Oct4 redox sensitivity potentiates iPSC reprogramming

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

[1]Department of Pathology, University of Utah School of Medicine, Salt Lake City, UT 84112, USA

[2]Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, UT 84112, USA

[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

Running head: Oct4 oxidative regulation in reprogramming

Words in abstract: 184

Abstract

The transcription factor Oct4/Pou5f1 is a key component of the regulatory circuitry governing pluripotency. Oct4 is also widely used to generate induced pluripotent stem cells (iPSCs) from somatic cells. These observations provide compelling rationale to understand Oct4's functions. Here we used domain swapping and mutagenesis to compare Oct4's reprogramming activity with the paralog Oct1/Pou2f1, identifying a DNA binding domain cysteine residue (Cys48) as a key determinant of both reprogramming and differentiation. In combination with the Oct4 N-terminus, Oct1^{S48C} is sufficient to confer strong reprogramming activity. Conversely, Oct4^{C48S} strongly reduces reprogramming potential. We find that Oct4^{C48S} confers DNA binding sensitivity to oxidative stress. Further, C48S sensitizes the protein to oxidative stress-mediated ubiquitylation and degradation. Engineering a *Pou5f1*^{C48S} point mutation in mouse embryonic stem cells (ESCs) has little effect on undifferentiated cells, but upon retinoic acid (RA)-mediated differentiation causes retention of Oct4 expression, decreased proliferation and increased apoptosis. *Pou5f1*^{C48S} ESCs also contribute poorly to adult somatic tissues. Collectively, the data support a model in which Oct4 redox sensing serves as a positive reprogramming determinant during one or more steps promoted by Oct4 downregulation during iPSC generation.

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

38 **Introduction**

39 The early mammalian embryo contains undifferentiated, pluripotent cells capable of generating all
40 embryonic tissue lineages. Cultured embryonic stem cells (ESCs) are derived from these cells and have
41 similar capabilities (Abranches et al., 2009). The transcription factor Oct4/Pou5f1 governs the
42 establishment of pluripotent, undifferentiated inner cell mass and epiblast cells, as well as primordial germ
43 cells, germline progenitor cells and ESCs (Morey et al., 2015). Both ESC differentiation in culture and
44 blastocyst implantation in animals are followed by Oct4 loss (Nichols et al., 1998). Together with
45 transcription factors such as Nanog and Sox2, Oct4 forms a network of pluripotency master transcriptional
46 regulators that bind their own and each other's genes to reinforce expression (Boyer et al., 2005).

47 Oct4 is also widely used to generate induced pluripotent stem cells (iPSCs) from somatic cells (Li
48 et al., 2011; Takahashi and Yamanaka, 2006). Reprogramming methods that omit Oct4 nevertheless rely
49 on activation of the endogenous *Pou5f1* (*Oct4*) gene (Gao et al., 2013; Heng et al., 2010). These findings
50 pinpoint Oct4 as a central reprogramming component and place priority on understanding its functions.
51 Reprogramming strategies that increase efficiency (Costa et al., 2013; Esteban et al., 2010; Rais et al.,
52 2013; Soufi et al., 2012; Vierbuchen and Wernig, 2012), or improve the quality of iPSC clones (Chen et
53 al., 2015; Velychko et al., 2019; Yuan et al., 2011) reveal difficult trade-offs: efficiently generating iPSC
54 ethat are of sub-optimal quality, or inefficiently generating iPSCs with improved developmental potential
55 (Buganim et al., 2014; Chen *et al.*, 2015; Velychko *et al.*, 2019).

56 Pluripotent cells co-express Oct4 together with the widely expressed prototypic POU transcription
57 factor Oct1/Pou2f1 (Okamoto et al., 1990; Perovanovic et al., 2023; Rosner et al., 1990; Shen et al., 2017;
58 Suzuki et al., 1990). The two proteins have similar DNA binding domains and recognize the same 8-bp
59 5'ATGCAAAT3' consensus octamer sequence as well as sequence variants (Tantin, 2013). The two
60 proteins also associate with many of the same target genes (Ferraris et al., 2011; Perovanovic *et al.*, 2023).
61 Nevertheless, it is established that Oct1 as well as other mammalian Oct4 paralogs do not function in
62 reprogramming (Feng et al., 2009; Takahashi and Yamanaka, 2006). In contrast, Sox2 and Klf4 can be

63 replaced by their close paralogs while c-Myc can be omitted (Nakagawa et al., 2008; Shi et al., 2008).
64 Structure/function studies comparing Oct4 with its paralogs have identified Oct4 transcriptional
65 determinants (Esch et al., 2013; Jerabek et al., 2017; Jin et al., 2016). An example is three residues present
66 in the Oct4 DNA binding domain but absent in Oct6/Pou3f1 needed to form Sox2 heterodimers.
67 Introducing these residues, plus the Oct4 N- and C-termini, results in efficient reprogramming (Jerabek *et*
68 *al.*, 2017). However, several Oct4 paralogs including Oct1 contain these residues but cannot reprogram,
69 indicating the presence of additional reprogramming determinants present in Oct4 (and likely Oct6) but
70 absent in Oct1.

71 Recent studies show that removal of Oct4 from reprogramming cocktails is highly inefficient,
72 requires higher-than-typical levels of the other factors, and only works with target cells of high mitotic
73 potential, but results in qualitatively superior iPSC clones (An et al., 2019; Velychko *et al.*, 2019). The
74 counterproductive activities of Oct4 are likely caused by transient promotion of counterproductive off-
75 target gene expression (Velychko *et al.*, 2019). The rare iPSCs derived in the absence of Oct4 have
76 enhanced developmental potential. For example, iPSCs reprogrammed without Oct4 are better able to
77 generate viable mice in a single generation using tetraploid complementation. A logical extension of these
78 findings is that transient down-regulation of counterproductive Oct4 activities can aid reprogramming.

79 Here, we identify a new, central Oct4 reprogramming determinant – a single DNA binding domain
80 cysteine residue (Cys48), that accounts for much of its differential reprogramming potential vis-à-vis
81 Oct1. A serine residue is present in Oct1 at this position In the presence of the Oct4 N-terminus, changing
82 the Oct1 serine to cysteine at this position is sufficient to confer levels of reprogramming activity
83 comparable to wild-type Oct4. Conversely, mutating this cysteine to serine in the context of Oct4 reduces
84 reprogramming efficiency by over 50%. Mechanistically, this cysteine confers DNA binding
85 hypersensitivity to oxidative stress, and induces Oct4 protein ubiquitylation via a redox regulation
86 mechanism. We then generated CRISPR-edited Oct4^{C48S} embryonic stem cell (ESC) lines, which are
87 phenotypically normal until differentiated. Upon differentiation, multiple edited ESCs abnormally retain

88 Oct4 expression, proliferate less and undergo programmed cell death at higher rates compared to their
89 unedited parental controls. When injected into preimplantation blastocysts, these cells contribute poorly
90 to the development of viable adult mice. Cumulatively, our findings identify a redox-sensitive Oct4
91 residue that restrains protein function, but nevertheless functions positively in reprogramming.

92 **Results**

93 **Oct4 vs. Oct1 differential reprogramming potential localizes to the DNA binding domain**

94 The POU-specific and POU-homeodomain (POU_S and POU_H) DNA binding subdomains of Oct1 and
95 Oct4 are similar, with 71% and 59% identity in mouse, respectively (Fig.1A). In contrast, the N- and C-
96 termini, and the linker between the subdomains, are highly divergent (Fig.1B). The human proteins are
97 structured similarly. Because the DNA binding domains are similar, we can standardize the amino acid
98 numbering by referring to the first amino acid of the POU_S DNA binding domain as residue 1 (Fig.1B).
99 To identify Oct4 determinants important for reprogramming, we generated Oct4/Oct1 chimeric constructs
100 containing the distinct domains. Because introducing different domains and epitopes can confound the
101 analyses with protein-specific antibodies, we included C-terminal FLAG-tags on Oct4, Oct1 and all
102 chimeric proteins to track protein expression. Unlike an N-terminal FLAG tag (Supplemental Fig.1),
103 engineering a C-terminal tag results in reprogramming efficiency similar to untagged Oct4 (Fig.1C). All
104 subsequent constructs therefore bore C-terminal tags. In addition, prior work comparing Oct4 and another
105 reprogramming-incompetent POU transcription factor, Oct6, identified a lysine residue within the POU_S
106 DNA binding subdomain (K40, of the Oct4 DNA binding domain, N320 in full-length Oct1) as being
107 important for reprogramming (Jin *et al.*, 2016). We therefore mutated N40 of the Oct1 DNA binding
108 domain to lysine in some constructs (Fig.1D-F).

109 To measure reprogramming efficiency, we transduce primary mouse embryonic fibroblasts
110 (MEFs) expressing GFP under the control of the endogenous *Pou5f1* gene (Lengner et al., 2007; Shakya
111 et al., 2015; Shen et al., 2018) with mSTEMCCA lentiviruses, which express all four Yamanaka factors
112 from an EF-1A promoter (Sommer et al., 2009) (Fig.1B). Substituting Oct1 for Oct4 eliminated all
113 reprogramming potential (Fig.1D,E). The other reprogramming factors – Sox2, Klf4 and c-Myc – can
114 mediate reprogramming without Oct4 (An *et al.*, 2019; Velychko *et al.*, 2019), however reprogramming
115 is inefficient. The results are therefore consistent with the original reprogramming findings showing that

116 Oct4 is essential for conventional reprogramming (Takahashi and Yamanaka, 2006), and set a useful
117 baseline for making domain-swaps and mutations to assess reprogramming potency.

118 We then replaced the Oct1 N- and C-termini with that of Oct4, and mutated asparagine 40 of the
119 Oct1 DNA binding domain to the lysine present in Oct4. Prior work (Jin *et al.*, 2016) showed that the
120 Oct4 lysine at this position (Fig.1A) is important for Oct4 reprogramming. Of these constructs, only fusion
121 of the Oct4 N-terminus generated detectable colony-forming potential, at ~5% the level of WT Oct4
122 (Fig.1D,E). These findings are consistent with work establishing the importance of the Oct4 N-terminus
123 (Boija et al., 2018). Similar data were obtained at multiple timepoints (not shown), indicating that this
124 result does not simply reflect kinetic differences in reprogramming. Anti-FLAG immunoblotting showed
125 equivalent expression (Fig.1F). Viral titers and expression of Klf4 from the mSTEMCCA vector were
126 also similar (not shown). Because a construct with both the Oct4 N- and C-termini displayed only ~5%
127 efficiency (Fig.1D,E), we conclude that determinants within the Oct4 DNA-binding domain confer the
128 bulk of reprogramming potential.

129

130 **Oct4 vs. Oct1 differential reprogramming potential concentrates in a DNA binding domain cysteine**

131 The POU DNA binding domain is comprised of two subdomains, connected by a peptide linker, that make
132 separate sequence-specific contacts with DNA (Tantin, 2013). These domains are termed POU_S and POU_H
133 (Fig.1A,B). We systematically mutagenized DNA binding domain amino acid residues that differ between
134 Oct1 and Oct4. We focused on possible Oct4 post-translationally modifiable residues: Lys, Ser, Thr and
135 Cys. This analysis identified 14 possible amino acid substitutions, not including the aforementioned
136 position lysine 40 (Fig.1A and not shown). Work from others has shown that Oct6 is capable of
137 reprogramming when modified to enable Sox2 dimerization (Kim et al., 2020). Filtering the data using
138 conservation between Oct4 and Oct6 resulted in just six residues. We replaced the Oct1 residues at these
139 positions with the corresponding Oct4 residue in the context of the "4N" construct, which contains the
140 Oct4 N-terminus. We reasoned that because this construct already has low but measurable activity, it

141 would be primed to show differences with single amino acid substitutions. Using this method, we found
142 that a single serine to cysteine substitution at position 48 in POU_S (Fig.1A, red box) conferred the majority
143 of the differential Oct4 activity, increasing activity >7-fold over the construct containing the Oct4 N-
144 terminus alone (Fig.2A,B). We term this the "1-4All" construct. iPSC colonies formed under these
145 conditions were robust morphologically and expressed GFP from the *Pou5f1* locus at levels comparable
146 to cells reprogrammed with Oct4 (Supplemental Fig.2). The remaining activity relative to Oct1 appeared
147 to be concentrated in the linker domain, as a construct containing the Oct4 N-terminus, Cys48 and the
148 Oct4 linker (4N-S48C-4L) had 100% activity. The Oct4 linker has been shown to be involved in
149 reprogramming via selective interactions with different cofactors (Esch *et al.*, 2013; Han et al., 2022). The
150 differences were maintained over time, with the 1-4All construct consistently generating ~60% activity
151 over multiple days, indicating that the differences were qualitative and not simply a function of different
152 colony formation rates (Supplemental Fig.3).

153 Replacing the Oct4 Cys with Ser at this position reduced activity by 60% (Fig.2A,B), indicating
154 that more than half of Oct4's differential reprogramming activity vis-à-vis Oct1 is lost by changing this
155 residue. The converse experiment, replacing Oct1 S48 with cysteine, resulted in no reprogramming
156 activity (Fig.2A,B, Oct1-S48C), indicating that Cys48 is important but not sufficient, and highlighting its
157 combined importance with the Oct4 N-terminus. The chimeric proteins were equivalently expressed
158 (Fig.2C). These results indicate that in the context of the Oct4 N-terminus, a cysteine residue at position
159 48 confers robust reprogramming activity. In the co-crystal structure of the Oct1 DNA binding domain
160 bound to DNA (Klemm et al., 1994), Ser48 hydrogen bonds with the DNA backbone (Fig.2D). In a recent
161 Oct4:Sox2:nucleosome structure (Michael et al., 2020), the Cys48 sulfhydryl also allows for hydrogen
162 bonding. Predictably, mutation of Cys48 to glycine greatly diminishes Oct4 DNA binding in vitro
163 (Marsboom et al., 2016). Oct6, which can reprogram when changes are made to POU_H allowing for Sox2
164 dimerization (Jerabek *et al.*, 2017), contains a cysteine at this position (Fig.2E). These structural studies

165 along with our findings suggest that Cys48 has minimal effects on DNA binding vis-à-vis a serine (but
166 not glycine) at this position, but nevertheless plays a role in reprogramming.
167
168 **Oct4 Cys48 is redox-sensitive and mediates degradation in response to oxidative stress**
169 Oct4 is known to be sensitive to oxidative stress (Guo et al., 2004; Lickteig et al., 1996; Marsboom *et al.*,
170 2016) in a manner reversed by thioredoxin (Guo *et al.*, 2004), implicating cysteine thiols in redox
171 sensitivity. We reproduced this result (Supplemental Fig.4), which with the results above suggested that
172 Cys48 may potentiate reprogramming through a mechanism involving redox sensitivity. We measured
173 wild-type and cysteine 48 mutant Oct4 DNA binding using electrophoretic mobility shift assays (EMSA)
174 with Cy5-labeled canonical octamer element containing DNA (Ferraris *et al.*, 2011). To measure redox
175 sensitivity, we used methods similar to Guo et al. (Guo *et al.*, 2004) in which Oct4 is pre-treated with the
176 oxidizing agent diamide prior to addition of the DNA probe.
177 We studied Oct4 Cys48 oxidation using a system with lysates from 293T cells transfected with
178 WT or C48S mutant Oct4. 293T cells do not express Oct4, but do express endogenous Oct1 as an internal
179 control (Fig.3A, lanes 2-3). Oct4 bound to octamer DNA specifically (Fig.3A, lanes 7-8) and was
180 hypersensitive to diamide treatment (lanes 9-11). Oct4^{C48S} bound DNA equivalently (lane 13) and also
181 conferred resistance to oxidation, in particular at higher concentrations (lanes 14-16). In contrast, little
182 sensitivity was observed with Oct1. Quantification from three independent replicate experiments is shown
183 in Fig.3B. The recombinant Oct4 contained a C-terminal FLAG tag, allowing for purification
184 (Supplemental Fig.5). Similar relative resistance to oxidative inhibition of DNA binding was observed
185 using purified C48S recombinant Oct4 protein (Fig.3C). Quantification from three independent replicates
186 is shown in Fig.3D. We conclude that Oct4 cysteine 48 sensitizes Oct4 to oxidative inhibition of DNA
187 binding in vitro.
188 Oct4 oxidation has been linked to its degradation (Marsboom *et al.*, 2016). In studying
189 recombinant Oct4 expressed in 293T cells, we noted in over-exposed immunoblots the presence of

190 multiple slow-migrating high molecular weight forms of wild-type Oct4. These high-molecular weight
191 forms were largely missing with C48S mutant protein (Fig.3E, lanes 2-3). These were increased by
192 hydrogen peroxide treatment specifically using wild-type protein (lanes 4-5), indicating that they are
193 increased by oxidative stress in a manner that requires cysteine 48. They were also specifically increased
194 using the proteasome inhibitor MG-132, strongly suggesting that the high-molecular weight forms are
195 ubiquitylated species (lanes 6-7). Interestingly, the combination of both H_2O_2 and MG-132 increased the
196 intensity of the banding pattern over that observed using either treatment individually (lanes 8-9). These
197 treatments alone and in combination did not affect short-term viability of the cells (Supplemental Fig.6).
198 To identify the nature of the ubiquitylated Oct4 species, we purified Oct4 expressed from 293T
199 cells using a C-terminal streptavidin tag, and subjected the isolated proteins to immunoblotting using
200 antibodies specific to different ubiquitin linked chains. Analysis using antibodies to pan linked Ub chains
201 revealed strong diminution with $Oct4^{C48S}$ as expected (Fig.3F, FK2). Strong diminution was observed with
202 K48 linkage-specific polyubiquitin antibodies (Fig.3F), while K63 linkage-specific antibodies however
203 showed more modest differences compared to total Oct4 high-molecular weight bands (Supplemental
204 Fig.7). We conclude that cysteine 48 promotes the degradation of Oct4 in response to oxidative stress
205 largely via catalysis of K48-linked polyubiquitin chains.
206
207 **ESCs with endogenous $Oct4^{C48S}$ mutation appear normal in a pluripotent state but differentiate**
208 **abnormally**
209 The ability of cysteine at position 48 to promote pluripotency, and the negative impact of Oct4 C48S
210 mutation on iPSC generation, suggested that sensitivity to oxidative stress mediated by Cys48 helps
211 potentiate reprogramming. To test if Oct4 oxidative stress sensitivity also regulates forward
212 developmental steps, we studied the role of Cys48 in ESC differentiation. Using CRISPR, we generated
213 ESCs with a single-base missense mutation that converts Cys48 to Ser (see Methods). Two silent point
214 mutations were also engineered in the gRNA to monitor targeting. Two low-passage parent ESC lines

215 were chosen for targeting, one on a pure C57BL/6 background and one on a mixed BL6/129 background.
216 Multiple independent clones were generated for each line. Homozygous mutation was confirmed by
217 HRMA (not shown) and resequencing (Fig.4A). Tested clones were confirmed to be karyotypically
218 normal and wild-type for *Tp53* (not shown). Undifferentiated ESCs were phenotypically identical to the
219 parent clone in terms of Oct4 expression and morphology (Fig.4B,C), as well as maintenance of
220 pluripotency and proliferative capacity (not shown). Cells were then differentiated in vitro using retinoic
221 acid (RA) to assess Oct4 expression, morphology and cell counts. Oct4 protein expression was abnormally
222 retained early during differentiation, mostly at day 2 (Fig.4D, lanes 3 and 7). Quantification from multiple
223 experiments revealed that the effect was transient, with a >2.8-fold increase of Oct4 at day 3 that was lost
224 by day 4 (Fig.4E). These results were recapitulated using a second clone from the same parental ESC line
225 (Supplemental Fig.8A,B), as well as a C48S clone from a different 129×C57BL/6 ESC line, G4-N, which
226 differentiates more slowly and shows even greater differences (Supplemental Fig.8C). RA-mediated
227 differentiation also revealed a significant depletion of cells later in the differentiation timecourse that was
228 retained through at least 12 days (Fig.4F,G).

229 The decreased cell numbers observed during differentiation of C48S-edited cells with RA could
230 be due to decreased proliferation, increased death or some combination of the two. In order to investigate
231 the nature of the effect on differentiation more closely, we studied proliferation using cell trace violet and
232 TUNEL staining to look at genomic DNA fragmentation during apoptosis. Adherent, viable differentiating
233 C48S cells showed increased apoptosis at day 2 (Supplemental Fig.9) and decreased proliferation at days
234 3 and 4 (Fig.4H) These results indicate that C48S-edited ESCs are phenotypically normal in the pluripotent
235 state, but differentiate abnormally in assays that necessitate Oct4 silencing.

236

237 **Oct4^{C48S} mutant ESCs contribute poorly to adult chimeric animals**

238 One essential hallmark of pluripotency is the ability to contribute to cells and tissues of the embryo (De
239 Los Angeles et al., 2015). We conducted chimerism assays to assess the ability of CRISPR-edited ESCs

240 to contribute to development. We used nine lines of mutant ESCs, derived from three independent wild-
241 type parent lines with different strain backgrounds to minimize background and clonal effects. Derived
242 clones had karyotypes similar to parent clones, with the proportion of diploid cells ranging between 71
243 and 93%. Compared to the parent cells, which contributed to offspring at a rate of 10-20% (Fig.5A,
244 chimerism ranging from 2-99%), none of the derived Oct4^{C48S} mutant ESCs were able to contribute to
245 adult mice (Fig.5). The results strongly suggest that Oct4C48S ESCs are defective in contributing to
246 embryonic development.

247 **Discussion**

248 Oct4 is a central mediator of pluripotency and is widely used to generate iPSCs from somatic cells (Feng
249 *et al.*, 2009; Nakagawa *et al.*, 2008; Okita et al., 2007; Takahashi et al., 2007). Together with other factors
250 it sustains pluripotency by activating "core" targets such as *Pou5f1* (which encodes Oct4), *Sox2* and *Nanog*
251 (Boyer *et al.*, 2005). Oct4 also maintains silent targets, including critical developmental regulators, in a
252 readily inducible "poised" state (Bernstein et al., 2006; Meissner et al., 2008). Yet the mechanisms by
253 which Oct4 is itself regulated, particularly at a post-transcriptional level, are poorly defined. Here, we
254 used Oct4/Oct1 domain-swapping to pinpoint Oct4 redox sensitivity as an important determinant in
255 reprogramming and differentiation. Oct1 itself appears to have no role in pluripotency establishment or
256 maintenance, but upon differentiation and Oct4 loss becomes important for guiding developmental-
257 mediated induction of targets shared with Oct4 (Perovanovic *et al.*, 2023).

258 Pluripotent cells co-express Oct4 together with the widely expressed prototypic Oct/POU factor
259 Oct1/Pou2f1 (Shen *et al.*, 2017). Despite having similar DNA binding domains and recognizing the same
260 8-bp consensus octamer sequence (5'ATGCAAAT3') (Kang et al., 2009b; Tantin, 2013), Oct1 and other
261 Oct4 paralogs do not function in reprogramming (Takahashi and Yamanaka, 2006). Structure/function
262 studies comparing Oct4 with these paralogs, in particular Oct6, have identified Oct4 transcriptional
263 determinants (Boija *et al.*, 2018; Esch *et al.*, 2013; Jerabek *et al.*, 2017; Jin *et al.*, 2016). An example is
264 three residues (K7,T22,151S) present in the Oct4 DNA binding domain (Fig.1A) but absent in Oct6
265 needed to form Sox2 heterodimers. Introducing these residues plus the Oct4 N- and C-termini into Oct6
266 results in efficient reprogramming (Jerabek *et al.*, 2017). However, several Oct4 paralogs including Oct1
267 contain these residues and readily heterodimerize with Sox2 but cannot reprogram, indicating that
268 additional reprogramming determinants may be present in Oct4 (and Oct6) but absent in Oct1. We identify
269 an additional central Oct4 reprogramming determinant – a single DNA binding domain cysteine residue
270 that confers redox sensitivity of Oct4.

271 We found that Oct4 C48S mutation maintains DNA binding and confers resistance of oxidation-
272 induced DNA binding inhibition in vitro, while reducing reprogramming efficiency by more than half.
273 Cys48 is located within helix 3 (the DNA recognition helix) of the POU-specific domain (Fig.1A). This
274 residue makes hydrogen bond contacts with the DNA phosphate backbone, predicting that any oxidation
275 event will block DNA binding. In the context of the Oct4 N-terminus, a single S48C mutation in the Oct1
276 DNA binding domain is sufficient to confer reprogramming activity. This result suggests that any unique
277 Oct4 cofactors required for reprogramming vis-à-vis Oct1 are likely to be recruited through the N-
278 terminus. The ability of the Oct4 N-terminus to mediate liquid-liquid phase transitions (Boija *et al.*, 2018)
279 may also play a role. Mutation of the endogenous Oct4 Cys48 to serine does not appear to affect the
280 undifferentiated state, but results in numerous phenotypes upon RA-mediated differentiation, including
281 abnormal retention of Oct4 expression, decreased proliferation and increased apoptosis. Prominent roles
282 for Oct4 in both cell survival and proliferation during post-implantation embryonic development have
283 been documented (DeVeale et al., 2013).

284 The redox sensitivity of Cys thiols is known to be enhanced by immediately adjacent basic amino
285 acid residues (Britto et al., 2002), and Cys48 is followed by a conserved Arg (Fig.1A). In Oct1, this residue
286 is a serine, strengthening hydrogen bonding due to the increased electronegativity of the oxygen in the
287 hydroxyl relative to the sulfur in the sulfhydryl. Redox-regulated transcription factors include Nrf2, Yap1,
288 p53 and HIF-1 (Brigelius-Flohe and Flohe, 2011; Delaunay et al., 2000; Marinho et al., 2014). A subset
289 of these are regulated by Cys oxidation, including disulfide bond formation (Brigelius-Flohe and Flohe,
290 2011; Delaunay *et al.*, 2000; Zheng et al., 1998). Oct4 is similarly known to be sensitive to oxidative stress
291 (Guo *et al.*, 2004; Lickteig *et al.*, 1996; Marsboom *et al.*, 2016) in a manner reversed by thioredoxin (Guo
292 *et al.*, 2004) (Supplemental Fig.4). Relative to other Oct proteins, mouse Oct4 has unusually high cysteine
293 content in the DNA binding domain, with two cysteines in each subdomain (Fig.1A), including cysteine
294 142 (Cys50 of the POU-homeodomain), which also lies in the recognition helix. Cys50 is the
295 distinguishing feature of POU-homeodomains relative to other members of the homeodomain family

296 (Kang *et al.*, 2009b; Tantin, 2013). All other homeodomains, from bacteria to flies to mammals, have
297 bulkier residues such as Q or K at this position that make base contacts (Tantin, 2013). Yet in Oct4,
298 mutating Cys142 to Gly has no effect on DNA binding (Marsboom *et al.*, 2016). In Oct1, C61A and
299 C142A mutation does not interfere with DNA binding, but rather improves crystallization properties
300 (Remenyi et al., 2001). These findings suggest that variable Cys oxidation interferes with uniform crystal
301 packing.

302 Prior studies have also identified a link between Oct4 oxidation and degradation (Marsboom *et al.*,
303 2016). In addition to mediating relative sensitivity to oxidative stress, we found that Cys48 increases
304 protein ubiquitylation, in particular following exposure to oxidative stress or MG-132. Oct4 is known to
305 be ubiquitylated via K63-linked chains by the Stub1 E3 Ub ligase to reduce protein half-life (Mamun et
306 al., 2022), however we found that Oct4 marked with K48-specific linkages were more strongly affected
307 with mutation of cysteine 48 to serine, suggesting a possibly related by distinct pathway.

308 Transcription factors with complex (both positive and negative) roles in reprogramming include
309 Oct4 and Smad3 (An *et al.*, 2019; Ruetz et al., 2017; Velychko *et al.*, 2019). Although endogenous Oct4
310 must be activated and expressed as a key component of the core pluripotency network, recent studies show
311 that removal of Oct4 from reprogramming cocktails is highly inefficient but results in qualitatively
312 superior iPSC clones that more closely resemble a pluripotent "ground state" (An *et al.*, 2019; Velychko
313 *et al.*, 2019). As a consequence, the few iPSC colonies derived in the absence of Oct4 have enhanced
314 developmental potential. For example, iPSCs reprogrammed without Oct4 are better able to generate
315 viable mice in a single generation using tetraploid complementation. The counter-productive activity of
316 Oct4 is likely due to transient promotion of off-target, ectoderm-specific gene expression (Velychko *et*
317 *al.*, 2019). The studies indicate that transient down-regulation of counterproductive Oct4 activities can be
318 an important step for efficient reprogramming. Our work identifying a role of Oct4 redox sensitivity in
319 reprogramming may explain this process and may provide a route to generating superior iPSC
320 reprogramming outcomes with high efficiency.

321 **Materials and Methods**

322

323 **Cell culture**

324 ESCs/iPSCs were cultured as described previously (Shakya *et al.*, 2015; Shen *et al.*, 2018) with "2iV"
325 supplementation: the MEK inhibitor PD03259010 (1 μM, LC laboratories) and the glycogen synthase
326 kinase-3β (GSK3β) inhibitor CHIR99021 (3 μM, LC Laboratories). Culture medium was also
327 supplemented with 50 μg/mL ascorbic acid (Sigma). H_2O_2 (Sigma) was supplied at 1 mM for 2 hr and
328 MG-132 (Sigma) was supplied at 10 μM for 2 hr. Cell trace violet proliferation assays were conducted
329 using a kit (Cell Trace Violet Cell Proliferation Kit for Flow Cytometry, ThermoFisher) according to the
330 manufacturer instructions. TUNEL assays were conducted using a kit (Apo-Direct, Becton-Dickinson)
331 according to the manufacturer instructions.

332

333 **FLAG-tagged and chimeric Oct4/Oct1 constructs for iPSC generation**

334 The "mSTEMCCA" 4-in-1 lentiviral vector expressing Oct4, Klf4, Sox2 and c-Myc was used for these
335 studies (Somers et al., 2010). To identify determinants of reprogramming, the Oct4 cDNA cassette was
336 replaced with c-terminally FLAG-tagged full-length mouse Oct4, mouse Oct1 or chimeric constructs.
337 FLAG-tagged Oct4 was generated by overlap PCR using overlapping primers containing the FLAG
338 peptide sequence: DYKDDDDK (Forward: 5'-
339 GATTACAAGGATGACGACGATAAGGGAAGTGGCGTGAAACAGA; Reverse: 5'-
340 CTTATCGTCGTCATCCTTGTAATCGTTTGAATGCATGGGAGAG). The first fragment was
341 amplified using a forward primer containing a *Not*I restriction enzyme site (5'-
342 AATGAAAAAAGCGGCCGCCATGGCTGGACACCT) and overlaping reverse primer containing the
343 FLAG sequence. A second fragment was amplified using the reverse primer containing an *Nde*I restriction
344 site (5'-GGGAATTCCATATGTGTGGCCATATTATCATCGTGT) and overlapping forward primer.
345 Mixed first fragment and second fragment as template and amplified the whole fragment by using forward

346 and reverse primers containing restriction enzymes (*NotI* and *Nde*I). After purification and digestion with
347 *Not*I and *Nde*I , the purified fragment was ligated with the mSTEMCCA vector digested by *Not*I and *Nde*I.
348 A similar strategy was applied to generate other tagged constructs. For point mutation, a site-directed
349 mutagenesis kit (New England Biolabs Q5 Site-Directed Mutagenesis kit) was used following
350 manufacturer instructions. When making fusions between the different linker domains and POU_H, we took
351 care to maintain the "RKRKR" sequence at the N-terminus of Oct4 POU_H, which has been shown to be
352 important for reprogramming (Jerabek *et al.*, 2017). Chimeric constructs between the Oct1 linker and Oct4
353 POU_H contained the amino acid sequence N-AEGLN/ARKRKR as shown in Fig.1A.
354

355 **iPSC generation**

356 293T cells were transfected by polycation polyethylenimine (PEI) with mSTEMCCA encoding FLAG-
357 tagged Oct4, Oct1 and different domain-swaps and mutants as well as the packaging plasmids
358 (pMDLg/pRRE and pRSV-Rev) and envelope plasmid(VSV-G). Lentiviral supernatants were filtered and
359 used to transduce passage 3 primary MEFs in which EGFP is expressed from the endogenous *Pou5f1*
360 locus (Oct4-GFP MEFs, Jackson Labs strain #008214). Cells were cultured in ESC medium (above).
361 Unless stated otherwise, green iPSC colony emergence was quantified after 12 days using an Olympus
362 IX-71 epifluorescence microscope.
363

364 **Immunoblotting**

365 Antibodies for immunoblotting were as follows: Oct4, Santa Cruz sc-5279; Oct1, Bethyl A301-
366 716A+A301-171A; Nanog, GeneTex GTX100863; Sox2, GeneTex GTX101507; GAPDH, EMD
367 Millipore MAB374; α-Tubulin, Santa Cruz sc-5286; β-actin, Santa Cruz sc-47778.
368

369 **Lentiviral Oct1 complementation**

370 The *Oct1* (*Pou2f1*) cDNA and IRES (internal ribosomal entry site) elements were amplified and cloned
371 together by overlap PCR. In the first PCR, primers to the 5' end of *Oct1* containing a *Not*I restriction site
372 and to the 3' end of *Oct1* that contained a 5' extension of IRES-complementary DNA were used. The
373 sequences were: Oct1-NotI-For: 5'-AATGAAAAAAGCGGCCGCCATGAATAATCCATCAGAAAC-
374 3'; Oct1-Rev-IRES: 5'-TTAGGGGGGGGGGAGGGATCTTCACTGTGCCTTGGAG-3'. In the second
375 PCR, an IRES sequence was amplified using primers to the 5' end of the IRES containing a 5' extension
376 of DNA complementary of the *Oct1* 3' end, and primers to the 3' end of the IRES containing an *Nde*I
377 restriction site. The sequences were: IRES-overlap-FOR: 5'
378 AGATCCCTCCCCCCCCCCTAACGTTACTGGCCGAA-3'; IRES-Rev-NdeI: 5'-
379 GGGAATTCCATATGTGTGGCCATATTATCATCGTGT-3'. The third PCR used as a template the
380 PCR products from the first two rounds, along with the Oct1-NotI-For and IRES-Rev-NdeI primers. This
381 process generated a DNA fragment containing an *Oct1* cDNA fused to an IRES at the 3' end, along with
382 a *Not*I site at the 5' terminus and an *Nde*I site at the 3' terminus. The fragment was cloned into the
383 optimized, self-inactivating, nonreplicative pHAGE lentiviral vector using the *Not*I and *Nde*I restriction
384 sites. To insert a *Puro* cassette after the IRES, the cDNA was amplified using primers containing 5' *Nde*I
385 and 3' *Cla*I restriction sites. The sequences were Puro-NdeI-For: 5'-
386 GGAATTCCATATGATGACCGAGTACAAGCCCACGGT-3'; Puro-ClaI-Rev: 5'
387 GGTTTATCGATTCAGGCACCGGGCTTGC-3'. Because the IRES apparently attenuated expression of
388 the *Puro* resistance cassette in this vector, puromycin selection was performed at 0.75 μg/mL. To generate
389 an empty vector control, the vector was cut with *Nde*I and *Not*I, filled in with Klenow fragment, and re-
390 ligated.
391

392 **Electrophoretic mobility shift assay**

393 EMSA was performed using published methods (Kang et al., 2009a; Tantin et al., 2008) Briefly, reactions
394 were assembled on ice in 0.6× buffer D (12 mM HEPES pH 7.9, 60 mM KCl, 0.06 mM EDTA, 12%

395 glycerol, all buffer reagents from Sigma) supplemented with 0.1 mM dithiothreitol (DTT, Sigma), 0.5
396 mM phenylmethylsulphonyl fluoride (PMSF, Sigma), 50 μg/mL poly-dI•dC (Sigma), 1 mg/mL bovine
397 serum albumin (BSA, Sigma) and added DNA probe and lysate or recombinant protein. Diamide (Sigma)
398 was added at final concentrations ranging from 0 to 300 μM on ice during reaction assembly. 5'Cy5-
399 labeled double-stranded DNA probes were as follows: wild-type octamer, 5'-
400 TGTCGAATGCAAATCACT-3'; mutant octamer, 5'-TGTCGAATGCAAGCCACT-3'. Reactions were
401 incubated for 30 min at room temperature prior to gel loading. For experiments involving reversal of
402 oxidative inhibition, thioredoxin (Sigma) was supplied at a final concentration of 0 to 60 μM after 15 min
403 incubation with diamide, and both thioredoxin-treated and control samples were incubated for a further
404 15 min at 37°C prior to gel loading. Reactions were resolved be electrophoresis through 0.5× TBE (45
405 mM Tris-Cl pH 8.3, 45 mM boric acid, 1 mM EDTA, all reagents from Sigma), 4% polyacrylamide gels
406 containing 1% glycerol. Gels were run for 90 min at room temperature. DNA was visualized using a
407 Molecular Dynamics Typhoon system.
408
409 **Recombinant Oct4 protein**
410 Wild-type and C48S mutant Oct4 constructs were generated bearing dual C-terminal FLAG and Twin-
411 Strep tags. A g-block (IDT) fragment coding for the Twin-Strep and FLAG epitope tags was inserted
412 between the *Nco*I and *Kpn*I sites in the transient expression vector pACE-MAM2 (Geneva Biotech). The
413 resultant construct was then digested with *Nco*I and *Pvu*II to insert the g-block fragment coding for either
414 wild-type or C48S mutant murine Oct4. After Sanger sequencing to confirm veracity, the plasmids were
415 transfected into Expi293F cells (ThermoFisher Scientific) using ExpiFectamine 293 Tranfection Kit
416 (ThermoFisher Scientific) following the manufacturer instructions. After 2 days, transfected cells were
417 harvested by centrifugation at 4°C for 5 min at 500×*g*, washed with ice-cold PBS, and centrifuged again
418 at 4°C for 5 min at 500×*g*. Cells were resuspended in Buffer A (20 mM HEPES pH 8.0, 1.5 mM $MgCl_2$,

419 10 mM KCl, 0.25% NP-40, 0.5 mM DTT, 2x protease inhibitor cocktail/Roche), incubated for 10 min on
420 ice, and homogenized using a Dounce homogenizer (Wheaton). The homogenate was centrifuged at 4°C
421 for 5 min at 3,800×*g*). Pelleted nuclei were resuspended in Buffer C (20 mM HEPES pH 8.0, 25% glycerol,
422 1.5 mM $MgCl_2$, 420 mM KCl, 0.25% NP-40, 0.2 mM EDTA, 0.5 mM DTT, 2x protease inhibitor cocktail)
423 and homogenized again using a Dounce homogenizer (Wheaton). Nuclei were extracted for 30 min at 4°C
424 with end-to-end rotation, then centrifuged at 4°C for 30 min at 20,000×*g*. The nuclear extract was then
425 subjected to immunoprecipitation with anti-FLAG (M2)-agarose beads (Sigma). Bead-bound proteins
426 were eluted with 200 ng/µL 3×FLAG peptide (Sigma) by incubating at 4°C for 30 min. The protein was
427 pure and monodisperse as verified using Coomassie blue staining and size exclusion chromatography.
428

429 **Generation of Oct4^{C48S} mutant mouse ESCs**

430 Mutant ESCs with a single-base missense mutation that converts Cys48 to Ser were generated by the
431 University of Utah Health Sciences Center Mutation Generation and Detection Core and the University
432 of Utah Transgenic and Gene Targeting Mouse Facilites using a templated CRISPR approach. Briefly, a
433 gRNA was designed to target Cas9 and cutting near the target nucleotide, 5'
434 GGCCTCGAAGCGA<u>C</u>AGATGG 3', where the underlined C indicates the target nucleotide. The sgRNA
435 was ordered as a modified synthetic sgRNA from Synthego and was complexed with Cas9 protein (IDT).
436 to form a ribonucleoprotein (RNP) complex that was electroporated with program CG-104 into into 6×10^4
437 mouse ESCs using a LONZA 4D-Nucleofector X-unit. The RNP was co-electroporated with a
438 corresponding single-stranded oligodeoxynucleotide (ssODN) with 5' and 3' phosphorothioate
439 modifications, 5'-
440 ctcactctgtttgatcggcctttcagGAAAGGTGTTCAGCCAGACCA**CCATCT<u>C</u>T<u>A</u>G<u>A</u>TTCGAGGCC**TTGC
441 AGCTCAGCCTTAAGAACATGTGTAAGCTG-3' where lower-case denotes intronic sequence. The
442 ssODN contained the target mutation (underlined), silent mutations in the seed region (double underlined)
443 to block CRISPR re-cutting after incorporation of the donor and to generate a unique *Xba*I restriction

444 enzyme site for screening purposes (the sgRNA site is in bold). Incorporation of the ssODN sequence was
445 identified via PCR and restriction enzyme digestion, and confirmed by sanger sequencing of the 709 bp
446 region flanking the mutation. Mutant clones were generated from a pure C57BL/6 ESC line, a 129/BL6
447 hybrid line expressing constitutive GFP and a reporter line used to de-differentiate ESCs into a 2C-like
448 state. Clones were tested by karyotype and any clones with <70% normal diploid content were excluded.
449

450 **Chimeric mouse assays**

451 Targeted or parental ESCs were disassociated into single cell suspension by treatment with 0.05% Trypsin
452 (ThermoFisher). After collection by centrifugation, the cells were resuspended in injection medium
453 consisting of DMEM with HEPES (ThermoFisher) supplemented with 10% FBS (Peak Serum) and 0.1
454 mM 2-mercaptoethanol (Sigma). C57B6 clones were injected into albino B6 blastocysts, while the G4N
455 and C4 lines (and derived lines, B6:129 and 129:129 in origin, respectively) were injected into B6
456 blastocysts. E3.5 blastocysts were used. 10-25 ESCs were injected into each blastocyst using
457 micromanipulators (Eppendorf) and a Leica inverted microscope. After culturing injected blastocysts in
458 KSOM (CytoSpring), approximately 12 to 18 blastocysts were surgically transferred into the uterine horns
459 of day 2.5 pseudopregnant mice or the oviduct of day 0.5 pseudopregnant mice. Mice were scored based
460 on coat and eye color by members of the transgenic core facility at weaning.

461 **Acknowledgements**

462 We thank S. Buckley for critical reading of the manuscript. We thank C. Davey and the University of Utah
463 Health Sciences Center Mouse Transgenic Core and Mutation Generation Core Facilities for assistance
464 with generating CRISPR-edited ESC lines, chimerism assays and karyotyping. This work was supported
465 by NIH/NIGMS grant (R01GM122778) to DT.

466

467 **Conflict of interest**

468 The authors declare that they have no conflicts of interest.

469 **Figure legends**

470

471 **Figure 1. Oct4's reprogramming activity vis-à-vis Oct1 maps to the DNA binding domain.**

472 (**A**) Mouse Oct1 vs Oct4 DNA binding domain alignment. Asparagine 40 mutated in some constructs to
473 lysine is shown in blue. The key Oct1 serine 48 residue (cysteine in Oct4) is shown in blue and boxed in
474 red. (**B**) Schematic of the reprogramming vector with C-terminal Oct4 FLAG tag. The different domains
475 of Oct4 and Oct1 are color coded. (**C**) Reprogramming activity of C-terminally tagged Oct4 relative to
476 the untagged control. Results show an average of experiments performed in biological triplicate. Error
477 bars show ±standard deviation. Anti-FLAG and anti-Oct4 immunoblots using tagged and untagged
478 constructs transiently transfected into 293T cells are shown at right. (**D**) iPSC formation assay using target
479 primary fibroblasts with an Oct4-GFP cassette. Reprogramming was assessed using the number of GFP-
480 expressing iPSC colonies. Oct4 was used as a positive control and Oct1 as a negative control. All
481 constructs bore C-terminal FLAG-tags. Experiments were performed in biological triplicate. Error bars
482 show ±standard deviation. (**E**) Schematic and summary of size and activity of the constructs used in (D).
483 The activity of Oct4 in (D) was set to 100%. C-terminal FLAG-tags are not shown for simplicity. (**F**)
484 Anti-FLAG immunoblot showing expression of the different constructs when transiently transfected into
485 293T cells. Cells were transduced with the modified STEMCCA reprogramming vectors similar to the
486 primary Oct4-GFP fibroblasts used in the reprogramming assays. Histone H3 is shown as a loading
487 control.

488

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

490 (**A**) iPSC formation assay using Oct1-Oct4 lentiviral constructs that manipulate cysteine 4. Oct4 was used
491 as a positive control and Oct1 as a negative control. All constructs bore C-terminal FLAG-tags.
492 Experiments were performed in biological triplicate. Error bars show ±standard deviation. (**B**) Schematic
493 and summary of size and activity of the constructs used in (A). The activity of Oct4 in (A) was set to

494 100%. (**C**) Anti-FLAG immunoblot showing expression of the different constructs when transiently
495 transfected into 293T cells. Cells were transduced with the modified STEMCCA reprogramming vectors
496 similar to the primary Oct4-GFP fibroblasts used in the reprogramming assays. Histone H3 is shown as a
497 loading control. (**D**) Detail from an Oct1:octamer DNA co-crystal structure (PDB ID: 1OCT) (Klemm *et*
498 *al.*, 1994) showing Ser48 making hydrogen bond contacts with the DNA backbone. (**E**) Paralog
499 comparison of mouse POU protein amino acid residues corresponding to α-helix 3 of the POU_S sub-
500 domain.
501
502 **Figure 3. Oct4 cysteine 48 confers DNA sensitivity to oxidative stress and oxidative stress-induced**
503 **ubiquitylation.**
504 (**A**) EMSA using canonical octamer DNA and wild-type of C48S mutant full-length Oct4. 293T cells were
505 transiently transfected with plasmids encoding C-terminally Twin-Strep-FLAG-tagged Oct4 constructs.
506 Protein lysates were made and immediately used for EMSA with increasing amounts of diamide (see
507 methods). The larger redox-resistant Oct1 protein that binds wild-type (WT) but not mutant (Mut) probe
508 is highlighted with an arrow. Asterisk shows a redox sensitive band with superior binding to mutant rather
509 than octamer DNA that is presumably nonspecific. Together these proteins serve as loading controls. (**B**)
510 Quantification of Oct4 band intensity relative to no diamide, which was set to 100%, from three
511 independent experiments. Error bars show ±standard deviation. (**C**) Oct4 containing a C-terminal
512 streptavidin tag was purified from 293T cell lysates using anti-FLAG beads (see methods). Isolated protein
513 was used in assays similar to (A). (**D**) Quantification of (C) and two other similar experiments. Error bars
514 show ±standard deviation. (**E**) Highly exposed Oct4 immunoblot using lysates prepared from 293T cells
515 transfected with constructs expressing C-terminally FLAG/Twin-Strep-tagged Oct4. Subsets of cells were
516 pre-treated with H_2O_2 (1 mM for 2 hr), MG-132 (10 μM for 2 hr) or both. All samples contained MG-132
517 in the lysis buffer to prevent degradation. Histone H3 is shown as a loading control. (**F**) Purified wild-
518 type and C48S mutant Oct4 was immunoblotted using antibodies against specific ubiquitin linkages. The

519 FK2 antibody (Enzo Life Sciences) detects pan-Ub linkages, and the antibody detecting Ub-K48-specific
520 linkages was obtained from Cell Signaling.
521
522
523 **Figure 4. Oct4^{C48S} mutant ESCs are phenotypically normal but retain Oct4 expression when**
524 **differentiated.**
525 (**A**) *Top:* Schematic for CRISPR-mediated *Pou5f1* (*Oct4*) mutation. A single nucleotide change (arrow)
526 converts the wild-type cysteine to the serine present in Oct1. Asterisk depict two silent point mutations
527 present in the guide RNA to monitor mutagenesis. *Bottom:* DNA sequencing trace of homozygous mutant
528 clone. (**B**) An example C57BL/6 ESC clone (#46) was immunoblotted for Oct4, Sox2 and Nanog. Histone
529 H3 was used as a loading control. (**C**) Example images of pluripotent colonies from a parental ESC line
530 and derived Oct4^{C48S} mutant ESC line (#46). Images were collected at 10× magnification. Scale bar = 0.2
531 mm. (**D**) Parent or derived Oct4^{C48S} mutant ESCs were differentiated without feeders using retinoic acid.
532 Lysates were collected at the indicated times and probed for Oct4 by immunoblotting. β-actin was used
533 as a loading control. (E) Average quantified Oct4/β-actin levels from three replicate immunoblots
534 conducted with the same ESC clone (#46). Error bars depict ±standard deviation. (**F**) The same cells were
535 used to collect phase-contrast images at longer retinoic acid timepoints. Images were collected at the
536 indicated times at 10× magnification. Scale bar = 0.2 mm. (**G**) Differentiating cells were trypsinized and
537 counted using trypan blue exclusion using a hemacytometer. Cell counts were averaged with two other
538 biological replicates (n=3) and plotted. Error bars depict ±standard deviation.
539
540 **Figure 5. Oct4^{C48S} mutant ESCs do not contribute to adult chimeric animals.**
541 Three different parent (left side) or CRISPR-mutated Oct4^{C48S} ESCs (right side) were injected into
542 recipient blastocysts and implanted in pseudo-pregnant animals. For each line, the proportion of chimeric

543 offspring at weaning is shown as a pie chart. All karyotypes were confirmed to be similar to parent lines

544 (between 72% and 92% normal).

545 **References**

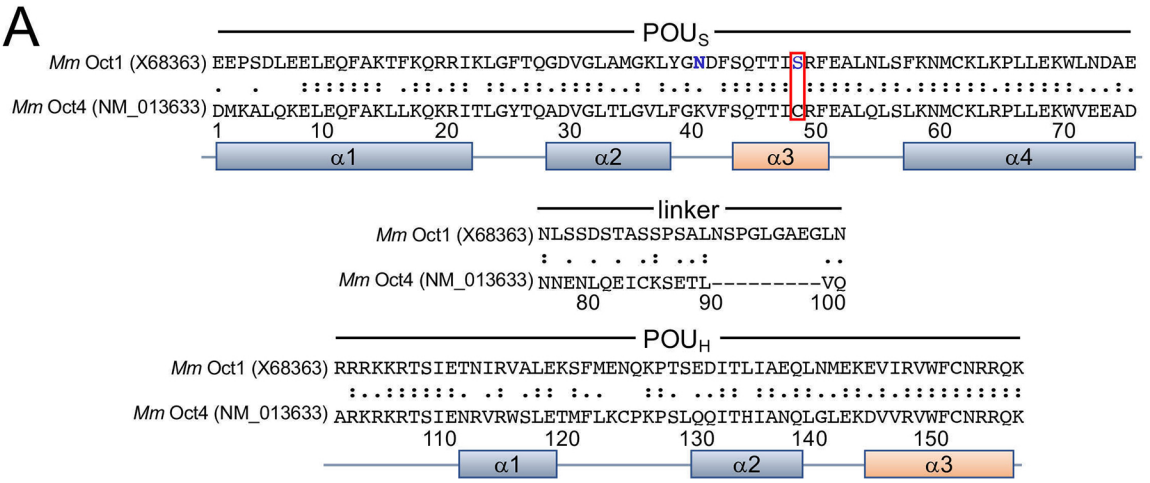
546 Abranches, E., Silva, M., Pradier, L., Schulz, H., Hummel, O., Henrique, D., and Bekman, E. (2009).
547 Neural differentiation of embryonic stem cells in vitro: a road map to neurogenesis in the embryo. PLoS
548 ONE *4*, e6286. 10.1371/journal.pone.0006286 [doi].
549 An, Z., Liu, P., Zheng, J., Si, C., Li, T., Chen, Y., Ma, T., Zhang, M.Q., Zhou, Q., and Ding, S. (2019).
550 Sox2 and Klf4 as the Functional Core in Pluripotency Induction Without Exogenous Oct4. Cell Rep *29*,
551 1986-2000. 10.1016/j.celrep.2019.10.026.
552 Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A.,
553 Wernig, M., Plath, K., et al. (2006). A bivalent chromatin structure marks key developmental genes in
554 embryonic stem cells. Cell *125*, 315-326. S0092-8674(06)00380-1 [pii]
555 10.1016/j.cell.2006.02.041 [doi].
556 Boija, A., Klein, I.A., Sabari, B.R., Dall'Agnese, A., Coffey, E.L., Zamudio, A.V., Li, C.H., Shrinivas, K.,
557 Manteiga, J.C., Hannett, N.M., et al. (2018). Transcription Factors Activate Genes through the Phase-
558 Separation Capacity of Their Activation Domains. Cell *175*, 1842-1855 e1816.
559 Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar,
560 R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry in human
561 embryonic stem cells. Cell *122*, 947-956.
562 Brigelius-Flohe, R., and Flohe, L. (2011). Basic principles and emerging concepts in the redox control of
563 transcription factors. Antioxid Redox Signal *15*, 2335-2381. 10.1089/ars.2010.3534.
564 Britto, P.J., Knipling, L., and Wolff, J. (2002). The local electrostatic environment determines cysteine
565 reactivity of tubulin. J Biol Chem *277*, 29018-29027. 10.1074/jbc.M204263200.
566 Buganim, Y., Markoulaki, S., van Wietmarschen, N., Hoke, H., Wu, T., Ganz, K., Akhtar-Zaidi, B., He,
567 Y., Abraham, B.J., Porubsky, D., et al. (2014). The developmental potential of iPSCs is greatly influenced
568 by reprogramming factor selection. Cell Stem Cell *15*, 295-309. 10.1016/j.stem.2014.07.003.
569 Chen, J., Gao, Y., Huang, H., Xu, K., Chen, X., Jiang, Y., Li, H., Gao, S., Tao, Y., Wang, H., et al. (2015).
570 The combination of Tet1 with Oct4 generates high-quality mouse-induced pluripotent stem cells. Stem
571 Cells *33*, 686-698. 10.1002/stem.1879.
572 Costa, Y., Ding, J., Theunissen, T.W., Faiola, F., Hore, T.A., Shliaha, P.V., Fidalgo, M., Saunders, A.,
573 Lawrence, M., Dietmann, S., et al. (2013). NANOG-dependent function of TET1 and TET2 in
574 establishment of pluripotency. Nature *495*, 370-374. nature11925 [pii]
575 10.1038/nature11925 [doi].
576 De Los Angeles, A., Ferrari, F., Xi, R., Fujiwara, Y., Benvenisty, N., Deng, H., Hochedlinger, K.,
577 Jaenisch, R., Lee, S., Leitch, H.G., et al. (2015). Hallmarks of pluripotency. Nature *525*, 469-478.
578 Delaunay, A., Isnard, A.D., and Toledano, M.B. (2000). H2O2 sensing through oxidation of the Yap1
579 transcription factor. EMBO J *19*, 5157-5166. 10.1093/emboj/19.19.5157.
580 DeVeale, B., Brokhman, I., Mohseni, P., Babak, T., Yoon, C., Lin, A., Onishi, K., Tomilin, A., Pevny, L.,
581 Zandstra, P.W., et al. (2013). Oct4 is required ~E7.5 for proliferation in the primitive streak. PLoS Genet
582 *9*, e1003957. 10.1371/journal.pgen.1003957.
583 Esch, D., Vahokoski, J., Groves, M.R., Pogenberg, V., Cojocaru, V., Vom Bruch, H., Han, D., Drexler,
584 H.C., Arauzo-Bravo, M.J., Ng, C.K., et al. (2013). A unique Oct4 interface is crucial for reprogramming
585 to pluripotency. Nat Cell Biol *15*, 295-301. ncb2680 [pii]
586 10.1038/ncb2680 [doi].
587 Esteban, M.A., Wang, T., Qin, B., Yang, J., Qin, D., Cai, J., Li, W., Weng, Z., Chen, J., Ni, S., et al.
588 (2010). Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. Cell Stem
589 Cell *6*, 71-79. S1934-5909(09)00624-9 [pii]
590 10.1016/j.stem.2009.12.001 [doi].

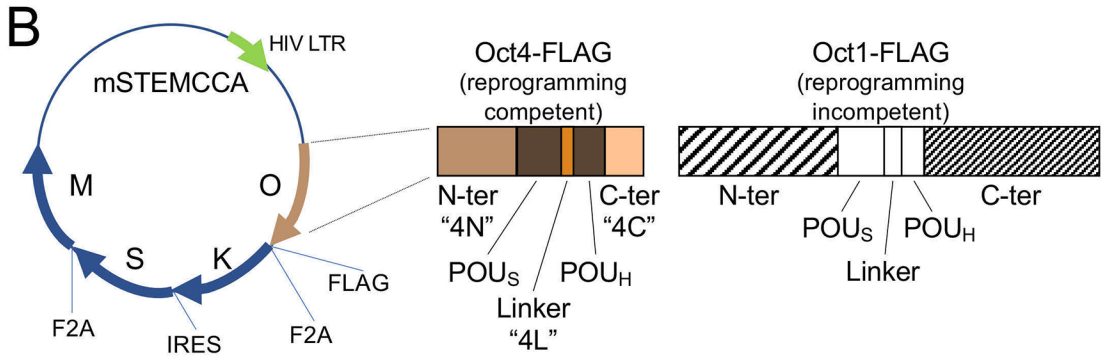
591 Feng, B., Jiang, J., Kraus, P., Ng, J.H., Heng, J.C., Chan, Y.S., Yaw, L.P., Zhang, W., Loh, Y.H., Han, J.,
592 et al. (2009). Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear
593 receptor Esrrb. Nat Cell Biol *11*, 197-203.
594 Ferraris, L., Stewart, A.P., Kang, J., DeSimone, A.M., Gemberling, M., Tantin, D., and Fairbrother, W.G.
595 (2011). Combinatorial binding of transcription factors in the pluripotency regions of the genome. Genome
596 Res. *21*, 1055-1064.
597 Gao, Y., Chen, J., Li, K., Wu, T., Huang, B., Liu, W., Kou, X., Zhang, Y., Huang, H., Jiang, Y., et al.
598 (2013). Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation
599 and hydroxymethylation in reprogramming. Cell Stem Cell *12*, 453-469. 10.1016/j.stem.2013.02.005.
600 Guo, Y., Einhorn, L., Kelley, M., Hirota, K., Yodoi, J., Reinbold, R., Scholer, H., Ramsey, H., and
601 Hromas, R. (2004). Redox regulation of the embryonic stem cell transcription factor Oct-4 by thioredoxin.
602 Stem Cells *22*, 259-264.
603 Han, D., Wu, G., Chen, R., Drexler, H.C.A., MacCarthy, C.M., Kim, K.P., Adachi, K., Gerovska, D.,
604 Mavrommatis, L., Bedzhov, I., et al. (2022). A balanced Oct4 interactome is crucial for maintaining
605 pluripotency. Sci Adv *8*, eabe4375. 10.1126/sciadv.abe4375.
606 Heng, J.C., Feng, B., Han, J., Jiang, J., Kraus, P., Ng, J.H., Orlov, Y.L., Huss, M., Yang, L., Lufkin, T.,
607 et al. (2010). The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells
608 to pluripotent cells. Cell Stem Cell *6*, 167-174. 10.1016/j.stem.2009.12.009.
609 Jerabek, S., Ng, C.K., Wu, G., Arauzo-Bravo, M.J., Kim, K.P., Esch, D., Malik, V., Chen, Y., Velychko,
610 S., MacCarthy, C.M., et al. (2017). Changing POU dimerization preferences converts Oct6 into a
611 pluripotency inducer. EMBO Rep *18*, 319-333.
612 Jin, W., Wang, L., Zhu, F., Tan, W., Lin, W., Chen, D., Sun, Q., and Xia, Z. (2016). Critical POU domain
613 residues confer Oct4 uniqueness in somatic cell reprogramming. Sci Rep *6*, 20818. 10.1038/srep20818.
614 Kang, J., Gemberling, M., Nakamura, M., Whitby, F.G., Handa, H., Fairbrother, W.G., and Tantin, D.
615 (2009a). A general mechanism for transcription regulation by Oct1 and Oct4 in response to genotoxic and
616 oxidative stress. Genes Dev *23*, 208-222. 10.1101/gad.1750709.
617 Kang, J., Shakya, A., and Tantin, D. (2009b). Stem cells, stress, metabolism and cancer: a drama in two
618 Octs. Trends Biochem Sci *34*, 491-499. 10.1016/j.tibs.2009.06.003.
619 Kim, K.P., Wu, Y., Yoon, J., Adachi, K., Wu, G., Velychko, S., MacCarthy, C.M., Shin, B., Ropke, A.,
620 Arauzo-Bravo, M.J., et al. (2020). Reprogramming competence of OCT factors is determined by
621 transactivation domains. Sci Adv *6*, eaaz7364. 10.1126/sciadv.aaz7364.
622 Klemm, J.D., Rould, M.A., Aurora, R., Herr, W., and Pabo, C.O. (1994). Crystal structure of the Oct-1
623 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. Cell *77*,
624 21-32.
625 Lengner, C.J., Camargo, F.D., Hochedlinger, K., Welstead, G.G., Zaidi, S., Gokhale, S., Scholer, H.R.,
626 Tomilin, A., and Jaenisch, R. (2007). Oct4 expression is not required for mouse somatic stem cell self-
627 renewal. Cell Stem Cell *1*, 403-415. 10.1016/j.stem.2007.07.020 [doi].
628 Li, Y., Zhang, Q., Yin, X., Yang, W., Du, Y., Hou, P., Ge, J., Liu, C., Zhang, W., Zhang, X., et al. (2011).
629 Generation of iPSCs from mouse fibroblasts with a single gene, Oct4, and small molecules. Cell Res *21*,
630 196-204. cr2010142 [pii]
631 10.1038/cr.2010.142 [doi].
632 Lickteig, K., Lamb, K., Brigman, K., and Rizzino, A. (1996). Effects of oxidation and reduction on the
633 binding of transcription factors to cis-regulatory elements located in the FGF-4 gene. Mol Reprod Dev *44*,
634 146-152.
635 Mamun, M.M.A., Khan, M.R., Zhu, Y., Zhang, Y., Zhou, S., Xu, R., Bukhari, I., Thorne, R.F., Li, J.,
636 Zhang, X.D., et al. (2022). Stub1 maintains proteostasis of master transcription factors in embryonic stem
637 cells. Cell Rep *39*, 110919. 10.1016/j.celrep.2022.110919.
638 Marinho, H.S., Real, C., Cyrne, L., Soares, H., and Antunes, F. (2014). Hydrogen peroxide sensing,
639 signaling and regulation of transcription factors. Redox Biol *2*, 535-562. 10.1016/j.redox.2014.02.006.

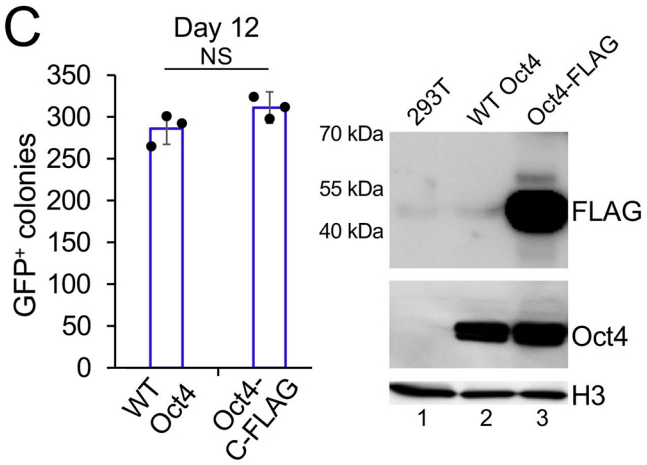
640 Marsboom, G., Zhang, G.F., Pohl-Avila, N., Zhang, Y., Yuan, Y., Kang, H., Hao, B., Brunengraber, H.,
641 Malik, A.B., and Rehman, J. (2016). Glutamine Metabolism Regulates the Pluripotency Transcription
642 Factor Oct4. Cell Rep *16*, 323-332. 10.1016/j.celrep.2016.05.089.
643 Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E.,
644 Nusbaum, C., Jaffe, D.B., et al. (2008). Genome-scale DNA methylation maps of pluripotent and
645 differentiated cells. Nature *454*, 766-770.
646 Michael, A.K., Grand, R.S., Isbel, L., Cavadini, S., Kozicka, Z., Kempf, G., Bunker, R.D., Schenk, A.D.,
647 Graff-Meyer, A., Pathare, G.R., et al. (2020). Mechanisms of OCT4-SOX2 motif readout on nucleosomes.
648 Science *368*, 1460-1465. 10.1126/science.abb0074.
649 Morey, L., Santanach, A., and Di Croce, L. (2015). Pluripotency and Epigenetic Factors in Mouse
650 Embryonic Stem Cell Fate Regulation. Mol Cell Biol *35*, 2716-2728.
651 Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y.,
652 Takizawa, N., and Yamanaka, S. (2008). Generation of induced pluripotent stem cells without Myc from
653 mouse and human fibroblasts. Nat Biotechnol *26*, 101-106.
654 Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and
655 Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU
656 transcription factor Oct4. Cell *95*, 379-391.
657 Okamoto, K., Okazawa, H., Okuda, A., Sakai, M., Muramatsu, M., and Hamada, H. (1990). A novel
658 octamer binding transcription factor is differentially expressed in mouse embryonic cells. Cell *60*, 461-
659 472.
660 Okita, K., Ichisaka, T., and Yamanaka, S. (2007). Generation of germline-competent induced pluripotent
661 stem cells. Nature *448*, 313-317.
662 Perovanovic, J., Wu, Y., Abewe, H., Shen, Z., Hughes, E.P., Gertz, J., Chandrasekharan, M.B., and Tantin,
663 D. (2023). Oct1 cooperates with Smad transcription factors to promote lineage specification. Sci. Signal.
664 *in press*.
665 Rais, Y., Zviran, A., Geula, S., Gafni, O., Chomsky, E., Viukov, S., Mansour, A.A., Caspi, I., Krupalnik,
666 V., Zerbib, M., et al. (2013). Deterministic direct reprogramming of somatic cells to pluripotency. Nature
667 *502*, 65-70. 10.1038/nature12587.
668 Remenyi, A., Pohl, E., Scholer, H.R., and Wilmanns, M. (2001). Crystallization of redox-insensitive Oct1
669 POU domain with different DNA-response elements. Acta Crystallogr D Biol Crystallogr *57*, 1634-1638.
670 Rosner, M.H., Vigano, M.A., Ozato, K., Timmons, P.M., Poirier, F., Rigby, P.W., and Staudt, L.M.
671 (1990). A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo.
672 Nature *345*, 686-692. 10.1038/345686a0 [doi].
673 Ruetz, T., Pfisterer, U., Di Stefano, B., Ashmore, J., Beniazza, M., Tian, T.V., Kaemena, D.F., Tosti, L.,
674 Tan, W., Manning, J.R., et al. (2017). Constitutively Active SMAD2/3 Are Broad-Scope Potentiators of
675 Transcription-Factor-Mediated Cellular Reprogramming. Cell Stem Cell *21*, 791-805 e799.
676 10.1016/j.stem.2017.10.013.
677 Shakya, A., Callister, C., Goren, A., Yosef, N., Garg, N., Khoddami, V., Nix, D., Regev, A., and Tantin,
678 D. (2015). Pluripotency transcription factor Oct4 mediates stepwise nucleosome demethylation and
679 depletion. Mol Cell Biol *35*, 1014-1025. 10.1128/MCB.01105-14.
680 Shen, Z., Formosa, T., and Tantin, D. (2018). FACT Inhibition Blocks Induction But Not Maintenance of
681 Pluripotency. Stem Cells Dev *27*, 1693-1701. 10.1089/scd.2018.0150.
682 Shen, Z., Kang, J., Shakya, A., Tabaka, M., Jarboe, E.A., Regev, A., and Tantin, D. (2017). Enforcement
683 of developmental lineage specificity by transcription factor Oct1. ELife *6*, e20937. 10.7554/eLife.20937.
684 Shi, Y., Desponts, C., Do, J.T., Hahm, H.S., Scholer, H.R., and Ding, S. (2008). Induction of pluripotent
685 stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. Cell
686 Stem Cell *3*, 568-574.

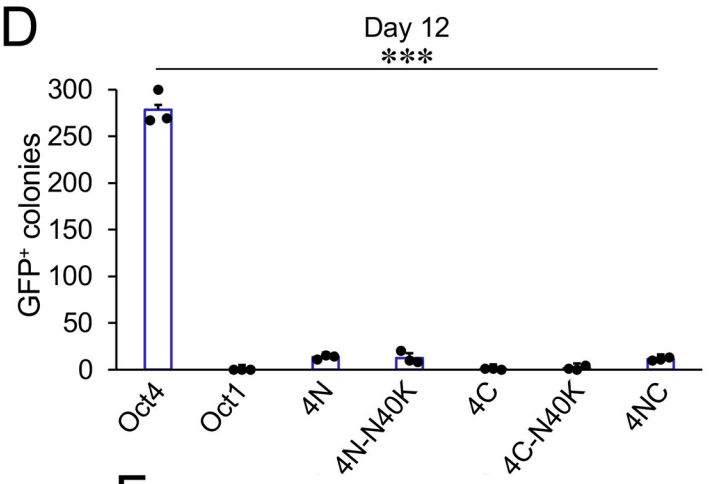
687 Shu, J., Wu, C., Wu, Y., Li, Z., Shao, S., Zhao, W., Tang, X., Yang, H., Shen, L., Zuo, X., et al. (2013).
688 Induction of pluripotency in mouse somatic cells with lineage specifiers. Cell *153*, 963-975.
689 10.1016/j.cell.2013.05.001.
690 Somers, A., Jean, J.C., Sommer, C.A., Omari, A., Ford, C.C., Mills, J.A., Ying, L., Sommer, A.G., Jean,
691 J.M., Smith, B.W., et al. (2010). Generation of transgene-free lung disease-specific human induced
692 pluripotent stem cells using a single excisable lentiviral stem cell cassette. Stem Cells *28*, 1728-1740.
693 10.1002/stem.495.
694 Sommer, C.A., Stadtfeld, M., Murphy, G.J., Hochedlinger, K., Kotton, D.N., and Mostoslavsky, G.
695 (2009). Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. Stem Cells *27*,
696 543-549.
697 Soufi, A., Donahue, G., and Zaret, K.S. (2012). Facilitators and impediments of the pluripotency
698 reprogramming factors' initial engagement with the genome. Cell *151*, 994-1004.
699 Suzuki, N., Rohdewohld, H., Neuman, T., Gruss, P., and Scholer, H.R. (1990). Oct-6: a POU transcription
700 factor expressed in embryonal stem cells and in the developing brain. EMBO J *9*, 3723-3732.
701 Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007).
702 Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861-872.
703 Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and
704 adult fibroblast cultures by defined factors. Cell *126*, 663-676. DOI 10.1016/j.cell.2006.07.024.
705 Tantin, D. (2013). Oct transcription factors in development and stem cells: insights and mechanisms.
706 Development *140*, 2857-2866. 10.1242/dev.095927.
707 Tantin, D., Gemberling, M., Callister, C., and Fairbrother, W. (2008). High-throughput Biochemical
708 Analysis of in-vivo Location Data Reveals Novel Distinct Classes of POU5F1(Oct4)/DNA complexes.
709 Genome Res *18*, 631-639.
710 Velychko, S., Adachi, K., Kim, K.P., Hou, Y., MacCarthy, C.M., Wu, G., and Scholer, H.R. (2019).
711 Excluding Oct4 from Yamanaka Cocktail Unleashes the Developmental Potential of iPSCs. Cell Stem
712 Cell *25*, 737-753. 10.1016/j.stem.2019.10.002.
713 Vierbuchen, T., and Wernig, M. (2012). Molecular roadblocks for cellular reprogramming. Mol Cell *47*,
714 827-838. S1097-2765(12)00784-8 [pii]
715 10.1016/j.molcel.2012.09.008 [doi].
716 Yuan, X., Wan, H., Zhao, X., Zhu, S., Zhou, Q., and Ding, S. (2011). Combined Chemical Treatment
717 Enables Oct4-Induced Reprogramming from Mouse Embryonic Fibroblasts. Stem Cells *29*, 549-553.
718 10.1002/stem.594 [doi].
719 Zheng, M., Aslund, F., and Storz, G. (1998). Activation of the OxyR transcription factor by reversible
720 disulfide bond formation. Science *279*, 1718-1721.
721

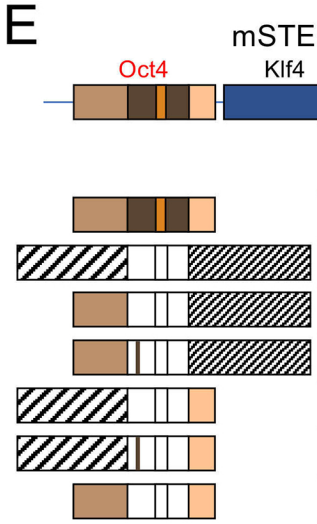
Shen et al. Fig.1



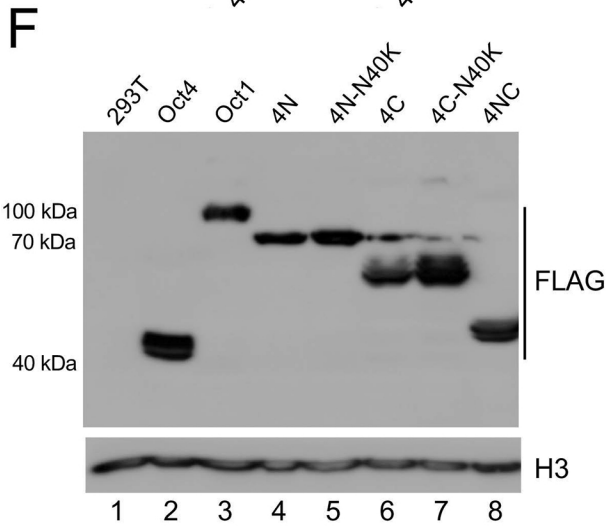




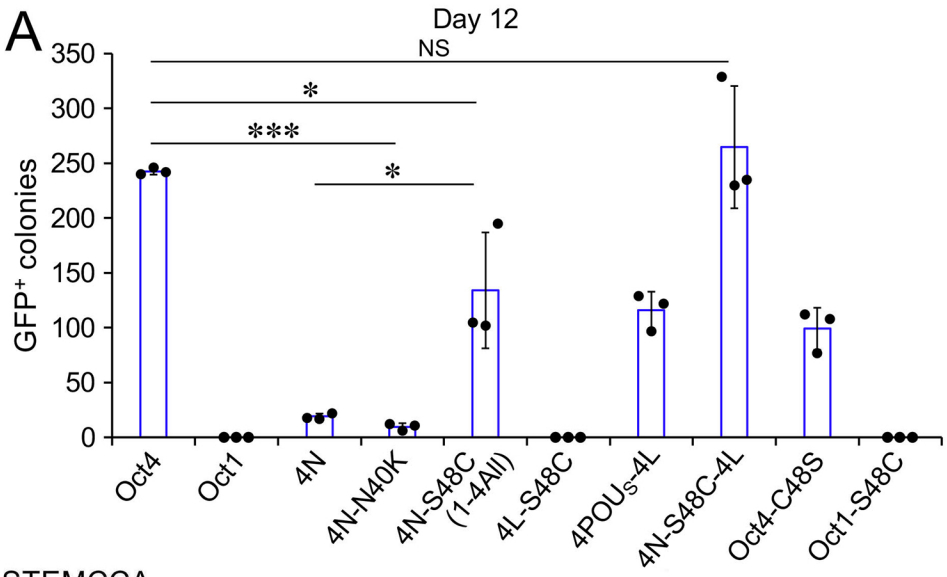


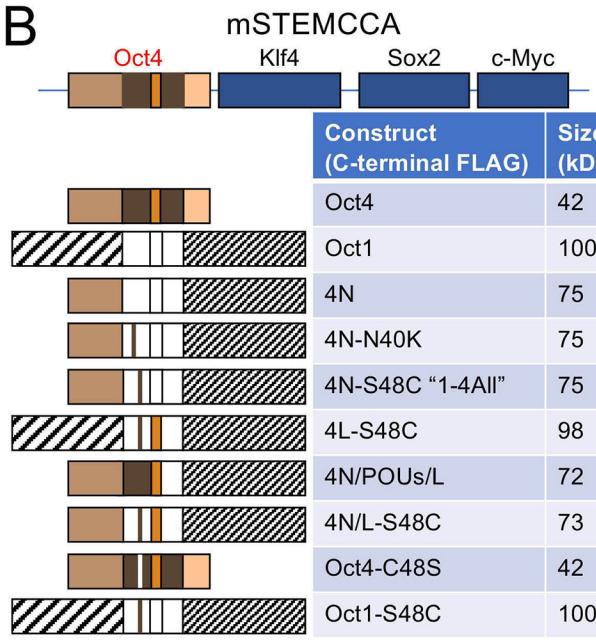


Construct (C-terminal FLAG)	Size (kDa)	Activity (%)
Oct4	42	100
Oct1	100	0.0
4N	75	4.8
4N-N40K	75	4.5
4C	60	0.4
4C-N40K	60	1.0
4NC	43	6.1

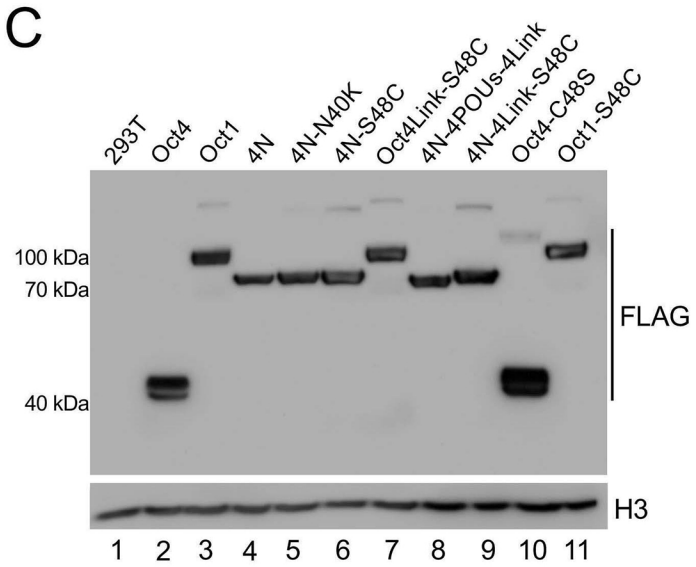


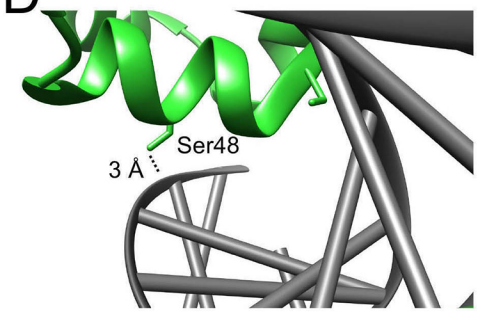
Shen et al. Fig.2





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

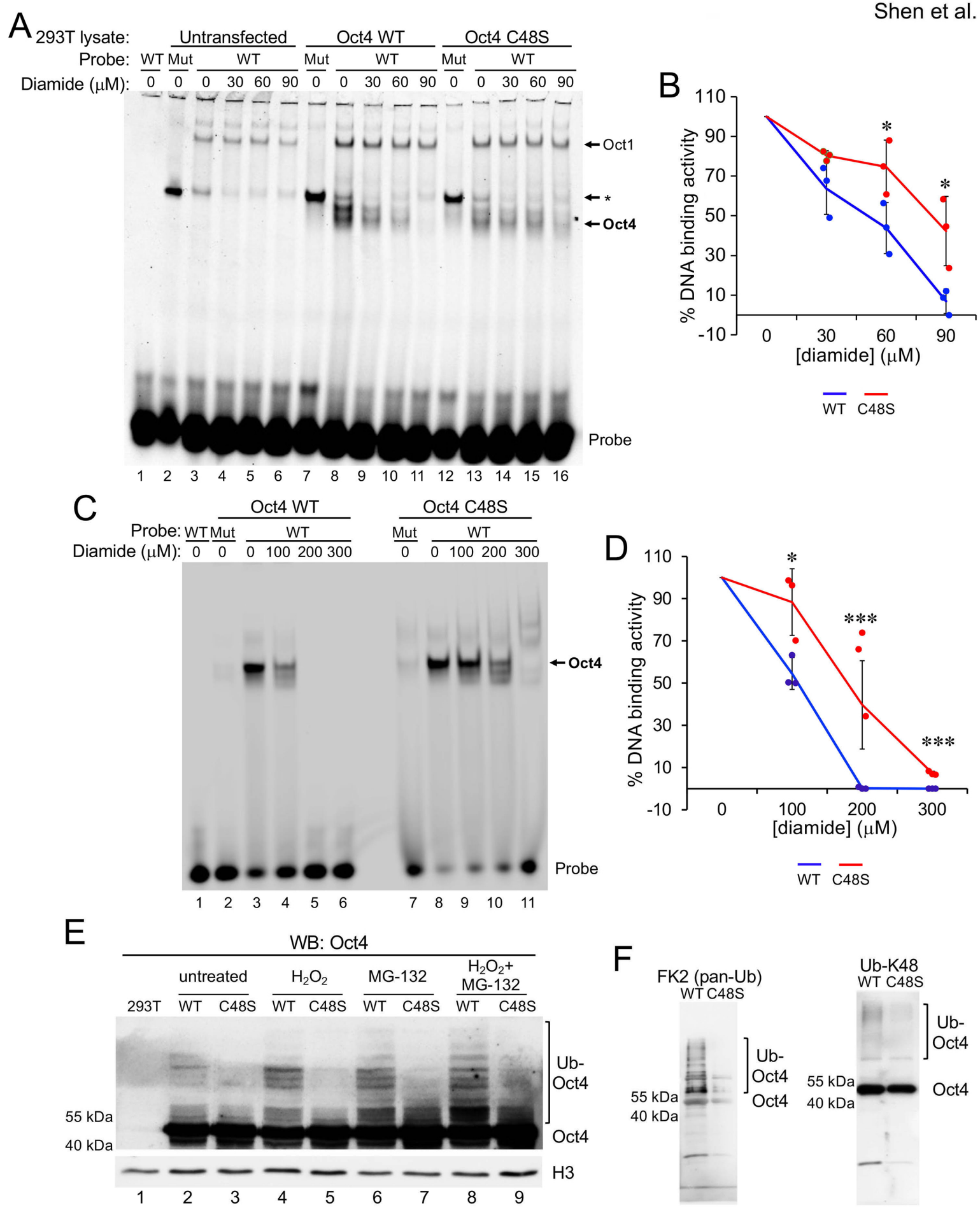




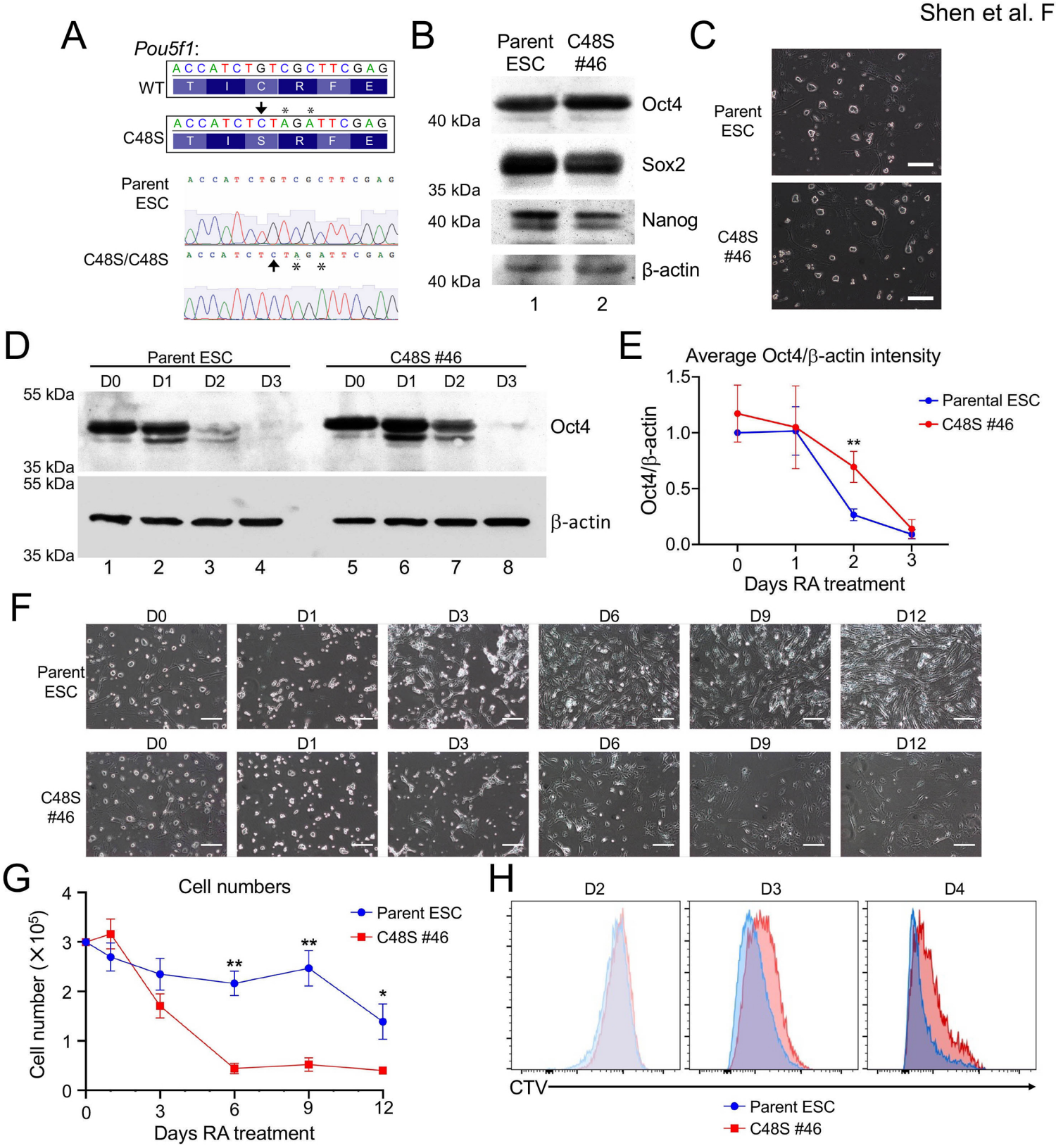
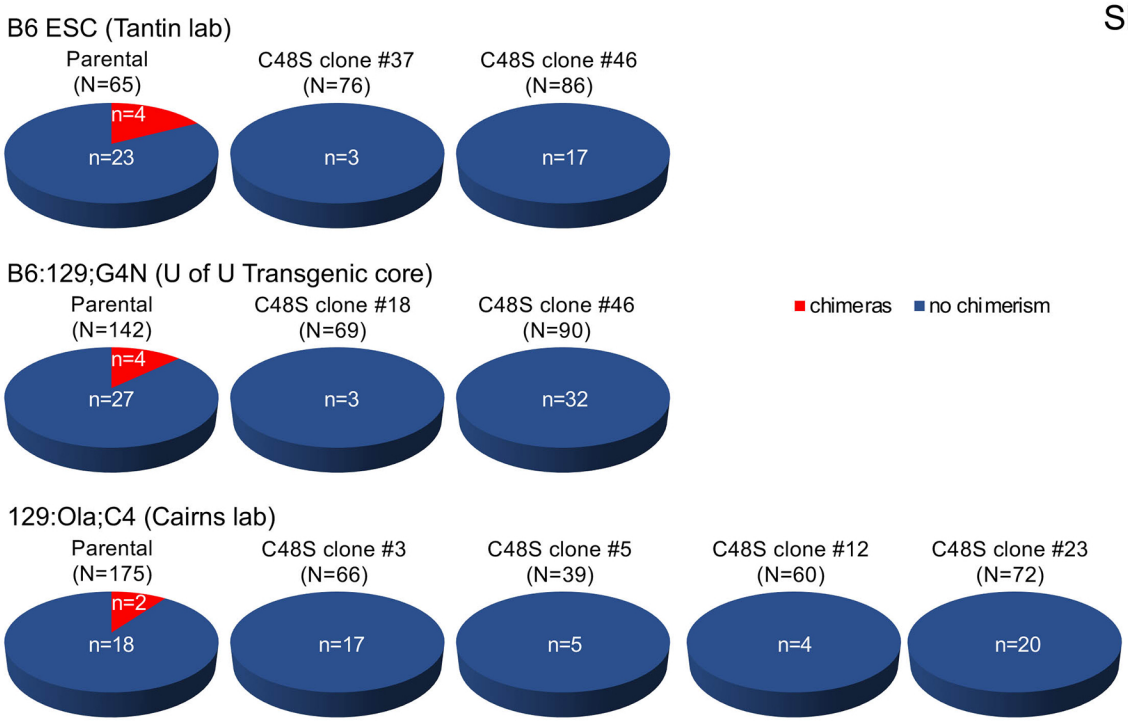
E

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



Shen et al. Fig.4



Shen et al. Fig.5