Direct reprogramming of somatic cells is promoted by maternal transcription factor Glis1

Momoko Maekawa[1,2], Kei Yamaguchi[3], Tomonori Nakamura[1,4], Ran Shibukawa[1,2], Ikumi Kodanaka[1,2], Tomoko Ichisaka[1,4], Yoshifumi Kawamura[3], Hiromi Mochizuki[3], Naoki Goshima[5] & Shinya Yamanaka[1,2,4,6]

[1]Center for iPS Cell Research and Application (CiRA), Kyoto University, Kyoto 606-8507, Japan.

[2]Yamanaka iPS Cell Special Project, JST, Kawaguchi 332-0012, Japan.

[3]Japan Biological Informatics Consortium, Tokyo 135-0064, Japan.

[4]Institute for Integrated Cell-Material Sciences, Kyoto University, Kyoto 606-8507, Japan.

[5]Biomedicinal Information Research Center, National Institute of Advanced Industrial Science and Technology, Tokyo 135-0064, Japan.

[6]Gladstone Institute of Cardiovascular Disease, San Francisco, California 94158, USA.

Induced pluripotent stem cells (iPSCs) are generated from somatic cells by the transgenic expression of three transcription factors collectively called OSK: Oct3/4 (also called Pou5f1), Sox2 and Klf4[1]. However, the conversion to iPSCs is inefficient. The proto-oncogene *Myc* enhances the efficiency of iPSC generation by OSK but it also increases the tumorigenicity of the resulting iPSCs[2]. Here we show that the Gli-like transcription factor Glis1 (Glis family zinc finger 1) markedly enhances the generation of iPSCs from both mouse and human fibroblasts when it is expressed together with OSK. Mouse iPSCs generated using this combination of transcription factors can form germline-competent chimaeras. Glis1 is enriched in unfertilized oocytes and in embryos at the one-cell stage. DNA microarray analyses show that

Glis1 promotes multiple pro-reprogramming pathways, including Myc, Nanog, Lin28, Wnt, Essrb and the mesenchymal–epithelial transition. These results therefore show that Glis1 effectively promotes the direct reprogramming of somatic cells during iPSC generation.

The generation of iPSCs is technically simple and highly reproducible[3,4] but only a small proportion of cells become iPSCs after introduction of the four transcription factors[5]. In addition, the generation of iPSCs is slow and requires multiple cell divisions[6]. Reprogramming towards pluripotency can also be achieved by nuclear transfer to meiotic oocytes[7] or mitotic zygotes[8]: this strategy is technically more demanding but it is efficient, rapid and independent of cell division. These differences may indicate that oocytes and zygotes contain factor(s) that promote reprogramming during the generation of iPSCs.

In this study, we initially evaluated a library of 1,437 human transcription factors for their ability to replace Kruppel-like factor 4 (Klf4) or POU domain, class 5, transcription factor 1 (Pou5f1, also known as Oct3/4) during iPSC generation from mouse skin fibroblasts containing a green fluorescent protein (GFP) reporter driven by the nanog homeobox (*Nanog*) promoter and enhancers[9] (Supplementary Table 1). We found that 18 factors could replace Klf4 reproducibly, although with much lower efficiencies of iPSC generation (Supplementary Table 2); we failed to identify any factors that replaced Oct3/4.

Among these 18 factors, we found that GLIS1, a GLI transcription factor[10], markedly increased the number of GFP-positive colonies when it was co-introduced with the 'OSK' transcription factors Oct3/4, SRY-box 2 (Sox2) and Klf4 into adult mouse skin fibroblasts (Fig. 1a). The effect of GLIS1 was comparable to that of MYC, as judged by

the number of GFP-positive colonies (Fig. 1b). We also observed a synergistic increase in the number of GFP-positive colonies when both GLIS1 and MYC were co-introduced with OSK. Notably, GLIS1 specifically promoted the generation of GFP-positive colonies, but not GFP-negative colonies, which represent either partially reprogrammed cells or transformed cells (Fig. 1c). In contrast, MYC increased the number of GFP-negative colonies more than the number of GFP-positive ones. This undesired effect of MYC was counteracted when GLIS1 was co-expressed.

Mouse iPSCs generated with OSK and GLIS1 showed morphologies similar to embryonic stem (ES) cells (Supplementary Fig. 1a). Pluripotency markers such as *Nanog* were expressed at comparable levels to those in ES cells (Supplementary Fig. 1b) and the iPSCs formed teratomas in nude mice (Supplementary Fig. 1c). Furthermore, they produced germline-competent chimaeras (Fig. 1d and Supplementary Table 3).

In human adult fibroblasts, GLIS1 showed a similar effect: it promoted the generation of ES-cell-like colonies to a comparable degree to MYC when it was co-introduced with OSK (Fig. 2a). Notably, GLIS1 specifically promoted the generation of ES-cell-like colonies with a flat, round shape and a distinct edge, but did not promote the generation of non-ES-cell-like colonies, which were granular with an irregular edge (Fig. 2b and Supplementary Fig. 2). In contrast, MYC increased the number of non-ES-cell-like colonies more than the number of ES-cell-like ones (Fig. 2b). The iPSCs generated with OSK and GLIS1 were similar to ES cells in morphology (Supplementary Fig. 3a) and in their expression of undifferentiated-ES-cell marker genes, such as *OCT3/4, SOX2, NANOG* and *ZFP42* (zinc finger protein 42 homolog (mouse), also known as *REX1*) (Supplementary Fig. 3b). DNA microarray analyses showed that human iPSCs established with OSK and GLIS1 had similar global gene expression to cells generated

with OSK and MYC (OSKM) (Fig. 2c). The promoter region of the *OCT3/4* gene showed a hypomethylation pattern (Supplementary Fig. 3c) and the iPSCs differentiated into various cells of the three germ layers in the embryoid body (Supplementary Fig. 3d) and also into teratomas (Fig. 2d). These results demonstrate that GLIS1 strongly and specifically promotes the generation of both mouse and human iPSCs by OSK.

We next studied the expression pattern of *Glis1* in mouse cells. Analyses of expressed sequence tag (EST) databases predicted that *Glis1* expression would be enriched in zygotes, especially in the fertilized ovum (http://www.ncbi.nlm.nih.gov/UniGene/ESTProfileViewer.cgi?uglist=Mm.331757 as of 7 December 2010). In addition, the gene expression data from reverse transcription PCR (RT–PCR), provided by the mouse genome database MGI, showed that there was moderate expression of *Glis1* in metaphase II oocytes and weak expression in two-cell embryos, but that expression was either absent or at trace levels in embryos at the four-cell to embryonic-day-4.5 stages (http://www.informatics.jax.org/searches/expression.cgi?32989 as of 7 December 2010, also reported in ref. 11 in their Supplementary Table 1). To confirm the specific expression of *Glis1* in oocytes and one-cell embryos, we isolated total RNA from oocytes, early embryos and several adult mouse tissues. Real-time PCR detected the highest expression of *Glis1* in one-cell embryos and unfertilized eggs. A modest level of expression was detected in two-cell embryos and placentas and weak expression was detected in several adult tissues (Fig. 3a). These data confirmed that *Glis1* RNA is enriched in unfertilized eggs and one-cell embryos.

We next examined whether endogenous Glis1 has a role during iPSC generation by OSK. We found that Glis1 is expressed at a low level in mouse fibroblasts before and

after the introduction of OSK (Supplementary Fig. 4a). We constructed retroviral vectors to express several *Glis1* small hairpin RNAs (shRNAs), as well as scrambled controls, and tested the knockdown efficiency of each shRNA retrovirus in skin fibroblasts. We found that shRNA2 and shRNA6 were effective (Supplementary Fig. 4b). We then introduced each of these shRNAs, together with OSK, into mouse embryonic fibroblasts (MEFs) containing the *Nanog*–GFP reporter. We found that both shRNA2 and shRNA6 significantly decreased the number of GFP-positive colonies (Supplementary Fig. 4c), in contrast to the scrambled control shRNA. These results show that endogenous Glis1 may have a supportive role during the generation of mouse iPSCs by OSK.

Finally, we tried to elucidate how Glis1 enhances iPSC generation by OSK. We previously reported that suppression of the p53 pathway markedly enhanced iPSC generation from both mouse and human cells[12]. We therefore hypothesized that Glis1 may enhance direct reprogramming by inhibiting p53. If this is the case, Glis1 should not be able to promote iPSC generation in cells with a p53-null background. To test this hypothesis, we introduced OSK plus mock (control) or OSK plus Glis1 into either wild-type or p53-knockout MEFs, both containing the *Nanog*–GFP reporter. Five days after transduction, we measured the proportion of *Nanog*–GFP-positive cells by flow cytometry. We found that even in p53-knockout MEFs, in which the generation of *Nanog*–GFP-positive cells by OSK was increased about 10-fold (to about 2%), the addition of Glis1 further increased the proportion of GFP-positive cells up to about 17% (Supplementary Fig. 5). These data indicate that Glis1 promotes iPSC generation irrespective of p53.

We then used the very high reprogramming efficiency in cells with the p53-null background to elucidate the function of Glis1. We sorted and collected

Nanog–GFP-positive cells 5 days after the transduction of OSK plus mock or OSK plus Glis1 into the p53-knockout MEFs. We then conducted microarray analysis to compare the gene expression levels of these cell populations undergoing reprogramming (Fig. 3b, c and Supplementary Table 4). We found that Glis1 markedly increased the expression of several genes whose products have been shown to enhance iPSC generation. These included estrogen-related receptor, beta (*Esrrb*)[13], several Wnt ligands (*Wnt3*, *Wnt6*, *Wnt8a* and *Wnt10a*)[14], lin-28 homologue A (*Lin28a*)[15], Nanog (ref. 16), *Mycn* and *Mycl1*(ref. 17). In contrast, the expression of *Myc* was suppressed by Glis1 (Fig. 3c). We have previously shown that Mycn and Mycl1 predominantly increase the numbers of ES-cell-like colonies, whereas Myc increases both ES-cell-like and non-ES-cell-like colonies[17]. Therefore, the altered balance between Mycn/Mycl1 and Myc should contribute, at least in part, to the specific promotion of iPSC generation by Glis1. Glis1 also markedly enhanced the expression of forkhead box A2 (*Foxa2*), a transcription factor that antagonizes the epithelial-to-mesenchymal transition. Because this transition is a prerequisite for iPSC generation[18,19], the activation of *Foxa2* should also have a role in the promotion of iPSC generation by Glis1. We confirmed the effect of Glis1 on *Nanog*, *Mycn*, *Myc*, neurogranin and tetraspanin 18 in a p53 wild-type background by quantitative PCR (Supplementary Fig. 6). Taken together, these data demonstrate that Glis1 promotes iPSC generation by activating multiple pro-reprogramming pathways.

We next performed chromatin immunoprecipitation assays to identify the direct transcriptional targets of Glis1. Cell lysates were isolated from p53-knockout MEFs transduced with OSK plus mock or OSK plus Glis1. Candidate target genes identified from the microarray analyses were amplified by PCR (Fig. 4a). We found that significantly higher amounts of *Mycn*, *Mycl1* and *Myc* were precipitated from the cells

transduced with OSK plus Glis1 than from those transduced with OSK plus mock. In contrast, no such specific precipitation was observed with *Esrrb*, *Lin28a*, *Foxa2* or *Nanog*. These results indicate that the three *Myc* genes are direct targets of Glis1, whereas *Esrrb*, *Lin28a*, *Foxa2* and *Nanog* may be indirect targets.

We next examined whether Glis1 physically associates with the OSK proteins. Using Flag-tagged Glis1, we saw that Oct3/4 and Sox2 co-purified with Glis1 (Fig. 4b), whereas co-purification was not observed with a Flag-tagged Venus protein. In addition, we observed the co-purification of Flag–Klf4 with Myc-tagged Glis1 (Fig. 4b). The zinc-finger domain of Glis1 and its N-terminal region were required for the interaction with Klf4 (Supplementary Fig. 7). The interaction between Klf4 and Glis1 was further confirmed with an *in vitro* protein fragment complementation assay (Supplementary Fig. 8). These data indicate that Glis1 can associate with OSK by a protein–protein interaction and thereby might promote the activation of OSK target genes.

In contrast to oocytes and one-cell-stage embryos, we found that the expression of Glis1 was very low in ES cells. We therefore examined the effects of forced expression of Glis1 in mouse ES cells[20] and found that this suppressed their proliferation (Supplementary Fig. 9). This effect may have contributed to the smaller number of partially reprogrammed cells observed with OSK plus Glis1, because such cells would fail to silence retroviruses and would still express Glis1 transgenes, which would suppress proliferation.

This study shows that the transcription factor Glis1, which is highly enriched in unfertilized eggs and one-cell-stage embryos, promotes iPSC generation effectively and specifically by activating multiple pro-reprogramming pathways. Glis1 might thus be a

link between reprogramming during iPSC generation and reprogramming after nuclear transfer. Furthermore, iPSCs generated by OSK and Glis1 did not cause a marked increase in mortality of chimaeric mice, although this did occur with iPSCs generated by Oct3/4, Sox2, Glis1 and Myc (Supplementary Fig. 10) and with iPSCs generated by OSK and Myc, as reported previously[17]. The identification of Glis1 might therefore be beneficial for future applications of iPSC technology.

METHODS SUMMARY

To screen transcription factors for their effects on iPSC generation, cDNAs were used from the human proteome expression resource (HuPEX) library[21]. Gateway entry clones of 1,437 human transcription factors were transferred to pMXs-GW retroviral expression vectors using the Gateway LR reaction. MEFs were isolated from 13.5 days post coitum (d.p.c.) embryos and adult skin fibroblasts were isolated from 20-week-old mice. The generation of mouse iPSCs with retroviruses was performed as described previously[2,9]. Human iPSCs were also generated as described previously[22]. The shRNA-mediated knockdown was performed as described in ref. 12. Retroviruses (pMXs) were generated with Plat-E packaging cells[23]. ES cells and iPSCs were cultured on SNL feeder cells[24]. The analyses of iPSCs, such as RT–PCR, alkaline phosphatase staining, DNA microarrays, *in vitro* differentiation, teratoma formation, bisulphite genomic sequencing and chimaera experiments, were performed as previously described[1,9,22]. Animal experiments were approved by committees of Kyoto University and the Japan Science and Technology Agency. To examine whether Glis1 is physically associated with the OSK proteins, immunoprecipitation and immunoblotting analyses were performed, as well as an *in vitro* protein fragment complementation assay[25]. In addition, a ChIP analysis was performed on Glis1 to identify its target genes. Sequences of primers and shRNAs

are listed in Supplementary Tables 5 and 6, respectively. Microarray data are available through GEO with accession number GSE26431.

1. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126, 663-676 (2006).
2. Nakagawa, M. et al. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat Biotechnol 26, 101-106 (2008).
3. Yamanaka, S. A fresh look at iPS cells. Cell 137, 13-7 (2009).
4. Yamanaka, S. Strategies and new developments in the generation of patient-specific pluripotent stem cells. . Cell Stem Cell 1, 39-49 (2007).
5. Yamanaka, S. Elite and stochastic models for induced pluripotent stem cell generation. Nature 460, 49-52 (2009).
6. Yamanaka, S. & Blau, H. M. Nuclear reprogramming to a pluripotent state by three approaches. Nature 465, 704-12 (2010).
7. Wilmut, I., Schnieke, A. E., McWhir, J., Kind, A. J. & Campbell, K. H. Viable offspring derived from fetal and adult mammalian cells. Nature 385, 810-3. (1997).
8. Egli, D., Rosains, J., Birkhoff, G. & Eggan, K. Developmental reprogramming after chromosome transfer into mitotic mouse zygotes. Nature 447, 679-85 (2007).
9. Kim, Y. S. et al. Identification of Glis1, a novel Gli-related, Kruppel-like zinc finger protein containing transactivation and repressor functions. J Biol Chem 277, 30901-13 (2002).
10. Guo, G. et al. Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. Dev Cell 18, 675-85 (2010).
11. Hong, H. et al. Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. Nature 460, 1132-5 (2009).
12. Feng, B. et al. Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. Nat Cell Biol 11, 197-203 (2009).
13. Marson, A. et al. Wnt signaling promotes reprogramming of somatic cells to pluripotency. Cell Stem Cell 3, 132-5 (2008).
14. Yu, J. et al. Induced pluripotent stem cell lines derived from human somatic cells. Science 318, 1917-20 (2007).
15. Silva, J. et al. Nanog is the gateway to the pluripotent ground state. Cell 138, 722-37 (2009).
16. Nakagawa, M., Takizawa, N., Narita, M., Ichisaka, T. & Yamanaka, S. Promotion of direct reprogramming by transformation-deficient Myc. Proc Natl Acad Sci U S A 107, 14152-7 (2010).

17. Samavarchi-Tehrani, P. et al. Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. Cell Stem Cell 7, 64-77 (2010).
18. Li, R. et al. A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. Cell Stem Cell 7, 51-63 (2010).
19. Niwa, H., Burdon, T., Chambers, I. & Smith, A. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. Genes Dev 12, 2048-60. (1998).
20. Goshima, N. et al. Human protein factory for converting the transcriptome into an in vitro-expressed proteome. Nat Methods 5, 1011-7 (2008).
21. Okita, K., Ichisaka, T. & Yamanaka, S. Generation of germ-line competent induced pluripotent stem cells. Nature 448, 313-7 (2007).
22. Takahashi, K. et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131, 861-72 (2007).
23. Morita, S., Kojima, T. & Kitamura, T. Plat-E: an efficient and stable system for transient packaging of retroviruses. Gene Ther 7, 1063-6 (2000).
24. McMahon, A. P. & Bradley, A. The Wnt-1 (int-1) proto-oncogene is required for development of a large region of the mouse brain. Cell 62, 1073-85. (1990).
25. Hashimoto, J. et al. Novel in vitro protein fragment complementation assay applicable to high-throughput screening in a 1536-well format. J Biomol Screen 14, 970-9 (2009).

Acknowledgements We thank T. Yamamoto, Y. Yamada and the members of our laboratory for valuable scientific discussions and administrative support. We thank M. Nakagawa, H. Seki, M. Murakami, A. Okada, M. Narita, M. Inoue, H. Shiga and T. Matsumoto for technical assistance ;and H. Suemori (Kyoto University) for human ESC. This work was supported in part by grants from the New Energy and Industrial Technology Development Organization (NEDO), the Leading Project of the Ministry of Education, Culture, Sports, Science and Technology (MEXT), the Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST Program) of the Japanese Society for the Promotion of Science (JSPS), Grants-in-Aid for Scientific Research from JSPS and MEXT, and the Program for Promotion of Fundamental Studies in Health Sciences of the National Institute of Biomedical Innovation (NIBIO). S.Y. is a member of scientific advisory boards of iPearian Inc. and iPS Academia Japan.

Author Contributions M.M. conducted most of the experiments in this study. K.Y. analysed the interactions of proteins. T.N. performed the computer analyses of the DNA microarray data, teratoma experiments, overexpression in ES cells and statistical analysis. R.S. generated mouse iPSCs and characterized mouse and human iPSCs. I.K. generated human iPSCs. T.I. performed the chimaera and teratoma experiments and maintained the mouse lines. Y.K. selected cDNA clones from HuPEX with bioinformatics. H.M. produced the retroviral expression clones. N.G. and S.Y. supervised the project. M.M. and S.Y. wrote the manuscript.

Author Information The microarray data are available from the Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) with the accession number GSE26431. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to S.Y. (yamanaka@cira.kyoto-u.ac.jp) and N.G. (n-goshima@aist.go.jp).

Figure 1 Promotion of mouse iPSC generation by GLIS1. a, Number of *Nanog*–GFP-positive colonies from mouse skin fibroblasts in a 100-mm dish, 28 d after infection. Three days after infection, fibroblasts were re-seeded on feeder cells. Exp, experiment. **b**, Number of *Nanog*–GFP-positive colonies from mouse skin fibroblasts in a 6-well plate, 22 d after infection. **c**, Proportion of *Nanog*–GFP-positive colonies to total number of colonies. Fig. 1c is derived directly from the experiments in 1b. **, *P*-values <0.01. Error bars, s.d.; $n = 3$. **d**, Upper panel: chimaeric mouse derived from iPSCs obtained by transfection of MEFs with OSK + GLIS1. Lower panel: coat colour of offspring, showing germline transmission.

Figure 2 Promotion of human iPSC generation by GLIS1. a, Number of ES-cell-like colonies from human dermal fibroblasts 30 d after infection. **b**, ES-cell-like colonies as a proportion of total colonies. **, $P < 0.01$ compared to cells expressing OSK alone. Error bars, s.d.; $n = 3$. **c**, Scatter plots comparing global gene expression patterns between iPSCs generated by OSK + GLIS1 and adult dermal fibroblasts (AHDF) (left panel), and between iPSCs from OSK + GLIS1 and iPSCs from OSKM (right panel). The green diagonal lines indicate twofold changes between the two samples. The correlation coefficient (R^2) is also shown. **d**, iPSCs generated by OSK + GLIS1 were subcutaneously transplanted into nude mice. Teratomas were analysed histologically with haematoxylin and eosin staining.

Figure 3 Characterization of Glis1: expression and roles during iPSC generation. a, Expression patterns of *Glis1* in different mouse tissues. Data are normalized to glyceraldehyde-3-phosphate dehydrogenase expression; *Glis1* expression in the kidney is set at a relative level of 1. Error bars, s.d.; $n = 4$. **b**, Ninety genes were found to be upregulated more than 20-fold in OSK + Glis1 cells compared to OSK + mock cells (upper panel). These included *Foxa2*, multiple Wnt-family genes and *Esrrb*. We also focused on 361 probes for which expression was more than 100-fold higher in ES cells than in fibroblasts. Among these, 32 probes showed an expression level that was more than three-fold higher in OSK + Glis1 cells than in OSK + mock cells (lower panel). These included *Esrrb, Oct3/4, Mycn, Lin28a* and *Nanog*. **c**, Expression levels of the Myc-family genes (C, *Myc*; N, *Mycn*; L, *Mycl1*) in OSK + Mock and OSK + Glis1 cells. The green diagonal lines indicate twofold changes between the two cell types.

Figure 4 Characterization of Glis1: target genes and protein–protein interactions. a, Chromatin immunoprecipitation and quantitative PCR analysis were conducted on the

basis of microarray data, using a Glis1-specific antibody and PCR primers specific for *Mycn*, *Mycl1*, *Myc*, *Nanog*, *Esrrb*, *Lin28a* and *Foxa2*. GATA binding protein 4 (*Gata4*) and NK2 transcription factor related, locus 5 (*Nkx2-5*) were used as negative controls. IP, immunoprecipitate. Error bars, s.d.; $n = 2$.*, P-values <0.05. **b**, Constructs encoding Flag-tagged Glis1 or Klf4 and untagged Oct3/4 (left panel), Flag-tagged Glis1 or Klf4 and untagged Sox2 (middle panel) or Flag-tagged Klf4 and Myc-tagged Glis1 (right panel) were transfected into HEK293T cells alone or in combination. Flag-tagged Venus was transfected as a negative control. The cell lysates were immunoprecipitated (IP) with an anti-Flag antibody, followed by an immunoblot analysis (IB). The expression levels in whole-cell lysates were determined by IB (bottom panels).

METHODS

cDNA library

cDNAs used to screen for novel factors that alter the efficacy of iPSC generation were obtained from the human proteome expression resource (HuPEX) library[21]. Among the 33,275 cDNAs, we selected those known to be transcription factors or identified by keyword searches of the Human Gene and Protein Database (HGPD, http://www.HGPD.jp/) and Entrez gene (http://www.ncbi.nlm.nih.gov/gene). We used cDNAs that covered more than 80% of the open reading frame reported in RefSeq and had identity with the reported protein sequence of more than 95% at the amino acid level. cDNAs encoding OCT3/4, SOX2, KLF4 or MYC were excluded. This resulted in 1,437 cDNAs (Supplementary Table 1), which were transferred to the pMXs-GW retroviral expression vector using the Gateway LR reaction.

Cell culture

Mouse iPSCs were maintained in ES cell medium (DMEM containing 15% fetal calf serum (FCS), 1× Non-Essential Amino Acids (NEAA), 1 mM sodium pyruvate, 5.5 mM 2-Mercaptoethanol (ME), 50 units ml^{-1} penicillin and 50 μg ml^{-1} streptomycin) on feeder layers of mitomycin-C-treated SNL cells stably expressing the puromycin-resistance gene[24]. As a source of leukaemia-inhibitory factor (LIF), we used the conditioned medium from Plat-E cell cultures that had been transduced with a LIF-expressing vector. Human iPSCs were generated and maintained in primate ES cell medium (ReproCELL), supplemented with 4 ng ml^{-1} recombinant human basic fibroblast growth factor, 50 units ml^{-1} penicillin and 50 μg ml^{-1} streptomycin. MEFs, mouse skin fibroblasts and human fibroblasts were maintained in DMEM containing 10% FCS, 50 units ml^{-1} penicillin and 50 μg ml^{-1} streptomycin. Plat-E cells[23] were maintained in DMEM containing 10% FCS, 50 units ml^{-1} penicillin, 50 μg ml^{-1} streptomycin, 1 μg ml^{-1} puromycin and 10 μg ml^{-1} blasticidin S. We used 13.5 d.p.c. embryos for MEF isolation and 20-week-old mice for the isolation of skin fibroblasts.

Mouse iPSC generation

The generation of mouse iPSCs with retroviruses was performed as previously described[2,9] with some modifications. Briefly, Plat-E cells were seeded at 2.5×10^6 cells per 100-mm dish. On the next day, pMXs-based retroviral vectors for each gene were independently introduced into Plat-E cells using the FuGENE 6 transfection reagent. After 24 h, the medium was replaced with 10 ml of DMEM containing 10% FCS. Fibroblasts were seeded at 8×10^5 cells per dish, in 100-mm dishes covered with a layer of gelatin or feeder cells. The next day, virus-containing supernatants from the Plat-E cultures were recovered and mixed, for example OCT3/4, SOX2, KLF4, and GLIS1.

Fibroblasts were incubated in the virus/polybrene-containing supernatants at a final concentration of 4 μg ml^{-1} for 24 h. Three days after infection, the medium was changed to ES cell medium supplemented with LIF. Fibroblasts on gelatin-coated dishes were then re-seeded onto dishes with feeder cells. The shRNA-mediated knockdown was performed as previously described[12].

Generation of human iPSCs

Human iPSCs were generated as previously described[22] with some modifications. Briefly, Plat-E cells were plated at 3.6×10^6 cells per 100-mm dish. The next day, pMXs-based retroviral vectors for each gene were independently introduced into the Plat-E cells using the FuGENE 6 transfection reagent. After 24 h, the medium was replaced with new medium. Human fibroblasts expressing the mouse *Slc7a1* (solute carrier family 7 (cationic amino acid transporter, y+ system), member 1) gene were seeded at 8×10^5 cells per 100-mm dish. The next day, virus-containing supernatants were recovered and mixed, for example OCT3/4, SOX2, KLF4, and GLIS1. Fibroblasts were incubated in the virus/polybrene-containing supernatants at a final concentration of 4 μg ml^{-1} for 24 h. Six days after transduction, fibroblasts were harvested by trypsinization and replated at 5×10^4 or 5×10^5 cells per 100-mm dish on SNL feeder cells. The next day, the medium was replaced with primate ES cell medium supplemented with 4 ng ml^{-1} basic fibroblast growth factor.

Characterization of iPSCs

The RT–PCR analyses, alkaline phosphatase staining, *in vitro* differentiation, teratoma formation, bisulphite genomic sequencing and chimaera experiments were performed as previously described[1,9,22]. The primers used for RT–PCR are listed in Supplementary

Table 5. In the *in vitro* differentiation assay, differentiated cells were stained positive for α-fetoprotein (endoderm), α-smooth muscle actin (mesoderm) and nestin (ectoderm). Nuclei were stained with Hoechst. For bisulphite genomic sequencing, the white circles indicate unmethylated CpG dinucleotides, whereas the black circles indicate methylated CpG dinucleotides.

DNA microarray

Total RNAs were labelled with Cy3 and hybridized to either a Whole Mouse Genome Microarray or a Whole Human Genome Microarray (Agilent) according to the manufacturer's protocol. Arrays were scanned using the G2505C Microarray Scanner System (Agilent). The data were analysed using the GeneSpring GX11.0.1 software program (Agilent). The microarray data are available from the Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) with the accession number GSE26431.

Chromatin immunoprecipitation assay

We used the Active Motif ChIP-IT Express kit for the chromatin immunoprecipitation assay. Genomic DNA and nuclear proteins were fixed with formaldehyde. Immunoprecipitation was performed with either anti-Glis1 (Santa Cruz) or purified goat IgG antibody and the elutes were used as templates for quantitative PCR. We selected DNA fragments containing putative Glis1-binding sites for PCR amplification. The primers used for quantitative PCR in the ChIP assay are listed in Supplementary Table 5.

Immunoprecipitation and immunoblotting analyses

Because the expression levels of *Glis1* in ES cells and fibroblasts are low, we were not able to elucidate whether there was an association among the endogenous proteins. HEK293T cells were therefore transfected with each cDNA clone in an expression vector

and were lysed in CytoBuster (Novagen). Cell lysates were incubated with an anti-Flag M2 Affinity Gel (Sigma) for 2 h and then removed. The gel suspensions were boiled in sample buffer and analysed by SDS–polyacrylamide gel electrophoresis and immunoblotting. The immunoblot analyses were performed using the following antibodies: anti-Flag M2 (Sigma), anti-Myc (Roche), anti-Oct3/4 (Santa Cruz) and anti-Sox2 (MBL).

In vitro protein fragment complementation assay

We prepared split monomeric Kusabira-Green protein (mKG) fragment proteins (Amalgaam) fused to Glis1 and Klf4 using a wheat-germ cell-free protein synthesis system (CellFree Sciences)[25]. Each protein solution was dispensed into a 384-well plate. After incubation at25℃ for 8h or 23h, the fluorescence was measured using the Typhoon 9200 (GE Healthcare).

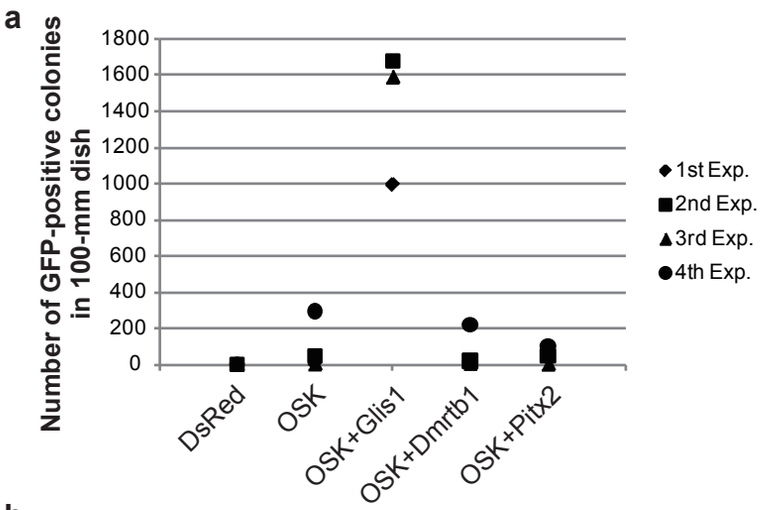
Overexpression of genes in ES cells

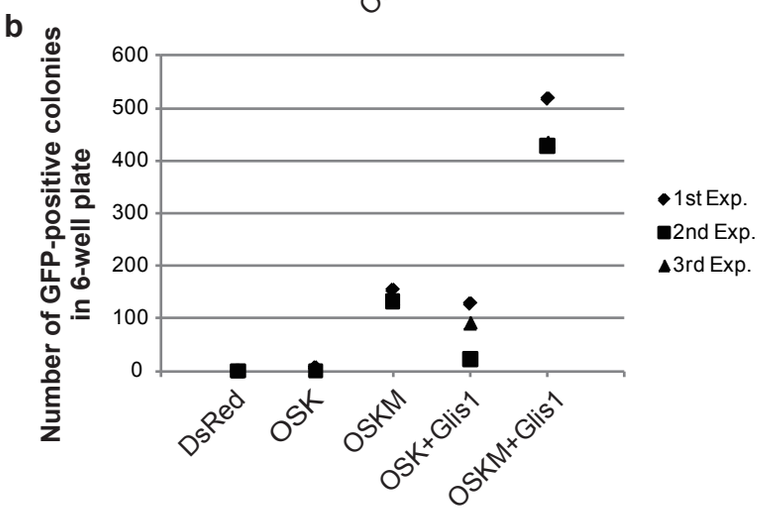
The mouse ES cell line MG1.19 was maintained in DMEM containing 10% FCS, 1× NEAA, 1 mM sodium pyruvate, 5.5 mM 2-ME, 50 units ml^{-1} penicillin, 50 μg ml^{-1} streptomycin and LIF. The vectors pCAG-IP (Mock) or pCAG-Glis1-IP were introduced into MG1.19 cells using Lipofectamine 2000 on day −1. On day zero, 1×10^5 cells were re-seeded on a gelatin-coated 6-well plate. On day 4, the cell number was counted.

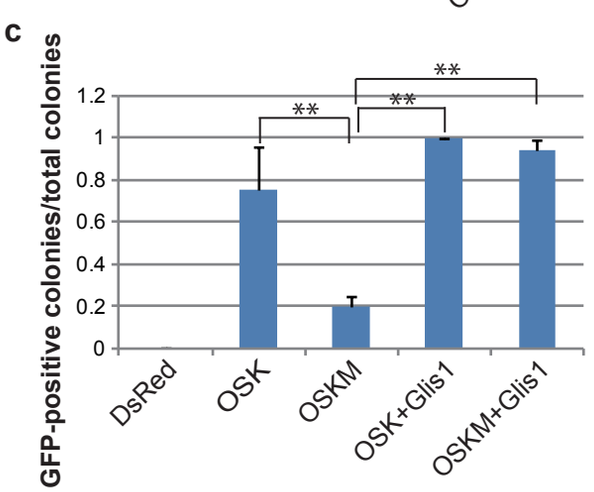
Statistical analyses

A one-way repeated-measures ANOVA and a post-hoc Bonferroni test were used for the analyses of the data in Figs 1c and 2b. The unpaired *t*-test was used for statistical analysis of the data shown in Fig. 4a (between OSK and OSKGlis1). Differences were considered to be statistically significant for *P*-values <0.05 (*), <0.01 (**) or <0.001 (***).

Figure 1


a
Number of GFP-positive colonies in 100-mm dish
1800
1600
1400
1200
1000
800
600
400
200
0
DsRed
OSK
OSK+Glis1
OSK+Dmrtb1
OSK+Pitx2
1st Exp.
2nd Exp.
3rd Exp.
4th Exp.


b
Number of GFP-positive colonies in 6-well plate
600
500
400
300
200
100
0
DsRed
OSK
OSKM
OSK+Glis1
OSKM+Glis1
1st Exp.
2nd Exp.
3rd Exp.


c
GFP-positive colonies/total colonies
1.2
1
0.8
0.6
0.4
0.2
0
**
**
**
DsRed
OSK
OSKM
OSK+Glis1
OSKM+Glis1

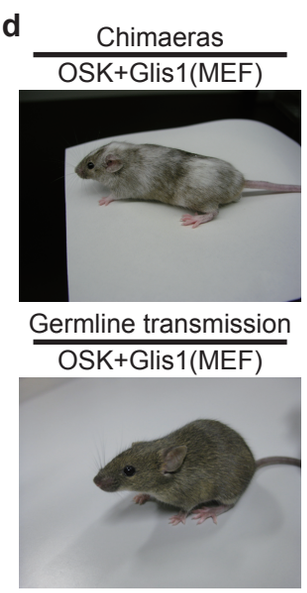
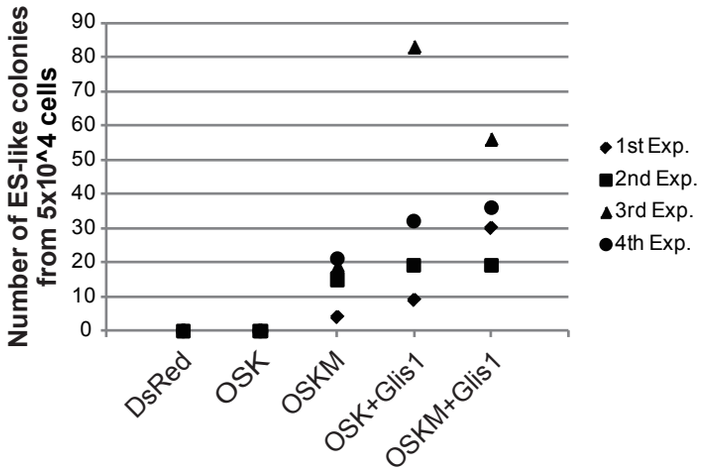
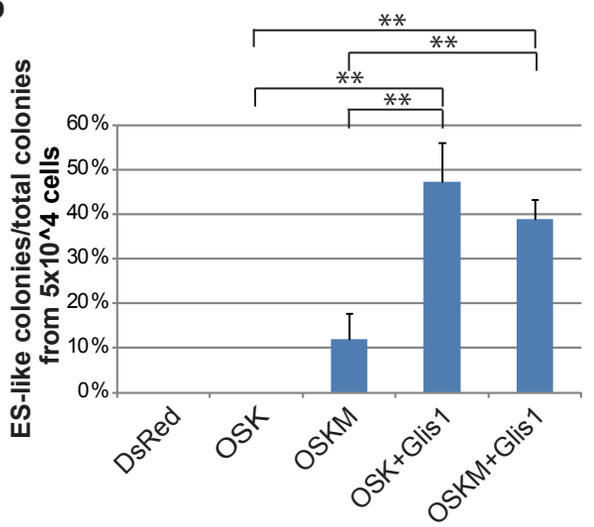

d
Chimaeras
OSK+Glis1(MEF)
Germline transmission
OSK+Glis1(MEF)

Figure 2

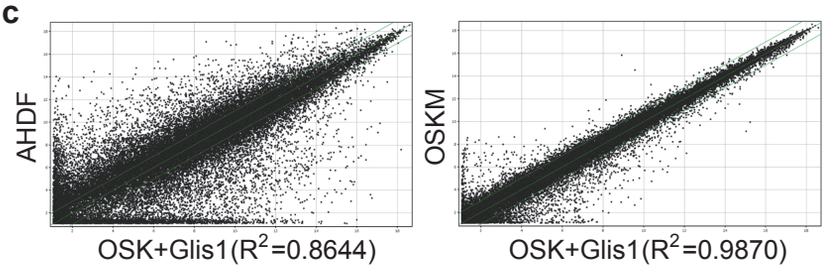
a



b



c



d

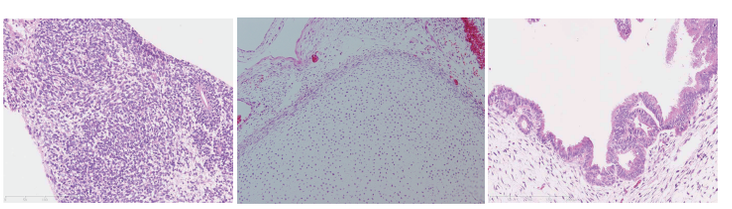
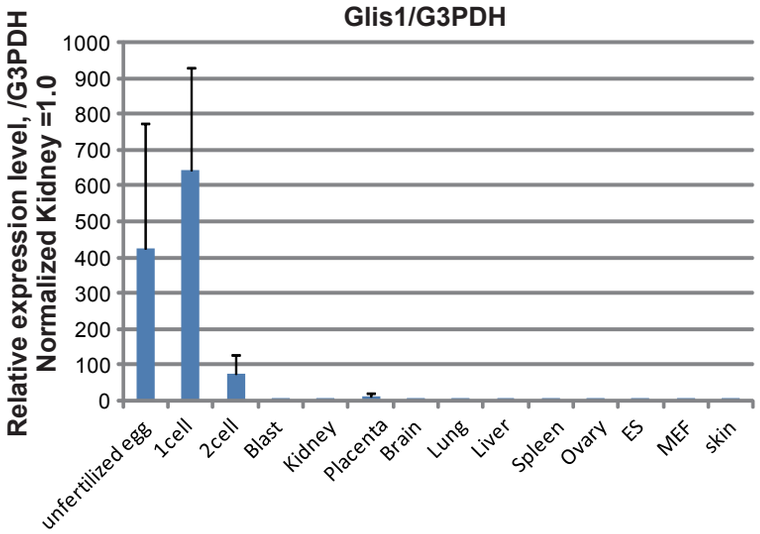
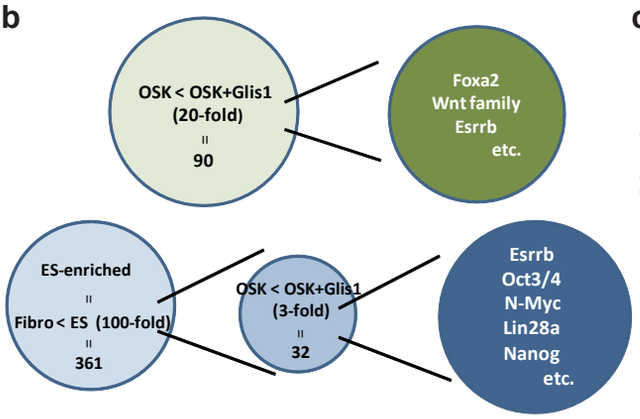


Figure 3

a



b



c Myc family

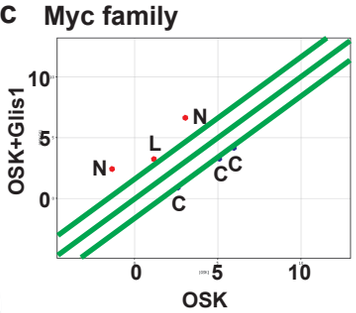
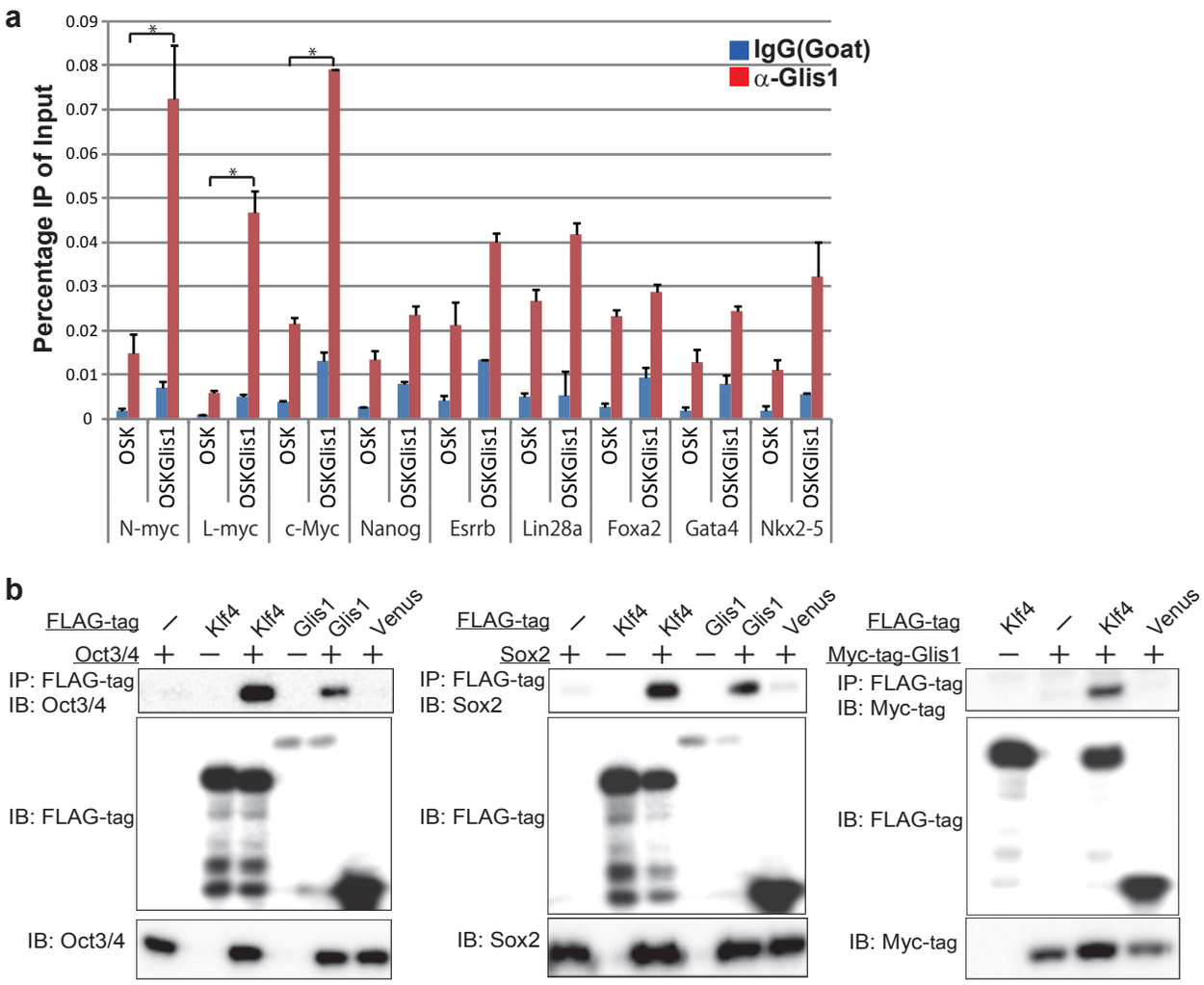
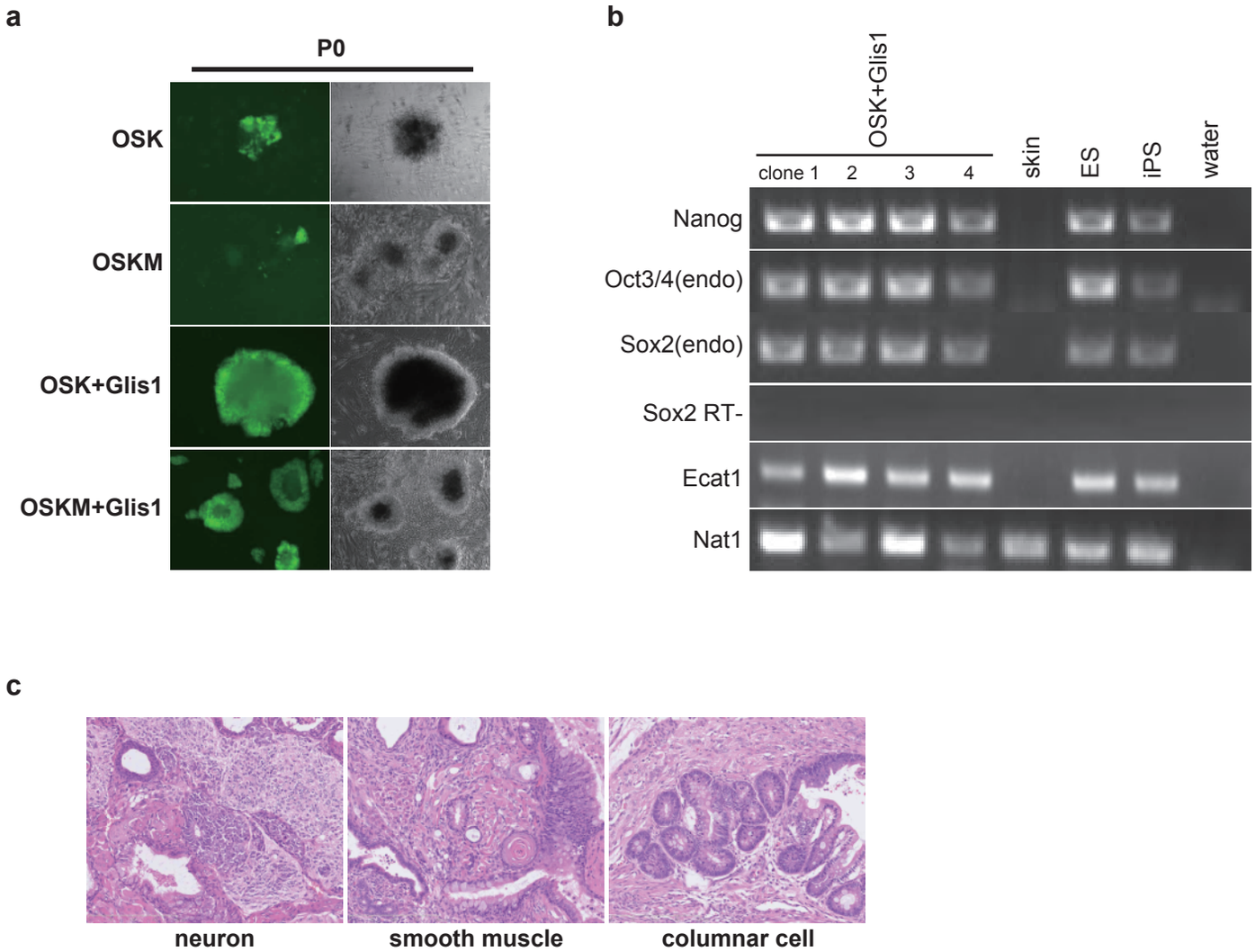


Figure 4

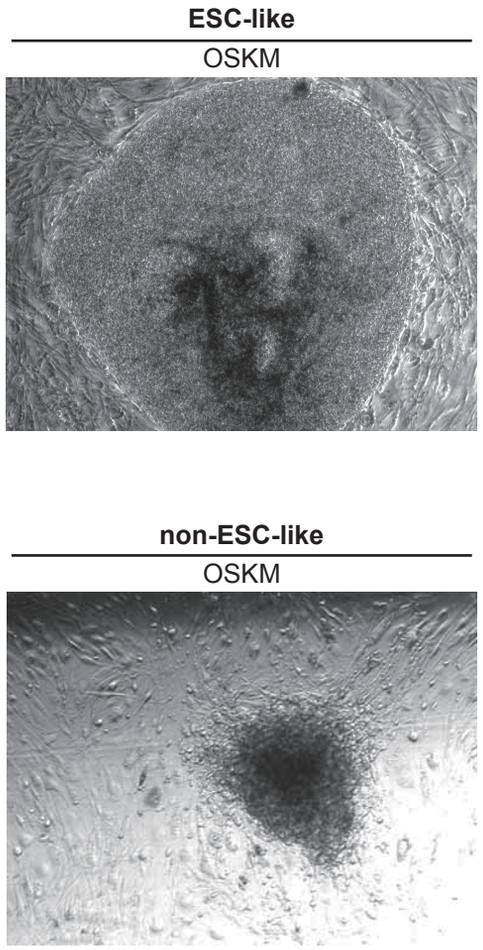

a
Percentage IP of Input
IgG(Goat)
α-Glis1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
*
*
*
OSK
OSKGlis1
N-myc
L-myc
c-Myc
Nanog
Esrrb
Lin28a
Foxa2
Gata4
Nkx2-5
b
FLAG-tag
Oct3/4
Sox2
Myc-tag-Glis1
Klf4
Glis1
Venus
IP: FLAG-tag
IB: Oct3/4
IB: Sox2
IB: Myc-tag
IB: FLAG-tag

Supplementary Figure 1



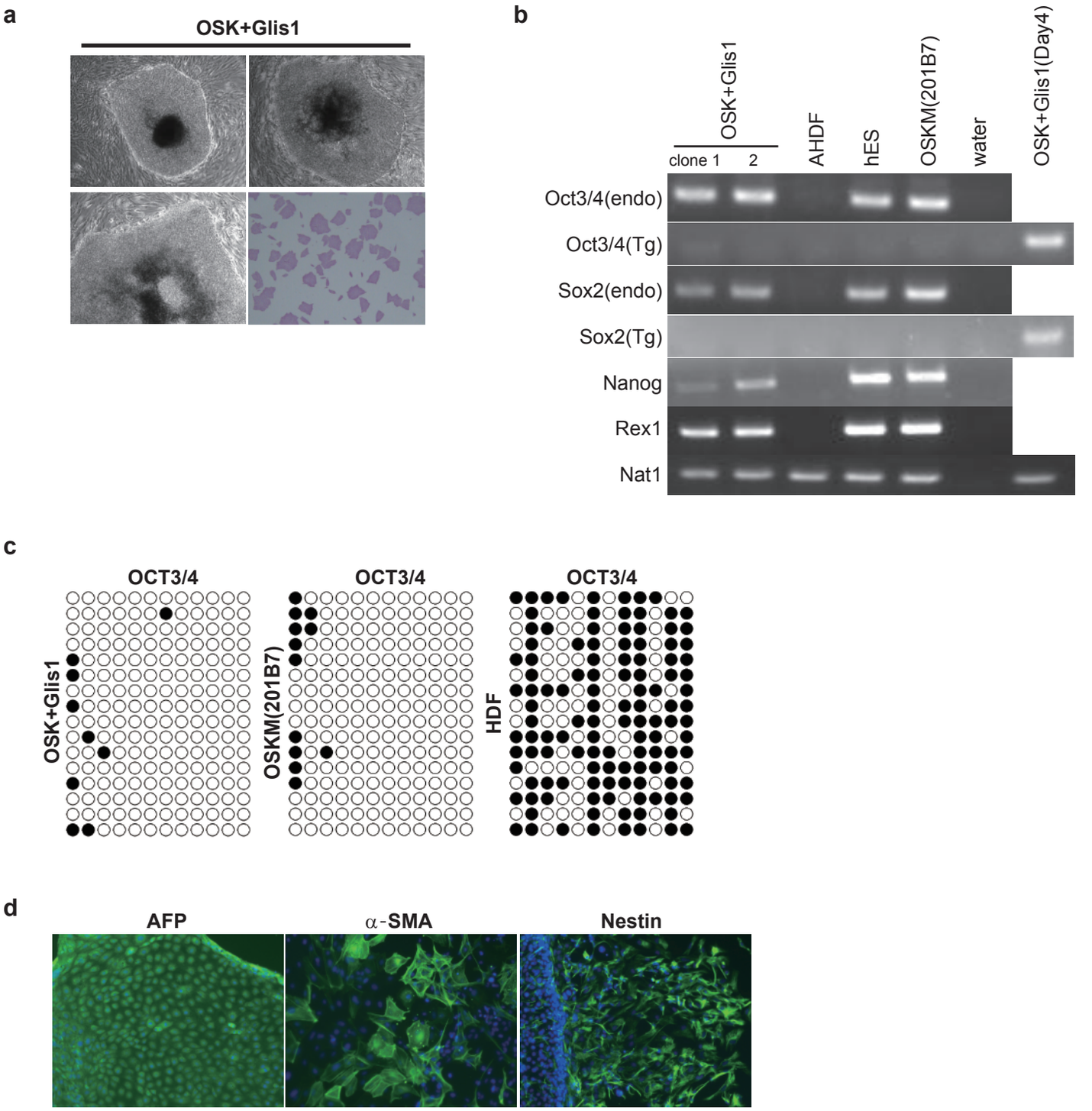
Supplementary Figure 1. (a) Phase contrast and fluorescent images of Nanog-GFP-positive colonies from mouse skin fibroblasts (P0; passage 0). (b) RT-PCR analyses of ESC-marker genes. (c) Teratoma formation from OSK+Glis1-iPSC.

Supplementary Figure 2



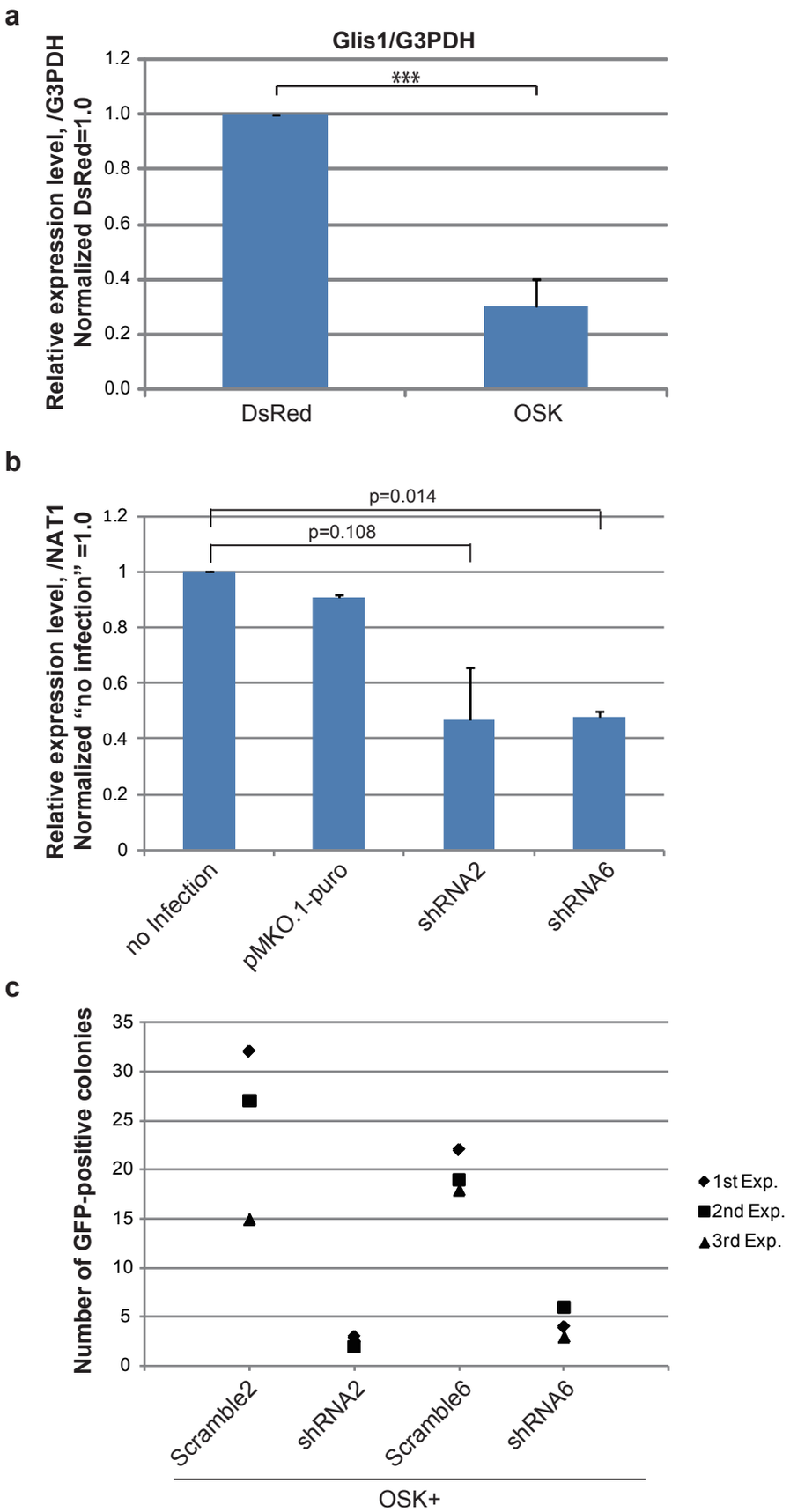
Supplementary Figure 2. ESC-like colonie with a flat, round shape and a distinct edge (upper), and non-ESC-like colonie, which were granulous with an irregular edge (lower).

Supplementary Figure 3



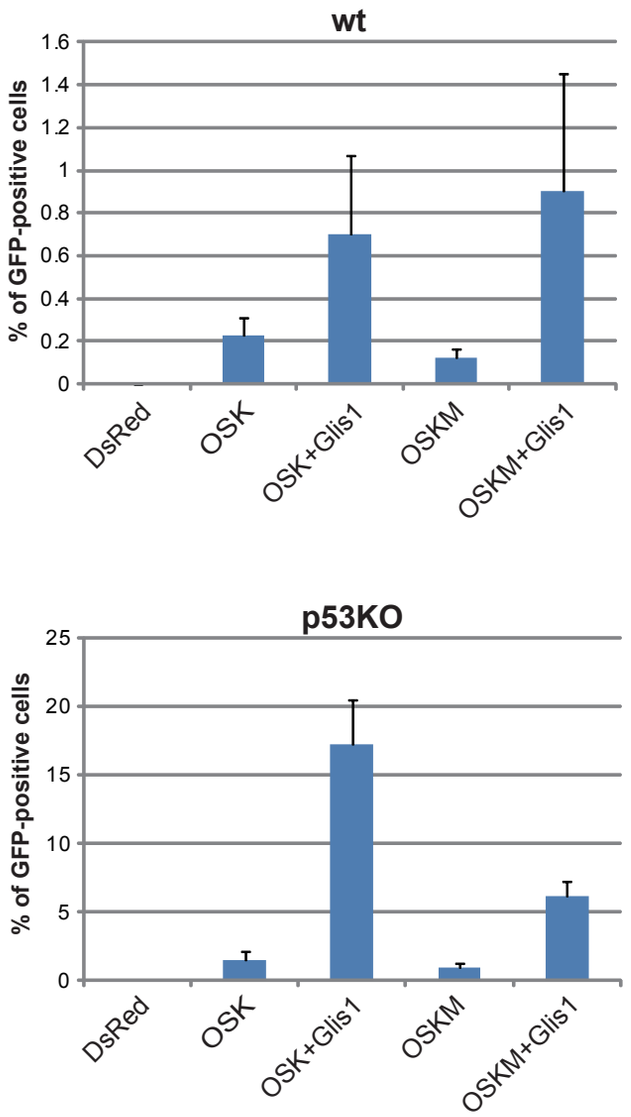
Supplementary Figure 3. (a) Phase contrast images and the results of alkaline phosphatase staining. (b) RT-PCR analyses of ESC-marker genes. (c) Bisulfite genomic sequencing of the promoter region of Oct3/4. (d) Embryoid body-mediated in vitro differentiation of OSK+Glis1-iPSC.

Supplementary Figure 4



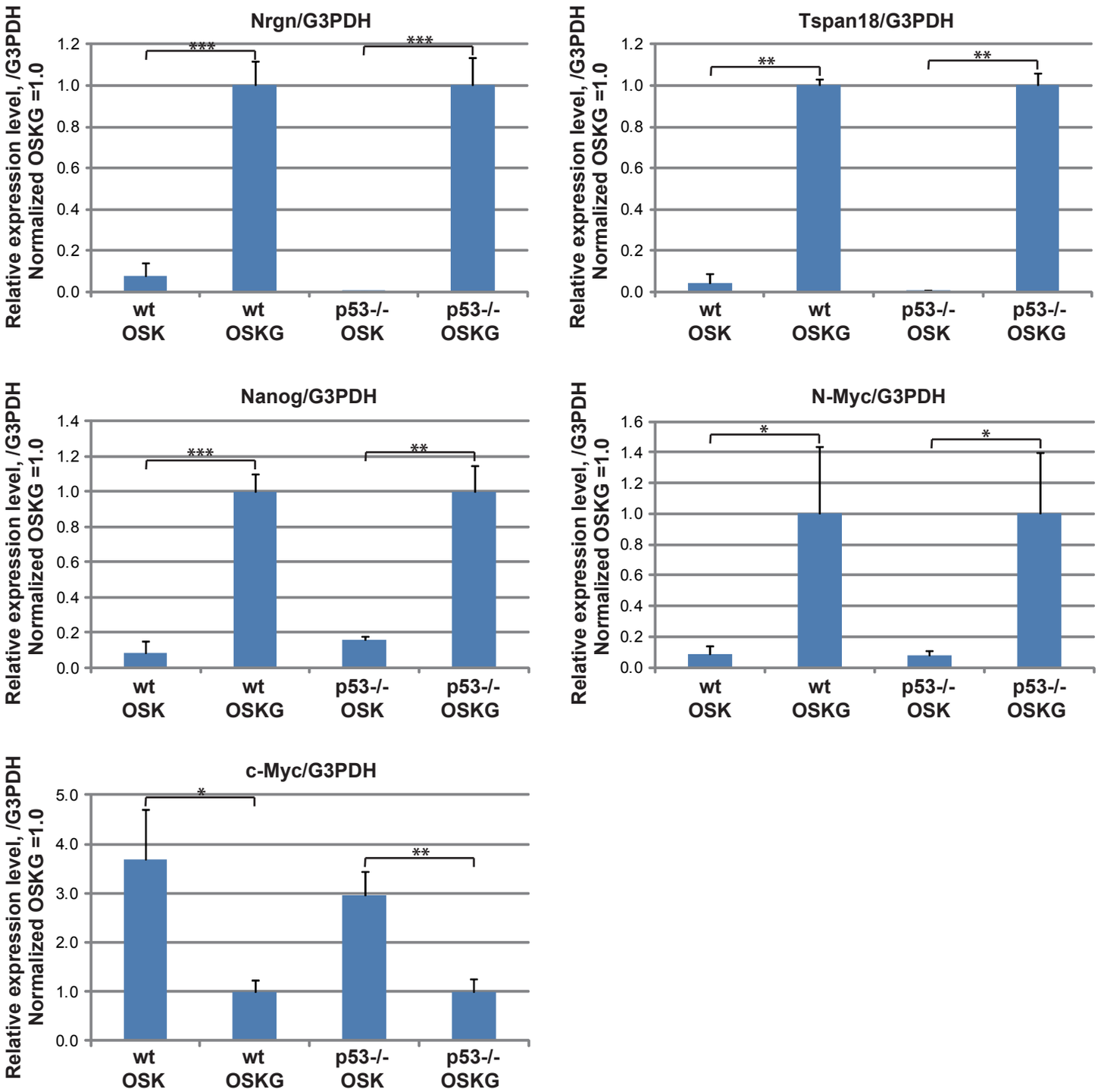
Supplementary Figure 4. (a) Glis1 expression levels in MEF three days after transfection with DsRed or OSK. The unpaired t-test was used for the statistical analyses. N=4. Error bars, s.d. (b) The quantitative RT-PCR analyses of endogenous Glis1 mRNA levels in skin fibroblasts exposed to Glis1 shRNAs. A paired t-test was used for the statistical analyses. N=2. Error bars, s.d. (c) Each of shRNAs or scrambled shRNAs was co-transfected with OSK into MEF. Three days after infection, the fibroblasts were reseeded on feeder cells (5,000 cells per 100-mm dish). About three weeks after transduction, the numbers of Nanog-GFP-positive colonies were counted. shRNA2 and shRNA6 significantly decreased the number of GFP-positive colonies. The actual values of three independent experiments are shown (1st, 2nd, and 3rd).

Supplementary Figure 5



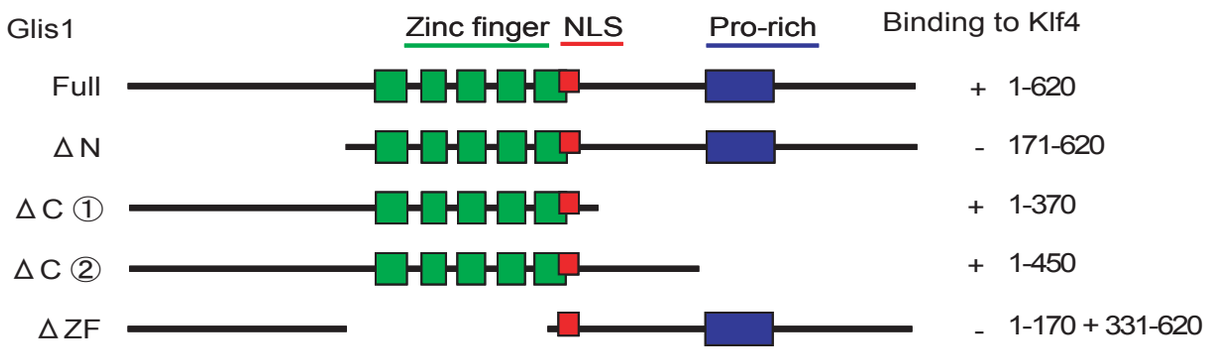
Supplementary Figure 5. Percentage of Nanog-GFP-positive cells from wt MEF or p53KO MEF five days after transduction with indicated factors. N=4. Error bars, s.d.

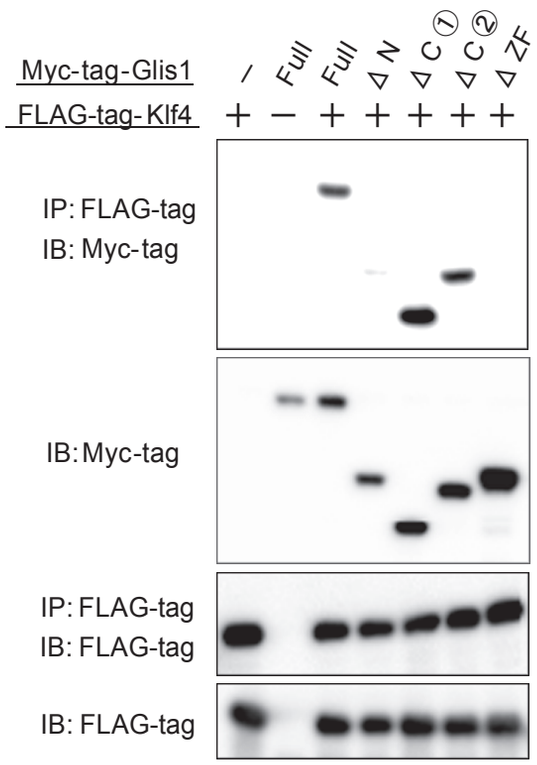
Supplementary Figure 6



Supplementary Figure 6. The expression levels of factors identified from microarray analysis. Nanog-GFP-positive cells were sorted from OSK or OSK+Glis1-transduced wt or p53KO MEF five days after infection. Expression levels of factors from microarray analysis were analyzed. These factors showed similar expression pattern between wt and p53KO MEF. The unpaired t-test was used for the statistical analyses. N=3 for Nrgn, Nanog, N-Myc, and c-Myc. N=2 for Tspan18. Error bars, s.d.

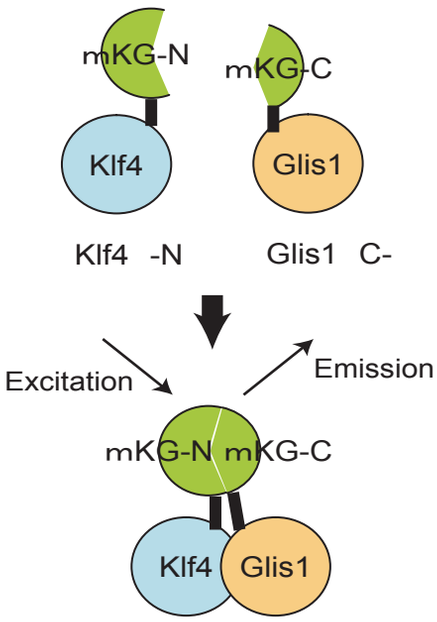
Supplementary Figure 7

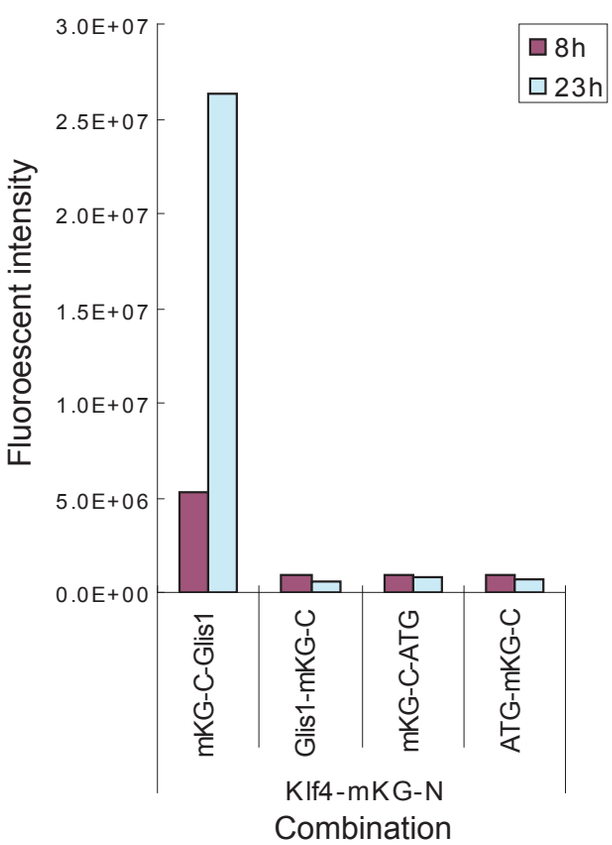




Supplementary 7. A schematic diagram to illustrate various Glis1 deletion mutants (upper). The zinc-finger domain and its N-terminal region of Glis1 interact with Klf4 when expressed in HEK293T cells (lower). Constructs encoding FLAG-tagged Klf4 and Myc-tagged Glis1 deletion mutant were transfected into 293T cells. The cell lysates were immunoprecipitated (IP) with anti-FLAG antibody, followed by an immunoblot analysis (IB). The expression level of whole cell lysates was determined by IB.

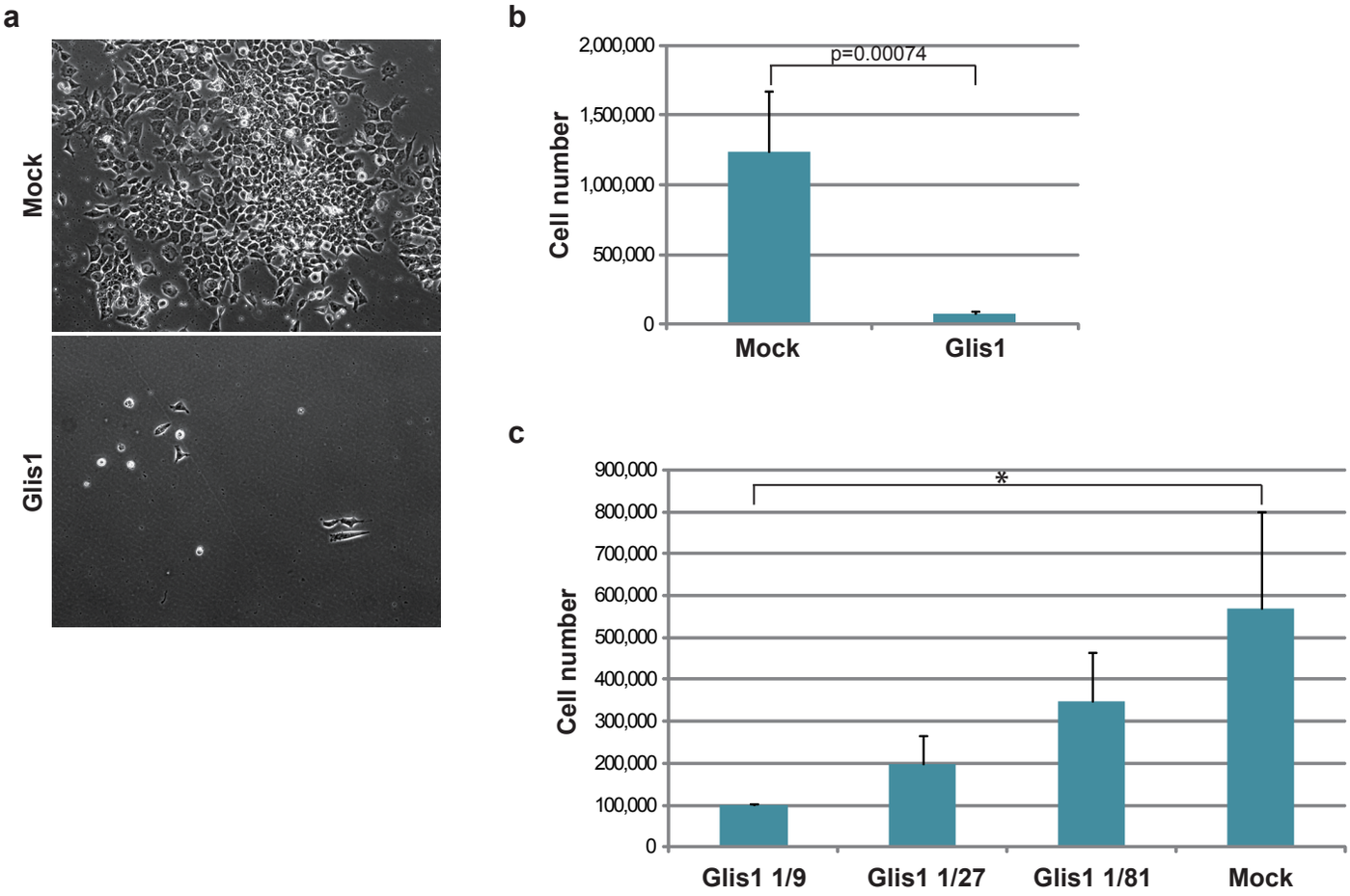
Supplementary Figure 8





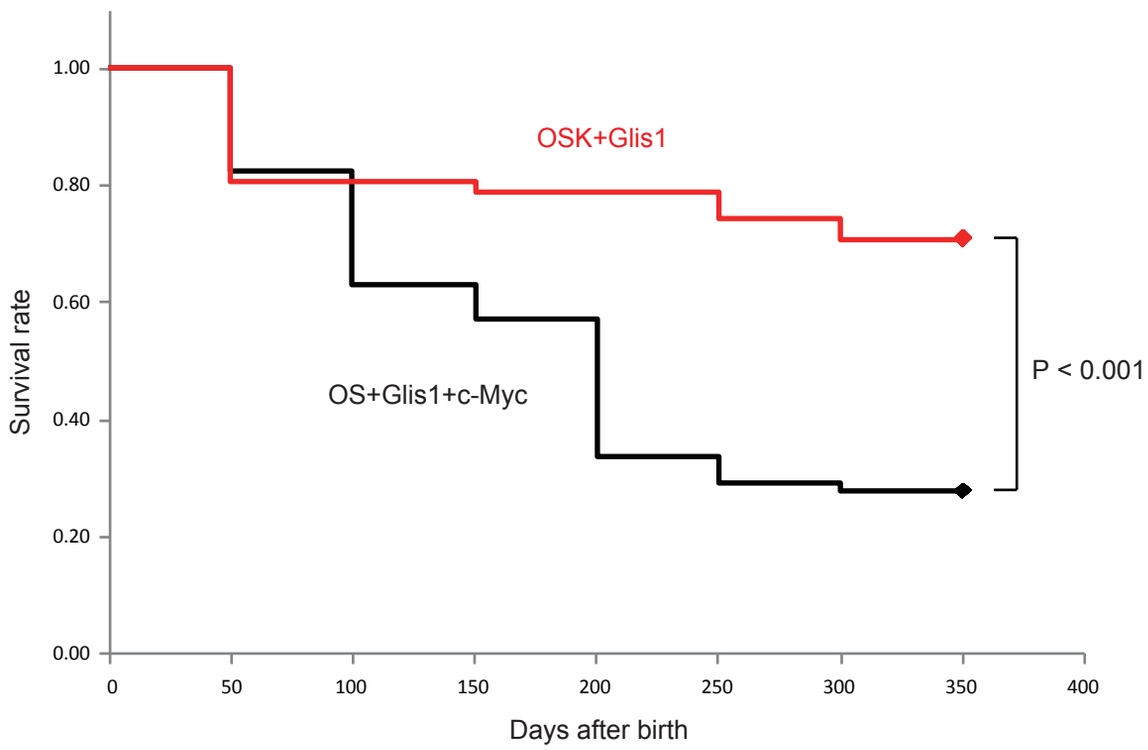
Supplementary 8. Outline of in vitro protein fragment complementation assay (PCA) with mKG (upper). Shown in the lower panel are fluorescent emissions of mKG combinations. N-terminal mKG (mKG-N)-fused Klf4 protein was combined with either C-terminal mKG proteins (mKG-C)-Glis1 fusion, Glis1-mKG-G fusion, or two negative controls (mKG-C-ATG or ATG-mKG-C).

Supplementary Figure 9



Supplementary Figure 9. We utilized the episomal expression system which allows the high and sustained expression of foreign genes in MG1.19 ESC. The overexpression of Glis1 in mouse ESC resulted in growth arrest or cell death. (a) The images of Mock or Glis1-introduced ESC on Day 4. (b) The graph shows number of Mock or Glis1-introduced ESC on Day 4, mean of five independent experiments and the unpaired t-test was used for the statistical analyses. Error bars, s.d. (c) Dilution of Glis1 plasmid resulted in increase of cell number. The graph shows the mean of three independent experiments and a one-way ANOVA test and a post-hoc Bonferroni test were used. Error bars, s.d.

Supplementary Figure 10



Supplementary Figure 10. Kaplan-Meier survival analysis showing survival rate of chimeric mice, which were derived from iPSC generated with OSK+Glis1 (red) or OS+Glis1+c-Myc (black). N=61 for OSK+Glis1. N=64 for OS+Glis1+c-Myc.

Supplementary Table S1. List of 1,437 human transcription factors which were selected from HuPEX.

Gene Symbol	RefSeq protein ID	Gene Symbol	RefSeq protein ID	Gene Symbol	RefSeq protein ID	Gene Symbol	RefSeq protein ID
AATF	NP_036270.1	CCNT2	NP_490595.1	DRAP1	NP_006433.2	FUBP1	NP_003893.2
ABRA	NP_631905.1	CD80	NP_005182.1	DYRK1B	NP_004705.1	GABPA	NP_002031.2
ABT1	NP_037507.1	CD86	NP_008820.2	E2F1	NP_005216.1	GABPB2	NP_005245.2
ACD	NP_075065.2	CD86	NP_787058.3	E2F2	NP_004082.1	GAS7	NP_003635.2
ADNP	NP_056154.1	CDC40	NP_056975.1	E2F4	NP_001941.2	GATA2	NP_116027.2
AEBP2	NP_694939.2	CDC5L	NP_001244.1	EAF2	NP_060926.2	GATAD1	NP_066990.3
AES	NP_001121.2	CDCA7L	NP_001120842.1	EBF1	NP_076870.1	GATAD1	NP_066990.3
AES	NP_945320.1	CDCA7L	NP_001120843.1	ECD	NP_009196.1	GATAD2A	NP_060130.3
AFAP1L2	NP_115939.1	CDH1	NP_004351.1	ECOP	NP_110423.3	GATAD2B	NP_065750.1
AHR	NP_001612.1	CDK7	NP_001790.1	ECOP	NP_110423.3	GCN5L2	NP_066564.2
ALX1	NP_008913.2	CEBPD	NP_005186.2	EED	NP_694536.1	GFI1	NP_005254.2
ALX4	NP_068745.2	CEBPE	NP_001796.2	EGR2	NP_000390.2	GFI1B	NP_004179.2
ANKRD1	NP_055206.2	CEBPG	NP_001797.1	EGR2	NP_000390.2	GLI4	NP_612474.1
ANKRD2	NP_065082.2	CEBPZ	NP_005751.2	EGR3	NP_004421.2	GLIS1	NP_671726.1
ANKRD49	NP_060174.2	CENPB	NP_001801.1	EGR4	NP_001956.2	GLIS2	NP_115964.2
ANKS1B	NP_064525.1	CHD4	NP_001264.2	EHF	NP_036285.2	GLIS3	NP_689842.3
APEX1	NP_001632.2	CHRAC1	NP_059140.1	EID1	NP_055150.1	GLP-1	NP_001096637.1
APLP1	NP_005157.1	CHURC1	NP_660148.2	EID1	NP_055150.1	GMCL1	NP_848526.1
APLP1	NP_001019978.1	CIAO1	NP_004795.1	EID2B	NP_689574.1	GMEB1	NP_006573.2
ARNT	NP_001659.1	CIAO1	NP_004795.1	ELF1	NP_758961.1	GMEB1	NP_077808.1
ARNT	NP_848514.1	CIR	NP_004873.3	ELF2	NP_006865.1	GMEB1	NP_077808.1
ARNT2	NP_055677.3	CITED1	NP_004134.1	ELF3	NP_001107781.1	GMEB2	NP_036516.1
ARNTL	NP_001025443.1	CITED2	NP_006070.2	ELF5	NP_001413.1	GRHL1	NP_055367.2
ARNTL	NP_001025444.1	CNBP	NP_003409.1	ELK1	NP_001107595.1	GRHL2	NP_079191.1
ARNTL2	NP_064568.3	CNBP	NP_003409.1	ELK3	NP_005221.2	GRHL3	NP_937816.1
ASCC2	NP_115580.2	CNOT2	NP_055330.1	ELK4	NP_001964.2	GRIP1	NP_066973.1
ASCL2	NP_005161.1	CNOT2	NP_055330.1	ELL3	NP_079441.1	GSC	NP_776248.1
ASXL1	NP_056153.2	CNOT3	NP_055331.1	EMX1	NP_004088.2	GTF2A2	NP_004483.1
ATF1	NP_005162.1	CNOT6L	NP_653172.2	EMX2	NP_004089.1	GTF2B	NP_001505.1
ATF2	NP_001871.2	CNOT7	NP_037486.2	ENO1	NP_001419.1	GTF2E2	NP_002086.1
ATF3	NP_001025458.1	CNOT7	NP_473367.2	ENO1	NP_001419.1	GTF2F1	NP_002087.2
ATF4	NP_001666.2	CNOT8	NP_004770.4	ENY2	NP_064574.1	GTF2F1	NP_002087.2
ATF5	NP_036200.2	CNTF	NP_000605.1	EOMES	NP_005433.2	GTF2F2	NP_004119.1
ATOH7	NP_660161.1	COBRA1	NP_056271.2	EPAS1	NP_001421.2	GTF2H1	NP_005307.1
ATOH8	NP_116216.1	COPS2	NP_004227.1	EPC1	NP_079485.1	GTF2H1	NP_005307.1
ATOH8	NP_116216.1	COPS3	NP_003644.2	ERCC8	NP_000073.1	GTF2H3	NP_001507.2
BACH1	NP_001177.1	COPS3	NP_003644.2	ERCC8	NP_001007234.1	GTF2H5	NP_997001.1
BAG1	NP_004314.4	COPS5	NP_006828.2	ERF	NP_006485.2	GTF2I	NP_127492.1
BANP	NP_060339.2	COPS7A	NP_057403.1	ERG	NP_891548.1	GTF2IRD1	NP_057412.1
BANP	NP_060339.2	CRABP2	NP_001869.1	ERH	NP_004441.1	GTF3C1	NP_001511.2
BANP	NP_524576.2	CREB1	NP_004370.1	ESR1	NP_000116.2	GTF3C3	NP_036218.1
BARD1	NP_000456.2	CREB1	NP_604391.1	ESR2	NP_001035365.1	GTF3C3	NP_036218.1
BARHL1	NP_064448.1	CREB3	NP_006359.3	ESRRA	NP_004442.3	GTF3C5	NP_036219.2
BARX1	NP_067545.3	CREB3L1	NP_443086.1	ESRRG	NP_001429.2	GTF3C6	NP_612417.1
BATF	NP_006390.1	CREB3L2	NP_919047.2	ESRRG	NP_996317.1	GZF1	NP_071927.1
BATF2	NP_612465.3	CREB3L3	NP_115996.1	ETS2	NP_005230.1	GZF1	NP_071927.1
BAZ2A	NP_038477.2	CREB3L4	NP_570968.1	ETS2	NP_005230.1	HAND1	NP_004812.1
BCL10	NP_003912.1	CREB5	NP_878901.2	ETV1	NP_004947.2	HBP1	NP_036389.2
BCL11A	NP_060484.2	CREB5	NP_001011666.1	ETV1	NP_004947.2	HCFC2	NP_037452.1
BCL11A	NP_075044.2	CREBL2	NP_001301.1	ETV2	NP_055024.2	HCLS1	NP_005326.2
BCL3	NP_005169.2	CREG1	NP_003842.1	ETV3	NP_005231.1	HDAC1	NP_004955.2
BCL6B	NP_862827.1	CREG1	NP_003842.1	ETV3L	NP_001004341.1	HDAC11	NP_079103.1
BHLHB2	NP_003661.1	CROP	NP_006098.2	ETV4	NP_001073143.1	HDAC3	NP_003874.2
BLZF1	NP_003657.1	CRY2	NP_066940.2	ETV5	NP_004445.1	HDAC4	NP_006028.2
BNC1	NP_001708.3	CSDA	NP_003642.2	ETV6	NP_001978.1	HDAC5	NP_005465.2
BNC2	NP_060107.3	CSDC2	NP_055275.1	ETV7	NP_057219.1	HDAC8	NP_060956.1
BRD8	NP_006687.3	CSDE1	NP_009089.4	EWSR1	NP_005234.1	HDAC9	NP_055522.1
BRF1	NP_663718.1	CTBP1	NP_001319.1	EZH1	NP_001982.2	HDX	NP_653258.2
BRF2	NP_060780.2	CTCFL	NP_542185.2	EZH2	NP_694543.1	HES1	NP_005515.1
BRF2	NP_060780.2	CTNNB1	NP_001091679.1	EZH2	NP_694543.1	HES2	NP_061962.2
BRMS1	NP_056214.1	CUX1	NP_001904.2	FAM136A	NP_116211.2	HES4	NP_066993.1
BTBD7	NP_060637.1	CXXC1	NP_055408.2	FAM90A1	NP_060558.3	HEXIM1	NP_006451.1
BTF3	NP_001198.2	DACH2	NP_444511.1	FARSA	NP_004452.1	HEXIM1	NP_006451.1
BTG1	NP_001722.1	DBP	NP_001343.2	FEV	NP_059991.1	HEXIM2	NP_653209.1
BUD31	NP_003901.2	DBX2	NP_001004329.2	FEZF1	NP_001019784.1	HEY1	NP_036390.3
C14orf166	NP_057123.1	DCP1A	NP_060873.3	FHL2	NP_001034581.1	HEY1	NP_036390.3
C14orf169	NP_078920.2	DDIT3	NP_004074.2	FHL3	NP_004459.2	HEY1	NP_001035798.1
C17orf56	NP_653280.1	DDX54	NP_076977.3	FHL5	NP_065228.3	HEY2	NP_036391.1
C19orf2	NP_604431.1	DEAF1	NP_066288.2	FIZ1	NP_116225.2	HEYL	NP_055386.1
C19orf33	NP_277055.1	DEDD2	NP_579874.1	FLI1	NP_002008.2	HHEX	NP_002720.1
C1D	NP_006324.1	DEK	NP_003463.1	FLJ36070	NP_872380.1	HIF1A	NP_001521.1
C1orf129	NP_079339.2	DIDO1	NP_071388.2	FLJ44894	NP_001034973.1	HIF1A	NP_851397.1
C20orf20	NP_060740.1	DIP2A	NP_996773.1	FOS	NP_005243.1	HIF3A	NP_071907.3
C2orf63	NP_689598.1	DLX1	NP_835221.2	FOSL1	NP_005429.1	HIF3A	NP_690007.1
CALCOCO1	NP_065949.1	DLX2	NP_004396.1	FOSL2	NP_005244.1	HIF3A	NP_690008.2
CAND1	NP_060918.2	DLX3	NP_005211.1	FOXA1	NP_004487.2	HLF	NP_002117.1
CARHSP1	NP_001035941.1	DLX4	NP_612138.1	FOXA3	NP_004488.2	HLX	NP_068777.1
CBFA2T2	NP_005084.1	DLX5	NP_005212.1	FOXD4L1	NP_036316.1	HLX	NP_068777.1
CBFA2T3	NP_005178.4	DLX6	NP_005213.2	FOXF1	NP_001442.2	HMG20A	NP_060670.1
CBFB	NP_001746.1	DMAP1	NP_001029195.1	FOXI2	NP_997309.1	HMG20B	NP_006330.2
CBX2	NP_116036.1	DMRT1	NP_068770.2	FOXJ1	NP_001445.2	HMGB2	NP_002120.1
CBX8	NP_065700.1	DMRT2	NP_006548.1	FOXL2	NP_075555.1	HMGN2	NP_005508.1
CCAR1	NP_060707.2	DMRTA1	NP_071443.1	FOXN1	NP_003584.2	HMGN4	NP_006344.1
CCDC17	NP_001108410.1	DMRTB1	NP_149056.1	FOXN2	NP_002149.2	HN1L	NP_653171.1
CCDC17	NP_001108410.1	DMRTC2	NP_001035373.1	FOXN4	NP_998761.1	HNRNPAB	NP_004490.2
CCDC72	NP_057017.1	DMTF1	NP_066968.2	FOXO3	NP_001446.1	HNRNPAB	NP_112556.2
CCDC96	NP_699207.1	DNAJC1	NP_071760.2	FOXP1	NP_116071.2	HNRNPD	NP_112737.1
CCNE1	NP_001229.1	DNMT3L	NP_787063.1	FOXP3	NP_001107849.1	HNRNPD	NP_112738.1
CCNH	NP_001230.1	DPPA2	NP_620170.2	FOXR2	NP_940853.1	HNRNPU	NP_114032.2
CCNL2	NP_001034666.1	DR1	NP_001929.1	FOXS1	NP_004109.1	HNRNPUL1	NP_653333.1

Gene Symbol	RefSeq protein ID	Gene Symbol	RefSeq protein ID	Gene Symbol	RefSeq protein ID	Gene Symbol	RefSeq protein ID
HNRPUL1	NP_008971.2	L3MBTL2	NP_113676.2	MLX	NP_937847.1	NR2E1	NP_003260.1
HOXA1	NP_005513.1	L3MBTL4	NP_775735.2	MLZE	NP_113603.1	NR2F1	NP_005645.1
HOXA10	NP_714926.1	LASS2	NP_071358.1	MNDA	NP_002423.1	NR2F2	NP_066285.1
HOXA11	NP_005514.1	LASS3	NP_849164.2	MORF4L1	NP_006782.1	NR2F6	NP_005225.2
HOXA3	NP_109377.1	LASS4	NP_078828.1	MORF4L1	NP_996670.1	NR3C1	NP_000167.1
HOXA3	NP_705896.1	LASS5	NP_671723.1	MORF4L2	NP_036418.1	NR4A1	NP_002126.2
HOXA5	NP_061975.2	LASS6	NP_982288.1	MRPL28	NP_006419.2	NR4A2	NP_006177.1
HOXA9	NP_689952.1	LBH	NP_112177.2	MSX1	NP_002439.2	NR5A1	NP_004950.2
HOXB13	NP_006352.2	LCOR	NP_115816.1	MSX2	NP_002440.2	NR5A2	NP_003813.1
HOXB3	NP_002137.4	LCORL	NP_710153.2	MTA3	NP_065795.1	NR5A2	NP_995582.1
HOXB3	NP_002137.4	LCORL	NP_710153.2	MTERF	NP_008911.1	NR6A1	NP_001480.3
HOXB5	NP_002138.1	LDB1	NP_003884.1	MTERF	NP_008911.1	NRAS	NP_002515.1
HOXB6	NP_061825.2	LETMD1	NP_056231.3	MTERFD3	NP_001028222.1	NRBF2	NP_110386.2
HOXB7	NP_004493.3	LGALS3	NP_002297.2	MTF1	NP_005946.2	NRG1	NP_039252.2
HOXB8	NP_076921.1	LHX2	NP_004780.3	MTF2	NP_031384.1	NRG1	NP_039253.1
HOXB9	NP_076922.1	LHX4	NP_203129.1	MXD1	NP_002348.1	NRG1	NP_039258.1
HOXC10	NP_059105.2	LHX6	NP_055183.2	MXD3	NP_112590.1	NRIP1	NP_003480.2
HOXC11	NP_055027.1	LHX6	NP_954629.2	MXD4	NP_006445.1	NRL	NP_006168.1
HOXC13	NP_059106.2	LHX8	NP_001001933.1	MXI1	NP_569157.2	NSBP1	NP_110390.1
HOXC4	NP_055435.2	LIN28	NP_078950.1	MYB	NP_005366.2	OASL	NP_003724.1
HOXC6	NP_710160.1	LIN28B	NP_001004317.1	MYBBP1A	NP_055335.2	OLIG1	NP_620450.1
HOXC8	NP_073149.1	LIN9	NP_775106.2	MYBL2	NP_002457.1	OLIG2	NP_005797.1
HOXC9	NP_008828.1	LITAF	NP_004853.2	MYCBP	NP_036465.2	OLIG3	NP_786923.1
HOXC9	NP_008828.1	LMO1	NP_002306.1	MYEF2	NP_057216.2	OPTN	NP_001008212.1
HOXD1	NP_078777.1	LMO2	NP_005565.1	MYF6	NP_002460.1	OSR1	NP_660303.1
HOXD3	NP_008829.3	LMO3	NP_001001395.1	MYNN	NP_061127.1	OSR2	NP_443727.1
HOXD4	NP_055436.2	LMO3	NP_001001395.1	MYOD1	NP_002469.2	OTP	NP_115485.1
HOXD8	NP_062458.1	LMO7	NP_005349.3	MYOG	NP_002470.2	OTX1	NP_055377.1
HSBP1	NP_001528.1	LMX1A	NP_796372.1	MYST2	NP_008998.1	OTX2	NP_758840.1
HSF2	NP_004497.1	LOC152485	NP_849157.2	MZF1	NP_003413.2	OVOL1	NP_004552.2
HSF2BP	NP_008962.1	LOC401898	NP_001013713.1	NAB2	NP_005958.1	OVOL2	NP_067043.2
HSFX1	NP_057237.1	LOC730394	NP_001035955.1	NANOG	NP_079141.2	PARP15	NP_689828.1
HSFY1	NP_149099.2	LOC91431	NP_001093246.1	NAT14	NP_065111.1	PAX9	NP_006185.1
HTATIP	NP_006379.2	LUZP4	NP_057467.1	NCOA4	NP_005428.1	PBX4	NP_079521.1
HTATIP2	NP_001091991.1	LYAR	NP_060286.1	NCOA7	NP_001116314.1	PBXIP1	NP_065385.2
ID1	NP_002156.2	LYL1	NP_005574.2	NEIL3	NP_060718.2	PCAF	NP_003875.3
ID2	NP_002157.2	LZTR1	NP_006758.2	NEURL	NP_004201.2	PCBD1	NP_000272.1
ID3	NP_002158.3	LZTR1	NP_006758.2	NEUROD1	NP_002491.2	PCGF2	NP_009075.1
ID4	NP_001537.1	MAF1	NP_115648.2	NEUROD4	NP_067014.2	PCGF6	NP_001011663.1
IER5	NP_057629.2	MAFB	NP_005452.2	NEUROD6	NP_073565.2	PCIF1	NP_071387.1
IGHMBP2	NP_002171.2	MAFF	NP_036455.1	NEUROG1	NP_006152.2	PDCD6	NP_037364.1
IKZF4	NP_071910.3	MAFG	NP_002350.1	NEUROG2	NP_076924.1	PDRG1	NP_110442.1
IKZF4	NP_071910.3	MAFK	NP_002351.1	NEUROG3	NP_066279.2	PELP1	NP_055204.2
IKZF5	NP_071911.3	MAML3	NP_061187.2	NFATC1	NP_006153.2	PEX14	NP_004556.1
ILF2	NP_004506.2	MAX	NP_002373.3	NFATC3	NP_775188.1	PEX14	NP_004556.1
ING1	NP_937860.1	MAX	NP_660092.1	NFATC4	NP_004545.2	PFDN1	NP_002613.2
ING2	NP_001555.1	MBD1	NP_056671.2	NFE2	NP_006154.1	PFDN5	NP_002615.2
ING3	NP_061944.2	MBD3L1	NP_660209.1	NFE2L2	NP_006155.2	PHB	NP_002625.1
ING4	NP_001121055.1	MBD6	NP_443129.3	NFE2L2	NP_006155.2	PHB2	NP_009204.1
INSM2	NP_115983.3	MBNL3	NP_060858.2	NFIB	NP_005587.2	PHF10	NP_060758.1
INTS12	NP_065128.2	MDFI	NP_005577.1	NFIC	NP_995315.1	PHF13	NP_722519.2
IRF1	NP_002189.1	MED1	NP_004765.2	NFIL3	NP_005375.2	PHF13	NP_722519.2
IRF2	NP_002190.2	MED15	NP_056973.2	NFKB1	NP_003989.2	PHF15	NP_056103.4
IRF2BP1	NP_056464.1	MED17	NP_004259.3	NFKBIA	NP_065390.1	PHF17	NP_079176.2
IRF3	NP_001562.1	MED18	NP_001120822.1	NFKBIB	NP_002494.2	PHF17	NP_955352.1
IRF6	NP_006138.1	MED20	NP_004266.2	NFKBIB	NP_001001716.1	PHF17	NP_955352.1
IRF8	NP_002154.1	MED21	NP_004255.2	NFKBIZ	NP_001005474.1	PHF19	NP_056466.1
IRX6	NP_077311.2	MED24	NP_001072986.1	NFRKB	NP_006156.2	PHF20	NP_057520.2
ISL1	NP_002193.2	MED25	NP_112235.2	NFXL1	NP_694540.3	PHF21A	NP_057705.3
ISX	NP_001008494.1	MED26	NP_004822.2	NFYA	NP_068351.1	PHF23	NP_077273.2
ITGB3BP	NP_055103.3	MED27	NP_004260.2	NFYB	NP_006157.1	PHF5A	NP_116147.1
JARID1C	NP_004178.2	MED29	NP_060062.1	NFYC	NP_055038.2	PHF6	NP_115711.2
JARID2	NP_004964.2	MED29	NP_060062.1	NHLH1	NP_005589.1	PHF7	NP_057567.3
JAZF1	NP_778231.2	MED30	NP_542382.1	NHLH2	NP_001104531.1	PHOX2A	NP_005160.2
JDP2	NP_569736.1	MED31	NP_057144.1	NIF3L1	NP_068596.2	PIAS1	NP_057250.1
JMJD2A	NP_055478.2	MED4	NP_054885.1	NKRF	NP_060014.2	PIAS1	NP_057250.1
JMY	NP_689618.2	MED6	NP_005457.2	NKX2-5	NP_004378.1	PIAS3	NP_006090.2
JRKL	NP_003763.2	MED8	NP_443109.2	NKX2-8	NP_055175.2	PIAS4	NP_056981.2
JUN	NP_002219.1	MED8	NP_963836.2	NKX6-3	NP_689781.1	PIBF1	NP_006337.2
JUNB	NP_002220.1	MED9	NP_060489.1	NME2	NP_001018147.1	PIR	NP_001018119.1
KBTBD8	NP_115894.1	MEF2C	NP_002388.2	NMI	NP_004679.2	PITX2	NP_700475.1
KCMF1	NP_064507.3	MEIS1	NP_002389.1	NMRAL1	NP_065728.1	PIWIL2	NP_060538.2
KCMF1	NP_064507.3	MEIS2	NP_733776.1	NOTO	XP_001719406.1	PKNOX2	NP_071345.2
KCNIP3	NP_038462.1	MEIS2	NP_758526.1	NPAS4	NP_849195.1	PLAG1	NP_001108106.1
KCTD7	NP_694578.1	MEIS2	NP_758527.1	NPM1	NP_002511.1	PLAG1	NP_001108107.1
KHDRBS1	NP_006550.1	MEIS3	NP_064545.1	NPM1	NP_002511.1	PLAGL1	NP_001074420.1
KLF1	NP_006554.1	MEN1	NP_570711.1	NR0B1	NP_000466.2	PLAGL1	NP_001074420.1
KLF10	NP_005646.1	MEOX2	NP_005915.2	NR0B2	NP_068804.1	PLAGL1	NP_001074424.1
KLF11	NP_003588.1	MESP1	NP_061140.1	NR1D1	NP_068370.1	PLRG1	NP_002660.1
KLF12	NP_009180.3	MGC21874	NP_689506.2	NR1D2	NP_005117.2	PMFBP1	NP_112583.1
KLF15	NP_054798.1	MGC46336	XP_001720872.1	NR1D2	NP_005117.2	PMFBP1	NP_112583.1
KLF17	NP_775755.3	MID1	NP_000372.1	NR1H2	NP_009052.3	POGZ	NP_665739.2
KLF3	NP_057615.3	MIER2	NP_060020.1	NR1H3	NP_005684.1	POLE3	NP_059139.2
KLF5	NP_001721.2	MIZF	NP_056332.2	NR1H4	NP_005114.1	POLR1E	NP_071935.1
KLF6	NP_001291.3	MKRN1	NP_038474.1	NR1H4	NP_005114.1	POLR2C	NP_116558.1
KLF7	NP_003700.1	MKRN1	NP_038474.1	NR1I2	NP_003880.3	POLR2K	NP_005025.1
KLF7	NP_003700.1	MKRN2	NP_054879.3	NR1I3	NP_005113.1	POU2AF1	NP_006226.2
KLF9	NP_001197.1	MKX	NP_775847.1	NR1I3	NP_001070945.1	POU2F1	NP_002688.2
KLHDC2	NP_055130.1	MLX	NP_733752.1	NR2C1	NP_003288.2	POU3F2	NP_005595.2
KRBA2	NP_998762.1	MLX	NP_937847.1	NR2C2	NP_003289.2	POU3F4	NP_000298.2

Gene Symbol	RefSeq protein ID	Gene Symbol	RefSeq protein ID	Gene Symbol	RefSeq protein ID	Gene Symbol	RefSeq protein ID
POU4F2	NP_004566.2	RUNX1	NP_001116079.1	STAT1	NP_009330.1	THRSP	NP_003242.1
POU5F2	NP_694948.1	RUNX1T1	NP_004340.1	STAT1	NP_644671.1	TIAL1	NP_001029097.1
POU6F1	NP_002693.2	RUVBL2	NP_006657.1	STAT3	NP_003141.2	TIGD1	NP_663748.1
PPARA	NP_001001928.1	RXRA	NP_002948.1	STAT5A	NP_003143.2	TIGD4	NP_663772.1
PPARD	NP_006229.1	RXRB	NP_068811.1	STAT5B	NP_036580.2	TIGD6	NP_112215.1
PPARG	NP_005028.4	RXRG	NP_008848.1	STAT6	NP_003144.3	TIGD7	NP_149985.2
PPARG	NP_056953.2	RYBP	NP_036366.3	SUB1	NP_006704.2	TLE1	NP_005068.2
PPARGC1B	NP_573570.2	SAFB	NP_002958.2	SUFU	NP_057253.2	TLE3	NP_065959.1
PQBP1	NP_001027553.1	SAP18	NP_005861.1	SUPT3H	NP_003590.1	TLE6	NP_079036.1
PRDM11	NP_064614.2	SAP30	NP_003855.1	SUPT3H	NP_852001.1	TLX2	NP_057254.1
PRDM14	NP_078780.1	SAP30BP	NP_037392.1	SUPT4H1	NP_003159.1	TLX3	NP_066305.2
PRDM4	NP_036538.3	SATB1	NP_002962.1	SUPT5H	NP_001104490.1	TNFAIP3	NP_006281.1
PREB	NP_037520.1	SATB2	NP_056080.1	SUPT7L	NP_055675.1	TNIP2	NP_077285.2
PRICKLE3	NP_006141.2	SAV1	NP_068590.1	TADA2L	NP_001479.3	TOX2	NP_001092268.1
PRICKLE3	NP_006141.2	SAV1	NP_068590.1	TADA2L	NP_597683.2	TP53	NP_000537.3
PRPF4B	NP_003904.3	SBNO2	NP_001093592.1	TADA3L	NP_006345.1	TP53INP1	NP_150601.1
PRPF6	NP_036601.2	SCAND1	NP_057642.1	TAF10	NP_006275.1	TP73	NP_001119712.1
PRPF6	NP_036601.2	SCMH1	NP_001026864.1	TAF11	NP_005634.1	TP73	NP_001119714.1
PRRX1	NP_073207.1	SCML2	NP_006080.1	TAF12	NP_005635.1	TRAPPC2	NP_001011658.1
PRRX2	NP_057391.1	SEC14L2	NP_036561.1	TAF13	NP_005636.1	TRERF1	NP_277037.1
PSMC3	NP_002795.2	SEC14L2	NP_036561.1	TAF15	NP_003478.1	TRIB3	NP_066981.2
PSMC5	NP_002796.4	SERTAD3	NP_037500.2	TAF15	NP_631961.1	TRIM16	NP_006461.3
PSMD9	NP_002804.2	SETDB2	NP_114121.1	TAF1A	NP_005672.1	TRIM24	NP_056989.2
PTRF	NP_036364.2	SF1	NP_004621.2	TAF1A	NP_647603.1	TRIM28	NP_005753.1
PTRF	NP_036364.2	SFRS2	NP_003007.2	TAF2	NP_003175.1	TRIM31	NP_008959.3
PTTG1	NP_004210.1	SFRS2B	NP_115285.1	TAF6	NP_005632.1	TRIM42	NP_689829.2
PURA	NP_005850.1	SHOX2	NP_003021.2	TAF6	NP_620835.1	TRIM45	NP_079464.1
PWP1	NP_008993.1	SHOX2	NP_006875.2	TAF6L	NP_006464.1	TRIM46	NP_079334.3
PYGO2	NP_612157.1	SIAH2	NP_005058.3	TAF7	NP_005633.2	TRIM52	NP_116154.1
RAB24	NP_001026847.1	SIX1	NP_005973.1	TAF9	NP_001015892.1	TRIM62	NP_060677.1
RABGEF1	NP_055319.1	SIX6	NP_031400.2	TAF9B	NP_057059.2	TRIM73	NP_944606.2
RABGEF1	NP_055319.1	SLC12A8	NP_078904.3	TARDBP	NP_031401.1	TRIP4	NP_057297.2
RAD18	NP_064550.2	SLC26A3	NP_000102.1	TAX1BP1	NP_006015.4	TRMT1	NP_060192.1
RAN	NP_006316.1	SLC30A9	NP_006336.3	TAX1BP1	NP_006015.4	TSC22D1	NP_006013.1
RARA	NP_000955.1	SMAD1	NP_001003688.1	TAX1BP1	NP_001073333.1	TSC22D3	NP_004080.2
RARB	NP_000956.2	SMAD2	NP_001003652.1	TAX1BP3	NP_055419.1	TSC22D3	NP_932174.1
RARB	NP_057236.1	SMAD2	NP_001003652.1	TBP	NP_003185.1	TSG101	NP_006283.1
RARG	NP_000957.1	SMAD3	NP_005893.1	TBPL1	NP_004856.1	TSHZ1	NP_005777.3
RARG	NP_001036193.1	SMAD5	NP_001001419.1	TBR1	NP_006584.1	TSHZ2	NP_775756.3
RB1	NP_000312.2	SMAD6	NP_005576.3	TBX15	NP_689593.2	TTRAP	NP_057698.2
RBBP4	NP_005601.1	SMAD9	NP_001120689.1	TBX21	NP_037483.1	TULP1	NP_003313.3
RBBP5	NP_005048.2	SMARCA5	NP_003592.2	TBX22	NP_001103349.1	TULP2	NP_003314.1
RBBP7	NP_002884.1	SMARCD2	NP_003068.3	TBX6	NP_542936.1	TULP3	NP_003315.2
RBBP8	NP_002885.1	SMARCD3	NP_001003802.1	TCEA1	NP_006747.1	TWIST2	NP_476527.1
RBM14	NP_006319.1	SMARCE1	NP_003070.3	TCEA1	NP_958845.1	TWISTNB	NP_001002926.1
RBM23	NP_060577.3	SMYD1	NP_938015.1	TCEA2	NP_942016.1	U2AF1	NP_006749.1
RBM23	NP_001070820.1	SNAI1	NP_005976.2	TCEA3	NP_003187.1	UBE2K	NP_005330.1
RBM39	NP_909122.1	SNAI2	NP_003059.1	TCEAL1	NP_001006640.1	UBP1	NP_001121632.1
RBM39	NP_909122.1	SNAPC1	NP_003073.1	TCEAL2	NP_525129.1	UHRF1	NP_001041666.1
RBM9	NP_001076048.1	SNAPC2	NP_003074.1	TCEB1	NP_005639.1	UIMC1	NP_057374.3
RBPJ	NP_056958.3	SNAPC5	NP_006040.1	TCEB2	NP_009039.1	UNKL	NP_075564.3
RC3H1	NP_742068.1	SND1	NP_055205.2	TCEB3	NP_003189.1	USF1	NP_009053.1
RC3H2	NP_061323.2	SNIP1	NP_078976.2	TCEB3B	NP_057511.2	USF2	NP_003358.1
RC3H2	NP_001094058.1	SNIP1	NP_078976.2	TCERG1L	NP_777597.2	UXT	NP_004173.1
RCAN1	NP_981963.1	SNW1	NP_036377.1	TCF21	NP_003197.2	VAV1	NP_005419.2
RCOR2	NP_775858.1	SOHLH1	NP_001012415.2	TCF4	NP_003190.1	VAX1	NP_954582.1
RCOR3	NP_060724.1	SOHLH2	NP_060296.1	TCF4	NP_001077431.1	VAX2	NP_036608.1
REL	NP_002899.1	SOX12	NP_008874.2	TCF7	NP_963963.1	VDR	NP_000367.1
RELB	NP_006500.2	SOX14	NP_004180.1	TCF7L2	NP_110383.2	VGLL2	NP_872586.1
REPIN1	NP_001093166.1	SOX15	NP_008873.1	TCP10L	NP_653260.1	VGLL4	NP_055482.1
REXO4	NP_065118.2	SOX17	NP_071899.1	TEAD2	NP_003589.1	VPS72	NP_005988.1
RFX2	NP_602309.1	SOX5	NP_008871.3	TEF	NP_003207.1	WBP2NL	NP_689826.1
RFX4	NP_115880.2	SOX5	NP_821078.1	TEF	NP_003207.1	WBP2NL	NP_689826.1
RFX4	NP_115880.2	SOX7	NP_113627.1	TFAM	NP_003192.1	WDR13	NP_060353.2
RFX5	NP_000440.1	SOX8	NP_055402.2	TFAM	NP_003192.1	WHSC1	NP_015627.1
RFXAP	NP_000529.1	SP1	NP_612482.2	TFAP2B	NP_003212.2	WHSC1	NP_579889.1
RFXDC1	NP_775831.1	SP3	NP_003102.1	TFAP2C	NP_003213.1	WHSC1L1	NP_060248.2
RFXDC2	NP_073752.5	SP6	NP_954871.1	TFAP2E	NP_848643.1	WHSC2	NP_005654.2
RHOXF2	NP_115887.1	SP7	NP_690599.1	TFAP4	NP_003214.1	WIZ	NP_067064.2
RHOXF2B	NP_001093155.1	SP8	NP_874359.2	TFB1M	NP_057104.2	WWTR1	NP_056287.1
RING1	NP_002922.2	SPDEF	NP_036523.1	TFB2M	NP_071761.1	YBX2	NP_057066.2
RNF103	NP_005658.1	SPIC	NP_689536.1	TFCP2	NP_005644.2	YEATS4	NP_006521.1
RNF113A	NP_008909.1	SPOP	NP_001007227.1	TFCP2L1	NP_055368.1	YWHAH	NP_003396.1
RNF113B	NP_849192.1	SPOP	NP_001007227.1	TFDP2	NP_006277.1	YY1	NP_003394.1
RNF12	NP_057204.2	SPRYD5	NP_116070.1	TFE3	NP_006512.2	YY2	NP_996806.2
RNF14	NP_004281.1	SPZ1	NP_115956.2	TFEC	NP_036384.1	ZBED2	NP_078784.2
RNF141	NP_057506.2	SPZ1	NP_115956.2	TFPT	NP_037474.1	ZBED3	NP_115743.1
RNF168	NP_689830.2	SRA1	NP_001030312.2	TGFB1	NP_000651.3	ZBTB1	NP_001116801.1
RNF24	NP_009150.1	SREBF1	NP_001005291.1	TGFB1I1	NP_057011.2	ZBTB11	NP_055230.2
RNF4	NP_002929.1	SRFBP1	NP_689759.2	TGIF1	NP_733796.2	ZBTB12	NP_862825.1
RNF6	NP_005968.1	SRFBP1	NP_689759.2	TGIF1	NP_775301.1	ZBTB16	NP_001018011.1
RORA	NP_599023.1	SS18	NP_001007560.1	TGIF2	NP_068581.1	ZBTB17	NP_003434.2
RORB	NP_008845.2	SS18L1	NP_945173.1	TGIF2LY	NP_631960.1	ZBTB17	NP_003434.2
RORC	NP_005051.2	SSBP2	NP_036578.2	TH1L	NP_945327.1	ZBTB20	NP_056457.2
RORC	NP_001001523.1	SSBP2	NP_036578.2	TH1L	NP_945327.1	ZBTB22	NP_005444.3
RQCD1	NP_005435.1	SSRP1	NP_003137.1	THEX1	NP_699163.2	ZBTB22	NP_005444.3
RRN3	NP_060897.3	SSX1	NP_005626.1	THOC1	NP_005122.2	ZBTB25	NP_008908.2
RRN3	NP_060897.3	SSX2	NP_003138.3	THRA	NP_003241.2	ZBTB26	NP_065975.1
RSBN1	NP_060834.2	SSX3	NP_783642.1	THRA	NP_955366.1	ZBTB3	NP_079060.1
RTF1	NP_055953.1	SSX5	NP_066295.3	THRB	NP_000452.2	ZBTB33	NP_006768.1

Gene Symbol	RefSeq protein ID	Gene Symbol	RefSeq protein ID	Gene Symbol	RefSeq protein ID	Gene Symbol	RefSeq protein ID
ZBTB37	NP_115911.1	ZNF189	NP_003443.2	ZNF438	NP_877432.1	ZNF641	NP_689533.1
ZBTB37	NP_115911.1	ZNF19	NP_008892.2	ZNF441	NP_689568.1	ZNF645	NP_689790.1
ZBTB4	NP_065950.1	ZNF19	NP_008892.2	ZNF442	NP_110451.1	ZNF648	NP_001009992.1
ZBTB43	NP_054726.1	ZNF193	NP_006290.1	ZNF443	NP_005806.1	ZNF649	NP_075562.2
ZBTB45	NP_116181.1	ZNF195	NP_009083.2	ZNF444	NP_060807.2	ZNF655	NP_001009958.1
ZBTB46	NP_079500.1	ZNF197	NP_008922.1	ZNF446	NP_060378.1	ZNF655	NP_001009960.1
ZBTB48	NP_005332.1	ZNF2	NP_066574.2	ZNF449	NP_689908.2	ZNF658B	NP_001027468.1
ZBTB7B	NP_056956.2	ZNF20	NP_066966.2	ZNF454	NP_872400.1	ZNF660	NP_775929.1
ZBTB8	NP_001035531.1	ZNF200	NP_003445.2	ZNF461	NP_694989.2	ZNF662	NP_997287.2
ZBTB9	NP_689948.1	ZNF202	NP_003446.2	ZNF467	NP_997219.1	ZNF664	NP_689650.1
ZC3H10	NP_116175.1	ZNF205	NP_001035893.1	ZNF480	NP_653285.1	ZNF665	NP_079009.3
ZC3H12A	NP_079355.2	ZNF211	NP_006376.2	ZNF483	NP_001007170.1	ZNF667	NP_071386.2
ZC3H15	NP_060941.2	ZNF212	NP_036388.2	ZNF484	NP_001007102.1	ZNF668	NP_078982.2
ZC3H7B	NP_060060.3	ZNF212	NP_036388.2	ZNF485	NP_660355.1	ZNF669	NP_079080.1
ZDHHC12	NP_116188.2	ZNF214	NP_037381.2	ZNF488	NP_694579.1	ZNF670	NP_149990.1
ZDHHC12	NP_116188.2	ZNF221	NP_037491.2	ZNF491	NP_689569.2	ZNF671	NP_079109.1
ZDHHC15	NP_659406.1	ZNF222	NP_037492.1	ZNF493	NP_787106.4	ZNF672	NP_079112.1
ZDHHC16	NP_115703.2	ZNF226	NP_001027544.1	ZNF497	NP_940860.1	ZNF675	NP_612203.2
ZDHHC21	NP_848661.1	ZNF227	NP_872296.1	ZNF500	NP_067678.1	ZNF678	NP_848644.1
ZDHHC23	NP_775841.2	ZNF227	NP_872296.1	ZNF501	NP_659481.2	ZNF680	NP_848653.1
ZDHHC4	NP_060576.1	ZNF228	NP_037512.3	ZNF506	NP_001092739.1	ZNF681	NP_612143.2
ZDHHC5	NP_056272.2	ZNF23	NP_666016.1	ZNF509	NP_660334.2	ZNF682	NP_149973.1
ZDHHC6	NP_071939.1	ZNF233	NP_861421.1	ZNF510	NP_055745.1	ZNF682	NP_001070817.1
ZDHHC7	NP_060210.1	ZNF24	NP_008896.1	ZNF512	NP_115810.2	ZNF683	NP_001108231.1
ZDHHC8	NP_037505.1	ZNF248	NP_066383.1	ZNF513	NP_653232.3	ZNF684	NP_689586.2
ZEB1	NP_110378.3	ZNF25	NP_659448.1	ZNF514	NP_116177.1	ZNF688	NP_660314.1
ZEB1	NP_110378.3	ZNF250	NP_066405.1	ZNF517	NP_998770.2	ZNF689	NP_612456.1
ZEB2	NP_055610.1	ZNF251	NP_612376.1	ZNF521	NP_056276.1	ZNF691	NP_056995.1
ZFAND3	NP_068762.1	ZNF253	NP_066385.2	ZNF524	NP_694951.1	ZNF691	NP_056995.1
ZFAND5	NP_001095890.1	ZNF257	NP_258429.2	ZNF526	NP_597701.1	ZNF692	NP_060335.2
ZFAND6	NP_061879.2	ZNF26	NP_062537.2	ZNF527	NP_115829.1	ZNF696	NP_112157.1
ZFAT	NP_065914.2	ZNF263	NP_005732.2	ZNF529	NP_066002.1	ZNF699	NP_940937.1
ZFP1	NP_710155.2	ZNF277	NP_068834.2	ZNF530	NP_065931.2	ZNF7	NP_003407.1
ZFP1	NP_710155.2	ZNF280A	NP_542778.1	ZNF532	NP_060651.2	ZNF70	NP_068735.1
ZFP161	NP_003400.2	ZNF280B	NP_542942.1	ZNF540	NP_689819.1	ZNF700	NP_653167.1
ZFP2	NP_085116.2	ZNF281	NP_036614.1	ZNF543	NP_998763.1	ZNF701	NP_060730.1
ZFP3	NP_694563.1	ZNF286A	NP_065703.1	ZNF545	NP_597723.1	ZNF704	NP_001028895.1
ZFP36	NP_003398.1	ZNF295	NP_001091872.1	ZNF547	NP_775902.2	ZNF705A	NP_001004328.1
ZFP36L1	NP_004917.2	ZNF3	NP_116313.3	ZNF549	NP_694995.1	ZNF707	NP_776192.2
ZFP36L2	NP_008818.3	ZNF3	NP_116313.3	ZNF550	NP_001034743.1	ZNF709	NP_689814.1
ZFP37	NP_003399.1	ZNF300	NP_443092.1	ZNF552	NP_079038.2	ZNF710	NP_940928.1
ZFP41	NP_776193.1	ZNF302	NP_001012320.1	ZNF553	NP_689865.1	ZNF713	NP_872439.1
ZFP42	NP_777560.2	ZNF32	NP_001005368.1	ZNF554	NP_001096121.1	ZNF714	NP_872321.2
ZFP64	NP_060667.2	ZNF321	NP_976052.2	ZNF556	NP_079243.1	ZNF74	NP_003417.2
ZFP64	NP_071371.3	ZNF322A	NP_078915.2	ZNF557	NP_001037852.1	ZNF74	NP_003417.2
ZFP64	NP_955459.2	ZNF322A	NP_078915.2	ZNF557	NP_001037853.1	ZNF740	NP_001004304.1
ZFP91	NP_444251.1	ZNF323	NP_665916.1	ZNF558	NP_653294.1	ZNF75A	NP_694573.1
ZIC4	NP_115529.2	ZNF324	NP_055162.1	ZNF560	NP_689689.2	ZNF763	NP_001012771.1
ZIC4	NP_115529.2	ZNF324B	NP_997278.1	ZNF561	NP_689502.1	ZNF764	NP_219363.1
ZIK1	NP_001010879.2	ZNF329	NP_078896.3	ZNF562	NP_060126.1	ZNF764	NP_219363.1
ZKSCAN1	NP_003430.1	ZNF331	NP_001073375.1	ZNF563	NP_660319.1	ZNF766	NP_001010851.1
ZKSCAN2	NP_001012999.3	ZNF333	NP_115809.1	ZNF564	NP_659413.1	ZNF768	NP_078947.3
ZKSCAN3	NP_077819.2	ZNF334	NP_060572.3	ZNF565	NP_001035939.1	ZNF77	NP_067040.1
ZKSCAN4	NP_061983.2	ZNF334	NP_060572.3	ZNF566	NP_116227.1	ZNF771	NP_057727.1
ZKSCAN5	NP_055384.1	ZNF334	NP_955473.1	ZNF567	NP_689816.2	ZNF773	NP_940944.1
ZMAT1	NP_115817.1	ZNF33A	NP_008905.1	ZNF569	NP_689697.2	ZNF774	NP_001004309.2
ZMAT5	NP_001003692.1	ZNF34	NP_085057.3	ZNF57	NP_775751.1	ZNF776	NP_775903.2
ZMYND11	NP_006615.1	ZNF341	NP_116208.3	ZNF572	NP_689625.1	ZNF780B	NP_001005851.1
ZMYND11	NP_006615.1	ZNF342	NP_660331.1	ZNF573	NP_689573.2	ZNF782	NP_001001662.1
ZMYND8	NP_898868.1	ZNF343	NP_077301.4	ZNF574	NP_073589.4	ZNF784	NP_976308.1
ZMYND8	NP_898868.1	ZNF345	NP_003410.1	ZNF575	NP_777605.1	ZNF785	NP_689671.2
ZMYND8	NP_898869.1	ZNF347	NP_115973.1	ZNF576	NP_077303.1	ZNF785	NP_689671.2
ZNF10	NP_056209.2	ZNF347	NP_115973.1	ZNF579	NP_689813.2	ZNF79	NP_009066.1
ZNF101	NP_149981.2	ZNF35	NP_003411.3	ZNF580	NP_057286.1	ZNF790	NP_996777.2
ZNF114	NP_705836.1	ZNF350	NP_067645.3	ZNF581	NP_057619.1	ZNF790	NP_996777.2
ZNF121	NP_001008727.1	ZNF354A	NP_005640.2	ZNF581	NP_057619.1	ZNF791	NP_699189.1
ZNF131	NP_003423.1	ZNF354B	NP_478137.1	ZNF582	NP_653291.1	ZNF793	NP_001013681.2
ZNF132	NP_003424.3	ZNF354C	NP_055409.1	ZNF583	NP_689691.1	ZNF799	NP_001074290.1
ZNF133	NP_001076799.1	ZNF358	NP_060553.4	ZNF585B	NP_689492.2	ZNF8	NP_066575.1
ZNF134	NP_003426.3	ZNF366	NP_689838.1	ZNF586	NP_060122.2	ZNF821	NP_060000.1
ZNF135	NP_003427.2	ZNF37A	NP_001007095.1	ZNF587	NP_116217.1	ZNF83	NP_001099019.1
ZNF136	NP_003428.1	ZNF382	NP_116214.2	ZNF593	NP_056955.2	ZNF83	NP_001099023.1
ZNF138	NP_006515.1	ZNF383	NP_689817.1	ZNF596	NP_001035880.1	ZNF84	NP_001120844.1
ZNF140	NP_003431.2	ZNF384	NP_001035005.1	ZNF597	NP_689670.1	ZNF92	NP_689839.1
ZNF143	NP_003433.3	ZNF385A	NP_056296.1	ZNF599	NP_001007249.1	ZNRD1	NP_055411.1
ZNF148	NP_068799.2	ZNF394	NP_115540.2	ZNF605	NP_899061.1	ZSCAN1	NP_872378.3
ZNF154	NP_001078853.1	ZNF395	NP_061130.1	ZNF607	NP_116078.3	ZSCAN10	NP_116194.1
ZNF155	NP_003436.2	ZNF396	NP_665699.1	ZNF610	NP_775801.1	ZSCAN16	NP_079507.1
ZNF155	NP_003436.2	ZNF397	NP_115723.1	ZNF613	NP_079116.2	ZSCAN18	NP_076415.2
ZNF16	NP_001025147.2	ZNF398	NP_065832.1	ZNF615	NP_940882.2	ZSCAN18	NP_076415.2
ZNF165	NP_003438.1	ZNF408	NP_079017.1	ZNF616	NP_848618.2	ZSCAN2	NP_060364.3
ZNF17	NP_008890.2	ZNF410	NP_067011.1	ZNF619	NP_775927.1	ZSCAN2	NP_870992.2
ZNF175	NP_009078.1	ZNF414	NP_115746.1	ZNF620	NP_787084.1	ZSCAN20	NP_660281.2
ZNF18	NP_653281.2	ZNF415	NP_060825.2	ZNF621	NP_001091884.1	ZSCAN22	NP_862829.1
ZNF180	NP_037388.1	ZNF415	NP_060825.2	ZNF622	NP_219482.1	ZSCAN29	NP_689668.3
ZNF180	NP_037388.1	ZNF418	NP_597717.1	ZNF623	NP_001075949.1	ZSCAN4	NP_689890.1
ZNF182	NP_008893.1	ZNF419	NP_001091964.1	ZNF626	NP_001070143.1	ZSCAN5	NP_077279.1
ZNF184	NP_009080.1	ZNF425	NP_001001661.1	ZNF627	NP_660338.1		
ZNF187	NP_689949.3	ZNF426	NP_077011.1	ZNF630	NP_001032824.2		
ZNF187	NP_001018854.2	ZNF433	NP_001073880.1	ZNF639	NP_057415.1		

Supplementary Table S2. List of 18 human transcription factors from the second screening.

Gene Symbol	RefSeq protein ID	Gene Symbol	RefSeq protein ID	Gene Symbol	RefSeq protein ID	Gene Symbol	RefSeq protein ID
GLIS1	NP_671726.1	ZBTB8	NP_001035531.1	ZSCAN4	NP_689890.1	OTX2	NP_758840.1
DMRTB1	NP_149056.1	ZBTB43	NP_054726.1	ZNF768	NP_078947.3	PRRX2	NP_057391.1
PITX2	NP_700475.1	ZNF202	NP_003446.2	PRPF4B	NP_003904.3	OTP	NP_115485.1
IRX6	NP_077311.2	ZNF383	NP_689817.1	NHLH1	NP_005589.1		
OVOL2	NP_067043.2	NR5A1	NP_004950.2	GRHL1	NP_055367.2		

Supplementary Table S3. Summary of blastocyst injection.

combination of genes	iPSC derived	number of born mice	number of chimeras (male)	number of chimeras mated for F1 offspring (germline contribution)
OS+Glis1 #1	skin fibroblasts	33	20 (8)	4(0)
OS+Glis1 #2	MEF	40	10 (6)	3(0)
OSM+Glis1 #1	skin fibroblasts	21	13 (9)	3(0)
OSM+Glis1 #2	MEF	80	32 (19)	8(1)
OSM+Glis1 #3	MEF	58	31 (19)	5(0)
OSK+Glis1 #1	skin fibroblasts	40	15 (10)	9(0)
OSK+Glis1 #2	MEF	51	31 (15)	14(1)
OSK+Glis1 #3	MEF	37	30 (16)	17(0)
OSKM #1	skin fibroblasts	27	10 (4)	0
OSKM #2	skin fibroblasts	42	27 (17)	5(1)

Supplementary Table S4. List of the 90(a) and 32(b) probes from microarray analysis.

a OSK<OSKG (20-fold) 90

ProbeName	GeneSymbol	ProbeName	GeneSymbol
A_51_P105480	Nanos3	A_51_P419047	Esrrb
A_51_P108581	Adrbk2	A_51_P439311	1810041L15Rik
A_51_P112932	Entpd2	A_51_P449824	
A_51_P127695	Greb1	A_51_P452714	Kcnmb4
A_51_P143162	Myh7	A_51_P457989	Rragd
A_51_P170725	1300002K09Rik	A_51_P462533	Syt7
A_51_P171616	Wnt10a	A_51_P477121	Pmaip1
A_51_P171832	Nrgn	A_51_P480136	Cryba2
A_51_P175988	Htr3a	A_51_P481221	Bace2
A_51_P204153	Igfbp5	A_51_P488819	4933400F03Rik
A_51_P210510	Sparcl1	A_51_P490337	Tmem190
A_51_P222467	Abcg1	A_51_P494037	Dgkg
A_51_P222773	Foxa2	A_51_P495986	Gmpr
A_51_P230175	Bcan	A_51_P503149	Tns4
A_51_P236483	Dcpp1	A_52_P1037027	
A_51_P236486	Dcpp1	A_52_P145415	Ptch2
A_51_P239601	Trpv5	A_52_P18299	Chd5
A_51_P240811	Wnt8a	A_52_P187058	Nptx2
A_51_P241319	Cilp	A_52_P203560	Fzd10
A_51_P262238	Tub	A_52_P257502	Igfbp4
A_51_P267783	Il11	A_52_P258116	Wnt3
A_51_P270997	Igfbpl1	A_52_P274496	Tspan18
A_51_P274223	Fgf17	A_52_P307860	Krt9
A_51_P296815	Gpr68	A_52_P361534	Wnt3
A_51_P297069	Tmod1	A_52_P373694	Jph4
A_51_P303217	Ucma	A_52_P403398	Ihh
A_51_P305003	Ntrk1	A_52_P415155	Wnt6
A_51_P306287		A_52_P416575	Trim61
A_51_P309754	LOC100046808	A_52_P419678	Serpina3f
A_51_P333253	Myo1g	A_52_P435561	Prr15l
A_51_P355427	Timp4	A_52_P448045	
A_51_P359173	Syt7	A_52_P469502	Cda
A_51_P359822	Sftpd	A_52_P490032	Rragd
A_51_P361150	Pcp4l1	A_52_P497392	Dcpp3
A_51_P367100	Itih3	A_52_P520037	Rimbp2
A_51_P367880	Krt84	A_52_P535962	Dcpp2
A_51_P372743	Frmpd3	A_52_P545132	Kcnc2
A_51_P377557	Cpsf4l	A_52_P54770	Fam19a4
A_51_P398971	Igfbp4	A_52_P577136	
A_51_P399305	Tnfrsf19	A_52_P633353	Igfbpl1
A_51_P401504	Col9a2	A_52_P64356	Sparcl1
A_51_P404193	Sp5	A_52_P70856	Frmpd1
A_51_P407984	Grifin	A_52_P71756	Prb1
A_51_P417720	Itga11	A_52_P88033	Myh7
A_51_P418820	Tcfap2c	A_52_P964651	Fam65c

b ES-enriched → OSK<OSKG (3-fold) 32

ProbeName	GeneSymbol
A_51_P146149	Napsa
A_51_P195044	Dppa3
A_51_P202340	Pou5f1
A_51_P246345	Myl7
A_51_P270997	Igfbpl1
A_51_P274223	Fgf17
A_51_P282538	Gad1
A_51_P294233	Nanog
A_51_P300657	Nefh
A_51_P306287	
A_51_P333253	Myo1g
A_51_P338278	Trh
A_51_P377557	Cpsf4l
A_51_P389885	Spic
A_51_P402617	Nkx6-2
A_51_P404193	Sp5
A_51_P407028	Car4
A_51_P418820	Tcfap2c
A_51_P419047	Esrrb
A_51_P433194	Bcas1
A_51_P450248	Esx1
A_51_P489935	
A_51_P497332	Mycn
A_52_P1004880	
A_52_P196161	Sh3gl2
A_52_P260659	Kcnj10
A_52_P294305	Lin28a
A_52_P488623	Fam169a
A_52_P536494	Mycn
A_52_P571780	Calb2
A_52_P617512	Camta1
A_52_P618417	

Supplementary Table S5. Primers list.

Primers for RT-PCR analysis

Gene	Forward primer	Reverse primer
mNanog	AGGGTCTGCTACTGAGATGCT	CAACACCTGGTTTTTCTGCCACCG
hNanog	CAGCCCCGATTCTTCCACCAGTCCC	CGGAAGATTCCCAGTCGGGTTCACC
mOct3/4(endo)	TCTTTCCACCAGGCCCCCGGCTC	TGCGGGCGGACATGGGGAGATCC
hOct3/4(endo)	GACAGGGGGAGGGGAGGAGCTAGG	CTTCCCTCCAACCAGTTGCCCCAAAC
Oct3/4(Tg)	CCCCAGGGCCCCATTTTGGTACC	CCCTTTTTCTGGAGACTAAATAAA
mSox2(endo)	TAGAGCTAGACTCCGGGCGATGA	TTGCCTTAAACAAGACCACGAAA
hSox2(endo)	GGG AAA TGG GAG GGG TGC AAA AGA GG	TTGCGTGAGTGTGGATGGGATTGGTG
Sox2(Tg)	GGCACCCCTGGCATGGCTCTTGGCTC	TTATCGTCGACCACTGTGCTGCTG
hRex1	CAGATCCTAAACAGCTCGCAGAAT	GCGTACGCAAATTAAAGTCCAGA
mEcat1	TGTGGGGCCCTGAAAGGCGAGCTGAGAT	ATGGGCCGCCATACGACGACGCTCAACT
Nat1	ATTCTTCGTTGTCAAGCCGCCAAAGTGGAG	AGTTGTTTGCTGCGGAGTTGTCATCTCGTC

Primers for qPCR analysis

Gene	Forward primer	Reverse primer
Glis1	CTCCAAGCATCCACACTGTT	GACAGGATGCCTGAAGCAAG
Nanog	AGGGTCTGCTACTGAGATGCT	CAACACCTGGTTTTTCTGCCACCG
Nrgn	TCCAAGCCAGACGACGATATT	CACACTCTCCGCTCTTTATCTTC
Tspan18	CAAGGAGCTTACCAAGCACTAC	GGCAGAGAAAACATCCGTATCG
N-Myc	CCTCACTCCTAATCCGGTCAT	GTGCTGTAGTTTTTCGTTCACTG
c-Myc	TCTCCATCCTATGTTGCGGTC	TCCAAGTAACTCGGTCATCATCT
G3PDH	ACC ACA GTC CAT GCC ATC AC	TCC ACC ACC CTG TTG CTG TA
Nat1	ATTCTTCGTTGTCAAGCCGCCAAAGTGGAG	AGTTGTTTGCTGCGGAGTTGTCATCTCGTC

Primers for ChIP analysis

Gene	Forward primer	Reverse primer
N-Myc	ACCTCCAGCGGCATCCAGGA	TCCAAACCGAGACCTCCCGCT
L-Myc	GGGAGGGGGAGGGGCTTGTC	CGCGATCTGCAGGCGCATTG
c-Myc	GAAACCCTGCAGCCCTGCCC	TGGCCACAGAGACCACAGCG
Nanog	TACTGAGTATAAGCTACTCAAGGCAACAG	CTTTTTAACGCAAGTCTGAAGAAAGAG
Esrrb	AGGCGCCTGGGGAGGAATGT	CCTGGCCATATGCAGGGTGGC
Lin28a	GGGAGGCAGCCAGGACAGGT	TCGCAGGCCCTCTCAGGGAC
Foxa2	GCAGTGCAGCCCACAGGCTT	GCGCACGCACACACAACAAGG
Gata4	CCCCGTAGATCTGAGGCTAGCAAGG	CCTACTCTCAGTGGTCCACGTCCAG
Nkx2-5	CACCACTCTCTGCTACCCACCTGG	GCTGCTGCTCCAGGTTCAGGATGTC

Supplementary Table S6. Sequence of hairpin of shRNAs

shRNA2	CCGGGGCCTCACCAACCCTGCACCTCTCGAGAGGTGCAGGGTTGGTGAGGCCTTTTTG
Scramble shRNA2	CCGGGCGGCACACACACTCTCTCCCCTCGAGGGGAGAGAGTGTGTGTGCCGCTTTTTG
shRNA6	CCGGGCCCTTCAATGCCCGCTACAACTCGAGTTGTAGCGGGCATTGAAGGGCTTTTTG
Scramble shRNA6	CCGGGCGCGCACACACACACTTTTCCTCGAGGAAAAGTGTGTGTGTGCGCGCTTTTTG