nature cell biology

Early role for IL-6 signalling during generation of induced pluripotent stem cells revealed by heterokaryon RNA-Seq

Jennifer J. Brady[1], Mavis Li[2,3], Silpa Suthram[1], Hui Jiang[2,3,4], Wing H. Wong[3] and Helen M. Blau[1,5]

Molecular insights into somatic cell reprogramming to induced pluripotent stem cells (iPS) would aid regenerative medicine, but are difficult to elucidate in iPS because of their heterogeneity, as relatively few cells undergo reprogramming (0.1–1%; refs 1,2). To identify early acting regulators, we capitalized on non-dividing heterokaryons (mouse embryonic stem cells fused to human fibroblasts), in which reprogramming towards pluripotency is efficient and rapid[3], enabling the identification of transient regulators required at the onset. We used bi-species transcriptome-wide RNA-seq to quantify transcriptional changes in the human somatic nucleus during reprogramming towards pluripotency in heterokaryons. During heterokaryon reprogramming, the cytokine interleukin 6 (*IL6*), which is not detectable at significant levels in embryonic stem cells, was induced 50-fold. A 4-day culture with IL6 at the onset of iPS reprogramming replaced stably transduced oncogenic c-Myc such that transduction of only Oct4, Klf4 and Sox2 was required. IL6 also activated another Jak/Stat target, the serine/threonine kinase gene *Pim1*, which accounted for the IL6-mediated twofold increase in iPS frequency. In contrast, LIF, another induced GP130 ligand, failed to increase iPS frequency or activate *c-Myc* or *Pim1*, thereby revealing a differential role for the two Jak/Stat inducers in iPS generation. These findings demonstrate the power of heterokaryon bi-species global RNA-seq to identify early acting regulators of reprogramming, for example, extrinsic replacements for stably transduced transcription factors such as the potent oncogene c-Myc.

Heterokaryons are ideally suited to elucidating regulators at the onset of reprogramming. When fused to human fibroblasts, mouse embryonic stem cells (mouse ESC) provide a panoply of stem cell factors in abundance, as the ratio is skewed to favour mouse ESC leading to a rapid induction of pluripotency transcripts in the somatic human fibroblast that can be distinguished on the basis of species differences[3]. To identify previously unrecognized early acting regulators of reprogramming, we monitored heterokaryon gene expression over a 3-day time course during which no cell division occurs[3]. Heterokaryons generated from an excess of mouse GFP$^+$ ESC and DsRed$^+$ human fibroblasts were enriched to 80% purity by fluorescence-activated cell sorting (FACS; Fig. 1a,b). The nuclear ratio (mouse/human) was determined by the relative abundance of mouse Dppa5 promoter copy number relative to human OCT4 promoter copy number by genomic quantitative PCR (qPCR), revealing an average nuclear ratio of 3.4 ± 0.3 mouse nuclei to 1 human nucleus (mean $\pm$ s.d.). In good agreement, FACS-sorted heterokaryons analysed singly by microscopy were found to have a mean content of 4 nuclei, when fitted to a Gaussian distribution (Fig. 1c).

Following heterokaryon isolation by FACS and Illumina library construction, sequencing revealed that the average heterokaryon RNA-seq library yielded 15.5 million reads that had an 87% mappable rate when aligned to mouse and human genomes (~23,400 transcripts for mouse and ~40,000 transcripts for human; Supplementary Table S1). In accordance with the nuclear ratio, the mappable reads consisted of 70% mouse, 26% human and 4% that were common to both species (Fig. 1d, workflow outlined in Fig. 1e). Of the total mappable reads, 4% were eliminated owing to species homology. However, importantly, no coding gene was entirely eliminated from the analysis. To quantify gene expression for transcripts with significant conservation between the mouse and human transcriptomes, we modified the original definition of RPKM (reads per kilobase of exon model per million mappable

[1]Baxter Laboratory for Stem Cell Biology, Department of Microbiology and Immunology, Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, California 94305, USA. [2]Institute for Computational and Mathematical Engineering, Stanford University, Stanford, California 94305, USA. [3]Department of Statistics, Stanford University, Stanford, California 94305, USA. [4]Present address: Department of Biostatistics, University of Michigan, Ann Arbor, Michigan 48109, USA.
[5]Correspondence should be addressed to H.M.B. (e-mail: hblau2@stanford.edu)

Received 16 December 2012; accepted 30 July 2013; published online 1 September 2013; DOI: 10.1038/ncb2835

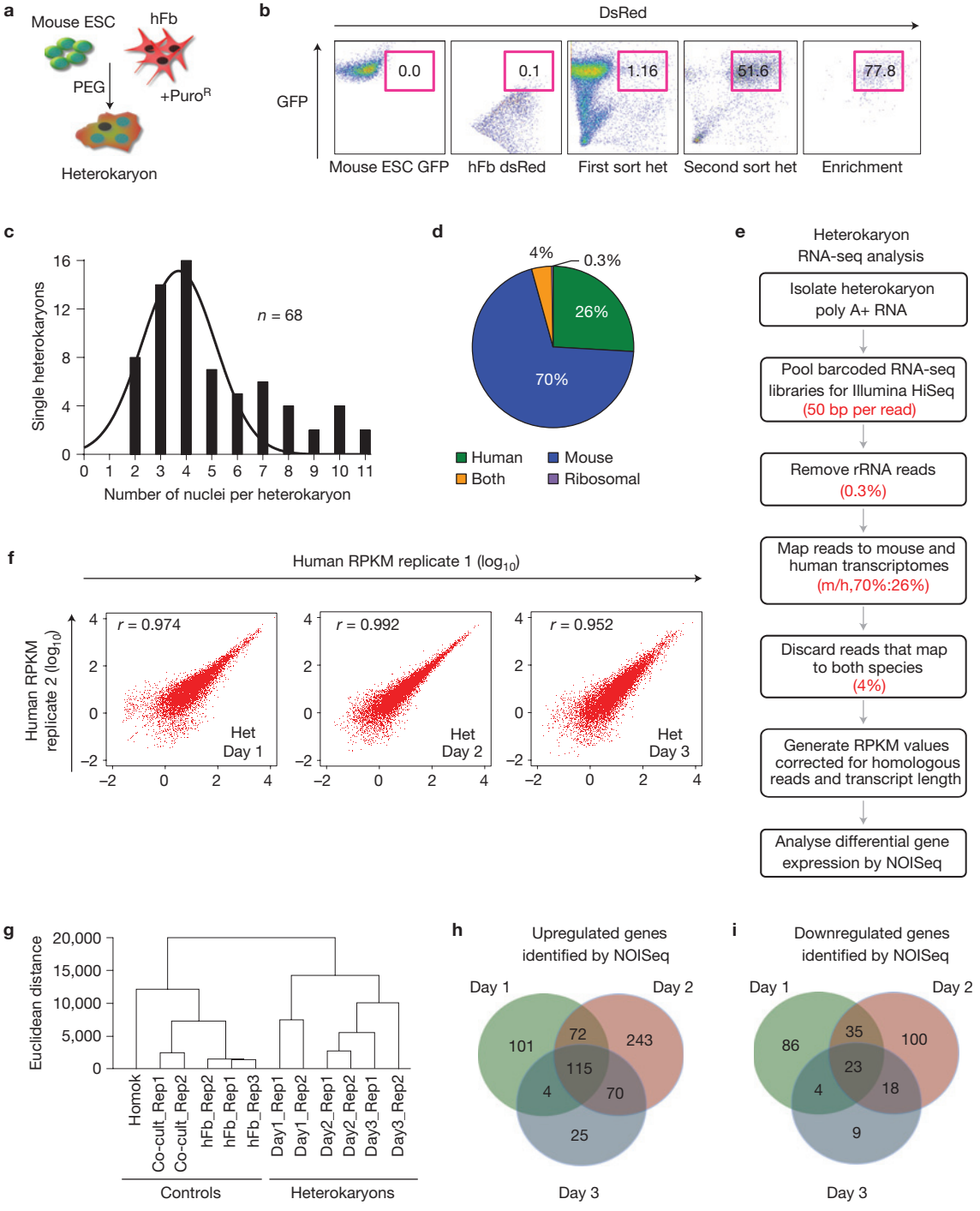


Figure 1 Heterokaryon global bi-species transcriptome analysis by RNA-seq identifies differentially expressed genes during somatic cell reprogramming towards pluripotency. (**a**) Scheme for heterokaryon generation. GFP mouse ESC were co-cultured with DsRed puro^R primary human fibroblasts (hFb) and then fused using PEG. (**b**) Heterokaryon enrichment and isolation by FACS. Representative FACS plots for heterokaryon isolation. From left to right: GFP$^+$ mouse ESC, DsRed$^+$ human fibroblasts, GFP$^+$DsRed$^+$ day 2 heterokaryons (het) first sort, second sort, enrichment. (**c**) Histogram of heterokaryon nuclear content (total number of nuclei per FACS-sorted heterokaryon assessed by single-cell microscopy, $n = 68$). A fitted Gaussian curve reveals an average nuclear content of 4 nuclei per heterokaryon. (**d**) Average mappable read summary for heterokaryon RNA-seq libraries. (**e**) Flow chart of heterokaryon RNA-seq analysis. (**f**) Comparison of biological replicates of human gene expression during heterokaryon reprogramming for the three days. Pearson's correlation values are shown for heterokaryon biological replicates. Replicate 1 for each day is shown on the *x* axis, and replicate 2 for each day on the *y* axis. (**g**) Hierarchical clustering of RNA-seq data for heterokaryon as well as control samples, using the Euclidean distance metric and complete linkage method. Homok, fused fibroblast homokaryons; Co-cult, unfused co-cultures of fibroblasts and mouse ESC. (**h**,**i**) Venn diagram of significantly upregulated and downregulated gene expression in heterokaryon reprogramming compared with the homokaryon, co-culture and unfused human fibroblast controls. Differential gene expression was identified using the NOISeq method with a cutoff probability threshold of 0.8. Source data are presented in Supplementary Table S1.

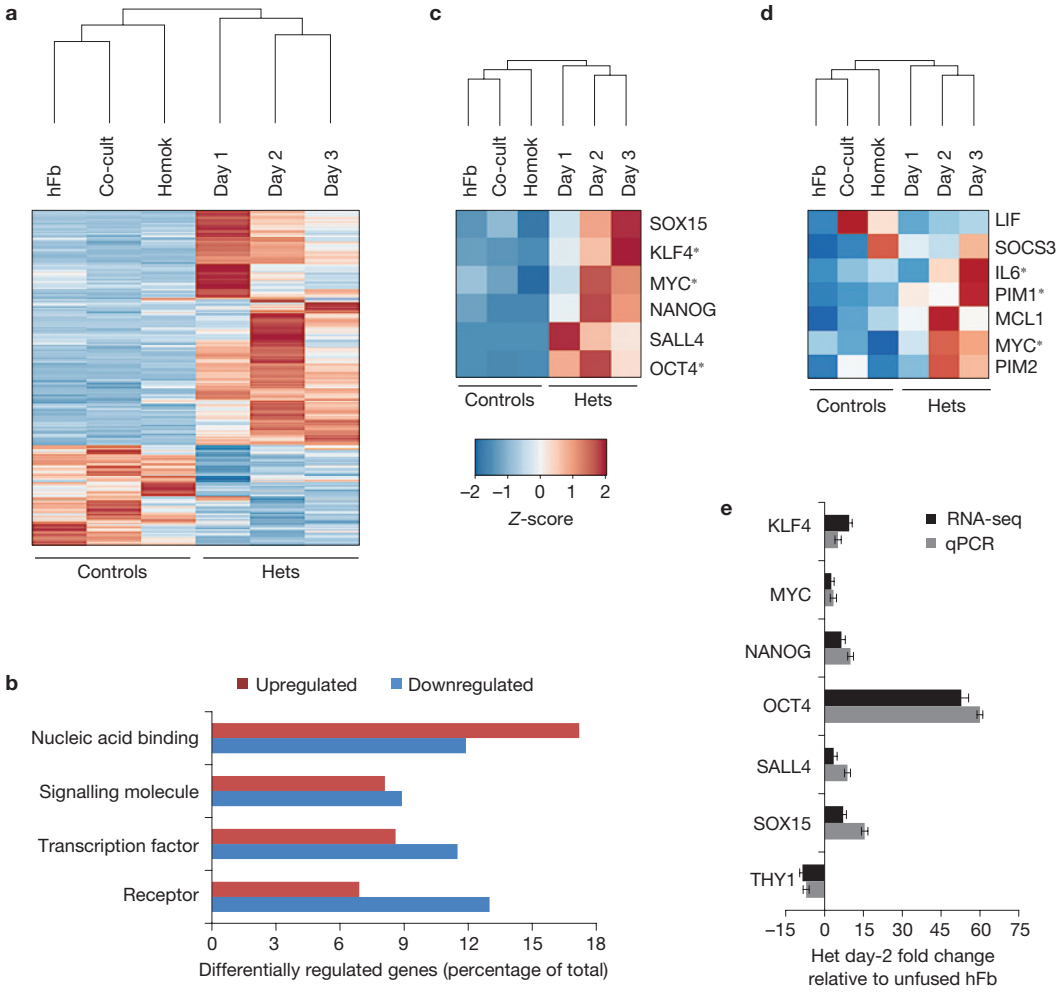


Figure 2 Heterokaryon bi-species RNA-seq identifies induction of GP130 signalling gene expression during somatic cell reprogramming. (**a**) Heat map of differentially expressed genes during heterokaryon reprogramming (day 1–3 post fusion) as compared with unfused human fibroblasts (hFb), co-culture (of fibroblasts and mouse ESC; Co-cult) and homokaryon (fibroblast–fibroblast fusion; Homok) controls. The gene expression (RPKM) values for each gene were normalized to the standard normal distribution to generate *Z*-scores. The *Z*-score colour bar is shown with the minimum expression value for each gene in blue and the maximum value in red. (**b**) Protein classification of the differentially expressed human genes ($n = 905$ genes) using the PANTHER (Protein Analysis Through Evolutionary Relationships) classification system with default settings. Select protein classes encompassing at least 5% of the total number of upregulated genes ($n = 630$) or downregulated genes ($n = 275$) are shown. (**c**,**d**) Heat map of human pluripotency transcription factors (**c**) and GP130 signalling gene pathway genes (**d**) during heterokaryon reprogramming. Asterisks indicate genes identified as significantly differentially expressed in the NOISeq analysis. Normalized gene expression values were used (see above). (**e**) Validation of heterokaryon RNA-seq gene expression levels by qPCR for six human pluripotency genes and the lineage marker THY1 (Day 2 post fusion), relative to unfused human fibroblasts. Data are represented as mean ± s.d. for qPCR data ($n = 3$) and mean with range ($n = 2$) for RNA-seq data.

reads, the standard measure of gene expression using RNA-seq)[4] to account for the length of the conserved regions (Supplementary Note S1). Heterokaryon human RPKM values showed a high level of reproducibility across biological replicates for days 1, 2 and 3, with Pearson correlation values of >0.95 (Fig. 1f). To focus our analysis on reprogramming specific to the somatic cell nucleus, we compared human gene expression over the heterokaryon time course with three control populations: human fibroblasts alone, fused fibroblast homokaryons, and unfused co-cultures of fibroblasts and mouse ESC. Hierarchical clustering of all of the samples showed that gene expression over the three days was closer in heterokaryon samples when compared with the controls, highlighting the quality of the data generated (Fig. 1g).

We identified a total of 905 differentially expressed genes over the three days using the NOISeq method analytic package[5], as shown in the Venn diagrams (Fig. 1h,i, and Supplementary Table S1). NOISeq analysis of heterokaryon reprogramming revealed 630 upregulated and 275 downregulated human genes unique to reprogramming (Fig. 2a and Supplementary Table S1). These genes separated into 28 PANTHER protein classes with transcription factors, nucleic acid binders, signalling molecules and receptors comprising 34% of the differentially expressed genes (Fig. 2b and Supplementary Table S1)[6]. Several of the expected pluripotency transcription factors were induced in human fibroblasts (Fig. 2c). In addition, we observed the induction of GP130 signalling pathway genes, including the transcription factor *MYC*, the pro-survival kinases *PIM1* and *PIM2*, the negative feedback regulator

SOCS3, the anti-apoptotic BCL2 family member *MCL1* (refs 7–9) and the GP130 ligands *LIF* and *IL6* (Fig. 2d), with *MYC*, *PIM1* and *IL6* identified as significantly differentially expressed in the NOISeq analysis. The GP130/STAT3-responsive gene subset was induced in heterokaryons concomitantly with most of the human embryonic stem cell gene signature, as defined by others[10–12]. The fold inductions observed by RNA-seq were comparable to those obtained by heterokaryon qPCR, validating the approach (Fig. 2e). On the basis of these findings, we postulated that IL6 plays a role in iPS reprogramming.

To determine whether IL6 or other IL6-type cytokine family members could be responsible for induction of the GP130 gene signature in heterokaryons, we examined the expression of the entire family of IL6 type cytokines, which consists of *IL6*, *LIF* (leukaemia inhibitory factor), *OSM* (oncostatin M), *IL11* (interleukin-11), *CNTF* (ciliary neurotrophic factor) and *CT-1* (cardiotrophin-1)[8]. The transcript levels in the human RNA-seq data set identified *IL6* as the prime candidate, as it was the most highly induced, increasing by 50-fold on day 3 relative to unfused fibroblasts (Fig. 3a and Supplementary Fig. S1a). The only family member other than IL6 that exhibited >2-fold induction was *LIF*, which was induced 4-fold by day 3, suggesting these two cytokines as potential inducers of the GP130 signature gene expression seen in heterokaryons (Figs 2d and 3a).

To determine whether in iPS generation, like heterokaryons, Il6 and Lif transcript levels were also induced, we examined their relative messenger RNA levels after transduction with the 4 factors (4F), which include Oct4, Klf4, Sox2, and c-Myc (also known as OKSM), transduction of mouse embryonic fibroblasts (MEFs), cultured in the absence of exogenous cytokines. Both were expressed at low levels before fusion. However, *Il6*, but not *Lif*, was induced >2-fold by day 4 of iPS reprogramming (Fig. 3b). In addition, we found that the Il6 receptor (*Il6r*) was induced 6.6-fold during 4F reprogramming whereas the levels of the Lif receptor (*Lifr*) did not exceed 1.4-fold induction (Supplementary Fig. S1b). To ensure that the MEFs used for reprogramming were responsive to IL6 treatment, we investigated whether exposure of MEFs to IL6 induced the Jak/Stat signal transduction pathway by assaying for phosphorylation of Stat3 Tyr 705 (ref. 8). Il6r receptor function was validated by detection of Stat3 Tyr-705 phosphorylation by western blotting (Fig. 3c). Fibroblasts treated with IL6 resulted in significantly higher levels of pStat3 than unphosphorylated Stat3 relative to α-tubulin, peaking 30 min post-stimulation (Fig. 3d). Thus, *Il6* but not *Lif* was deemed the most likely candidate capable of enhancing iPS reprogramming through Gp130.

To determine whether IL6 played a functional role in reprogramming somatic cells to iPS, the potential of IL6 to augment MEF reprogramming to iPS was investigated in combination with lentiviral transduction of 4F (Fig. 3e). iPS colonies were assessed by immunostaining for the pluripotency markers Nanog and SSEA-1 (Fig. 3f,g and Supplementary Fig. S2a,b). Inclusion of LIF had no effect above 4F on iPS colony yield, eliminating this cytokine as an enhancer of reprogramming to iPS (Fig. 3h). In contrast, IL6 addition to the medium resulted in a 2.2 ± 0.2-fold increase in mouse iPS colony yield (mean $\pm$ s.e.m., $n = 3$).

As IL6-mediated Jak/Stat signalling is known to induce c-Myc, we investigated whether transient extrinsic IL6 could replace stably integrated oncogenic viral c-Myc, which is frequently associated with human cancers and is one of the four transduced factors routinely used to generate iPS. This possibility was also suggested by the heterokaryon NOISeq analysis, which revealed induction of *MYC* among the GP130 signature genes (Fig. 2d). To determine whether IL6 induces endogenous *c-Myc* expression, we treated MEFs with IL6 and assessed c-Myc transcript levels by qPCR. We detected a 3.1 ± 0.7-fold induction of endogenous *c-Myc* transcripts (mean $\pm$ s.e.m., $n = 3$), which peaked at 60 min post IL6 stimulation (Fig. 4a). Importantly, control MEFs treated with LIF, or medium change alone (mock), did not induce significant changes in *c-Myc* expression, demonstrating that the observed gene induction was specific to IL6 addition (Fig. 4a). To investigate whether IL6 could replace viral c-Myc, we generated a three-factor lentivirus containing Oct4, Klf4 and Sox2 (OKS, 3F) and a second virus encoding only c-Myc for direct comparison of 3F reprogramming with and without exogenous c-Myc (Fig. 4b). At day 12 post 3F transduction, no colonies were obtained for 3F only, or 3F + mock virus. Notably, daily addition of IL6 to the reprogramming medium after transduction with 3F virus resulted in generation of $Nanog^+$ colonies, to levels comparable to the 3F + c-Myc control (Fig. 4c). These results show that addition of IL6 to the medium can replace the requirement for stable integration of the oncogene c-Myc.

As c-Myc is known to function early during iPS reprogramming[13], we postulated that IL6 might also play an early role. To explore this hypothesis, we assessed the minimum time window required for IL6 exposure to achieve maximum colony yield. Addition of IL6 to 3F-transduced MEFs sufficed to achieve maximal iPS colony yield when added to the medium daily during the first four days of reprogramming and no increase in yield was seen if daily exposure to IL6 continued beyond this time point (Fig. 4d and Supplementary Table S1). Hence, IL6 stimulation is necessary only transiently during an early time window at the onset of reprogramming. Exogenous addition of IL6 to MEF cultures is dispensable 4 days post viral transduction for iPS generation and subsequent iPS maintenance. These results establish IL6 as an early regulator of iPS generation that does not require stable integration and acts transiently through addition to the media.

To characterize the pluripotency quality of IL6 + 3F-derived iPS clones, we first determined whether they expressed a similar panel of mouse ESC pluripotency genes[14,15]. All 15 genes analysed by PCR with reverse transcription were expressed in IL6 + 3F-derived iPS clones, 4F-derived iPS clones and mouse ESC, but not expressed in MEF controls (Supplementary Fig. S2c). IL6 + 3F-derived iPS clones were confirmed to be free of exogenous viral c-Myc integration using genomic PCR (Supplementary Fig. S2d). A more rigorous test of pluripotent potential entailed an *in vivo* analysis of the 3F + IL6-derived iPS clones by standard teratoma derivation. Staining and histological analyses of the resulting teratomas confirmed that IL6 + 3F-derived iPS clones could generate tissues representing all three germ layers (Fig. 4e). Global transcriptome analyses of 3F + IL6 iPS clones, 4F iPS clones and mouse ESC confirmed these observations, as gene expression of all iPS clones (with and without transduced c-Myc) clustered closely with mouse ESC (Pearson correlation $r = 0.96$; Fig. 4f). Together, these results demonstrate that the pluripotent potential of 3F + IL6-derived iPS clones is comparable to that of 4F iPS.

The demonstration that IL6 could augment reprogramming efficiency 2.2-fold in the presence of the 4F where c-Myc levels are in excess (Fig. 3h) suggested that IL6 has functional roles in addition to c-Myc induction that lead to enhanced iPS generation.

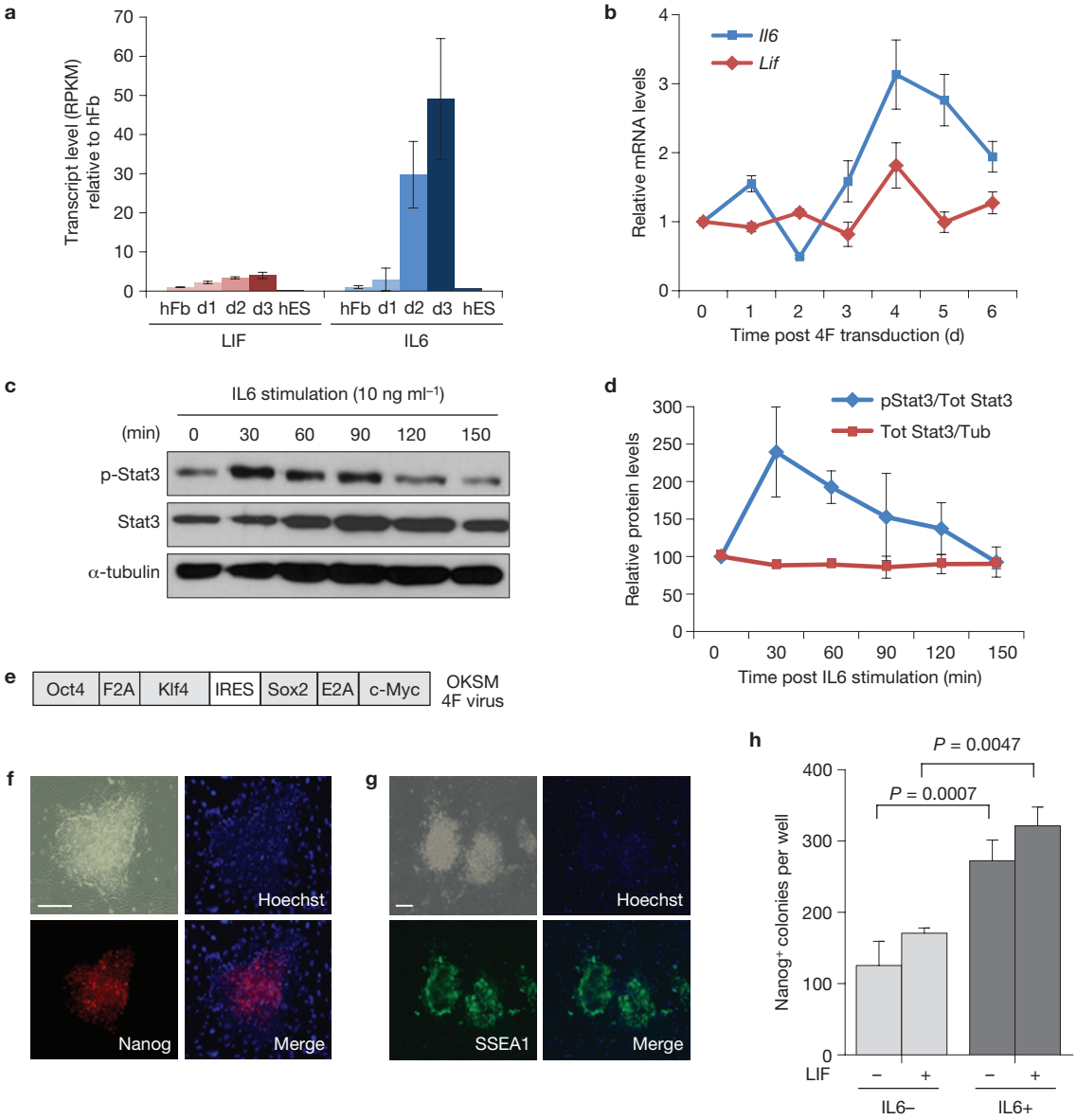


Figure 3 IL6, but not LIF, enhances 4F iPS generation. (**a**) Relative transcript levels of LIF and IL6 during heterokaryon reprogramming, normalized to unfused human fibroblasts (hFb). hES, human ESC. Data are represented as mean RPKM with range ($n = 2$). (**b**) Relative mRNA levels for Il6 and Lif gene expression during 4F iPS generation in MEFs, normalized to mouse *Hprt* and mouse *Gapdh*, as determined by qPCR. Data are represented as mean ± s.e.m., $n = 4$. (**c**) Induction of pStat3 Tyr-705 in fibroblasts on IL6 stimulation. A representative western blot over a 2.5-h time course is shown. Full blot is shown in Supplementary Fig. S5. (**d**) Quantification of pStat3 induction on IL6 stimulation, normalized to α-tubulin protein levels. Mean with range is shown ($n = 2$). Maximal pStat3 is observed 30 min post-IL6 stimulation. (**e**) Schematic of the 4F lentiviral vector used to generate 4F virus for mouse iPS induction. (**f**) Phase and immunofluorescence images of a representative iPS colony stained for the mouse pluripotency marker Nanog (red) and nuclear Hoechst (blue). Scale bar, 200 μm. (**g**) Representative iPS colonies, presented as in **f** with staining for the mouse pluripotency marker SSEA-1 (green). (**h**) Nanog$^+$ iPS colony formation from MEFs transduced with mock or 4F virus in the absence of additional cytokines and cultured with or without IL6 and or LIF. Data are represented as mean ± s.e.m., $n = 3$. P values were calculated using Student's unpaired two-tailed t-test. Source data are presented in Supplementary Table S1.

Therefore, we explored the role of the pro-survival kinase Pim1, an additional downstream target of IL6-mediated GP130 signalling identified as significantly upregulated in heterokaryons (Fig. 2d). *Pim1* was induced 3.4 ± 0.5-fold in MEFs (mean ± s.e.m., $n = 3$), peaking at 90 min post-IL6 stimulation (Fig. 5a). As with *c-Myc*, induction of *Pim1* was specific to IL6 stimulation and was not observed in LIF- or mock-treated MEFs. The role of Pim1 activity in IL6-mediated reprogramming was assessed by loss of function using two distinct and highly specific Pim1 inhibitors[16,17] (Supplementary Fig. S3). Inhibition of Pim1 activity during 3F + IL6 reprogramming resulted in a significant reduction in Nanog$^+$ colonies (Fig. 5b), establishing a functional role for Pim1 downstream of IL6 in enhancing iPS reprogramming frequency. Pharmacological inhibition of Pim1 in 4F + IL6 reprogramming (where c-Myc levels are in excess) reduced the IL6-mediated 2-fold increase in iPS colony formation back to 4F-only levels (Fig. 5c). 4F reprogramming in the presence of Pim1

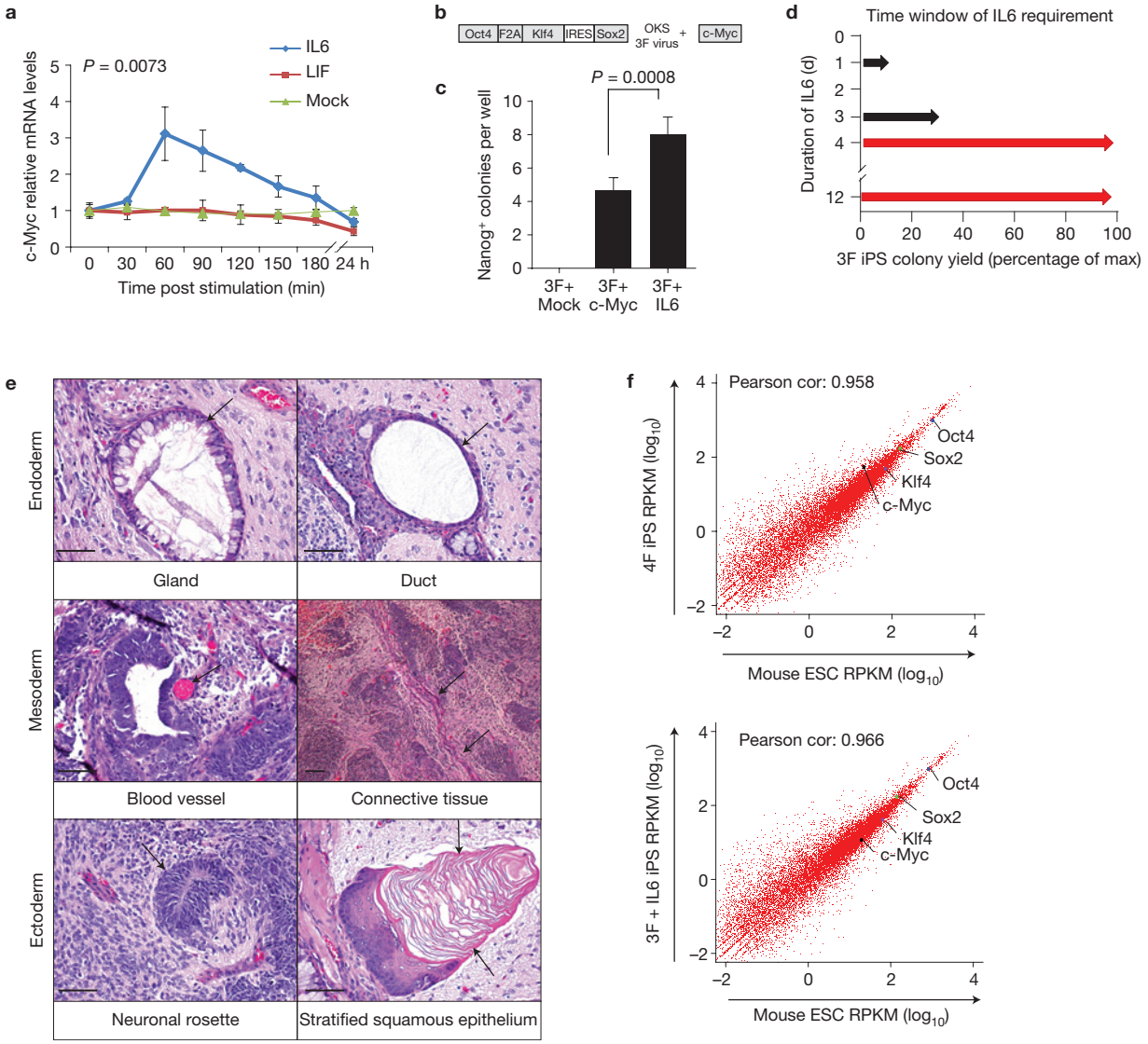


Figure 4 Transient exposure to IL6 in medium replaces viral c-Myc in reprogramming to iPS and is required only at the onset of 3F reprogramming. (**a**) Time course of relative mRNA expression levels for endogenous c-Myc in MEFs after stimulation with IL6, LIF or mock (medium change alone), normalized to mouse *Actb* and $t = 0$ min. Data are represented as mean ± s.d., $n = 3$. *P* values were calculated using Student's unpaired two-tailed *t*-test. Source data are presented in Supplementary Table S1. (**b**) Schematic of the 3F lentiviral vector. c-Myc was restored using a second virus when indicated. (**c**) Nanog$^+$ iPS colony formation from MEFs transduced with 3F virus with or without IL6 or co-transduction with mock or c-Myc virus. Data are represented as mean ± s.e.m., $n = 6$. *P* values were calculated using Student's unpaired two-tailed *t*-test. Source data are presented in Supplementary Table S1. (**d**) Nanog$^+$ iPS colony formation from MEFs on IL6 withdrawal from the medium on the day indicated. Maximum colony yield is obtained when IL6 is added to the medium for the first 4 days of reprogramming. Treatment beyond day 4 did not increase iPS colony number. (**e**) Teratoma histological analysis from 3F+IL6 iPS clones 1 and 4 showing contribution to all three germ layers. Magnification is ×200. Scale bars, 50 μm. (**f**) Comparison of global gene expression between individually generated iPS (4F and 3F+IL6) clones and mouse ESC (average of $n = 2$ for each). The Pearson's correlation values are shown for iPS versus mouse ESC gene expression.

inhibitors also reduced colony formation. These results establish Pim1 as a downstream target of IL6 that augments iPS generation and also plays an endogenous role in reprogramming.

To confirm that the induction of *c-Myc* and *Pim1* was indeed attributable to IL6, we performed loss of function experiments in MEFs using short hairpin RNAs (shRNAs) targeting the Il6 receptor (Supplementary Fig. S4). Knockdown of *Il6r* in MEFs by 4 distinct shRNAs blocked IL6-mediated phosphorylation of Stat3 (Fig. 5d). Moreover, loss of Il6r function resulted in a decreased induction of the Il6/Gp130 target genes *c-Myc*, *Pim1* and *Mcl1* (Fig. 5e), establishing the specificity of the shRNAs.

To assess whether Pim1 could enhance 4F reprogramming, we overexpressed Pim1 in MEFs, confirmed at the protein and mRNA level (Fig. 5f,g). Pim1 overexpression led to an increase in 4F iPS colony generation, demonstrating that Pim1 can substitute for exogenous IL6 during the reprogramming process (Fig. 5h). It has recently been reported that the TLR3/NFκB innate immunity pathway increases the efficiency of reprogramming[18], and that Pim1 can activate NFκB

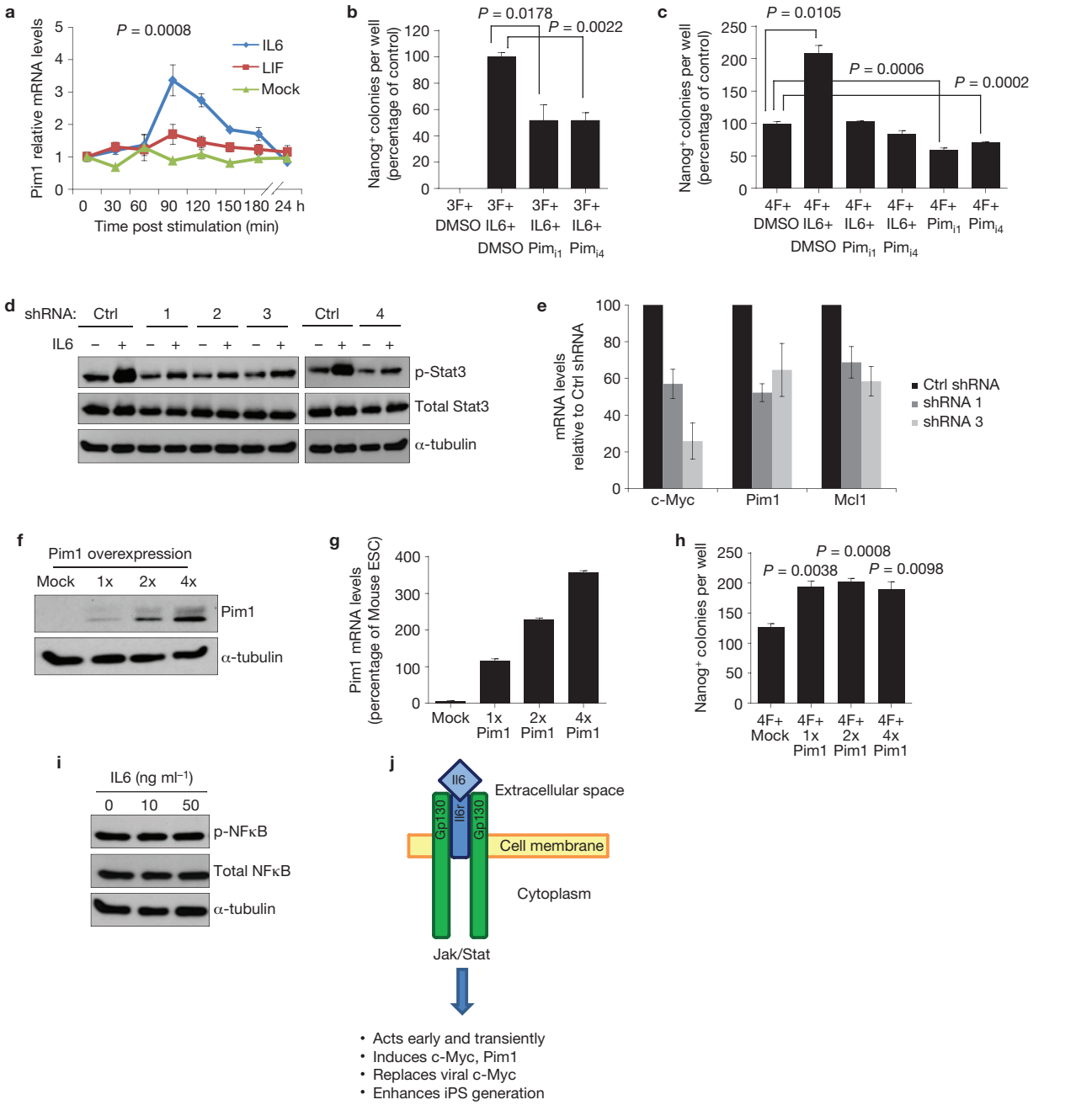


Figure 5 Pim1, a pro-survival gene, acts downstream of IL6 to enhance reprogramming to iPS. (**a**) Time course of Pim1 mRNA expression levels in MEFs after stimulation with IL6, LIF or mock (medium change alone). Gene expression was normalized to *Actb* and $t = 0$ min, as determined by qPCR (mean ± s.d., $n = 3$). Data are presented in Supplementary Table S1. (**b**) Nanog$^+$ iPS colony formation from MEFs transduced with 3F virus with or without addition of IL6, Pim1 inhibitor or dimethylsulphoxide (DMSO; mean ± s.e.m., $n = 3$, shown as a percentage of the control). Data are presented in Supplementary Table S1. (**c**) Nanog$^+$ iPS colony formation from MEFs transduced with 4F virus, with or without IL6, Pim1 inhibitor or dimethylsulphoxide as indicated, shown as a percentage of the control (mean ± s.e.m., $n = 3$). Data are presented in Supplementary Table S1. (**d**) Western blots of p-Stat3 induction in MEFs transduced with shRNAs against Il6r or control shRNA, assessed 30 min post stimulation with IL6 or mock. MEFs were stimulated 3 days after shRNA virus transduction. Total Stat3 levels and α-tubulin levels are shown as a reference and loading control, respectively. (**e**) Mouse c-Myc, Pim1 and Mcl1 mRNA levels in MEFs transduced with shRNA virus targeting Il6r or control shRNA virus and stimulated with IL6, as assessed by qPCR. mRNA expression levels were normalized to mouse *Actb* and assessed at peak control shRNA levels post IL6 stimulation occurring at 60, 90 and 120 min post stimulation, respectively (mean ± s.d., $n = 3$). Data are presented in Supplementary Table S1. (**f**) Western blot for Pim1 overexpression in MEFs. (**g**) Pim1 mRNA levels in MEFs transduced with Pim1 virus, as determined by qPCR, normalized to *Actb* and *Pim1* levels in mouse ESC (mean ± s.d., $n = 3$). Data are presented in Supplementary Table S1. (**h**) Nanog$^+$ iPS colony formation from MEFs transduced with 4F and Pim1 virus or mock virus (mean ± s.e.m., $n = 3$). Data are presented in Supplementary Table S1. (**i**) Western blot for p-NFκB protein levels in IL6-treated MEFs, 3 h post-stimulation with IL6. NFκB and α-tubulin are shown as controls. (**j**) Model for early IL6-mediated enhancement of iPS generation. The yellow rectangle indicates the cell membrane containing the Gp130 and Il6 receptor. Il6 ligand is depicted as a diamond. Extracellular space and cytoplasm are indicated. *P* values were calculated using Student's unpaired two-tailed *t*-test. Full blots are shown in Supplementary Fig. S5.

(ref. 19). We investigated whether exogenous IL6 could induce activation of NFκB; however, we saw no further induction of p-NFκB in MEFs on IL6 stimulation, suggesting that IL6/Pim1 signalling may be acting through an alternative mechanism (Fig. 5i). Together these results demonstrate that IL6 acts early to enhance reprogramming and that this effect is mediated by Pim1 in a manner distinct from that of LIF (Fig. 5j).

In summary, we demonstrate here that heterokaryon global RNA profiling can reveal previously unrecognized genes and pathways that can be translated to iPS generation and that these reprogramming mechanisms share common regulators. This contrasts with several reports suggesting that the reprogramming of cells towards pluripotency by different approaches is likely to entail distinct and unrelated mechanisms in which oocytes and heterokaryons are thought to similarly reprogram in the absence of cell division, unlike iPS (refs 21–23).

Notably, heterokaryons can reveal regulators that are transiently expressed. These regulators would typically be missed, as they are not highly expressed in either parental cell type or end-stage pluripotent cells (embryonic stem or iPS). Although heterokaryons differ from iPS in many ways, they have noteworthy advantages over nuclear transfer (higher throughput) or iPS (more rapid and efficient), for elucidating molecules mediating reprogramming towards pluripotency. We first showed this for a candidate regulator with a role early in DNA demethylation, the deaminase AID (refs 3,20) and now show a transient early role for IL6. The data presented here demonstrate that candidates identified in heterokaryons can be successfully translated to iPS.

Our data provide the first evidence that in reprogramming to iPS, IL6 can functionally replace c-Myc, an oncogene found in most human cancers[24]. The demonstration that IL6 can augment reprogramming in the presence of the 4F implies that the effects mediated by IL6 are not solely due to the induction of c-Myc alone. IL6-mediated induction of the serine/threonine kinase Pim1 accounts, at least in part, for the increased efficiency of iPS generation. Notably, like the action of IL6, the induction of c-Myc and Pim1 is transient. A corollary to our finding is the reported stage-specific role of IL6 signalling in regulating survival and senescence in lymphomagenesis[25]. Others have shown that c-Myc promotes both reprogramming and apoptosis during iPS generation[26]. Pim1 may counteract the apoptotic effects of excess c-Myc by phosphorylating and inactivating the mitochondrial pro-apoptotic protein BAD (ref. 27). As IL6 signalling has pleiotropic effects, it is likely that other downstream effectors (in addition to Pim1, Stat3, Mcl1 and c-Myc) play a contributing role.

As in development, our findings highlight that stage-specific short-term responses to cytokines can play a major role in promoting cellular reprogramming. Exogenous secreted factors such as IL6 are involved in reprogramming to pluripotency during embryogenesis, initiated by the fusion between spermatozoa and oocytes (the quintessential fusion reprogramming scenario). Notably, IL6 also plays a role in the peri-implantation environment post-fusion[28]. Others have used heterokaryons; however, no attempts were made to translate the findings to iPS or reprogramming in general[29], even when bi-species transcriptome profiling was performed[30]. We show here that RNA-seq expands the use of heterokaryons as a discovery tool, enabling unprecedented molecular analysis of reprogramming leading to the identification and validation of unanticipated regulators.

A search for additional early regulators of iPS generation is now possible, including both facilitators of reprogramming and factors that impede the process and constitute barriers to reprogramming such as those that maintain the somatic cell phenotype[31–33]. Indeed, heterokaryons are particularly well suited to discerning molecular regulators that act at the onset of reprogramming, both in a positive and negative manner, including non-coding RNA species, such as microRNAs (ref. 34) which could enhance the efficiency of the reprogramming process. The heterokaryon approach should have broad applications for gene discovery and understanding the mechanisms underlying reprogramming, including directed differentiation[1,2,31,35,36]. Biological ligands, such as IL6, which induce endogenous signalling pathways are particularly advantageous as they can be added transiently to the medium, overcoming safety issues that plague genetic manipulations and the stable integration of genes, such as c-Myc (ref. 37). Such studies will advance our knowledge of cellular reprogramming and improve the safety and efficacy of iPS or other reprogrammed cell types and their differentiated derivatives for use in regenerative medicine. □

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Supplementary Information is available in the online version of the paper

ACKNOWLEDGEMENTS

We thank M. Wernig and S. Marro for providing expert guidance with iPS protocols and for the STEMCAA 4F lentiviral construct and R. Tran-Bussat for isolating MEFs. We thank S. L. Oliver, S. Corbel and D. Burns for their insightful comments on the manuscript, the ENCODE Project, in particular Barbara Wold who was very helpful in the initiation of this work, and the Stanford FACS Facility, PAN Facility, and Stem Cell Institute Genome Facility. This work was supported by NSF and Bio-X Graduate Research Fellowships to J.J.B., by NIH grant R01 HG005717 and HG006018 to W.H.W., and by CIRM grant RB1-01292, NIH U01 HL100397, NIH R01 AG009521 and the Baxter Foundation to H.M.B.

AUTHOR CONTRIBUTIONS

J.J.B. and H.M.B. designed the experiments, interpreted the results and wrote the manuscript. J.J.B. performed all of the experiments, H.J. designed SeqMap and M.L. generated the RPKM values under the direction of W.H.W. S.S. performed the NOISeq analysis.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at www.nature.com/doifinder/10.1038/ncb2835
Reprints and permissions information is available online at www.nature.com/reprints

1. Vierbuchen, T. & Wernig, M. Molecular roadblocks for cellular reprogramming. *Mol. Cell* **47,** 827–838 (2012).
2. Yamanaka, S. & Blau, H. M. Nuclear reprogramming to a pluripotent state by three approaches. *Nature* **465,** 704–712 (2010).
3. Bhutani, N. *et al.* Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* **463,** 1042–1047 (2010).
4. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5,** 621–628 (2008).
5. Tarazona, S., Garcia-Alcalde, F., Dopazo, J., Ferrer, A. & Conesa, A. Differential expression in RNA-seq: a matter of depth. *Genome Res.* **21,** 2213–2223 (2011).
6. Thomas, P. D. *et al.* PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucleic Acids Res.* **31,** 334–341 (2003).
7. Hirano, T., Ishihara, K. & Hibi, M. Roles of STAT3 in mediating the cell growth, differentiation and survival signals relayed through the IL-6 family of cytokine receptors. *Oncogene* **19,** 2548–2556 (2000).

8. Heinrich, P. C. *et al*. Principles of interleukin (IL)-6-type cytokine signalling and its regulation. *Biochem. J.* **374,** 1–20 (2003).
9. Puthier, D., Bataille, R. & Amiot, M. IL-6 up-regulates mcl-1 in human myeloma cells through JAK/STAT rather than ras/MAP kinase pathway. *Eur. J. Immunol.* **29,** 3945–3950 (1999).
10. Assou, S. *et al*. A meta-analysis of human embryonic stem cells transcriptome integrated into a web-based expression atlas. *Stem Cells* **25,** 961–973 (2007).
11. Bhattacharya, B. *et al*. Gene expression in human embryonic stem cell lines: unique molecular signature. *Blood* **103,** 2956–2964 (2004).
12. Yang, J., Gao, C., Chai, L. & Ma, Y. A novel SALL4/OCT4 transcriptional feedback network for pluripotency of embryonic stem cells. *PLoS One* **5,** e10766 (2010).
13. Sridharan, R. *et al*. Role of the murine reprogramming factors in the induction of pluripotency. *Cell* **136,** 364–377 (2009).
14. Buganim, Y. *et al*. Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell* **150,** 1209–1222 (2012).
15. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126,** 663–676 (2006).
16. Cheney, I. W. *et al*. Identification and structure-activity relationships of substituted pyridones as inhibitors of Pim-1 kinase. *Bioorganic Med. Chem. Lett.* **17,** 1679–1683 (2007).
17. Pierce, A. C., Jacobs, M. & Stuver-Moody, C. Docking study yields four novel inhibitors of the protooncogene Pim-1 kinase. *J. Med. Chem.* **51,** 1972–1975 (2008).
18. Lee, J. *et al*. Activation of innate immunity is required for efficient nuclear reprogramming. *Cell* **151,** 547–558 (2012).
19. Nihira, K. *et al*. Pim-1 controls NF-kappaB signalling by stabilizing RelA/p65. *Cell Death Differ.* **17,** 689–698 (2010).
20. Pasque, V., Jullien, J., Miyamoto, K., Halley-Stott, R. P. & Gurdon, J. B. Epigenetic factors influencing resistance to nuclear reprogramming. *Trends Genet.* **27,** 516–525 (2011).
21. Byrne, J. A. *et al*. Producing primate embryonic stem cells by somatic cell nuclear transfer. *Nature* **450,** 497–502 (2007).
22. Hanna, J. *et al*. Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* **462,** 595–601 (2009).
23. Bhutani, N. *et al*. A critical role for AID in the initiation of reprogramming to induced pluripotent stem cells. *FASEB J.* **27,** 1107–1113 (2013).
24. Lin, C. Y. *et al*. Transcriptional amplification in tumor cells with elevated c-Myc. *Cell* **151,** 56–67 (2012).
25. Gilbert, L. A. & Hemann, M. T. Context-specific roles for paracrine IL-6 in lymphomagenesis. *Genes Dev.* **26,** 1758–1768 (2012).
26. Soufi, A., Donahue, G. & Zaret, K. S. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* **151,** 994–1004 (2012).
27. Aho, T. L. *et al*. Pim-1 kinase promotes inactivation of the pro-apoptotic Bad protein by phosphorylating it on the Ser112 gatekeeper site. *FEBS Lett.* **571,** 43–49 (2004).
28. Van Mourik, M. S., Macklon, N. S. & Heijnen, C. J. Embryonic implantation: cytokines, adhesion molecules, and immune cells in establishing an implantation environment. *J. Leukocyte Biol.* **85,** 4–19 (2009).
29. Tsubouchi, T. *et al*. DNA synthesis is required for reprogramming mediated by stem cell fusion. *Cell* **152,** 873–883 (2013).
30. Foshay, K. M. *et al*. Embryonic stem cells induce pluripotency in somatic cell fusion through biphasic reprogramming. *Mol. Cell* **46,** 159,–170, (2012).
31. Holmberg, J. & Perlmann, T. Maintaining differentiated cellular identity. *Nat. Rev. Genet.* **13,** 429–439 (2012).
32. Silva, J., Chambers, I., Pollard, S. & Smith, A. Nanog promotes transfer of pluripotency after cell fusion. *Nature* **441,** 997–1001 (2006).
33. Pasque, V., Halley-Stott, R. P., Gillich, A., Garrett, N. & Gurdon, J. B. Epigenetic stability of repressed states involving the histone variant macroH2A revealed by nuclear transfer to *Xenopus* oocytes. *Nucleus* **2,** 533–539 (2011).
34. Huang, X. A. & Lin, H. The miRNA regulation of stem cells. *WIREs Membr. Transp. Signal.* **1,** 83–95 (2012).
35. Graf, T. Historical origins of transdifferentiation and reprogramming. *Cell Stem Cell* **9,** 504–516 (2011).
36. Salero, E. *et al*. Adult human RPE can be activated into a multipotent stem cell that produces mesenchymal derivatives. *Cell Stem Cell* **10,** 88–95 (2012).
37. Marson, A. *et al*. Wnt signaling promotes reprogramming of somatic cells to pluripotency. *Cell Stem Cell* **3,** 132–135 (2008).

DOI: 10.1038/ncb2835

METHODS

Heterokaryon generation and isolation. Procedures for the generation of the GFP ESC and dsRed human fibroblasts were described previously[2]. Briefly, 500,000 human DsRed puro^R fibroblasts and 3×10^6 GFP mouse ESC were co-cultured overnight in embryonic stem cell medium and fused using PEG 1500 (Roche). After fusion, puromycin was added at 1 μg ml^{-1} to select against the non-resistant GFP mouse ESC. For sorting, plates were trypsinized on the day indicated, washed in FACS buffer (2.5% goat serum and 2 mM EDTA in PBS) and sorted on a modified BD FacsVantage running Diva hardware and software. The experiments generating the representative images in Fig. 1b were successfully repeated 12 times.

Preparation of RNA-seq libraries and sequencing. To obtain a complete profile of heterokaryon reprogramming, heterokaryon samples of $n > 30{,}000$ were isolated in duplicate over a 3-day time course. Control samples included human fibroblasts co-cultured with mouse ESC to distinguish transcripts induced by secreted factors compared with those requiring cell fusion, and a homokaryon (human fibroblast–fibroblast fusion) sample to identify transcriptional changes induced by PEG-mediated fusion. Unfused human fibroblast samples (Day 0) were included as a baseline reference point for all normalizations and calculations of gene induction. Total RNA was prepared from heterokaryons double sorted directly into Buffer RLT (RNeasy Micro Kit, Qiagen). Poly-A+ RNA was isolated from total RNA using Sera-Mag oligo dT beads (Thermo Scientific) and used for first strand synthesis (Superscript III, Life Technologies). Libraries were constructed according to standard Illumina protocols. The integrity and quality of RNA and complementary DNA were monitored using an Agilent Bioanalyzer 2100. Ultrahigh-throughput sequencing was performed on the Illumina HiSeq 2000 or 2500 for 50-nucleotide reads.

RNA-seq read mapping and gene expression calculations. The first base was trimmed from the 51-base-pair (bp) reads. The reads were then mapped to the refSeq hg19 and mm10 annotations, as well as mouse ribosomal RNA and human rRNA. Mapping was done using SeqMap (ref. 38), allowing for a maximum of 3 mismatches. All reads mapping to rRNA were removed from the analysis, as were any reads that mapped to both human and mouse annotations. In calculating the gene expression (RPKM), we used the appropriate reduced length (effective length) of the transcript, as described below.

$$\mathrm{RPKM} = \frac{(10^9 \times \text{Number of reads mapped to a particular transcript})}{(\text{Total mappable reads} \times \text{Effective transcript length})}$$

Effective transcript length (in base pairs) was calculated by taking the full transcript length and reducing that length by 50 bp (the length of a read) and then by 1 additional bp for every 50 bp conserved section that matched perfectly between mouse and human (based on a sliding window across the transcript).

NOISeq differential gene expression analysis. The R-package previously implemented by the authors of NOISeq (ref. 5) was used to identify significantly differentially expressed genes. NOISeq is a non-parametric approach for the identification of differentially expressed genes from RNA-seq count data that aims to be robust against the number of available reads. Essentially, NOISeq creates a null or noise distribution of count changes by contrasting fold-change differences (M) and absolute expression differences (D) for all of the genes in samples within the same condition. This reference distribution is then used to assess whether the (M, D) values computed between two conditions for a given gene are likely to be part of the noise or represent a true differential expression. A probability threshold of 0.8 was used for the analysis. Candidate genes were identified by their induction or repression over the course of reprogramming compared with unfused fibroblast, homokaryon and co-culture controls. Homokaryon and co-culture controls were used to eliminate genes that were differentially expressed owing to the shock of fusion, or to account for gene expression changes induced by fibroblast co-culture with mouse ESC alone, respectively. In particular, we identified significantly differentially expressed genes in the heterokaryons when compared with unfused fibroblasts as well as co-culture controls. We then filtered genes that were differentially expressed in homokaryons when compared with co-culture controls.

Viral constructs and virus production. The 4F STEMCAA vector was a gift from M. Wernig (Stanford University, USA). The 3F virus was cloned by removing the Sox2–E2A–c-Myc cassette and replacing it with Sox2 lacking the terminal E2A tag and including a stop codon. Mouse c-Myc was amplified from the STEMCAA vector and cloned into pBabe-puro. Pim1 was amplified from mouse cDNA and cloned into pBabe-puro. Mock virus was generated by removing the Oct4–Klf4 and Sox2–c-Myc STEMCAA inserts and replacing them with the 2xFlag sequence and the puromycin resistance gene. pBabe-puro was used as mock virus where appropriate. Control and Il6r shRNA viral constructs were purchased from Sigma (TRCN0000305259, TRCN0000375089, TRCN0000068293, TRCN0000068294). For virus production, 5×10^6 HEK293T cells were seeded into 10 cm dishes and transfected with the vector of interest and appropriate packaging plasmids. Medium was changed 24 h later. Supernatants were collected at 48 and 72 h, pooled, passed through a 45 μm filter and concentrated ×40 by centrifugation at 18,000 r.p.m. for 1.5 h at 4 °C using an SW 41 Ti Rotor (Beckman Coulter). Concentrated virus was aliquoted and stored at −80 °C until use.

Real-time, genomic and RT–PCR. cDNA was synthesized using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems 4368814). Real-time PCR was performed using an ABI 7900HT Real-time PCR system using SYBR Green PCR Master Mix (#4309155, Applied Biosystems). Samples were cycled at 55 °C for 2 min, 95 °C for 10 min, and 40× (95 °C for 30 s, 60 °C for 60 s). For PCR with reverse transcription, cDNA was amplified for 30 cycles using the GoGreenTaq 2× Master Mix (M7122, Promega). Genomic DNA was prepared using the Dneasy Kit (Qiagen). For PCR using the GoGreen Taq 2x Master Mix (M7122, Promega), 100 ng of genomic DNA was used. All primer sequences are listed in Supplementary Table S2.

iPS generation and propagation. Into each well of a 6-well plate, 60,000 mouse embryonic fibroblasts (P1) were seeded. The next day, cells were transduced using concentrated viral supernatant. MEF medium (high-glucose DMEM, 15% FBS, penicillin/streptomycin, sodium pyruvate and non-essential amino acids) was changed daily, with or without the addition of 10 ng ml^{-1} of recombinant IL6 or LIF (1,000 units ml^{-1}). iPS were scored on day 7 for 4F-derived iPS and on day 12 for 3F-derived iPS. Pim1 inhibitor 1 and 4, where relevant, were used at 100 nM and 20 nM in dimethylsulphoxide, respectively and daily only for the first 4 days of reprogramming. Inhibitor or dimethylsulphoxide was added 30 min before IL6 stimulation. To establish iPS clones, single colonies were picked and expanded onto irradiated feeders and expanded for several passages.

Immunocytochemistry and fluorescent imaging. Cells were washed twice in PBS and fixed at the times indicated for 15 min at room temperature using 4% paraformaldehyde. Cells were washed 3 times with PBS and stored at 4 °C until staining. For staining, cells were blocked for 1 h with blocking buffer (20% goat serum and 0.3% Triton X-100). The antibodies used were anti-Nanog (Bethyl Labs A300-398A, Montgomery) and anti-SSEA-1 (MC-480 MAB4301, EMD Millipore, Billerica) used at 1:200. Secondary antibodies were DyLight 549 goat anti-Mouse IgG (115-506-003, Jackson ImmunoResearch) and Alexa-Fluor 488 Goat anti-Mouse IgM (A-11032 and A-21042, Invitrogen), used at 1:500. Hoechst 33242 was added at 1:10,000. Images were acquired using an Axio Observer.A1 microscope, a Zeiss Neofluar ×10/0.3 objective lens, an AxioCam MRC camera and an AxioVision 4.8.1 (Carl Zeiss MicroImaging). The experiments generating the representative images in Fig. 3f,g were successfully repeated 57 times.

Western blotting. All antibodies were from Cell Signaling Technologies, unless otherwise noted. Blotting was done using standard methods using antibodies for pStat3 (Cell Signaling #9145, 1:1,000), total Stat3 (Cell Signaling #9132, 1:1,000), alpha tubulin (Sigma, T6074, 1:10,000), pNFκB (Cell Signaling #3033, 1:1,000), total NFκB (Cell Signaling #6956, 1:1000), Pim1 (Santa Cruz Biotechnologies #13513, 1:500), pBad (Cell Signaling #5284, 1:1,000) and total Bad (Cell Signaling #9292, 1:1,000). Secondary antibodies were anti-mouse or rabbit HRP (GE Healthcare Life Sciences, NXA931 and NA9340V, 1:10,000). The experiments generating the representative image in Fig. 3c were successfully repeated 2 times.

Teratoma formation assay. Athymic nude 4-week-old male mice (Hsd:Athymic Nude-Foxn1nu; Harlan Laboratories) were used for teratoma formation assays. Confluent iPS in a 6-well dish were trypsinized, washed twice in PBS, and 2 million

iPS per clone were subcutaneously injected into the mice. Teratomas were allowed to grow for 4 weeks at which point the mice were euthanized. Teratomas were washed in PBS and fixed in 4% PFA overnight at room temperature before sectioning into 5 μm sections and staining with haematoxylin and eosin. All experiments were performed in accordance with Stanford University Animal Care and Use Committee guidelines.

Statistical analyses. Data are presented as mean ± s.e.m., unless otherwise noted. Comparisons between groups used the Student's *t*-test assuming two-tailed distributions. No statistical method was used to predetermine sample size. The R programming language was used to generate replicate comparison plots (Fig. 1f), sample hierarchical clustering (Fig. 1g) and heat maps (Fig. 2a,c,d). Hierarchical clustering was performed using the Euclidean distance metric and complete linkage method. The NOISeq method has been previously described[5], with further details provided in the Methods under the subheading NOISeq differential gene expression analysis.

Accession numbers. The complete data sets of the heterokaryon and iPS RNA-seq analyses are available from the NCBI GEO repository (GSE43549 and GSE46104).

38. Jiang, H. & Wong, W. H. SeqMap: mapping massive amount of oligonucleotides to the genome. *Bioinformatics* **24,** 2395–2396 (2008).

DOI: 10.1038/ncb2835

nature cell biology

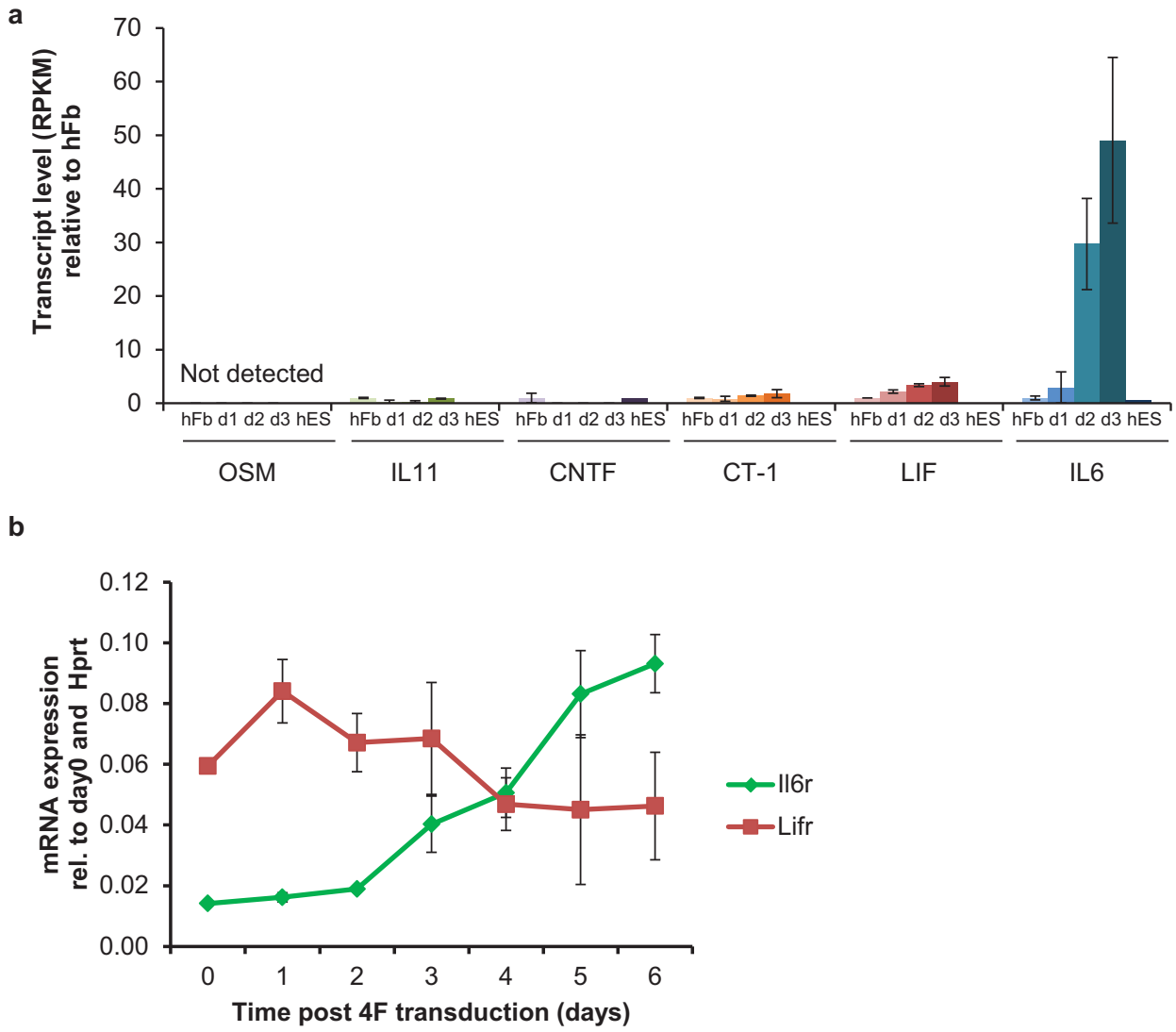


Figure S1 Expression of IL6 type cytokine family members, Il6r, and Lifr in heterokaryon or 4F virus induced reprogramming. **a,** Relative transcript levels of IL6 type cytokine family members OSM, IL11, CNTF, CT-1, LIF, and IL6 during heterokaryon reprogramming, normalized to unfused human fibroblasts. Data are represented as mean with range (n=2). **b,** MEFs were transduced with 4F virus and Il6 and Lif receptor transcript levels were monitored over a 6 day reprogramming timecourse by qPCR. *Il6r* was maximally induced 6.6 ± 0.7-fold while *Lifr* was maximally induced 1.4 ± 0.2-fold (mean ± s.d. for n=4 experiments). Gene expression was calculated relative to Day 0 and *Hprt* levels.

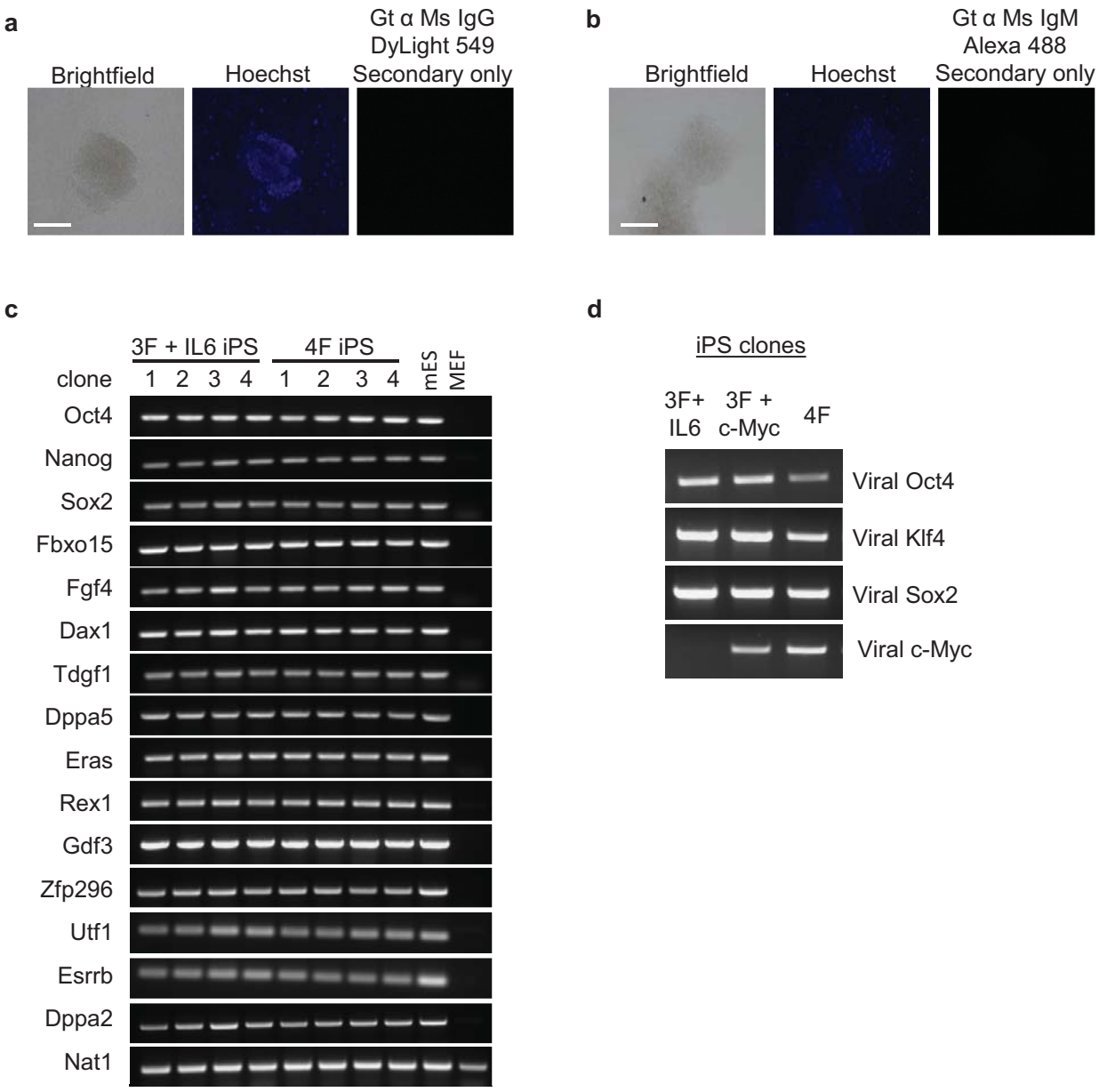
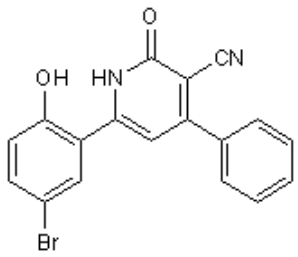


Figure S2 Characterization of iPS colonies and clones. **a, b,** iPS colonies were stained for Hoechst and the secondary antibody indicated. Scale bar = 200 μm. **c,** RT-PCR for mouse pluripotency genes from 3F + IL6 and 4F derived mouse iPS clones. Mouse ES cells and untreated MEFs are shown as positive and negative controls for gene expression, respectively. Full gels are shown in Supplementary Figure 5 . **d,** Genomic PCR for viral integration of the reprogramming factors Oct4, Klf4, Sox2 and c-Myc from 3F + IL6, 3F + c-Myc, and 4F virus derived iPS clones. Full gels are shown in Supplementary Figure 5.

a

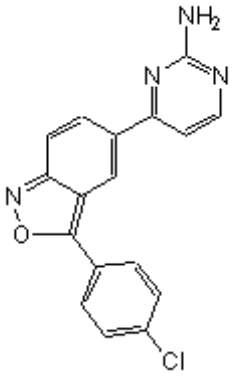
Pim1 inhibitor 1 (P_{i1})
(ATP competitor)



3-Cyano-4-phenyl-6-(3-bromo-6-hydroxy) phenyl-2(1*H*)-pyridone

b

Pim1 inhibitor 4 (P_{i4})
(blocks active site)



4-(3-(4-Chlorophenyl)-2,1-benzisoxazol-5-yl)-2-pyrimidinamine

c

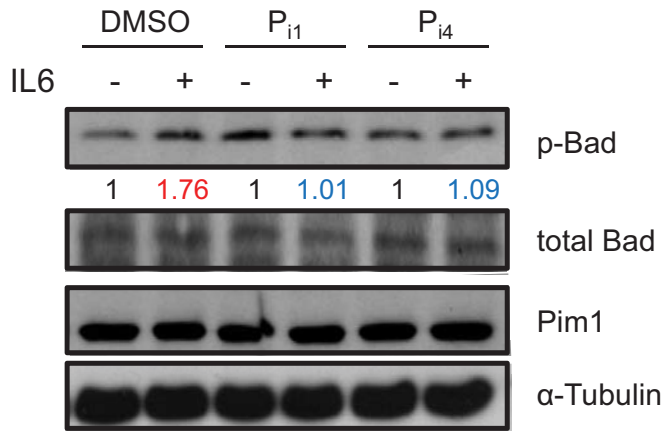
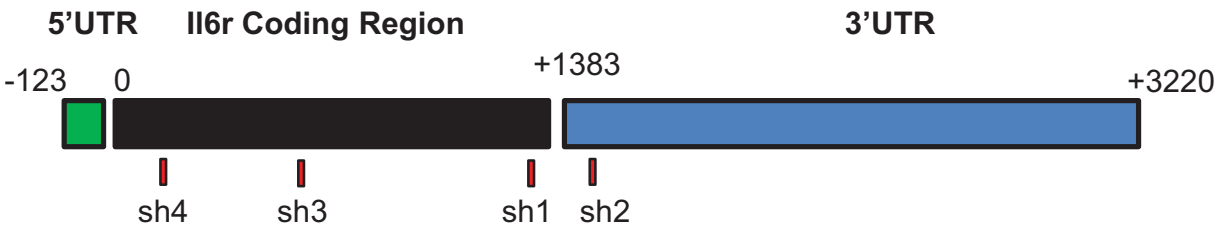


Figure S3 Pim1 inhibitors used in this study. **a,** 3-Cyano-4-phenyl-6-(3-bromo-6-hydroxy) phenyl-2(1*H*)-pyridone. **b,** 4-(3-(4-Chlorophenyl)-2,1-benzisoxazol-5-yl)-2-pyrimidinamine. **c,** Western blot for Bad phosphorylation (a Pim1 substrate) upon IL6 stimulation in the presence of DMSO or Pim1 inhibitors. Quantification is presented as levels of p-Bad normalized to levels of total Bad. Pim1 and α-Tubulin are shown as a reference and loading control. Full blots are shown in Supplementary Figure 5.

a



b

shRNA	Mouse Il6r target sequence
1	AGCAACAGAGACTACTTATTC
2	GTGGAAGCTTGGTTCCGATTT
3	CGAAGCGTTTCACAGCTTAAA
4	GCAGCAGGCAATGTTACCATT

c

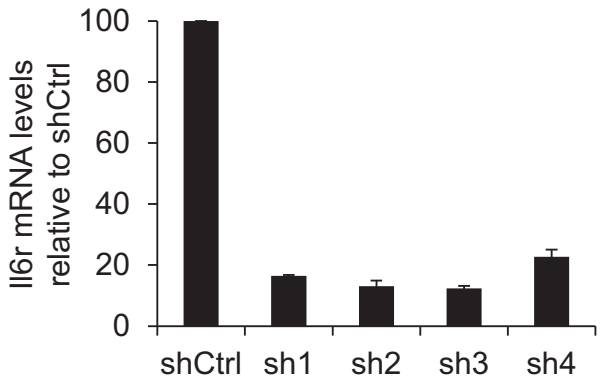


Figure S4 Il6r loss of function and validation. **a,** Diagram of the mouse Il6 receptor transcript and shRNA target regions. **b,** Target sequences of Il6r shRNAs. **c,** Knockdown of mIl6r in MEFs. Transcript levels of Il6r mRNA are shown relative to Actin and the shControl, assessed 3 days post transduction (mean ± s.d., n=3).

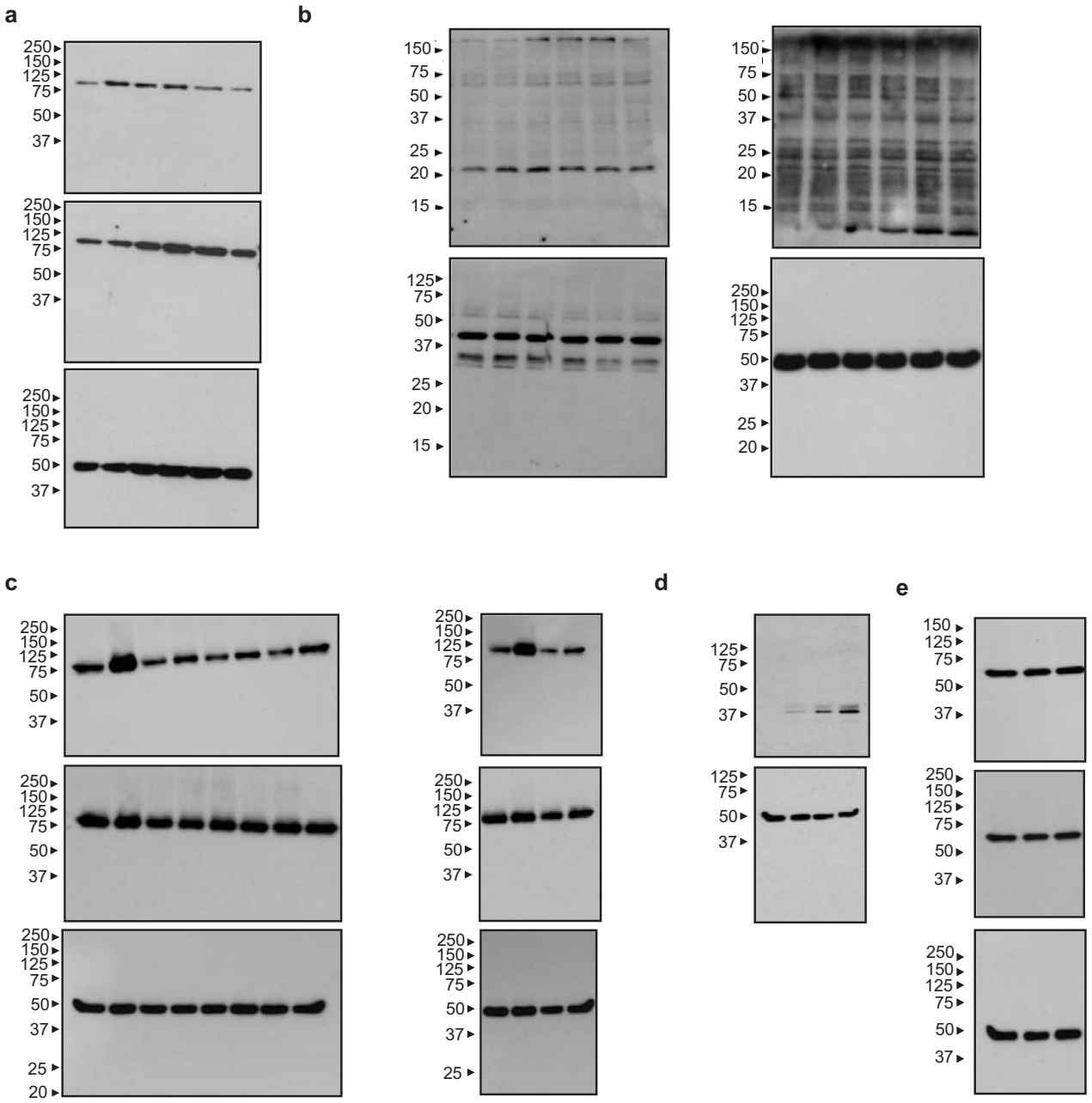


Figure S5 Full scans of western blots and gels. **a,** Western blots for Figure 3c. **b,** Western blots for Figure S3c. **c,** Western blots for Figure 5d. **d,** Western blots for Figure 5f. **e,** Western blots for Figure 5i. **f,** Gels for Fig S2c. **g,** Gels for Fig S2d.

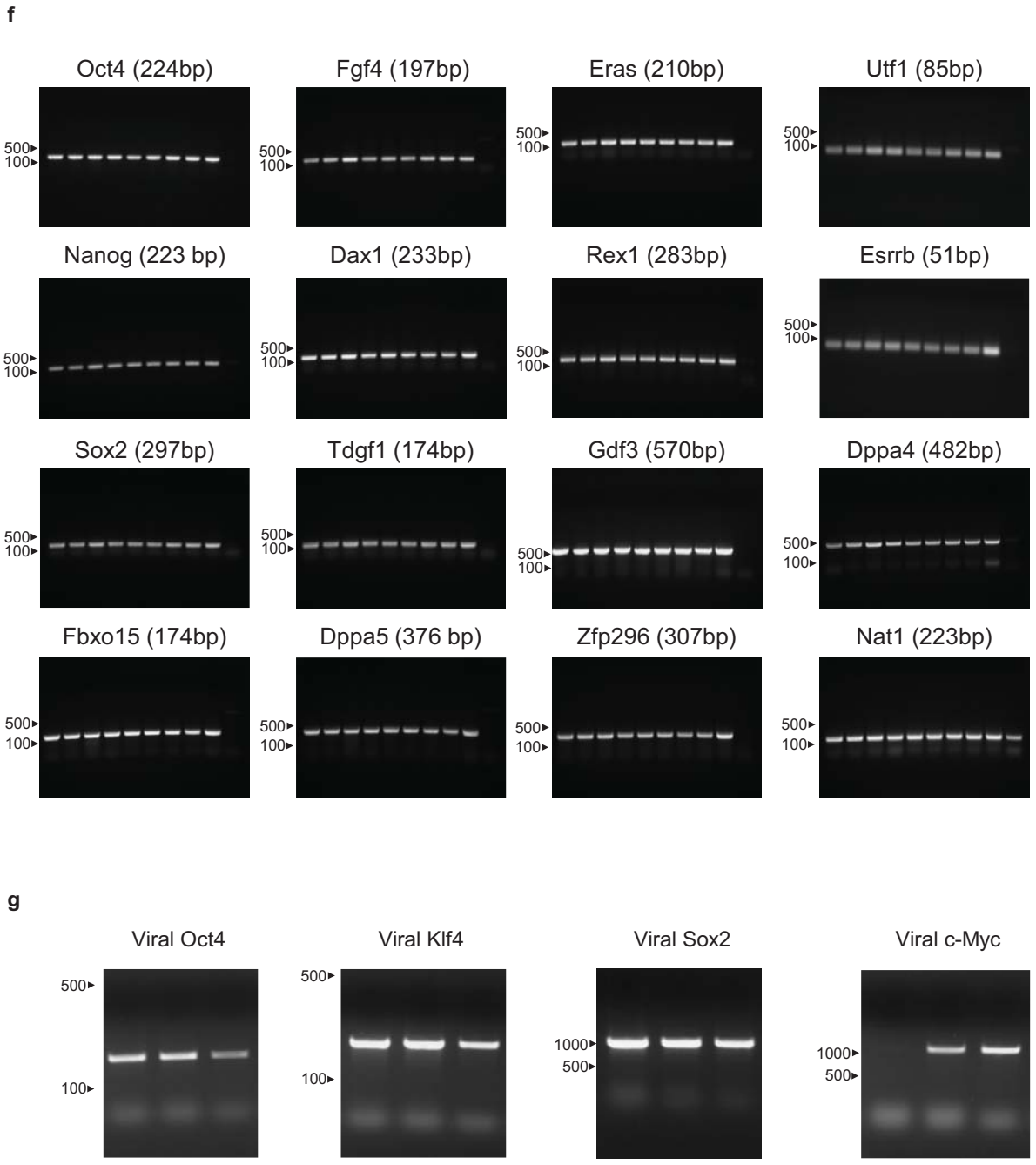


Figure S5 continued

Supplementary Table Legends:

Table S1 Source data
Table S2 Primer Sequences

Supplementary Note 1 - Blau

a

① Discard rRNA reads from analysis (0.3% of total reads)

② Map 50bp reads to mouse and human transcriptomes (mm10 and hg19) allowing for a maximum of 3 mismatches per read. (Total reads: ~15.5 * 10^6)

③ Conserved reads align to both transcriptomes. Unable to distinguish whether read originated from mouse or human transcriptome. Conserved reads were eliminated from the analysis. (4% of total reads)

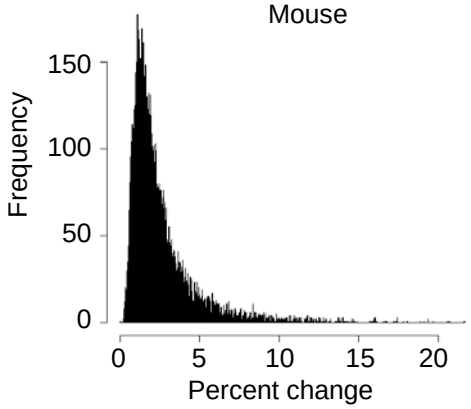
④ To calculate the gene expression (RPKM) for each transcript, the number of reads that map to a particular transcript was reduced (where applicable) to account for the elimination of any conserved reads from the data. The molar concentration of the reads per transcript was normalized to the appropriate reduced length (effective length) of the transcript, according to the equation below:

$$\text{RPKM} = \frac{(10\text{^}9 * Number\ of\ reads\ mapped\ to\ a\ particular\ transcript)/}{(Total\ mappable\ reads * Effective\ transcript\ length)}$$

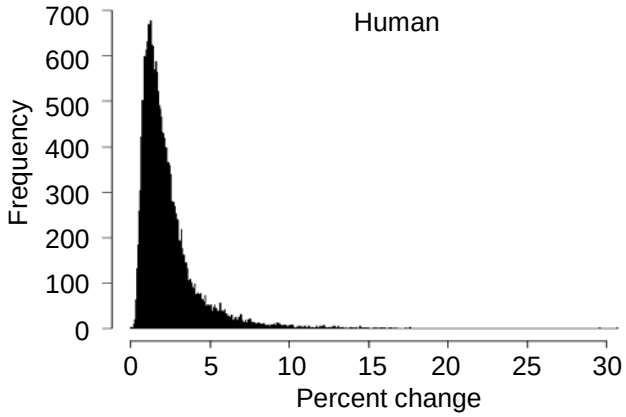
Effective transcript length (in base pairs) was calculated by taking the full transcript length and reducing that length by 50bp (the length of a read) and then by 1 bp for every 50bp conserved section which matched perfectly between mouse and human (based on a sliding window across the transcript).

For example, an hg19 transcript that is 2000bp has an effective length of 1950 if no 50bp window in it maps to mm10. If there are 25 windows of 50bp that have mappings to mm9, then the effective transcript length is 1950-25 = 1925.

b



c



Supplementary Note 1. Calculation of RPKM from bi-species RNA-Seq data

a, Correction for transcript length. **b, c,** Histogram of percent change in RPKM value for mouse and human transcripts. The correction results in an average change of 3.3% for human RPKM values and 2.5% for mouse RPKM values.