001, >2.5-fold, Fig. 4H,
201 Supplemental Table S3). Intersecting these genes with the set of genes normally decreased upon
202 differentiation of parent cells at this timepoint (1536 genes) showed that 40% of the 421 genes with
203 retained expression in mutant cells (169 genes) are normally down-regulated with differentiation. These
204 included *Oct4* itself (*Pou5f1*), as well as *Sox2*, *Nanog* and *Klf4* (Fig. 4H). Comparing these genes with
205 established Oct4 ChIP-seq (King and Klose 2017) identified 114 of 169 genes (67%) as direct Oct4 targets
206 (Supplemental Table S3). The complete set of 169 genes were enriched for gene ontology (GO) terms

207 associated with pluripotency such as "response to leukemia inhibitory factor" (including *Pou5f1*, *Klf4*,
208 and *Fgf4*) and "stem cell population maintenance" (including *Nanog*, *Pou5f1*, *Sox2* and *Fgf4*, Fig. 4I).
209 Similar results were obtained using a hierarchical clustering approach. We utilized the 100 most
210 differentially expressed genes upon differentiation for both genotypes (150 total genes). Although the
211 differentiating parental and mutant replicates clustered together, there was a cohort of pluripotency
212 associated genes (including *Pou5f1*) whose down-regulation was blunted in mutant cells (Supplemental
213 Fig. S10B). A second cohort of differentiation-specific genes showed poorer expression in mutant cells
214 upon differentiation. These included *Krt18*, *Mest*/*Peg1* and *Dag1* (Supplemental Fig. S10B). Comparing
215 the complete set of 150 genes to Oct4 ChIP-seq (King and Klose 2017) identified 117 (78%) as direct
216 Oct4 targets (Supplemental Table S3). These results indicate that Oct4^{C48S} ESCs inappropriately delay
217 loss of pluripotency gene expression signatures during differentiation.

218 The level and duration of retained Oct4 expression during differentiation can affect lineage
219 specification (Wang et al. 2012; Aksoy et al. 2013; DeVeale et al. 2013; Radzisheuskaya et al. 2013). To
220 determine the effects of Oct4^{C48S} mutation on differentiation, we visualized cells treated with RA
221 microscopically. These experiments revealed a significant depletion of cells beginning at RA
222 differentiation day 3 (Fig. 5A,B, D6-D12). The decreased cell numbers could be due to decreased
223 proliferation and/or increased cell death. We studied proliferation using cell trace violet (CTV) and
224 apoptosis using TUNEL staining. Adherent, viable differentiating C48S cells showed decreased
225 proliferation at days 3 and 4 (Fig. 5C) and increased apoptosis at day 2 (Fig. 5D). ESCs differentiating
226 with RA also experience an increase in global reactive oxygen species levels at day 3, though specific
227 oxidizing molecules at specific subcellular locations may be higher (Fig. 5E). Finally, we used mutant
228 ESCs to form embryoid bodies (EBs) and subsequently generate cardiomyocytes. Differentiated
229 cardiomyocytes form gap junctions, with functionality scorable by spontaneous and synchronous beating
230 of groups of cells (Lynch et al. 2018). We found that unlike parent cells which formed beating
231 cardiomyocytes robustly, two different mutant ESC lines completely failed to form beating cells when

232 differentiated (Fig. 5F and Supplemental Movies S1-S6). Beating was not rescued by extended culture,
233 indicating that the phenotype was not kinetic in nature (not shown). These results define a functional role
234 for Oct4 Cys48 in ESC differentiation.
235
236 *Oct4^{C48S} ESCs form less differentiated teratomas and contribute poorly to embryonic development*
237 We used mutant and control ESC lines to form teratomas in mice, which were monitored for two weeks
238 before tumor excision. Tumors were fixed, H&E stained, and pathologically scored in a blinded fashion.
239 Parent tumors contained areas of maturing structures such as squamous tissue, respiratory epithelium (cilia
240 and/or goblet cells), maturing neural components, gastric foveolar epithelium and fat. In contrast, tumors
241 formed with Oct4^{C48S} ESCs injected into the contralateral flanks of the same mice were more
242 homogeneous with few maturing components, instead harboring immature neuroepithelial elements,
243 immature cartilage, embryonal carcinoma and yolk sac (Fig. 6A,B). Additional images are shown in
244 Supplemental Fig. S11. Similar results were obtained using a wild type and mutant teratomas derived from
245 a different ESC line (Fig. 6C). At later time points, larger mutant teratomas began to accumulate mature
246 structures and the differences largely normalized (Supplemental Fig. S12).
247 One hallmark of pluripotency is the ability to contribute to cells and tissues of the embryo (De Los
248 Angeles et al. 2015). We conducted chimerism assays to assess the ability of Oct4^{C48S} ESCs to contribute
249 to development. We used eight lines of mutant ESCs, derived from three parent lines of different strain
250 backgrounds to minimize background and clonal effects. Compared to the parent cells, which contributed
251 to offspring at a rate of 10-15% (Fig. 6C), none of the derived Oct4^{C48S} ESCs were able to contribute to
252 adult mice (Fig. 6D). The results indicates that Oct4^{C48S} ESCs contribute poorly to embryonic
253 development. These results identify a role for Oct4 redox sensitivity in differentiation and development.
254
255 *Oct4^{C48S} homozygous mutant mice show developmental lethality after the blastocyst stage*
256 We generated germline mutant mouse lines with the Oct4^{C48S} mutation (dubbed *Janky* mice) using

257 CRISPR modification of the endogenous *Pou5f1* gene. A heterozygous male founder was generated,
258 which transmitted the *Jky* allele normally, as did its male and female offspring (Fig. 6E). In contrast,
259 heterozygote intercrosses were associated with decreased litter sizes and a significant deviation from 1:2:1
260 Mendelian ratios, with homozygous *Jky* offspring severely attenuated (Fig. 6E). One *Jky* homozygous
261 male was born to these crosses, which was phenotypically indistinguishable from its littermates at weaning
262 but was sterile with underdeveloped testes (Supplemental Fig. S13A). Microscopic examination revealed
263 significant heterogeneity in the size and diameter of the seminiferous tubules (Supplemental Fig. S13B).
264 Immunofluorescence using spermatogonia and spermatocyte markers showed that a minority of
265 seminiferous tubules had normal physiology (Supplemental Fig. S13C, Box 1), but many displayed
266 abnormal localization of spermatocytes and germ cell loss. Interestingly, the DNA damage marker γ-
267 H2AX, which is normally associated with meiotic sex chromosome inactivation in pachytene
268 spermatocytes (Fernandez-Capetillo et al. 2003), was not detected in *Jky*/*Jky* LIN28A^{+} spermatogonia,
269 suggesting that loss of Oct4 redox sensing may regulate germ cell development but not spermatogonia
270 maintenance. The developmental lethality associated with most embryos occurred after embryonic day
271 4.5, as blastocysts collected at this time genotyped as *Jky*/*Jky* at normal ratios (Fig. 6E). Collectively, these
272 results identify Oct4 Cys48 as redox-sensitive and uncover dual roles for this residue in reprogramming and
273 development.

274 **Discussion**

275 Oct4 proteins levels must be precisely controlled for efficient iPSC reprogramming (Ilia et al. 2023) and
276 exit from pluripotency (Wang et al. 2012; Aksoy et al. 2013; DeVeale et al. 2013). Nevertheless, the
277 mechanisms by which Oct4 is post-transcriptionally regulated are poorly defined. Here, we pinpoint Oct4
278 redox sensitivity as an important determinant of both reprogramming and differentiation. Pluripotent cells
279 co-express Oct4 and Oct1 (Shen et al. 2017). Despite having similar DBDs and recognizing the same 8-
280 bp consensus octamer sequence (Tantin 2013), Oct1 and other Oct4 paralogs do not function in
281 reprogramming (Takahashi and Yamanaka 2006). Structure/function studies comparing Oct4 with these
282 paralogs have identified different Oct4 reprogramming determinants (Esch et al. 2013; Jin et al. 2016;
283 Jerabek et al. 2017; Boija et al. 2018). We identify Cys48 of the Oct4 DBD as a critical reprogramming
284 determinant that sensitizes the protein to oxidative inhibition and degradation. In Oct1, this residue is a
285 serine. Both residues hydrogen bond with the DNA phosphate backbone.

286 Transcription factors with both positive and negative roles in reprogramming include Oct4 and
287 Smad3 (Ruetz et al. 2017; An et al. 2019; Velychko et al. 2019). Recent studies show that removal of Oct4
288 from reprogramming cocktails is inefficient but results in superior iPSC clones with enhanced
289 developmental potential (An et al. 2019; Velychko et al. 2019). The counter-productive activity of Oct4
290 is due to promotion of off-target, ectoderm-specific gene expression (Velychko et al. 2019). These
291 findings suggest that transient down-regulation of counterproductive Oct4 activities may be an important
292 step for efficient reprogramming. Our work suggests that Oct4 redox sensitivity fulfills such a role. The
293 findings may provide a route to generating superior iPSC reprogramming outcomes with higher efficiency.

294 Oct4 is hypersensitive to oxidative stress (Lickteig et al. 1996; Guo et al. 2004; Marsboom et al.
295 2016) in a manner reversed by thioredoxin (Guo et al. 2004). Relative to other Oct proteins, Oct4 has
296 unusually high cysteine content in the DBD, with two cysteines in each subdomain. Cys142 (Cys50 of the
297 POU-homeodomain) lies in the recognition helix of POU_H and is the distinguishing feature of POU-
298 homeodomains (Tantin 2013). Cys48 is located within the DNA recognition helix of the other DNA

299 binding subdomain, POU_S. The hydrogen bond contact between Cys48 and the DNA predicts that cysteine
300 oxidation will block DNA binding. Consistently, $Oct4^{C48S}$ mutation maintains DNA binding and confers
301 resistance to oxidation-induced DNA binding inhibition in vitro. The same mutation reduces
302 reprogramming efficiency by more than half. Cysteine redox sensitivity is known to be enhanced by
303 adjacent basic amino acid residues (Britto et al. 2002), and Cys48 is followed by a conserved Arg. Oct1
304 S48C mutation is sufficient to confer reprogramming activity, but only in the context of the Oct4 N-
305 terminus. The Oct4 N-terminus is known to mediate liquid-liquid phase transitions to control target gene
306 transcription (Boija et al. 2018).

307 Prior studies have identified a link between Oct4 oxidation and degradation (Marsboom et al.
308 2016). Cys48 strongly promotes protein ubiquitylation, in particular following exposure to oxidative stress
309 or MG-132. Oct4 is known to be polyubiquitylated by K63-linked chains via the action of the Stub1 E3
310 Ub ligase (Mamun et al. 2022), however we found that K48-specific linkages were more strongly affected
311 with $Oct4^{C48S}$ mutation, suggesting a distinct pathway.

312 Endogenous $Oct4^{C48S}$ mutation only minimally affects undifferentiated ESCs, but results in
313 abnormalities upon RA-mediated differentiation, including retention of Oct4 expression, deregulated gene
314 expression, decreased proliferation and increased apoptosis. Prominent roles for Oct4 in cell survival and
315 proliferation during post-implantation embryonic development have been documented (DeVeale et al.
316 2013). When injected into adult mice, ESCs form teratomas containing differentiating structures
317 representing all three germ layers. When injected into blastocysts, they contribute to adult mouse tissues
318 (De Los Angeles et al. 2015). We found that both features become defective with Cys48 mutation. Despite
319 being gain-of-function, mice with this germline mutation (dubbed *Janky* mice) have no observed defects
320 in the heterozygous condition. This may be due to a necessity for increased Oct4 activity above a certain
321 threshold to trigger a phenotype. Homozygous *Jky* mutant mice undergo developmental lethality between
322 E4.5 and birth. Given the timing of Oct4 expression during development, we predict that defects arise in
323 these embryos at the blastocyst and/or peri-implantation stages, though this still awaits testing. Rare *Jky*

324 mice obtained from these crosses appear normal at weaning, suggesting that mice circumventing this
325 developmental restriction subsequently develop normally. However, the single male mouse we obtained
326 from these crosses was sterile. These findings are compatible with known roles for Oct4 in the germline
327 (Scholer et al. 1990a; Kehler et al. 2004). The abnormal seminiferous tubule architecture of this mouse
328 resembles the Pou5f1 TNAP-Cre conditional knockout (Kehler et al. 2004), suggesting a potentially
329 central role for Oct4 redox regulation in the survival and migration of primordial germ cells. These results
330 indicate that Cys48 is as important for differentiation and exit of pluripotency as it is for reprogramming
331 and indicate a physiological role for Oct4 redox regulation through Cys48 during development.

332 **Materials and methods**

333

334 *Cell culture*

335 ESCs/iPSCs were cultured in Dulbecco's modified Eagle medium (Sigma) with "2iV" supplementation
336 as described previously (Shakya et al. 2015; Shen et al. 2018). Media contained 15% fetal calf serum
337 (FCS, Avantor), 2 mM L-glutamine (ThermoFisher), 50 μM β-mercaptoethanol (2ME, Sigma), 1,000
338 U/mL leukemia inhibitor factor (LIF, Chemicon), 0.1 mM minimal essential medium nonessential amino
339 acids (NEAA, ThermoFisher), 1 mM sodium pyruvate (ThermoFisher), the MEK inhibitor PD03259010
340 (1 μM, LC laboratories) and the glycogen synthase kinase-3β (GSK3β) inhibitor CHIR99021 (3 μM, LC
341 Laboratories). Culture medium was also supplemented with 50 μg/mL ascorbic acid (Sigma). H_2O_2
342 (Sigma) was supplied at 1 mM for 2 hr and MG-132 (Sigma) at 10 μM for 2 hr. RA differentiation media
343 omitted LIF, MEK inhibitor and GSK3β inhibitor, and contained retinoic acid (Sigma) at 1.5 μg/mL.
344 Cardiomyocytes were generated using established protocols in which EBs are first generated via hanging
345 drops and transferred to 12-well plates with differentiation medium (DMEM, 15% FCS, 2 mM L-
346 glutamine, 0.1 mM NEAA, 1 mM sodium pyruvate, 0.1 mM 2ME, 50 μg/mL ascorbic acid) (Lynch et al.
347 2018). After 8 days in differentiation culture, cells were observed using an Olympus IX-51 inverted phase-
348 contrast microscope. CTV proliferation assays were conducted using a kit (CTV Proliferation Kit for Flow
349 Cytometry, ThermoFisher). TUNEL assays were conducted using a kit (Apo-Direct, Becton-Dickinson).
350 Alkaline phosphatase staining was performed using a kit (System Biosciences) according to the
351 manufacturer instructions. DCFDA staining was performed using a kit (AbCam), following the
352 manufacturer protocol.

353

354 *FLAG-tagged lentiviral constructs for iPSC generation*

355 The mSTEMCCA lentiviral vector expressing Oct4, Klf4, Sox2 and c-Myc was used for these studies
356 (Somers et al. 2010). The Oct4 cDNA cassette was replaced with c-terminally FLAG-tagged full-length

357 mouse Oct4, mouse Oct1 or chimeric constructs. FLAG-tagged Oct4 was generated by overlap PCR using
358 primers containing the DYKDDDDK FLAG peptide sequence: (Fwd: 5'-
359 GATTACAAGGATGACGACGATAAGGGAAGTGGCGTGAAACAGA; Rev: 5'-
360 CTTATCGTCGTCATCCTTGTAATCGTTTGAATGCATGGGAGAG). The first fragment was
361 amplified using a Fwd primer containing a *Not*I restriction enzyme site (5'-
362 AATGAAAAAAGCGGCCGCCATGGCTGGACACCT) and overlapping Rev primer containing the
363 FLAG sequence. A second fragment was amplified using the Rev primer containing an *Nde*I restriction
364 site (5'-GGGAATTCCATATGTGTGGCCATATTATCATCGTGT) and overlapping Fwd primer. The
365 first and second fragments were mixed as template and amplified Fed and Rev primers containing *NotI*
366 and *Nde*I. After purification and digestion with *Not*I and *Nde*I, the fragment was ligated with the
367 mSTEMCCA vector digested by *Not*I and *Nde*I. Similar strategies were applied to generate other
368 constructs. For point mutation, a site-directed mutagenesis kit (New England Biolabs Q5) was used. When
369 making fusions between the different linker domains and POU_H, we maintained the "RKRKR" sequence
370 at the N-terminus of Oct4 POU_H, which has been shown to be important for reprogramming (Jerabek et
371 al. 2017).
372
373 *iPSC generation*
374 293T cells were transfected with mSTEMCCA encoding FLAG-tagged Oct4, Oct1 and different domain-
375 swaps and mutants as well as the packaging plasmids (pMDLg/pRRE and pRSV-Rev) and envelope
376 plasmid (VSV-G) using polyethylenimine. Lentiviral supernatants were filtered, 8 μg/mL polybrene
377 (Sigma) was added, and the mixture used to transduce passage 3 primary MEFs expressing EGFP from
378 the endogenous *Pou5f1* locus. The MEFs were prepared from E12.5 mouse embryos (Jackson Labs strain
379 #008214, B6;129S4-Pou5f1^{tm2Jae}/J) and were seeded in 6-well plates the night before transduction at a
380 density of 1×10^5 cells/well. Cells were cultured in ESC medium (above). GFP^+ iPSC colony emergence
381 was quantified after 12 days using an Olympus IX-71 epifluorescence microscope. Lentivirus titers were

382 established using two techniques. For the experiment shown in Supplemental Fig. S2B, a kit measuring
383 p24 was used (QuickTiter Lentivirus Titer Kit, Cell Biolabs). For the experiment shown in Supplemental
384 Fig. S4C, a kit measuring conserved areas of viral RNA was used (Lenti-X qRT-PCR, Takara Bio).
385
386 *Immunoblotting and immunofluorescence*
387 Antibodies for immunoblotting were: Oct4, α-tubulin and β-actin, Santa Cruz; Oct1, Bethyl; Nanog and
388 Sox2, GeneTex; GAPDH, EMD Millipore; FK2 (pan-Ub), Enzo Life Sciences; Ub-K48, Cell Signaling.
389 All immunoblots are shown in uncropped form in Supplemental Fig. S14-17. For immunofluorescence,
390 testes from both *Oct4*$^{Jky/Jky}$ mouse and an *Oct4*$^{+/+}$ littermate control were weighed and fixed using 4%
391 formaldehyde. The paraffinized testes were sectioned into 5 μm slices. Antibodies for
392 immunofluorescence were: mouse anti-SYCP3, Santa Cruz; goat anti-LIN28A, R&D Systems; rabbit anti-
393 phospho-H2AX, Cell Signaling; Alex Fluor 647 donkey anti-mouse IgG, Alex Fluor 488 donkey anti-
394 rabbit IgG and Alex Fluor 594 donkey anti-goat IgG, ThermoFisher Scientific. Data were imaged using a
395 Nikon ECLIPSE Ti inverted microscope with Nikon Elements AR software.
396
397 *Electrophoretic mobility shift assay*
398 EMSA was performed using published methods (Kang et al. 2009; Ferraris et al. 2011). Diamide (Sigma)
399 was added at final concentration of up to 300 μM on ice during reaction assembly. 5'Cy5-labeled double-
400 stranded DNA probes were as follows: wild type octamer, 5'-TGTCGAATGCAAATCACT-3'; mutant
401 octamer, 5'-TGTCGAATGCAAGCCACT-3'. Reactions were incubated for 30 min at room temperature
402 prior to gel loading. For experiments involving reversal of oxidative inhibition, thioredoxin (Sigma) was
403 supplied at a final concentration of 0 to 60 μM after 15 min incubation with diamide, and both thioredoxin-
404 treated and control samples were incubated for a further 15 min at 37°C prior to gel loading. DNA was
405 visualized using a Molecular Dynamics Typhoon system.
406

407 *Recombinant Oct4 protein*

408 Wild type and C48S Oct4 constructs were generated bearing dual C-terminal FLAG and Twin-Strep tags.
409 A g-block (IDT) fragment coding for the Twin-Strep and FLAG epitope tags was inserted between the
410 *Nco*I and *Kpn*I sites in the transient expression vector pACE-MAM2 (Geneva Biotech). The resultant
411 construct was then digested with *Nco*I and *Pvu*II to insert the g-block fragment coding for either wild type
412 or C48S murine Oct4. After Sanger sequencing to confirm veracity, the plasmids were transfected into
413 Expi293F cells (ThermoFisher) using ExpiFectamine 293 Tranfection Kit (ThermoFisher) following the
414 manufacturer instructions. After 2 days, transfected cells were harvested and washed with ice-cold PBS.
415 Cells were resuspended in Buffer A (20 mM HEPES pH 8.0, 1.5 mM $MgCl_2$, 10 mM KCl, 0.25% NP-40,
416 0.5 mM DTT, 2x protease inhibitor cocktail/Roche), incubated for 10 min on ice, and homogenized using
417 a Dounce homogenizer (Wheaton). The homogenate was centrifuged at 4°C for 5 min at 3,800×*g*). Pelleted
418 nuclei were resuspended in Buffer C (20 mM HEPES pH 8.0, 25% glycerol, 1.5 mM $MgCl_2$, 420 mM
419 KCl, 0.25% NP-40, 0.2 mM EDTA, 0.5 mM DTT, 2x protease inhibitor cocktail) and Dounce
420 homogenized again. Nuclei were extracted for 30 min at 4°C with end-to-end rotation, then centrifuged at
421 4°C for 30 min at 20,000×*g*. The nuclear extract was then subjected to immunoprecipitation with anti-
422 FLAG (M2)-agarose beads (Sigma). Bead-bound proteins were eluted with 200 ng/µL 3×FLAG peptide
423 (Sigma) by incubating at 4°C for 30 min. The protein was pure and monodisperse as verified using
424 Coomassie blue staining and size exclusion chromatography.

425

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

427 Mutant ESCs with a single-base missense mutation that converts Cys48 to Ser were generated using
428 templated CRISPR. A modified synthetic sgRNA (Synthego) was designed to target Cas9 and cutting near
429 the target nucleotide, 5' GGCCTCGAAGCGA<u>C</u>AGATGG 3', where the underlined C indicates the target
430 nucleotide. The sgRNA was complexed with Cas9 protein (IDT) to form a ribonucleoprotein (RNP)
431 complex that was electroporated with program CG-104 into 6×10^4 mouse ESCs using a LONZA 4D-

432 Nucleofector X-unit. The RNP was co-electroporated with a corresponding single-stranded
433 oligodeoxynucleotide (ssODN) with 5' and 3' phosphorothioate modifications, 5'-
434 ctcactctgtttgatcggcctttcagGAAAGGTGTTCAGCCAGACCA**CCATCTCTAGATTCGAGGCC**TTGC
435 AGCTCAGCCTTAAGAACATGTGTAAGCTG-3' where lower-case denotes intronic sequence and the
436 sgRNA site is in bold. The ssODN contained the target mutation (underlined), silent mutations in the seed
437 region (double underlined) to block CRISPR re-cutting after incorporation of the donor and to generate a
438 unique *Xba*I restriction enzyme site for screening purposes. Incorporation of the ssODN sequence was
439 identified via PCR and restriction enzyme digestion and confirmed by sequencing of the 709 bp region
440 flanking the mutation. Mutant clones were generated from a pure C57BL/6 ESC line, a 129/BL6 hybrid
441 line expressing constitutive GFP and a reporter line used to de-differentiate ESCs into a 2C-like state.
442 Clones were tested by karyotype and any clones with <70% normal diploid content were excluded.
443
444 *RNA-seq*
445 ESCs were first depleted by feeder cells by incubation on plates coated with gelatin. Parental and derived
446 Oct4^{C48S} ESC lines were differentiated in media lacking LIF and containing 1.5mg/L RA for 2 days.
447 Undifferentiated and RA-differentiated cells were collected, and RNA was prepared using a kit (RNeasy,
448 Qiagen). RNA poly-A selected for mRNA using NEBNext rRNA Depletion Kit v2 (Human, Mouse, Rat)
449 (E7400). Bulk RNA was processed and sequenced as reported previously (Perovanovic et al. 2023).
450 Briefly, RNASeq reads were aligned to GRCh38/hg38 using STAR version 2.7.9a. Read counts were
451 normalized using the DESeq2 analysis package. The DESeq2-regularized log transformation was applied
452 to counts, which preferentially shrinks the overall variance among low abundant transcripts. The high-
453 throughput data are available on the gene expression omnibus (GEO) server under series record
454 GSE241332. The data are accessible for download through private token gtyfyseqljufjwz while in private
455 status. Differential genes were identified using 2.5-fold change and $p<0.001$ as cutoffs. GO terms were
456 identified using PANTHER (Mi et al. 2013).

457

458 *Teratoma formation*

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

463

464 *Chimeric mouse assays*

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

474

475 *Janky* mice

476 Mutant C57BL6/J mice with a single-base missense mutation that converts Cys48 to Ser (*Jky* mice) were
477 generated using the same templated CRISPR strategy used to generate mutant ESCs. CRISPR RNPs were
478 introduced into single cell C57BL6/J mouse embryos using pronuclear injection. Founder mice that
479 carried the targeted mutation were backcrossed to WT and the N1 generation was sequenced to confirm
480 the presence of the heterozygous allele.

481 **Competing interest statement**

482 The authors declare no competing interests.

483

484 **Acknowledgements**

485 We thank S. Buckley for critical reading of the manuscript. We thank C. Davey, H. Lan and the University
486 of Utah Mouse Transgenic and Gene Targeting Core as well as the Mutation Generation and Detection
487 Core for assistance with CRISPR-edited ESC lines, chimerism assays, karyotyping and the generation of
488 germline mutant mice. We thank B. Dalley and the High-Throughput Genomics Core facility, and Q. Li
489 and the Bioinformatics Analysis core facility for assistance with RNA-seq. We thank D. Lum and the
490 Preclinical Research Resource core facility for assistance with teratoma formation. We thank J. Marvin
491 and the Flow Cytometry Core, and X. Wang and the Cell Imaging Core. We thank B. Weaver for the gift
492 of the C4 ESC line. This work was supported by NIH/NIGMS grant (R01GM122778) to DT.

493

494 *Author contributions*: D.T. conceived the study. Z.S., Y.W. and D.T. designed experiments. Z.S., Y.W.
495 and M.B.C. generated new reagents for the study. Z.S., Y.W., A.M., C.Y., B.R.C. and K.J.E. performed
496 experiments and/or analyzed data. K.J.E. performed blinded pathological analysis of teratoma samples.
497 Z.S., Y.W. and A.M. generated figures. All authors contributed to the writing of the manuscript and read
498 and approved the final version of the manuscript prior to submission for publication.

499 **Figure legends**

500

501 **Figure 1.** Oct4 reprogramming activity vis-à-vis Oct1 maps to the DBD. (*A*) Mouse Oct1 vs Oct4 DBD
502 alignment. POU_S and POU_H refer to the two DNA binding sub-domains. Asparagine 40 (blue) was
503 mutated to lysine in some constructs. The key Oct1 serine 48 residue (cysteine in Oct4) is boxed in red.
504 (*B*) Schematic of the reprogramming vector with C-terminal Oct4 FLAG tag. HIV=human
505 immunodeficiency virus. LTR=long terminal repeat. O=Oct4. K=Klf4. S=Sox2. M=c-Myc. The different
506 domains of Oct4 and Oct1 are color coded. Chimeric constructs are shown at right. The construct name
507 indicates which domain of Oct4 is included separated by slashes, e.g. 4N/C=the Oct4 N- and C-termini
508 fused to the rest of Oct1. Point mutations are shown with a dash. (*C*) Reprogramming activity of C-
509 terminally tagged Oct4 relative to the untagged control. Results show an average of N=3 independent
510 experiments. Error bars show ±standard deviation. Anti-FLAG and anti-Oct4 immunoblots using tagged
511 and untagged constructs transiently transfected into 293T cells are shown at right. (*D*) iPSC formation
512 assay using target Oct4-GFP MEFs. Reprogramming was assessed using the number of GFP-expressing
513 iPSC colonies. Oct4 was used as a positive control and Oct1 as a negative control. All constructs bore C-
514 terminal FLAG-tags. N=3 experiments were performed. Error bars show ±standard deviation. (*E*)
515 Schematic and summary of size and activity of the constructs used in D. To determine % activity, the
516 averaged colony count using Oct4 in D was set to 100%. C-terminal FLAG-tags are not shown for
517 simplicity. (*F*) Anti-FLAG immunoblot showing expression of the different constructs when transiently
518 transfected into 293T cells transduced with the reprogramming vectors. Histone H3 is shown as a loading
519 control.

520

521 **Figure 2.** Activity of Oct4 cysteine 48 in reprogramming. (*A*) iPSC formation assay using Oct1-Oct4
522 lentiviral constructs that manipulate cysteine 4. Oct4 was used as a positive control and Oct1 as a negative
523 control. All constructs bore C-terminal FLAG-tags. N=3 experiments. Error bars show ±standard

524 deviation. (*B*) Schematic and summary of size and activity of the constructs used in A. The activity of
525 Oct4 in A was set to 100%. (*C*) Anti-FLAG immunoblot showing expression of the different constructs
526 when transiently transfected into 293T cells transduced with the reprogramming vectors. Histone H3 is
527 shown as a loading control. (*D*) Detail from an Oct1:octamer DNA co-crystal structure (PDB ID:
528 1OCT)(Klemm et al. 1994) showing Ser48 making hydrogen bond contacts with the DNA backbone. (*E*)
529 Paralog comparison of mouse POU protein amino acid residues corresponding to α-helix 3 of the POU_S
530 sub-domain.
531
532 **Figure 3.** Oct4 cysteine 48 confers DNA sensitivity to oxidative stress and oxidative stress-induced
533 ubiquitylation. (*A*) EMSA using canonical octamer DNA and wild type of C48S full-length Oct4. 293T
534 cells were transiently transfected with plasmids encoding C-terminal FLAG-tagged Oct4 constructs.
535 Protein lysates were made and immediately used for EMSA with increasing amounts of diamide. The
536 larger redox-resistant Oct1 protein that binds wild type (WT) but not mutant (Mut) probe is highlighted
537 with an arrow. Asterisk shows a redox sensitive band with superior binding to mutant rather than octamer
538 DNA. (*B*) Quantification of Oct4 band intensity relative to no diamide, which was set to 100%. N=3
539 independent experiments. Error bars show ±standard deviation. (*C*) Oct4 containing a C-terminal
540 streptavidin tag was purified from 293T cell lysates using anti-FLAG beads. Isolated protein was used in
541 assays similar to A. (*D*) Quantification of C and two similar experiments. Error bars show ±standard
542 deviation. (*E*) Highly exposed Oct4 immunoblot using lysates prepared from 293T cells transfected with
543 constructs expressing C-terminally FLAG/Twin-Strep-tagged Oct4. Subsets of cells were pre-treated with
544 H_2O_2 (1 mM for 2 hr), MG-132 (10 μM for 2 hr) or both. All samples contained MG-132 in the lysis
545 buffer to prevent degradation. Histone H3 is shown as a loading control. (*F*) Purified wild type and C48S
546 Oct4 was immunoblotted using antibodies against specific ubiquitin linkages. (*G*) Oct4 and control

547 immunoprecipitates from wild type ESCs were immunoblotted using FK2 antibodies. Cells were treated
548 with 10 μM MG-132 for 2 hr or 2.5 mM H_2O_2 for 2 hr.
549
550 **Figure 4.** Oct4^{C48S} ESCs are phenotypically normal but retain Oct4 expression when differentiated. (*A*)
551 *Top:* Schematic for CRISPR-mediated *Pou5f1* (*Oct4*) mutation. A single nucleotide change (arrow)
552 converts the wild type cysteine to the serine present in Oct1. Asterisks depict two silent point mutations
553 present in the guide RNA to monitor mutagenesis. *Bottom:* DNA sequencing trace of homozygous mutant
554 clone. (*B*) An example C57BL/6 ESC clone (#46) was immunoblotted for Oct4, Sox2, Nanog and β-actin
555 as a loading control. (*C*) Quantification of Oct4 protein levels from N=4 immunoblots. Levels were
556 normalized to β-actin. Error bars depict ±standard deviation. (*D*) Example images of pluripotent colonies
557 from a parental ESC line and derived Oct4^{C48S} ESC line (#46). Images were collected at 10×
558 magnification. Scale bar = 0.2 mm. (*E*) Pan-Ub and control Oct4 immunoblots of Oct4
559 immunoprecipitates from parent and Oct4^{C48S} (clone #2) ESC lines. Cells were grown on gelatin without
560 feeders. Normal culture medium was replaced with medium lacking supplemental glutamine and β-
561 mercaptoethanol 2 hr prior to the experiment. Where indicated, cells were treated with 2.5 mM H_2O_2 for
562 a further 2 hr. Oct4 was immunoprecipitated using a rabbit polyclonal antibody (Abcam) and
563 immunoblotted using monoclonal mouse antibodies against pan-Ub and Oct4. (*F*) Parent or derived
564 Oct4^{C48S} ESCs were differentiated using RA. Lysates were collected at the indicated times and probed for
565 Oct4 by immunoblotting. β-actin was used as a loading control. (*G*) Average quantified Oct4/β-actin
566 levels from N=3 immunoblots conducted with mutant ESC clone #46. Error bars depict ±standard
567 deviation. (*H*) mRNA from parental and #46 mutant cells differentiated for 2 days with RA, and
568 undifferentiated controls, were transcriptionally profiled using RNA-seq. Significantly increased genes in
569 differentiating C48S relative to parent cells (421 genes) were intersected with the set of genes down

570 regulated in parental cells upon differentiation (1536 genes), to yield 169 genes. (*I*) PANTHER GO terms
571 associated with 169 genes from G identified using the "biological process" ontology.
572
573 **Figure 5.** Oct4^{C48S} ESCs differentiate abnormally. (*A*) Phase-contrast images of parent and derived
574 Oct4^{C48S} ESCs (clone #46) differentiated using RA for 12 days. Images were collected at the indicated
575 times at 10× magnification. Scale bar = 0.2 mm. (*B*) Differentiating cells were trypsinized and counted
576 using trypan blue exclusion using a hemacytometer. Cell counts (N=3) were averaged and plotted. Error
577 bars depict ±standard deviation. (*C*) Proliferation was measured in parent and C48S ESCs differentiating
578 with RA at days 2, 3 and 4. Cells were loaded with cell trace violet (CTV) at day 2. (*D*) TUNEL assay
579 showing Oct4^{C48S} ESCs differentiated using RA. Cells were differentiated for 2 days and incubated with
580 FITC-labeled dUTP for 1 hour at 37°C prior to flow cytometry. (*E*) Undifferentiated and 3-day
581 differentiating parent ESCs were incubated with DCFDA to assess ROS levels. (*F*) Parent and derived
582 Oct4^{C48S} ESCs (clones #2 and 46) were differentiated into EBs, mesoderm lineage cells and
583 cardiomyocytes to assess beating. Average number of beating events per well is shown for N=36 wells,
584 with 10 EBs plated per well. Error bars depict ±standard deviation.
585
586 **Figure 6.** Oct4^{C48S} ESCs form teratomas lacking maturing structures and contribute poorly to adult
587 chimeras. (*A*) 1×10^6 parental or Oct4^{C48S} C57BL/6:129 ESCs (clone G4-N) were injected into contralateral
588 flanks or NRG mice. After 15 days, tumors were harvested, fixed, sectioned, stained with H&E,
589 deidentified and scored pathologically. The percent of cells within the tumor containing undifferentiated
590 components was averaged and plotted. N=5 independent tumors for each condition combined from two
591 separate experiments. For each tumor, values represent the average from a pair of slides. Two nonadjacent
592 sections from each tumor were used. Error bars depict ±standard deviation. *P*-value was generated using
593 an unpaired one-tailed T-test. (*B*) Representative images are shown. Mature (differentiated) components,

594 e.g. airway epithelium in teratomas derived from parental ESCs are highlighted by yellow arrows. Asterisk
595 indicates area of necrosis. (*C*) Similar data as in (A), except generated using an independently derived
596 mutant from a different ESC line. N=5 independent tumors. (*D*) Three different parent ESC lines and 2-4
597 derived lines of CRISPR-mutated Oct4^{C48S} ESCs were injected into recipient blastocysts (N=number
598 injected) and implanted into pseudo-pregnant animals. For each line, the proportion of chimeric offspring
599 at weaning is shown as a pie chart. All karyotypes were confirmed to be similar to parent lines (between
600 72% and 92% normal).
601
602 **Supplemental figure legends**
603
604 **Figure S1.** Effect of N-terminal FLAG-tag on iPSC formation. (*A*) Similar to Fig. 1C, except cells were
605 transduced with lentiviral vectors expressing untagged and N-terminal rather than C-terminal FLAG-
606 tagged Oct4. (*B*) Anti-FLAG and Oct4 immunoblots showing expression of tagged and untagged Oct4.
607 α-tubulin is shown as a loading control.
608
609 **Figure S2.** GFP^{+} iPSC numbers over time using different Oct1-Oct4 chimeric constructs. (*A*) Similar to
610 Fig. 1D except multiple time points are shown. Wild type Oct4 is shown as a positive control. Results
611 show an average of experiments performed in biological triplicate. Error bars show ±standard deviation.
612 (*B*) Titers for six of the viral supernatants used in Fig.1 and in (A) were tested for p24 levels using a kit.
613 Averages of N=2 independent experiments are shown. TU=transduction units.
614
615 **Figure S3.** iPSC colony morphology and GFP intensity using different chimeric and mutant
616 reprogramming constructs. (*A*) Phase-contrast and epifluorescence images are shown of example iPSC
617 colonies generated using the constructs shown in Fig. 2. Scale bar=200 μm. (B) For each construct in (A),
618 GFP intensity from N=X colonies from three different wells was quantified by dividing the total GFP

619 pixel intensity by the area of each colony, generating an average GFP intensity per unit area (arbitrary
620 units). Values for each well were then averaged. Error bars denote ±standard deviation.
621
622 **Figure S4.** The effect of Cys48 on iPSC colony formation is not dependent on kinetics or on using Oct4-
623 GFP MEFs. (*A*) GFP^+ iPSC numbers over time using different Oct1-Oct4 chimeric constructs. Similar to
624 Fig. 2A except multiple time points are shown. Wild type Oct4 is shown as a positive control. Results
625 show an average of experiments performed in biological triplicate. Error bars show ±standard deviation.
626 (*B*) An iPSC colony formation assay using non-GFP primary MEFs and alkaline phosphatase staining
627 rather than GFP positivity and pluripotent morphology to score iPSC formation. N=3. Error bars show
628 ±standard deviation. (*C*) Titers of the viral supernatants used in (B) were measured based on viral RNA
629 levels using a kit. Averages of N=3 replicates are shown. Error bars show ±standard deviation.
630 IFU=infectious units.
631
632 **Figure S5.** Oxidative inhibition of Oct4 DNA binding can be reversed with the addition of thioredoxin.
633 EMSA is shown similar to Fig. 3C, except that after 15 min incubation with diamide, thioredoxin was
634 added at the indicated concentrations. Reactions were incubated for a further 15 min at 37°C prior to
635 native electrophoresis.
636
637 **Figure S6.** Purification of recombinant wild type and C48S mutant Oct4 from expi-293T cells. Oct4 was
638 FLAG-tagged at the C-terminus. Protein was purified using beads covalently coupled to anti-FLAG
639 antibodies. Lysates were prepared 48 hr post-transfection.
640
641 **Figure S7.** H_2O_2 and MG-132 treatment of primary MEFs largely maintains viability. Phase-contrast
642 images are shown of Oct4-GFP MEFs treated with the same concentrations and timepoints of H_2O_2 and
643 MG-132 used in Fig. 3E. Scale bar=200 μm.

644

645 **Figure S8.** Oct4 C48S promotes Oct4 ubiquitylation. Similar to Fig. 3F except additional antibodies were
646 used for immunoblotting: an antibody detecting K63-linked Ub chains (Cell Signaling) and a pan-Oct4
647 antibody as a positive control (Enzo Life Sciences). Oct4 was Twin-Strep- and FLAG-tagged at the C-
648 terminus.

649

650 **Figure S9.** The effects of Oct4 C48S mutation are not parent line- or clone-specific. (*A*) High-resolution
651 melting analysis of parental, homozygous mutant and Indel clones generated by CRISPR targeting of
652 *Pou5f1* in mouse ESCs. The majority of clones were Indels and a smaller number of clones were
653 homozygous mutant. Heterozygous and unedited (wild-type) lines were not recovered by the procedure.
654 (*B*) Similar to Fig. 4D except using alkaline phosphatase staining. (C) Similar to Fig. 4B except a different
655 clone from the same parent line (#2) was tested. (*C*) Similar to Fig. 4F except a different clone from the
656 same parent line (#2) was tested. (*D*) Quantification from N=3 immunoblots. Background was subtracted
657 from individual bands, and intensity for each protein was divided by H3 loading control to generate a
658 measure of relative protein level in arbitrary units. Error bars denote ±standard deviation. Quantification
659 was performed with Image J software. (*E*) Similar to Fig. 4F except the same clone was tested against
660 parent ESCs that had been simultaneously electroporated without CRISPR RNP ("zap" control). (*F*)
661 Similar to Fig. 4F except a different clone from the same parent line (#2) was tested. (*G*) Similar to Fig.
662 4F except a different clone (#18) from a different parent line (G4-N) was tested.

663

664 **Figure S10.** RNA-seq analysis of differentiating parental and Oct4^{C48S} ESCs. (*A*) Similarity between bulk
665 RNA-seq replicates and conditions. Clustering of the RNA-seq replicates using Euclidean distance is
666 shown. (*B*) The union of the top 100 most differentially expressed genes upon differentiation for both
667 genotypes (150 genes total) was subjected to hierarchical clustering and displayed as a heat map. Example
668 genes are shown at right.

669

670 **Figure S11.** Additional G4-N clone #18 teratoma images similar to Fig.6D. Asterisk represents necrotic
671 regions.

672

673 **Figure S12.** The decreased differentiation of teratomas from Oct4^{C48S} ESCs normalizes in larger tumors
674 at later timepoints. (*A*) Representative H&E-stained sections of tumors dissected from parental and mutant
675 teratomas at day 28. Regions surrounded by black lines highlight immature neuroepithelial elements. Blue
676 lines highlight mature epithelium (squamous, respiratory, and/or gastroinstestinal). Red asterisk represents
677 mature neural regions. Green lines show mature cartilage. Yellow lines show mature fat. Orange lines
678 show yolk sac (immature). (*B*) Average percent immature component for day 28 parental homozygous
679 mutant teratomas. N=3 individual tumors. Error bars denote ±standard deviation. (*C*) Similar analysis
680 performed for a different parent and homozygous Oct4^{C48S} mutant BL/6 ESC pair grown as teratomas for
681 28 days. N=5.

682

683 **Figure S13.** Abnormal testes in *Jky* homozygous mice. (*A*) Gross images of dissected testes from a 2.5
684 month-old male *Jky*/*Jky* mouse and +/+ littermate control. Averaged weights of two testes from each
685 animal are also shown. (*B*) H&E staining of FFPE testes sections from *Jky*/*Jky* and +/+ littermate control
686 testes. (*C*) Representative immunofluorescence images showing abnormal spermatogenesis and germ cell
687 loss in testis sections from the *Oct4*$^{Jky/Jky}$ mouse: gamma H2AX (green); LIN28A (Red); SYCP3 (violet);
688 and DNA (Cyan). Box #1 highlights examples of superficially normal spermatogenesis, while #2
689 highlights example abnormal localization of spermatocytes and germ cell loss. Arrow points to abnormal
690 localization of spermatocytes within interstitial regions.

691

692 **Figure S14.** Uncropped immunoblots for Figs. 1-3.

693

694 **Figure S15.** Uncropped immunoblots for Fig. 4.

695

696 **Figure S16.** Uncropped immunoblots for Suppl. Fig. S1 and S8.

697

698 **Figure S17.** Uncropped immunoblots for Suppl. Fig. S9.

699

700 **Movie S1.** Example of beating cells derived from parent EBs plated and differentiated into
701 cardiomyocytes.

702

703 **Movie S2.** Example of beating cells derived from parent EBs plated and differentiated into
704 cardiomyocytes.

705

706 **Movie S3.** Example of non-beating cells derived from mutant #2 EBs plated and differentiated into
707 cardiomyocytes.

708

709 **Movie S4.** Example of non-beating cells derived from mutant #2 EBs plated and differentiated into
710 cardiomyocytes.

711

712 **Movie S5.** Example of non-beating cells derived from mutant #46 EBs plated and differentiated into
713 cardiomyocytes.

714

715 **Movie S6.** Example of non-beating cells derived from mutant #46 EBs plated and differentiated into
716 cardiomyocytes.

717 **References**

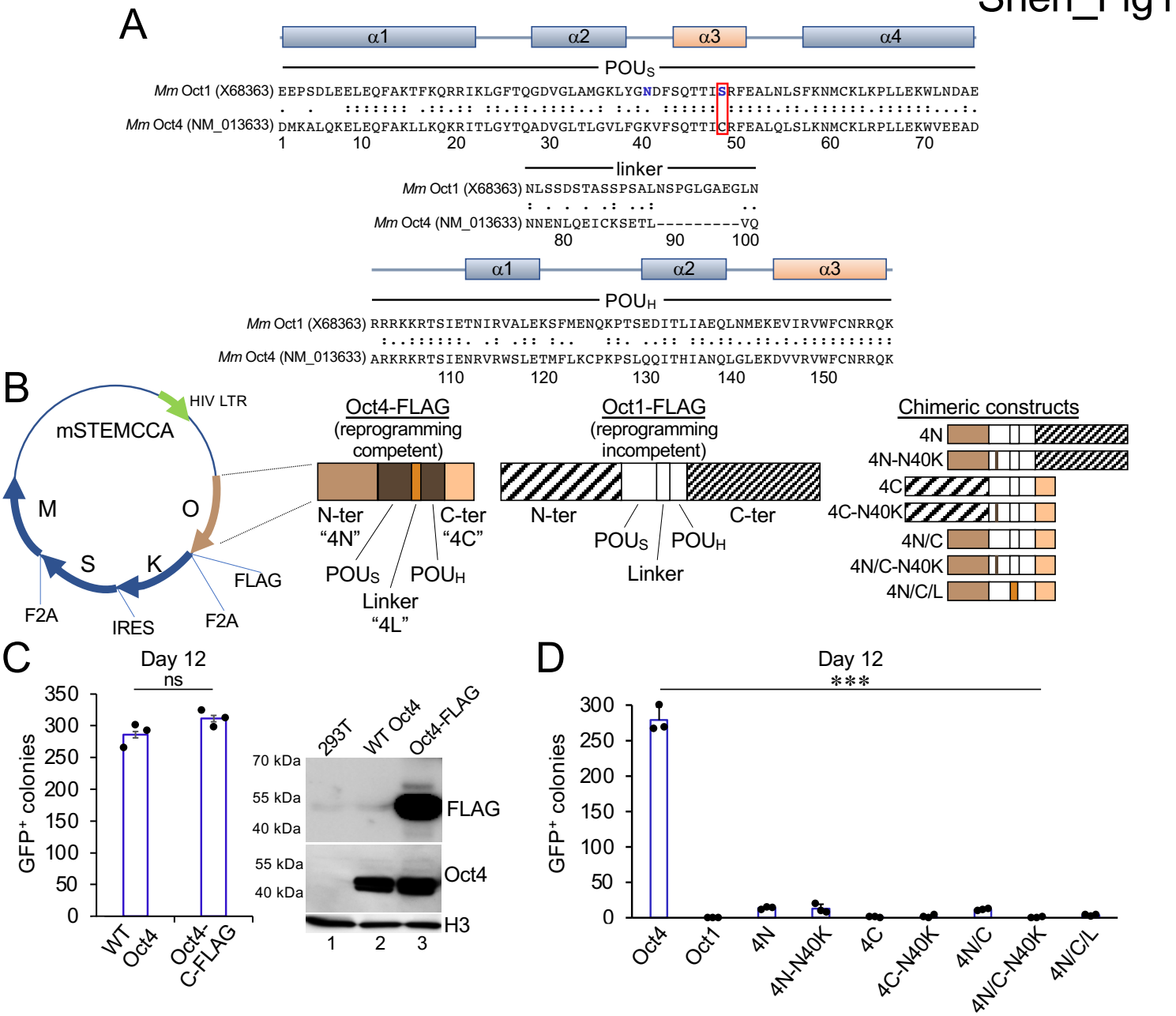
718 Aksoy I, Jauch R, Chen J, Dyla M, Divakar U, Bogu GK, Teo R, Leng Ng CK, Herath W, Lili S et al.
719 2013. Oct4 switches partnering from Sox2 to Sox17 to reinterpret the enhancer code and specify
720 endoderm. *EMBO J* **32**: 938-953.
721 An Z, Liu P, Zheng J, Si C, Li T, Chen Y, Ma T, Zhang MQ, Zhou Q, Ding S. 2019. Sox2 and Klf4 as the
722 Functional Core in Pluripotency Induction Without Exogenous Oct4. *Cell Rep* **29**: 1986-2000.
723 Boija A, Klein IA, Sabari BR, Dall'Agnese A, Coffey EL, Zamudio AV, Li CH, Shrinivas K, Manteiga
724 JC, Hannett NM et al. 2018. Transcription Factors Activate Genes through the Phase-Separation
725 Capacity of Their Activation Domains. *Cell* **175**: 1842-1855 e1816.
726 Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray
727 HL, Jenner RG et al. 2005. Core transcriptional regulatory circuitry in human embryonic stem
728 cells. *Cell* **122**: 947-956.
729 Britto PJ, Knipling L, Wolff J. 2002. The local electrostatic environment determines cysteine reactivity of
730 tubulin. *J Biol Chem* **277**: 29018-29027.
731 Buganim Y, Markoulaki S, van Wietmarschen N, Hoke H, Wu T, Ganz K, Akhtar-Zaidi B, He Y,
732 Abraham BJ, Porubsky D et al. 2014. The developmental potential of iPSCs is greatly influenced
733 by reprogramming factor selection. *Cell Stem Cell* **15**: 295-309.
734 Chen J, Gao Y, Huang H, Xu K, Chen X, Jiang Y, Li H, Gao S, Tao Y, Wang H et al. 2015. The
735 combination of Tet1 with Oct4 generates high-quality mouse-induced pluripotent stem cells. *Stem*
736 *Cells* **33**: 686-698.
737 Costa Y, Ding J, Theunissen TW, Faiola F, Hore TA, Shliaha PV, Fidalgo M, Saunders A, Lawrence M,
738 Dietmann S et al. 2013. NANOG-dependent function of TET1 and TET2 in establishment of
739 pluripotency. *Nature* **495**: 370-374.
740 De Los Angeles A, Ferrari F, Xi R, Fujiwara Y, Benvenisty N, Deng H, Hochedlinger K, Jaenisch R, Lee
741 S, Leitch HG et al. 2015. Hallmarks of pluripotency. *Nature* **525**: 469-478.
742 DeVeale B, Brokhman I, Mohseni P, Babak T, Yoon C, Lin A, Onishi K, Tomilin A, Pevny L, Zandstra
743 PW et al. 2013. Oct4 is required ~E7.5 for proliferation in the primitive streak. *PLoS Genet* **9**:
744 e1003957.
745 Esch D, Vahokoski J, Groves MR, Pogenberg V, Cojocaru V, Vom Bruch H, Han D, Drexler HC, Arauzo-
746 Bravo MJ, Ng CK et al. 2013. A unique Oct4 interface is crucial for reprogramming to
747 pluripotency. *Nat Cell Biol* **15**: 295-301.
748 Esteban MA, Wang T, Qin B, Yang J, Qin D, Cai J, Li W, Weng Z, Chen J, Ni S et al. 2010. Vitamin C
749 enhances the generation of mouse and human induced pluripotent stem cells. *Cell Stem Cell* **6**: 71-
750 79.
751 Evans MJ, Kaufman MH. 1981. Establishment in culture of pluripotential cells from mouse embryos.
752 *Nature* **292**: 154-156.
753 Feng B, Jiang J, Kraus P, Ng JH, Heng JC, Chan YS, Yaw LP, Zhang W, Loh YH, Han J et al. 2009.
754 Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor
755 Esrrb. *Nat Cell Biol* **11**: 197-203.
756 Fernandez-Capetillo O, Mahadevaiah SK, Celeste A, Romanienko PJ, Camerini-Otero RD, Bonner WM,
757 Manova K, Burgoyne P, Nussenzweig A. 2003. H2AX is required for chromatin remodeling and
758 inactivation of sex chromosomes in male mouse meiosis. *Dev Cell* **4**: 497-508.
759 Ferraris L, Stewart AP, Kang J, DeSimone AM, Gemberling M, Tantin D, Fairbrother WG. 2011.
760 Combinatorial binding of transcription factors in the pluripotency regions of the genome. *Genome*
761 *Res* **21**: 1055-1064.
762 Gao Y, Chen J, Li K, Wu T, Huang B, Liu W, Kou X, Zhang Y, Huang H, Jiang Y et al. 2013. Replacement
763 of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and
764 hydroxymethylation in reprogramming. *Cell Stem Cell* **12**: 453-469.

765 Guo Y, Einhorn L, Kelley M, Hirota K, Yodoi J, Reinbold R, Scholer H, Ramsey H, Hromas R. 2004.
766 Redox regulation of the embryonic stem cell transcription factor Oct-4 by thioredoxin. *Stem Cells*
767 **22**: 259-264.
768 Han D, Wu G, Chen R, Drexler HCA, MacCarthy CM, Kim KP, Adachi K, Gerovska D, Mavrommatis
769 L, Bedzhov I et al. 2022. A balanced Oct4 interactome is crucial for maintaining pluripotency. *Sci*
770 *Adv* **8**: eabe4375.
771 Heng JC, Feng B, Han J, Jiang J, Kraus P, Ng JH, Orlov YL, Huss M, Yang L, Lufkin T et al. 2010. The
772 nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to
773 pluripotent cells. *Cell Stem Cell* **6**: 167-174.
774 Ilia K, Shakiba N, Bingham T, Jones RD, Kaminski MM, Aravera E, Bruno S, Palacios S, Weiss R, Collins
775 JJ et al. 2023. Synthetic genetic circuits to uncover the OCT4 trajectories of successful
776 reprogramming of human fibroblasts. *Sci Adv* **9**: eadg8495.
777 Jerabek S, Ng CK, Wu G, Arauzo-Bravo MJ, Kim KP, Esch D, Malik V, Chen Y, Velychko S, MacCarthy
778 CM et al. 2017. Changing POU dimerization preferences converts Oct6 into a pluripotency
779 inducer. *EMBO Rep* **18**: 319-333.
780 Jin W, Wang L, Zhu F, Tan W, Lin W, Chen D, Sun Q, Xia Z. 2016. Critical POU domain residues confer
781 Oct4 uniqueness in somatic cell reprogramming. *Sci Rep* **6**: 20818.
782 Kang J, Gemberling M, Nakamura M, Whitby FG, Handa H, Fairbrother WG, Tantin D. 2009. A general
783 mechanism for transcription regulation by Oct1 and Oct4 in response to genotoxic and oxidative
784 stress. *Genes Dev* **23**: 208-222.
785 Kehler J, Tolkunova E, Koschorz B, Pesce M, Gentile L, Boiani M, Lomeli H, Nagy A, McLaughlin KJ,
786 Scholer HR et al. 2004. Oct4 is required for primordial germ cell survival. *EMBO Rep* **5**: 1078-
787 1083.
788 Kim KP, Wu Y, Yoon J, Adachi K, Wu G, Velychko S, MacCarthy CM, Shin B, Ropke A, Arauzo-Bravo
789 MJ et al. 2020. Reprogramming competence of OCT factors is determined by transactivation
790 domains. *Sci Adv* **6**: eaaz7364.
791 King HW, Klose RJ. 2017. The pioneer factor OCT4 requires the chromatin remodeller BRG1 to support
792 gene regulatory element function in mouse embryonic stem cells. *Elife* **6**.
793 Klemm JD, Rould MA, Aurora R, Herr W, Pabo CO. 1994. Crystal structure of the Oct-1 POU domain
794 bound to an octamer site: DNA recognition with tethered DNA-binding modules. *Cell* **77**: 21-32.
795 Lengner CJ, Camargo FD, Hochedlinger K, Welstead GG, Zaidi S, Gokhale S, Scholer HR, Tomilin A,
796 Jaenisch R. 2007. Oct4 expression is not required for mouse somatic stem cell self-renewal. *Cell*
797 *Stem Cell* **1**: 403-415.
798 Lickteig K, Lamb K, Brigman K, Rizzino A. 1996. Effects of oxidation and reduction on the binding of
799 transcription factors to cis-regulatory elements located in the FGF-4 gene. *Mol Reprod Dev* **44**:
800 146-152.
801 Lynch AT, Mazzotta S, Hoppler S. 2018. Cardiomyocyte Differentiation from Mouse Embryonic Stem
802 Cells. *Methods Mol Biol* **1816**: 55-66.
803 Mamun MMA, Khan MR, Zhu Y, Zhang Y, Zhou S, Xu R, Bukhari I, Thorne RF, Li J, Zhang XD et al.
804 2022. Stub1 maintains proteostasis of master transcription factors in embryonic stem cells. *Cell*
805 *Rep* **39**: 110919.
806 Marsboom G, Zhang GF, Pohl-Avila N, Zhang Y, Yuan Y, Kang H, Hao B, Brunengraber H, Malik AB,
807 Rehman J. 2016. Glutamine Metabolism Regulates the Pluripotency Transcription Factor Oct4.
808 *Cell Rep* **16**: 323-332.
809 Mi H, Muruganujan A, Casagrande JT, Thomas PD. 2013. Large-scale gene function analysis with the
810 PANTHER classification system. *Nat Protoc* **8**: 1551-1566.
811 Michael AK, Grand RS, Isbel L, Cavadini S, Kozicka Z, Kempf G, Bunker RD, Schenk AD, Graff-Meyer
812 A, Pathare GR et al. 2020. Mechanisms of OCT4-SOX2 motif readout on nucleosomes. *Science*
813 **368**: 1460-1465.

814 Morey L, Santanach A, Di Croce L. 2015. Pluripotency and Epigenetic Factors in Mouse Embryonic Stem
815 Cell Fate Regulation. *Mol Cell Biol* **35**: 2716-2728.
816 Nakagawa M, Koyanagi M, Tanabe K, Takahashi K, Ichisaka T, Aoi T, Okita K, Mochiduki Y, Takizawa
817 N, Yamanaka S. 2008. Generation of induced pluripotent stem cells without Myc from mouse and
818 human fibroblasts. *Nat Biotechnol* **26**: 101-106.
819 Nichols J, Zevnik B, Anastassiadis K, Niwa H, Klewe-Nebenius D, Chambers I, Scholer H, Smith A.
820 1998. Formation of pluripotent stem cells in the mammalian embryo depends on the POU
821 transcription factor Oct4. *Cell* **95**: 379-391.
822 Okamoto K, Okazawa H, Okuda A, Sakai M, Muramatsu M, Hamada H. 1990. A novel octamer binding
823 transcription factor is differentially expressed in mouse embryonic cells. *Cell* **60**: 461-472.
824 Perovanovic J, Wu Y, Abewe H, Shen Z, Hughes EP, Gertz J, Chandrasekharan MB, Tantin D. 2023.
825 Oct1 cooperates with the Smad family of transcription factors to promote mesodermal lineage
826 specification. *Sci Signal* **16**: eadd5750.
827 Radzisheuskaya A, Le Bin Chia G, Dos Santos RL, Theunissen TW, Castro LF, Nichols J, Silva JC. 2013.
828 A defined Oct4 level governs cell state transitions of pluripotency entry and differentiation into all
829 embryonic lineages. *Nat Cell Biol*.
830 Rais Y, Zviran A, Geula S, Gafni O, Chomsky E, Viukov S, Mansour AA, Caspi I, Krupalnik V, Zerbib
831 M et al. 2013. Deterministic direct reprogramming of somatic cells to pluripotency. *Nature* **502**:
832 65-70.
833 Rosner MH, Vigano MA, Ozato K, Timmons PM, Poirier F, Rigby PW, Staudt LM. 1990. A POU-domain
834 transcription factor in early stem cells and germ cells of the mammalian embryo. *Nature* **345**: 686-
835 692.
836 Ruetz T, Pfisterer U, Di Stefano B, Ashmore J, Beniazza M, Tian TV, Kaemena DF, Tosti L, Tan W,
837 Manning JR et al. 2017. Constitutively Active SMAD2/3 Are Broad-Scope Potentiators of
838 Transcription-Factor-Mediated Cellular Reprogramming. *Cell Stem Cell* **21**: 791-805 e799.
839 Scholer HR, Dressler GR, Balling R, Rohdewohld H, Gruss P. 1990a. Oct-4: a germline-specific
840 transcription factor mapping to the mouse t-complex. *EMBO J* **9**: 2185-2195.
841 Scholer HR, Ruppert S, Suzuki N, Chowdhury K, Gruss P. 1990b. New type of POU domain in germ line-
842 specific protein Oct-4. *Nature* **344**: 435-439.
843 Shakya A, Callister C, Goren A, Yosef N, Garg N, Khoddami V, Nix D, Regev A, Tantin D. 2015.
844 Pluripotency transcription factor Oct4 mediates stepwise nucleosome demethylation and
845 depletion. *Mol Cell Biol* **35**: 1014-1025.
846 Shen Z, Formosa T, Tantin D. 2018. FACT Inhibition Blocks Induction But Not Maintenance of
847 Pluripotency. *Stem Cells Dev* **27**: 1693-1701.
848 Shen Z, Kang J, Shakya A, Tabaka M, Jarboe EA, Regev A, Tantin D. 2017. Enforcement of
849 developmental lineage specificity by transcription factor Oct1. *ELife* **6**: e20937.
850 Shi Y, Desponts C, Do JT, Hahm HS, Scholer HR, Ding S. 2008. Induction of pluripotent stem cells from
851 mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. *Cell Stem Cell*
852 **3**: 568-574.
853 Somers A, Jean JC, Sommer CA, Omari A, Ford CC, Mills JA, Ying L, Sommer AG, Jean JM, Smith BW
854 et al. 2010. Generation of transgene-free lung disease-specific human induced pluripotent stem
855 cells using a single excisable lentiviral stem cell cassette. *Stem Cells* **28**: 1728-1740.
856 Sommer CA, Stadtfeld M, Murphy GJ, Hochedlinger K, Kotton DN, Mostoslavsky G. 2009. Induced
857 pluripotent stem cell generation using a single lentiviral stem cell cassette. *Stem Cells* **27**: 543-
858 549.
859 Soufi A, Donahue G, Zaret KS. 2012. Facilitators and impediments of the pluripotency reprogramming
860 factors' initial engagement with the genome. *Cell* **151**: 994-1004.
861 Suzuki N, Rohdewohld H, Neuman T, Gruss P, Scholer HR. 1990. Oct-6: a POU transcription factor
862 expressed in embryonal stem cells and in the developing brain. *EMBO J* **9**: 3723-3732.

863 Takahashi K, Yamanaka S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult
864 fibroblast cultures by defined factors. *Cell* **126**: 663-676.
865 Tantin D. 2013. Oct transcription factors in development and stem cells: insights and mechanisms.
866 *Development* **140**: 2857-2866.
867 Velychko S, Adachi K, Kim KP, Hou Y, MacCarthy CM, Wu G, Scholer HR. 2019. Excluding Oct4 from
868 Yamanaka Cocktail Unleashes the Developmental Potential of iPSCs. *Cell Stem Cell* **25**: 737-753.
869 Vierbuchen T, Wernig M. 2012. Molecular roadblocks for cellular reprogramming. *Mol Cell* **47**: 827-838.
870 Wang Z, Oron E, Nelson B, Razis S, Ivanova N. 2012. Distinct lineage specification roles for NANOG,
871 OCT4, and SOX2 in human embryonic stem cells. *Cell Stem Cell* **10**: 440-454.
872 Yuan X, Wan H, Zhao X, Zhu S, Zhou Q, Ding S. 2011. Combined Chemical Treatment Enables Oct4-
873 Induced Reprogramming from Mouse Embryonic Fibroblasts. *Stem Cells* **29**: 549-553.

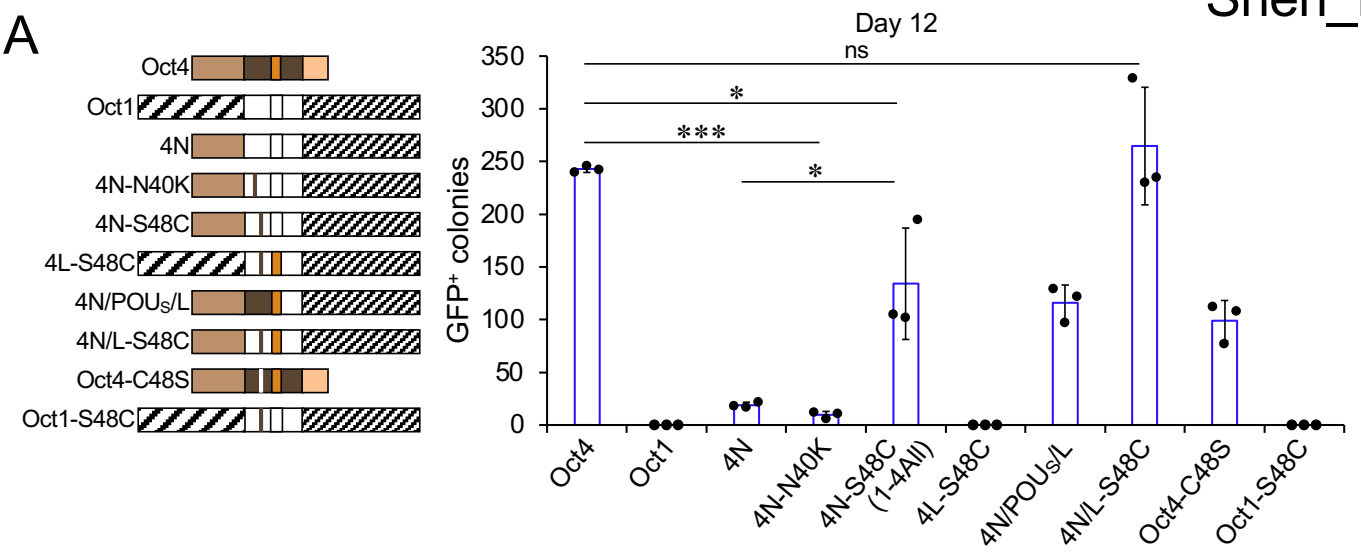
Shen_Fig1



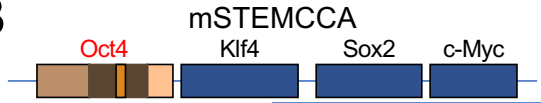
Construct (C-terminal FLAG)	Size (kDa)	Activity (%)
Oct4	42	100
Oct1	100	0.0
4N	75	4.8
4N-N40K	75	4.5
4C	60	0.4
4C-N40K	60	1.0
4N/C	43	6.1
4N/C-N40K	42	0.1
4N/C/L	42	1.1

Shen_Fig2

A

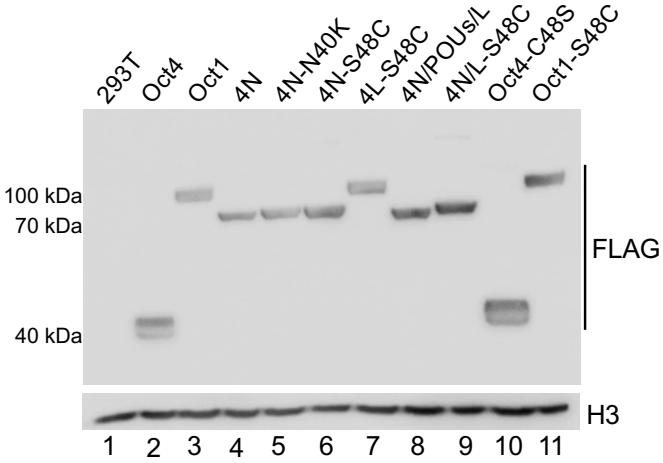


B

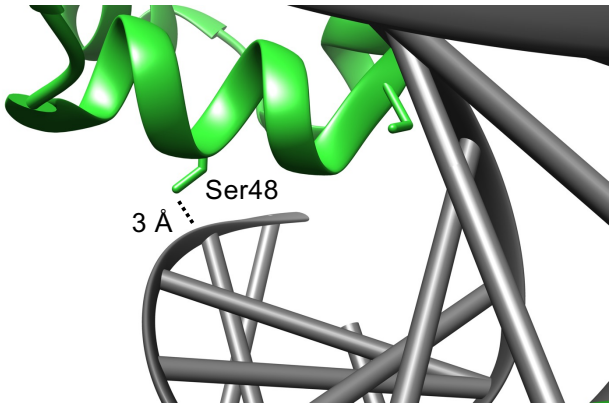


Construct (C-terminal FLAG)	Size (kDa)	Activity (%)
Oct4	42	100
Oct1	100	0.0
4N	75	7.8
4N-N40K	75	4.0
4N-S48C “1-4All”	75	55.2
4L-S48C	98	0
4N/POUs/L	72	47.8
4N/L-S48C	73	109
Oct4-C48S	42	40.8
Oct1-S48C	100	0

C



D

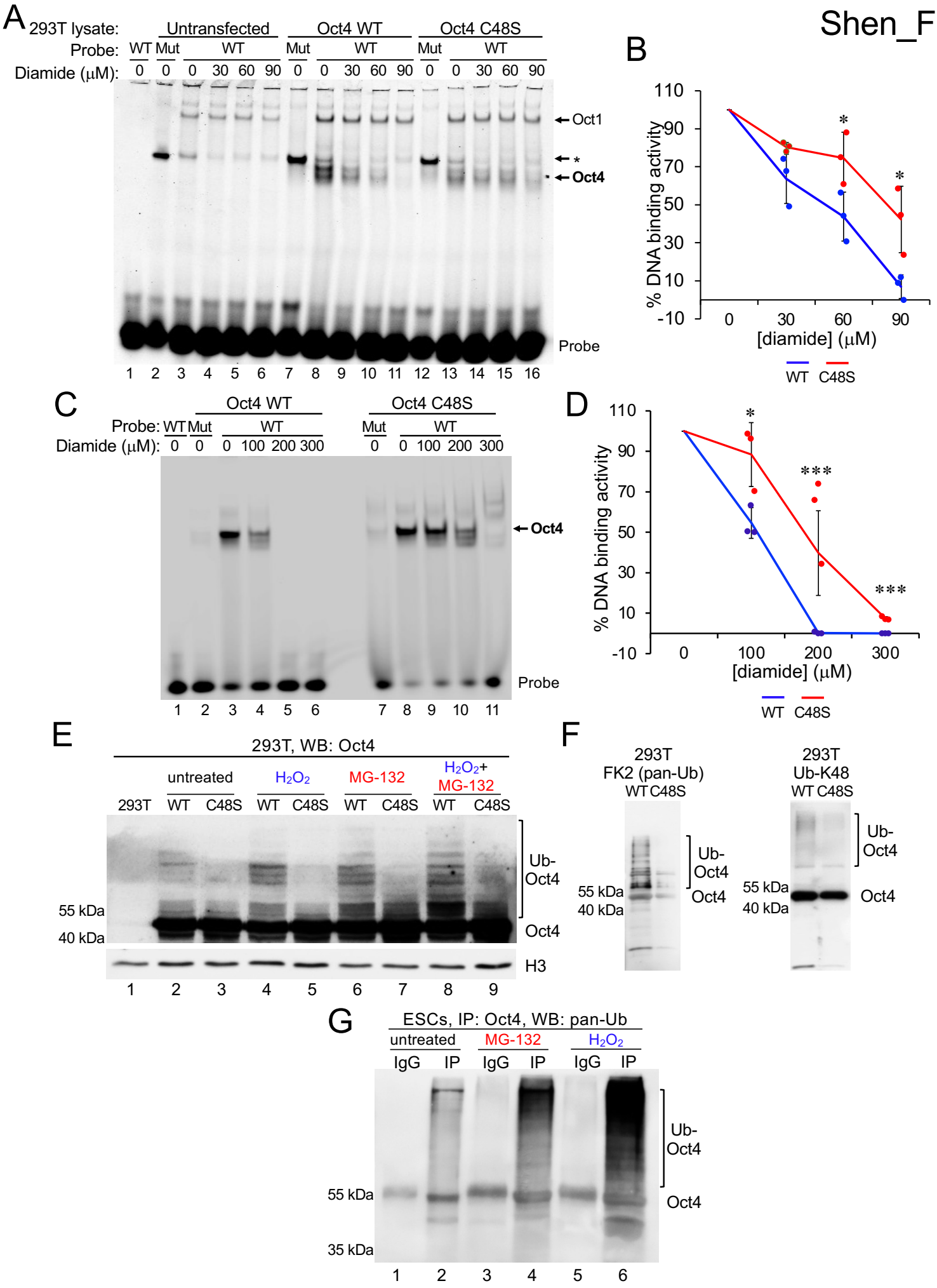


E

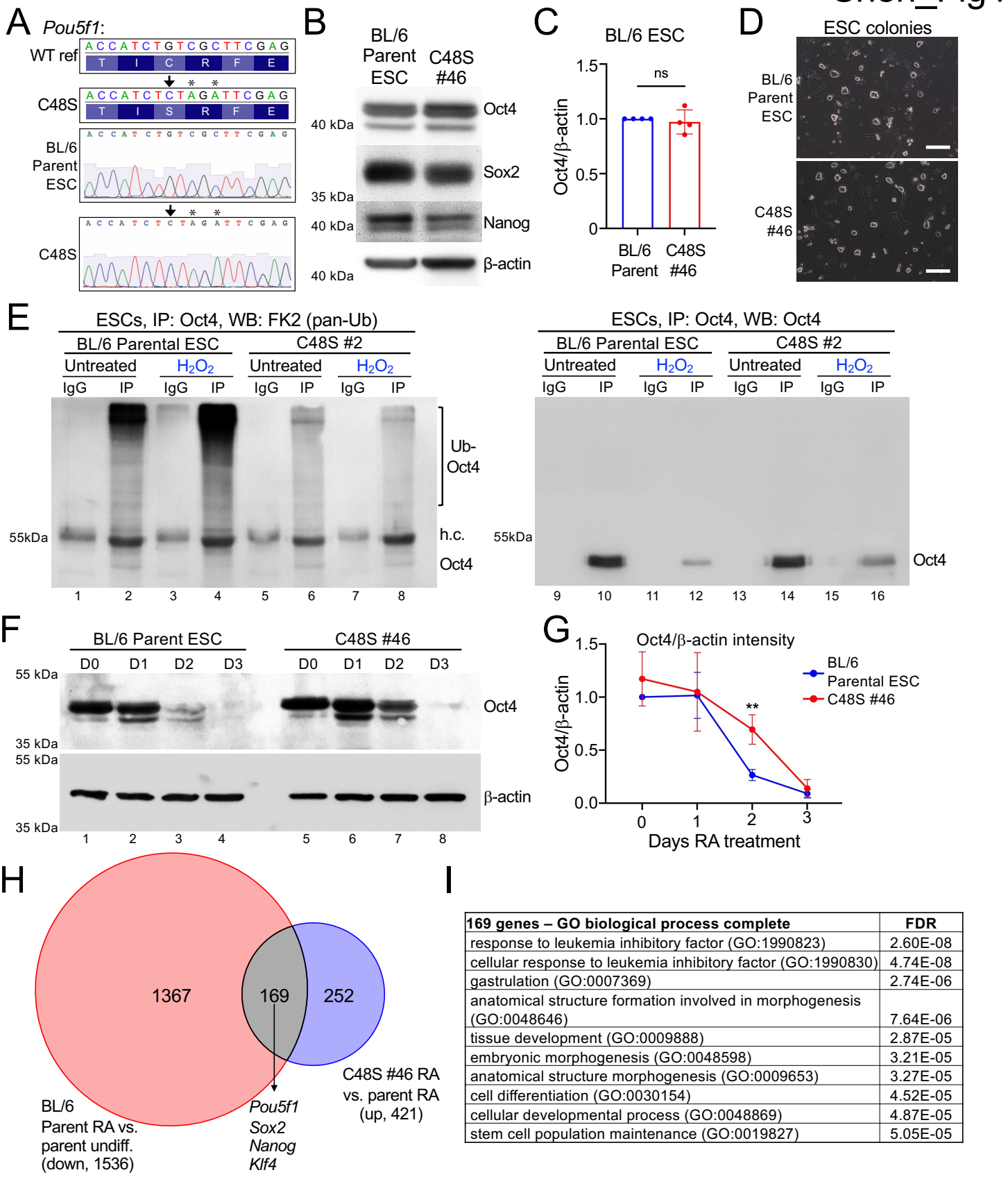
```
42        51
FSQTTISRFE  Mm Oct1/Pou2f1 (NP_035267)
FSQTTICRFE  Mm Oct4/Pou5f1 (NP_038661)
FSQTTISRFE  Mm Oct2/Pou2f2 (Q00196-1)
FSQTTICRFE  Mm Oct6/Pou3f1 (P21952)
FSQTTISRFE  Mm Oct11/Pou2f3 (XP_011240722)
FSQTTICRFE  Mm Brn2/Pou3f2 (XP_021017060)
FSQTTICRFE  Mm Pit1/Pou1f1 (Q00286)

FSQTTICRFE  Hs OCT4/POU5F1 (NP_002692.2 )
FSQTTICRFE  Rr Oct4/Pou5f1 (XP_032744950)
FSQTTICRFE  Dr Pou2 (BAA05901.1)
```

Shen_Fig3

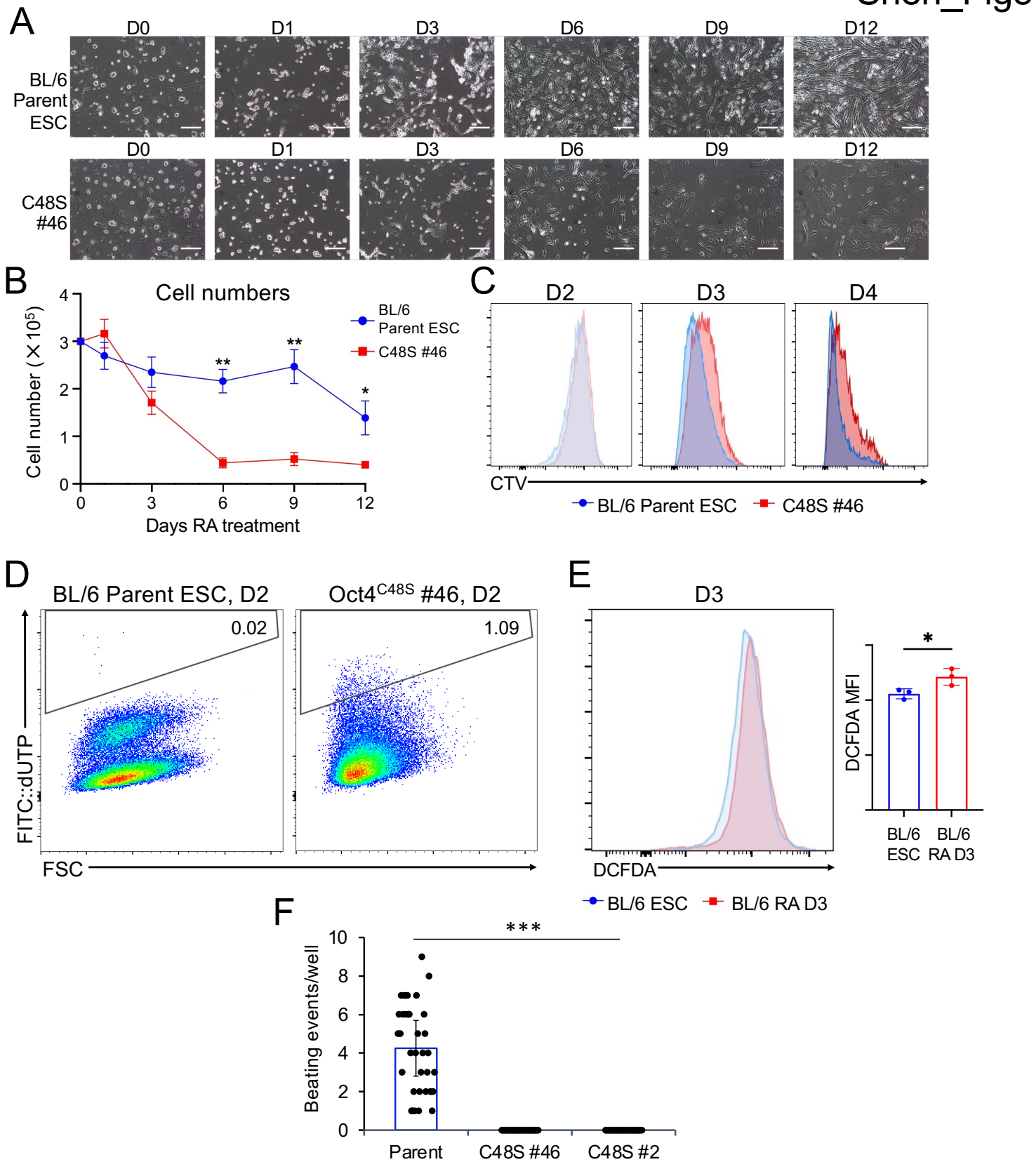

A
293T lysate: Untransfected Oct4 WT Oct4 C48S
Probe: WT Mut WT Mut WT Mut WT
Diamide (μM): 0 0 0 30 60 90 0 0 30 60 90 0 0 30 60 90
Oct1
*
Oct4
Probe
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
B
% DNA binding activity
[diamide] (μM)
WT C48S
C
Oct4 WT Oct4 C48S
Probe: WT Mut WT Mut WT
Diamide (μM): 0 0 0 100 200 300 0 0 100 200 300
Oct4
Probe
1 2 3 4 5 6 7 8 9 10 11
D
% DNA binding activity
[diamide] (μM)
WT C48S
E
293T, WB: Oct4
untreated H2O2 MG-132 H2O2+ MG-132
293T WT C48S WT C48S WT C48S WT C48S
Ub-Oct4
55 kDa
40 kDa
Oct4
H3
1 2 3 4 5 6 7 8 9
F
293T FK2 (pan-Ub) WT C48S
293T Ub-K48 WT C48S
Ub-Oct4
55 kDa
40 kDa
Oct4
G
ESCs, IP: Oct4, WB: pan-Ub
untreated MG-132 H2O2
IgG IP IgG IP IgG IP
Ub-Oct4
55 kDa
35 kDa
Oct4
1 2 3 4 5 6

Shen_Fig4

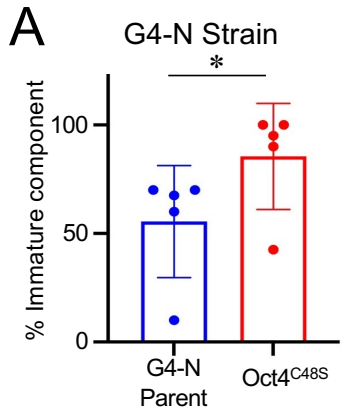


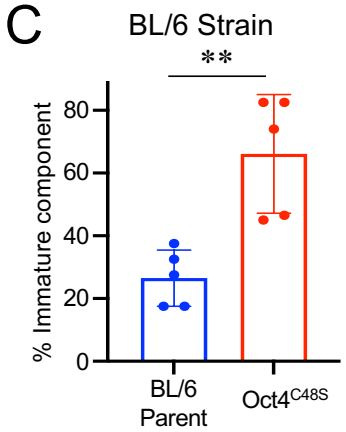
169 genes – GO biological process complete	FDR
response to leukemia inhibitory factor (GO:1990823)	2.60E-08
cellular response to leukemia inhibitory factor (GO:1990830)	4.74E-08
gastrulation (GO:0007369)	2.74E-06
anatomical structure formation involved in morphogenesis (GO:0048646)	7.64E-06
tissue development (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

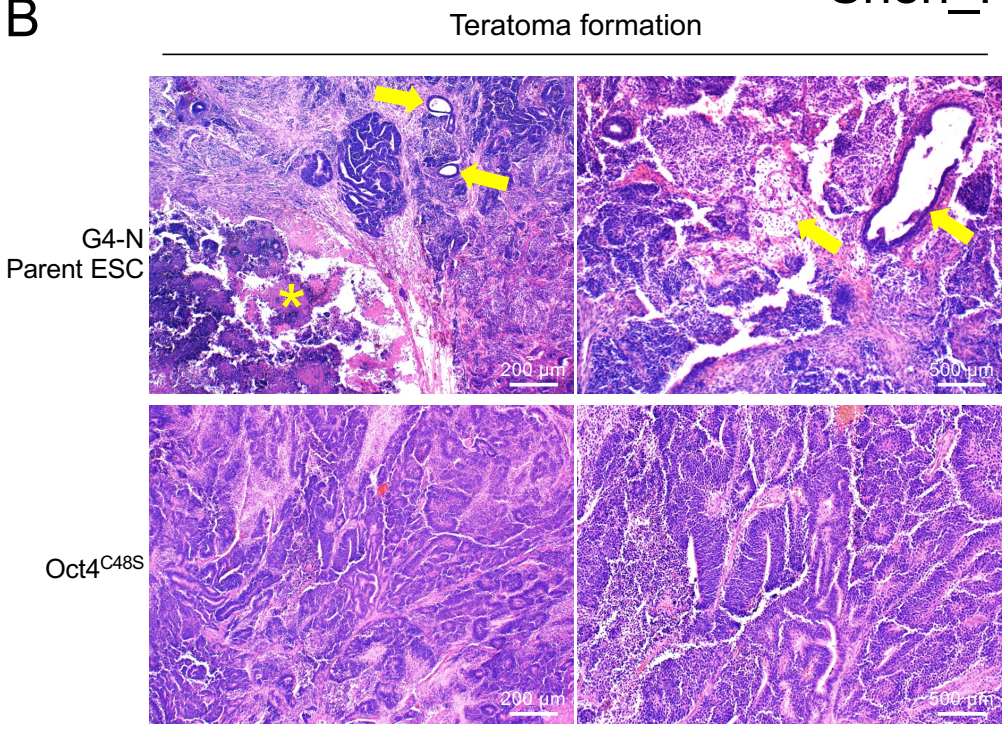
Shen_Fig5

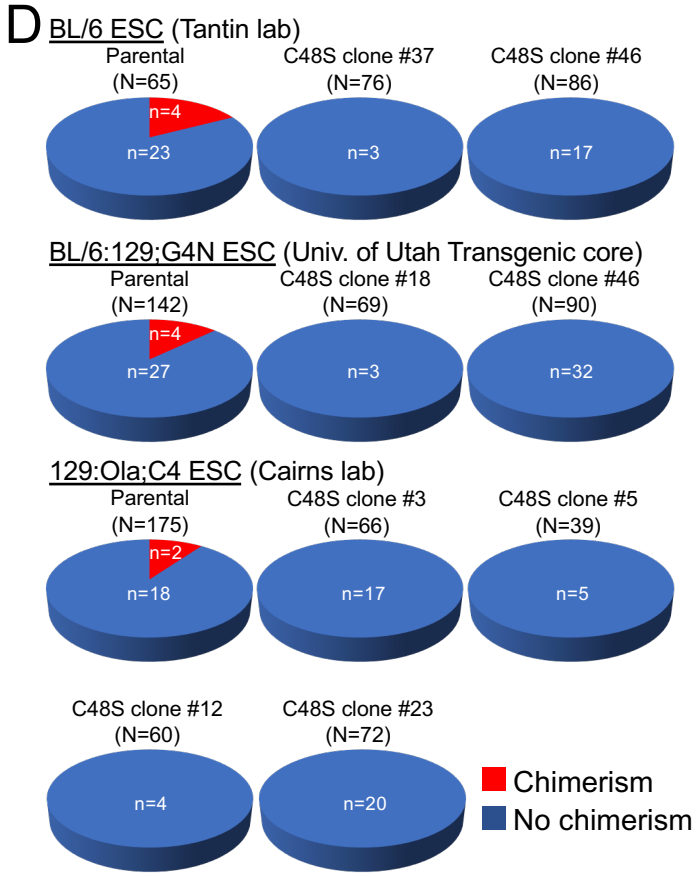

A
D0
D1
D3
D6
D9
D12
BL/6 Parent ESC
C48S #46
B
Cell numbers
Cell number (X 10^5)
Days RA treatment
BL/6 Parent ESC
C48S #46
C
D2
D3
D4
CTV
BL/6 Parent ESC
C48S #46
D
BL/6 Parent ESC, D2
Oct4^C48S #46, D2
0.02
1.09
FITC::dUTP
FSC
E
D3
DCFDA
DCFDA MFI
BL/6 ESC
BL/6 RA D3
F
Beating events/well
Parent
C48S #46
C48S #2

Shen_Fig6









E

(♀) +/+ x (♂) +/*Jky* at weaning						
Litters (N)	+/+	+/*Jky*	*Jky*/*Jky*	Total	*P*-value*	Mean litter size
6	21	22	N/A	43	0.879	7.2

(♀) +/*Jky* x (♂) +/+ at weaning						
Litters (N)	+/+	+/*Jky*	*Jky*/*Jky*	Total	*P*-value*	Mean litter size
3	11	7	N/A	18	0.346	6.0

(♀) +/*Jky* x (♂) +/*Jky* at weaning						
Litters (N)	+/+	+/*Jky*	*Jky*/*Jky*	Total	*P*-value*	Mean litter size
6	12	19	1	32	0.013	5.3

(♀) +/*Jky* x (♂) +/*Jky* E4.5 blastocysts						
Dams (N)	+/+	+/*Jky*	*Jky*/*Jky*	Total	*P*-value*	Mean litter size
4	19	73	27	119	0.027	29.8

*Two-tailed Chi-square test