NAC1 Regulates Somatic Cell Reprogramming by Controlling Zeb1 and E-cadherin Expression

Francesco Faiola,1,3,4,5,6,* Nuoya Yin,4,5,6 Miguel Fidalgo,1,3 Xin Huang,1,3 Arven Saunders,1,2,3 Junjun Ding,1,3 Diana Guallar,1,3 Baoyen Dang,1,3 and Jianlong Wang1,2,3,*

1Black Family Stem Cell Institute
2Graduate School of Biomedical Sciences
3Department of Cell, Developmental and Regenerative Biology
Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
4State Key Laboratory of Environmental Chemistry and Ecotoxicology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, 100085, China
5College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China
6Co-first author
*Correspondence: faiola@rcees.ac.cn (F.F.), jianlong.wang@mssm.edu (J.W.)
http://dx.doi.org/10.1016/j.stemcr.2017.07.002

SUMMARY

Reprogramming somatic cells to induced pluripotent stem cells (iPSCs) is a long and inefficient process. A thorough understanding of the molecular mechanisms underlying reprogramming is paramount for efficient generation and safe application of iPSCs in medicine. While intensive efforts have been devoted to identifying reprogramming facilitators and barriers, a full repertoire of such factors, as well as their mechanistic actions, is poorly defined. Here, we report that NAC1, a pluripotency-associated factor and NANOG partner, is required for establishment of pluripotency during reprogramming. Mechanistically, NAC1 is essential for proper expression of E-cadherin by a dual regulatory mechanism: it facilitates NANOG binding to the E-cadherin promoter and fine-tunes its expression; most importantly, it downregulates the E-cadherin repressor ZEB1 directly via transcriptional repression and indirectly via post-transcriptional activation of the miR-200 miRNAs. Our study thus uncovers a previously unappreciated role for the pluripotency regulator NAC1 in promoting efficient somatic cell reprogramming.

INTRODUCTION

The discovery of induced pluripotent stem cells (iPSCs) marked a milestone in the development of strategies in regenerative medicine (Takahashi and Yamanaka, 2006). However, the generation of iPSCs is a lengthy and inefficient procedure that requires many processes such as global remodeling of chromatin and resetting of the epigenome (Apostolou and Hochedlinger, 2013; Papp and Plath, 2013; Watanabe et al., 2013). In recent years, many efforts have been focused on the identification of important players that could either facilitate (Theunissen and Jaenisch, 2014) or hinder (Winzi et al., 2014) the reprogramming process, leading to the discovery of NANOG as one of the reprogramming facilitators (Silva et al., 2006, 2009). Although NANOG accelerates the induction of pluripotency, its mechanisms of action are only partially understood (reviewed in Saunders et al., 2013).

In our pursuit to identify pluripotency and reprogramming factors that may modulate NANOG functions in reprogramming, we examined additional components of the NANOG interactome (Costa et al., 2013; Wang et al., 2006). In particular we identified nucleus accumbens-associated protein 1 (NAC1), a stem cell-enriched factor that also interacts with OCT4 (Ding et al., 2012) and SOX2 (Ding et al., 2015). NAC1 belongs to the bric-a-brac tram-trac broad complex/pox virus and zinc-finger (BTB/POZ) family of transcription factors (Mackler et al., 2003), and it is a ubiquitously expressed protein originally identified in the nucleus accumbens of the rat brain as a cocaine-inducible gene (Cha et al., 1997). Subsequently, NAC1 has been shown to play a role in the behavioral responses to psychostimulants (Mackler et al., 2000). In ESCs, NAC1 is a common interacting partner (Wang et al., 2006) of, and upstream modulator (Kim et al., 2008) for, many pluripotency factors and epigenetic regulators. However, its mechanistic actions in pluripotency are not defined. Besides being upregulated in pluripotent cells, NAC1 overexpression is also a hallmark of several type of cancers, including ovarian, cervical, and uterine (Ishikawa et al., 2010; Shih et al., 2011; Yeasmin et al., 2012). At the molecular level, NAC1 possesses a POZ domain N-terminally, and a BEN domain at the C terminus. The NAC1 POZ domain interacts with many factors, but is unique in that it does not contain a zinc-finger DNA-binding domain such as other POZ transcription factors. Therefore, it is believed that the NAC1 C-terminal BEN domain can mediate its binding to chromatin similarly to other BEN-containing transcriptional repressors (Dai et al., 2013).

We have begun to investigate the role of NAC1 in the maintenance and establishment of pluripotency and demonstrated that Nac1 was surprisingly dispensable for early embryo development (Yap et al., 2013). Not unexpectedly, thereafter we were able to derive Nac1 knockout
Figure 1. Nac1 Is Required for Somatic Cell Reprogramming

(A) Images of AP-stained wells for MEF-derived iPSCs upon control and Nac1 KD.
(B) Images of AP-stained iPS colonies upon control and Nac1 KD.
(C) Quantification of control and Nac1 KD iPS colonies scored based on intensity of AP staining.
(D) Images in bright field and GFP fluorescence for iPS colonies upon control and Nac1 KD MEF reprogramming.
(E) Quantification of control and Nac1 KD iPS colonies scored for GFP expression.
(F) Representative pictures of wells of AP-stained iPS derived from Nac1 WT (+/+), het (+/-), and null (-/-) MEFs.

(legend continued on next page)
(KO) mouse embryonic stem cells (mESCs), which undergo normal self-renewal and maintain pluripotency (our unpublished data). In this study, we dissected the functional contribution of NAC1 in establishing pluripotency during somatic cell reprogramming. We identified a critical role for NAC1 in transcriptionally and post-transcriptionally modulating E-cadherin and Zeb1 expression during the generation of iPSCs. In the absence of NAC1 functions, reprogramming is diverted to an anomalous state that can be fully rescued with the re-expression of E-CADHERIN, but not NANOG or ESRRB. Our data thus uncover a previously unappreciated reprogramming factor that plays an indispensable role, beyond the mesenchymal-to-epithelial transition (MET), in controlling E-cadherin expression and establishing the bona fide pluripotency of iPSCs.

RESULTS

NAC1 Depletion Impairs Somatic Cell Reprogramming

Several pluripotency factors, including NANOG, TET1, and TET2, are essential for somatic cell reprogramming, while dispensable for stem cell maintenance once pluripotency is established (Golipour et al., 2012). Although NAC1 functions in the maintenance of pluripotency in ESCs were mostly superfluous (our unpublished data), we decided to explore whether NAC1 could play a role in the establishment of pluripotency during somatic cell reprogramming.

To test the effects of NAC1 on reprogramming, we knocked down its expression in mouse embryonic fibroblasts (MEFs) harboring an Oct4 distal enhancer-driven GFP reporter that is only expressed in fully pluripotent iPSCs (Yeom et al., 1996). Subsequently, we transduced the four Yamanaka factors, as depicted in Figure S1A. Nac1 knockdown (KD) was efficient (Figure S1D, top) and minimally altered MEF proliferation (Figure S1B). However, it drastically affected the total number and morphology of alkaline phosphatase (AP) positively stained iPS colonies, as well as the intensity of the staining (Figures 1A–1C). When scoring for GFP-positive colonies, we found that NAC1 downregulation not only diminished total GFP-positive populations (Figure S1C), but also compromised the morphology of iPSC colonies, compared with scramble small hairpin RNA (shRNA) control (shSCR) (Figure 1D). Data from three independent reprogramming experiments revealed that the majority of the iPS colonies upon Nac1 KD were GFP negative (Figure 1E).

Since the Nac1 KO mouse was not embryonic lethal, we were able to derive Nac1 wild-type (WT), heterozygous (het), and null MEFs (Figure S1D, bottom). We then employed these fibroblasts in our reprogramming assays. As shown in Figures 1F and 1G, there was minimal difference in total number of iPS colonies upon AP staining among WT, het, and null cells. However, Nac1 null colonies stained less efficiently for AP, due to their pre-iPS-like morphology (Figures 1G and 1H) compared with WT and het cells. We also crossed our Nac1−/− mice with the Oct4-GFP reporter mouse and derived Nac1 mutant MEFs harboring the GFP reporter (Figure S1E, top). Consistent with Nac1 KD experiments, Oct4-GFP expression in reprogrammed colonies, which were less compact and with a disintegrated morphology, was lower in the absence of Nac1 (Figure S1E, bottom).

To assess whether Nac1-depleted iPSCs were indeed not fully reprogrammed, we switched the medium from serum/leukemia inhibitory factor (LIF) to 2i/LIF and allowed the reprogrammed cells to grow further for 10 days, to select for fully reprogrammed iPSC colonies and kill partially reprogrammed cells (Silva et al., 2008). As depicted in Figure 1I, about 50% of Nac1 WT iPSCs survived in the 2i/LIF medium. In contrast, null cells showed significantly lower rates of survival, suggesting that the vast majority of Nac1 null colonies were not fully reprogrammed (Figure 1I). In addition, the typical pre-iPS morphology of Nac1 null iPSCs was not due to a slower reprogramming process, because this morphology persisted for more than 15 passages in serum/LIF conditions (data not shown).

To analyze the effects of Nac1 depletion at the gene expression level during reprogramming, we picked several morphologically good Nac1 WT and het iPS colonies, and abnormal null iPS colonies under serum/LIF culture. We then investigated the expression of markers for pluripotency, early and late reprogramming, typical pre-iPS to iPS, and MET/cell adhesion, by qRT-PCR analyses. We found incomplete upregulation of a number of late-acting pluripotency genes including Nanog, Lin28, Tcl1, Dmmt3l, and Rex1, when Nac1 was deleted (Figure 1J). When we examined cell adhesion and MET genes, we found that epithelial cell-adhesion markers such as E-cadherin and

(G) Quantification of Nac1 WT, het, and null iPS colonies based on AP staining.
(H) Images of representative Nac1 WT, het, and null iPS colonies in bright field (top panel) and after AP staining (bottom panel).
(I) Pictures of duplicated wells for Nac1 WT, het, and null iPS colonies stained with AP upon incubation in serum/LIF or 2i/LIF medium.
(J) Average qPCR gene expression profiling for three Nac1 WT, three het, and nine null clonal iPSC lines. Indicated are selected pluripotency markers, late reprogramming markers, and MET/cell-adhesion genes. E-/N-cad stands for E-/N-cadherin.

Results in (C), (E), and (G) are averages ±SD from three independent experiments; *p < 0.05, **p < 0.01, ***p < 0.001. Statistical significance is relative to shSCR AP High (C), shSCR GFP+ (E) and WT AP+ (G). See also Figure S1.
Cdh3 were downregulated, whereas mesenchymal markers such as N-cadherin and Zeb1 were upregulated in cells depleted of Nac1, relative to WT control (Figure 1J). Many other late reprogramming markers previously defined (Sridharan et al., 2009) were also not properly regulated other late reprogramming markers previously defined and not other unknown reasons, we overexpressed human NAC1 (hNAC1) in null atypical iPS clones that had been passaged extensively but yet retained their aberrant iPS morphology. As shown in Figure 2A, re-introduction of the hNAC1 protein quickly and completely rescued the iPS morphology, suggesting that incomplete or abnormal reprogramming was due to the absence of NAC1. Since NAC1 can also function as a cytosolic protein in cancer (Wu et al., 2011; Yap et al., 2012), we tested whether this could be also true in pluripotent cells. Nuclear/cytosolic fractionation of mESCs, followed by western blot analyses, indeed revealed that NAC1 was abundantly expressed in both nuclear and cytosolic extracts (Figure S2B). To examine whether cytoplasmic functions of NAC1 may have contributed to the observed rescue of Nac1 null iPS morphology, we overexpressed a form of hNAC1 (hNAC1 Cyt), based on overall colony morphology, seemed to only slightly rescue the phenotype, despite its expression being higher than the WT

**Figure 2. NAC1 Re-introduction Rescues the Null iPS Phenotype**

(A) Bright-field images of Nac1 null iPSCs transfected with empty vector (EV), hNAC1 WT, and hNAC1 Cyt (cytosolic only).

(B) Western blot (WB) analyses showing overexpression of NAC1 WT and mutant.

(C) Quantification of NAC1 WT and mutant rescue efficiencies based on iPS colony morphology. Data are average percentages ±SD of three independent experiments; ***p < 0.001. Statistical significance is relative to EV control.

(D) Heatmap of time course microarray analyses for two Nac1 null atypical iPS clones during hNAC1 overexpression rescue. Indicated are genes known to have a role in reprogramming, and/or involved in MET/cell adhesion. Underlined are putative NAC1 target genes in mESCs.

(E) Gene ontology (GO) analyses of differentially regulated genes during the hNAC1 rescue. CC stands for cellular compartment. (F) qRT-PCR analyses of selected pluripotency and MET markers during the rescue. Data for Dppa3 are extracted from Figure S2D. Results are from three independent experiments with two different iPS lines and plotted as average ± SD; *p < 0.05, **p < 0.01. Statistical significance is relative to day 0. See also Figure S2.

**hNAC1 Overexpression Rescues the Null iPS Phenotype**

To confirm that the Nac1 null abnormal iPS phenotype we obtained was due to lack of Nac1 functional contribution and not other unknown reasons, we overexpressed human NAC1 (hNAC1) in null atypical iPS clones that had been passaged extensively but yet retained their aberrant iPS morphology. As shown in Figure 2A, re-introduction of
counterpart (Figure 2B). However, a closer examination of the individual colonies revealed that hNAC1 Cyt-rescued colonies were more like the empty vector (EV) condition than the WT one (Figure S2A). These results indicate that NAC1-dependent nuclear activities are required and mostly responsible for NAC1 functions during somatic cell reprogramming. To identify downstream target genes that could mediate NAC1 nuclear functions, we performed a time course hNAC1 rescue of the null phenotype and analyzed the gene expression profiles from day 0 to day 5 in two different clonal lines. After 5 days, the rescue was evident morphologically (Figure S2C). To assess the validity of the approach, we also checked the expression of a known NAC1 target gene in ESCs (Kim et al., 2008), Dppa3, during the time course of rescue by qRT-PCR. As shown in Figure S2D, the Dppa3 expression pattern was consistent with it being a NAC1 target and demonstrated that the hNAC1-dependent rescue became evident around day 2 and increased over time.

To examine gene expression at a global level during rescue, we selected day 0, 2, and 5 samples for microarray analyses in biological duplicates. As shown in Figure 2D, hundreds of genes were differentially regulated during the rescue (see Table S3 for a list of all the genes). We noticed that a number of genes differentially regulated in Nac1 WT and null iPS clones analyzed by qRT-PCR (Figure 1J) were not present in the heatmap. This likely reflects the dynamic regulation of these genes during the reprogramming process. Alternatively, the 2-fold cutoff stringency of the heatmap and/or the increased sensitivity of qRT-PCR analyses may have caused this discrepancy. Nevertheless, it is important to point out that two mesenchymal genes (Mmp2 and Vim), previously reported to be repressed during MET (reviewed in Esteban et al., 2012) and several pre-iPS genes (Akt1, Muc1, Ptgfrn, Rnase4, Creb3, and Iftm3), known to be downregulated during the pre-iPS to iPSC transition (Sridharan et al., 2009), are all repressed upon the rescue of Nac1 null atypical iPSCs with WT hNAC1. Conversely, many genes which were reported to be stimulated during the transition from pre-iPS to iPSCs (Golipour et al., 2012; Samavarchi-Tehrani et al., 2010; Sridharan et al., 2009) are also upregulated upon WT hNAC1 rescue. These include Acss1 and Kit, the adhesion genes Tjp3, Prol1, Gjb1, and Cdh3, and the pluripotency genes Nr5a2, Dppa3, Satb1, and Klf4, many of which are putative NAC1 target genes in mESCs (Kim et al., 2008) (underlined in Figure 2D). Interestingly, gene ontology (GO) assessments revealed that the main categories were factors involved in cell-cell adhesion and membrane functions, and several pluripotency and reprogramming factors known to be putative NAC1 targets or regulators of MET genes. Consistent with the microarray results and/or data in Figure 1J, we found upregulation of Nanog, Klf4, Satb1, Dppa3, E-cadherin, and Occludin (Ocln), and downregulation of Zeb1 (Figure 2F). To complement the time course gene expression during the ectopic hNAC1 rescue, we also checked a number of pluripotency, late reprogramming, and cell-adhesion markers in four additional Nac1 null iPSC clones, stably transfected with EV or WT hNAC1. As presented in Figure S2E, among the pluripotency factors tested, endogenous Oct4 expression was not significantly dependent on hNAC1 presence. In contrast, Nanog and Rex1 were appreciably upregulated upon ectopic hNAC1 expression. More interestingly, late reprogramming markers such as Dppa3 and Lefty2, and the two most abundantly expressed cadherins in ESCs, E-cadherin and Cdh3, were considerably activated upon hNAC1 rescue. These data further imply that hNAC1 might rescue the reprogramming phenotypes by completing the requisite MET process (reviewed in Shu and Pei, 2014) started in the early stage of reprogramming and/or activating late-acting pluripotency gene expression during the final stages of reprogramming. To test the requirement of NAC1 function beyond the MET stage in reprogramming, we generated NAC1 null neural progenitor cells (NPCs) from ESCs for reprogramming assays (see Figures S2F and S2G for NPC characterization). We found that most of the Nac1 null NPC-derived iPSC clones still retained the Nac1 null atypical phenotype observed during MEF reprogramming (see Figure S2H for morphology, and Figure S2I for quantification).

Together, these results suggest that the Nac1 null abnormal iPS morphology is not simply due to an incomplete MET process, but more likely to the partial/failed reactivation of both MET genes as well as pluripotency genes beyond the MET process.

**Ectopic Expression of E-CADHERIN Is Sufficient to Rescue the Nac1 Null iPSC Phenotype**

To further investigate the molecular mechanisms by which NAC1 regulates reprogramming, we tested whether known or potential NAC1-downstream target genes could recapitulate NAC1 functions. Among the pluripotency factors we chose Nanog, Esrrb, and Klf4 for the following reasons. Nanog was previously demonstrated to be a NAC1-regulated gene in mESCs (Kim et al., 2008). Moreover, our gene expression analyses during reprogramming revealed that Nanog could not be fully activated in the absence of NAC1 (Figures 1J, 2F, and S2E). ESRRB has been shown to recapitulate NANOG functions during reprogramming (Festuccia et al., 2012). KLF4 is the major transcription factor regulating genes involved in cell-cell adhesion during...
reprogramming (Li et al., 2010). Interestingly, ectopic expression of none of the transcription factors mentioned above was able to rescue the Nac1 null abnormal iPS phenotype, compared with ectopic hNAC1-dependent rescue (Figure 3A). We confirmed that those ectopic factors were properly expressed (Figures 3B and 3C). These data suggest that during reprogramming Nac1 lies downstream of Nanog, Esrrb, and Klf4 action in promoting bona fide pluripotency. Alternatively, there may exist NAC1-specific targets that are not controlled by those pluripotency factors. In addition, our results indicate that Nac1−/− aberrant iPSCs are distinct from Nanog−/− pre-iPSCs (Festuccia et al., 2012; Silva et al., 2009), and that Nac1 has regulatory functions other than, or downstream of, the activation of pluripotency genes during reprogramming.

To uncover such unique NAC1 functions in reprogramming, we decided to test the transgenic rescue of the Nac1−/− atypical iPS phenotype by overexpressing E-CADHERIN and OCLN, two factors involved in cell-cell adhesion. This choice was based on our findings that membrane- and cell-adhesion-related terms were enriched in the GO analysis of the differentially expressed genes upon hNAC1 rescue (Figure 2E), and that E-cadherin and Ocln were fully upregulated under our reprogramming settings only in the presence of NAC1 (Figures 1J, 2F, and S2E). In addition, E-CADHERIN has been shown to be vital
during somatic cell reprogramming (Chen et al., 2010; Redmer et al., 2011). Remarkably, E-CADHERIN overexpression alone phenocopied hNAC1 overexpression to generate typical dome-shaped compact iPSC colonies (Figure 3A), despite its ectopic expression being much lower than hNAC1 overexpression (Figure 3C). In contrast, OCLN did not rescue the Nac1 pre-iPS phenotype (Figure 3A). qRT-PCR analyses showed that Nanog (Figure 3D) together with E-cadherin (Figure 3E) were upregulated during hNAC1 and E-CADHERIN-mediated rescues, as expected. In contrast, expression of other pluripotency genes, such as endogenous Oct4, Klf4, and Sall4, was not affected (Figure 3D). Additional pluripotency and adhesion factors, i.e., Dppa3, Cdhl3, and Ocln, were rescued only by hNAC1, whereas Lefty2 was rescued by both hNAC1 and E-CADHERIN overexpressions (Figure 3D). These results establish E-cadherin as a critical target of NAC1 in reprogramming and further suggest that proper regulation of E-cadherin expression beyond the early MET stage can be a critical molecular event leading to efficient and complete somatic cell reprogramming.

NAC1 Protects Reprogramming Cells from Acquiring Abnormal iPSC States in Response to High Transgene Expression

During the reprogramming experiments described in Figures 1 and S1, we noticed the appearance of a few Nac1 WT iPSC colonies with abnormal morphology (Figures 3A and S1). Remarkably, Nac1 WT iPSC colonies displayed a distinct morphology compared to the majority of Nac1 null iPSC colonies, which exhibited typical dome-shaped compact colonies (Figure 3A). To assess whether these rare WT colonies were pre-iPS or colonies with an atypical pluripotent state similar to the majority of Nac1 null iPSC colonies, we harvested two of them (together with two iPSC colonies with normal morphology) and performed global gene expression analyses by RNA sequencing (RNA-seq). As shown in Figure 4A, many genes were differentially regulated between morphologically good (G1 and G2) and bad (B1 and B2) WT iPSC colonies. When we scored these genes for GO, we found similar categories as those for Nac1 null iPSC colonies after and before reprogramming.
hNAC1-dependent rescue (Figure 4B versus Figure 2E). Moreover, we analyzed the expression of a few pluripotency, late reprogramming, and MET/adhesion markers, and found the trend of their expression between late reprogramming, and MET/adhesion markers, and found that the Nac1 WT and null colonies (Figure 4C versus Figure 1J), and between Nac1 null colonies after and before hNAC1-driven rescue (Figure 4C versus Figures 2F, S2E, 3D, and 3E). These results suggest that the Nac1 WT and null aberrant iPS colonies underwent similar pathways toward alternative pluripotent states, reminiscent of both pre-iPS and F-class cells (Ionge et al., 2014). However, Nac1 WT MEFs required higher expression of transgenes to be diverted toward those morphologically abnormal iPS colonies, compared with null cells (data not shown), indicating a protective role of NAC1 for proper iPS formation. In addition, hNAC1 overexpression did not rescue the WT bad morphology phenotype (Figure S3A), but rather triggered differentiation, indicating that very high levels of NAC1, in conjunction with high expression of the Yamanaka factors, may be deleterious for the self-renewal abilities of iPSCs. Also, unsuccessful rescue by hNAC1 was not due to failed hNAC1 overexpression but likely to the inadequate increase of the levels of endogenous E-cadherin (Figures S3B and S3C). Conversely, ectopic E-CADHERIN rescue was successful, confirming a major role for E-CADHERIN in controlling iPS morphology.

We next assessed whether the atypical morphology in Nac1 WT and null iPSCs was associated with similar gene expression signatures. We performed RNA-seq analyses of Nac1 WT and null iPS colonies with abnormal morphology, upon rescue with EV, hNAC1, or E-CADHERIN. We also included the four RNA-seq samples analyzed in Figure 4 (WT iPSCs with good morphology, G1 and G2, and bad morphology, B1 and B2). As depicted in Figures S3D and S3E, the heatmap of differentially regulated genes did not illustrate striking differences among the samples. To our surprise, samples appeared to cluster according to the amount of the STEMCCA reprogramming viruses employed during iPS generation, more than the overall iPS colony morphology, or the genetic background (WT versus null). Moreover, principal-component (PC) analyses clearly demonstrated a correlation between the position in the PC 2D space and the amount of the viruses over other parameters (Figure 4D). However, the presence of NAC1 significantly reduced the number of iPS colonies with atypical morphology for each amount of the virus (data not shown and Figures 1G and 1H), indicating a protecting role for NAC1 in preventing reprogramming toward abnormal iPS states.

NAC1 Collaborates with NANOG in Regulating E-cadherin Expression during Reprogramming

To dissect how NAC1 might transcriptionally control E-cadherin expression during reprogramming, we tested whether E-cadherin was a direct transcriptional target of NAC1. We transfected Nac1 null aberrant iPSCs with hNAC1 and successfully confirmed its binding to the E-cadherin promoter (Figure 5A) by chromatin immunoprecipitation (ChIP)-qPCR assays. However, reporter assays in both Nac1 null ESCs and heterologous HEK293T cells indicate a minimal or negative effect of ectopic hNAC1 expression on E-cadherin promoter activity (Figure 5B). These results suggest that the transcriptional action of NAC1 on target gene regulation during reprogramming may be different from the one in self-renewing ESCs and/or require additional factors that are not present in 293T cells.

Since E-cadherin upregulation is a hallmark of reprogramming, and it has been detected as a downstream target of other pluripotency regulators (see the ChEA website (http://amp.pharm.mssm.edu/lib/chea.jsp) and references therein), we thus hypothesized that NAC1 may control E-cadherin expression in cooperation with other stem cell factors. To identify such potential players, we turned our attention to Nanog due to: (1) its upregulation upon NAC1 rescue (Figures 2F and 3D); (2) the previous identification of Nanog as a downstream target of NAC1 in ESCs (Kim et al., 2008); and (3) the interaction between NANOG and NAC1 (Costa et al., 2013; Wang et al., 2006). Indeed, luciferase reporter assays in ESCs showed a NANOG-dependent activation of the E-cadherin promoter, which was counteracted by concomitant hNAC1 expression (Figure 5B, top). We also confirmed NANOG binding at the E-cadherin locus (Figure 5C blue bars). Importantly, we detected an enhancement in NANOG recruitment at the E-cadherin proximal promoter upon hNAC1 expression, compared with negative control (EV) and E-CADHERIN-mediated rescues (Figure 5C, green bars with pound signs versus blue and red bars). This was despite Nanog being similarly upregulated by E-CADHERIN and hNAC1 ectopic expression (Figure 3D). In addition, we detected NAC1 binding at the ~4.7 kb enhancer region of the Nanog locus in rescued iPSCs (Figure 5D, green bars). In contrast, E-CADHERIN did not bind there (Figure 5D red bars). The specific requirement of NAC1 for enhanced NANOG binding to the E-cadherin promoter, together with the transcriptional activation of Nanog (Figures S2E and 3D) and direct NAC1 binding to its regulatory locus, may explain why E-CADHERIN or hNAC1 can, but NANOG alone cannot rescue the reprogramming defect (Figure 3A).

Downregulation of Zeb1 Can Fully Reprogram Stalled Nac1 Null Pre-iPSCs

Since we noticed that Zeb1, a major E-cadherin repressor (Eger et al., 2005; Shirakihara et al., 2007), was not completely downregulated in the absence of Nac1 (compare Figure 1J with 2F for Zeb1 expression) during

920 Stem Cell Reports | Vol. 9 | 913–926 | September 12, 2017
reprogramming, we wondered whether NAC1 could also regulate E-cadherin indirectly via Zeb1. We first tested whether Zeb1 depletion would mimic NAC1 and E-CADHERIN overexpression in rescuing the Nac1 null pre-iPS phenotype. Indeed, as depicted in Figure 6A, two independent shRNAs against Zeb1, both of which significantly downregulated its expression (Figure 6B), rescued the Nac1 null aberrant iPS morphology. Gene expression analyses upon shZEB1 rescues revealed that E-cadherin itself (Figure 6B), and Nanog (Figure S4A), were drastically upregulated compared with controls. Additional qRT-PCR experiments showed that the endogenous pluripotency genes, Oct4 and Klf4, were not significantly upregulated (Figure S4A). These results confirm that OCT4 is not involved in any rescue and that KLF4 expression levels by themselves cannot completely explain the levels of E-cadherin expression before and after the rescues, despite the fact that KLF4 has been previously shown to be a major transcription factor regulating E-cadherin during reprogramming (Li et al., 2010) and in cancer cells (Koopmansch et al., 2013). In contrast, the pluripotency markers Sall4, Dppa3, and Lefty2, as well as the cell-adhesion molecule Cdh3, were significantly activated upon the rescues (Figures S4A), similarly to hNAC1- and E-CADHERIN-dependent rescues. This further highlights the critical functions of NAC1 in the transcriptional regulation of the MET-EMT-related genes for efficient reprogramming.

NAC1 Directly and Indirectly Represses Zeb1 during Reprogramming

Our finding that knockdown of Zeb1 was sufficient to recapitulate the NAC1 rescue of the null abnormal iPS phenotype (Figure 6A) prompted us to postulate a direct NAC1 role in repressing the Zeb1 promoter. We first assessed the ability of NAC1 to bind to the Zeb1 locus by ChIP. As shown in Figure 6C, hNAC1 was enriched at the Zeb1 proximal promoter in Nac1 null atypical iPSCs upon rescue. We then employed a 600-bp fragment of the human ZEB1 promoter fused to the luciferase gene to determine whether NAC1 could repress ZEB1 in HEK293T cells that are devoid of stem cell-specific factors. As shown in Figure 6D, hNAC1 repressed the ZEB1 proximal promoter...
efficiently, demonstrating that NAC1 can directly repress Zeb1 to favor E-cadherin expression.

The miR-200 family of microRNAs (miRNAs) has been demonstrated to downregulate Zeb1/2 during somatic cell reprogramming (reviewed in Leonardo et al., 2012). Therefore, we inspected whether NAC1 could also repress Zeb1 indirectly by acting on those miRNAs. As shown by ChIP assays in Figures 6E and 6f, overexpressed hNAC1 bound to the miR-200 family loci in rescued Nac1 null pre-iPSCs, which correlated with higher expression of those pri-miRNAs during the reprogramming process (Figure 6G). These findings clearly indicate that both the transcriptional (direct) and post-transcriptional (indirect) regulations of Zeb1 expression could contribute to NAC1 nuclear functions for efficient reprogramming. To assess the importance of the role of NAC1 in regulating the miR-200 family during reprogramming, we attempted to rescue the abnormal Nac1 null reprogrammed cells by overexpressing all the members of the miR-200 family. As shown in Figure S4B, however, we were not able to rescue the null phenotype. These data suggest that the major role of NAC1 during reprogramming is to regulate E-cadherin expression via Zeb1 repression.

**DISCUSSION**

Cell-cell adhesion, particularly the one mediated by E-CADHERIN, is fundamental for pluripotent stem cell biology because it regulates the degree of stemness,
differentiation, and somatic cell reprogramming (reviewed in Pieters and van Roy, 2014). It has been also well established that a crucial step in IPS generation is the MET process (Esteban et al., 2012), exemplified by the upregulation of E-cadherin. In fact, if E-cadherin is not expressed, MET cannot complete and reprogramming is halted (Chambers et al., 2007; Ding et al., 2012; Kim et al., 2008; Wang et al., 2006). However, its functions and molecular mechanisms involved in the establishment and maintenance of pluripotency were poorly defined. Remarkably, here we found that NAC1 could function as a reprogramming factor and was critical for ground state pluripotency in reprogramming during and beyond the early MET stage. This reinforces the notion that many regulators of the late-maturation phase during somatic cell reprogramming can be dispensable for early development or stem cell maintenance, as reported previously (Golipour et al., 2012), which is exemplified by Nanog (Chambers et al., 2007), Esrrb (Martello et al., 2012), and Klf4 (Katz et al., 2002; Segre et al., 1999).

During reprogramming NAC1 could not be replaced by NANOG, ESRRB, and KLF4, although NAC1 was thought to regulate their expression in pluripotent cells (Kim et al., 2008). Rather, NAC1 was essential for Zeb1 repression and proper expression of E-cadherin to reach full pluripotency manifested by typical compact dome-shaped IPS morphology and reactivation of the Oct4 distal enhancer. Our data indicate that proper induction of E-cadherin cannot be reached in the absence of NAC1, even when KLF4 is overexpressed (Figure 3), despite the fact that KLF4 can promote epithelial gene expression, and is essential for the MET process (Li et al., 2010). Since it has been previously shown that high levels of ZEB1/2 in cancer cells can displace KLF4 from the E-cadherin promoter with concomitant repression of transcription (Koopmansch et al., 2013), the regulatory action of KLF4 on E-cadherin expression during reprogramming may require NAC1-dependent downregulation of Zeb1. Our findings also argue against Nac1 null atypical iPSCs being similar to the pre-iPSCs generated by the Silva group (Silva et al., 2009). Unlike Silva pre-iPSCs, our Nac1 null abnormal iPSCs already expressed NANOG (Figure 3B, the EV lane), and their reprogramming to full pluripotency was not dependent on NANOG overexpression or the 2i/LIF medium. Another interesting aspect of NAC1 function during reprogramming is its protecting role in preventing reprogramming cells to be diverted toward altered pluripotent states, reminiscent of pre-iPS and F-like states, particularly when reprogramming transgene expression is high (Figure 4D).

In conclusion, our study identifies NAC1 as a reprogramming factor, critical for proper expression of E-cadherin during iPSC generation with a multifaceted regulatory mechanism (Figure 7). First, NAC1 binds and transcriptionally represses Zeb1, one of the main repressors of E-cadherin. Second, it stimulates the expression of the miR-200 family of miRNAs to downregulate Zeb1 post-transcriptionally. Third, NAC1 directly binds to the E-cadherin promoter and regulates co-factor (e.g., NANOG) binding to fine-tune its expression. These results enlighten our knowledge of the molecular mechanisms of somatic cell reprogramming and bring us a step closer to more efficient generation of iPSCs. Finally, our study will benefit our understanding of the role of NAC1 in cancer progression and metastasis. In that respect, our findings suggest that, in cancers where NAC1 is overexpressed, its repressor functions may be the driving force in the down-regulation of E-cadherin, leading to enhanced EMT, cancer cell migration, and metastasis, as described previously (Gao et al., 2014).

**EXPERIMENTAL PROCEDURES**

**Cell Culture**

iPSCs were grown in standard serum/LIF condition unless otherwise specified.

**qRT-PCR Assays**

RNAs were extracted with the QIAGEN RNeasy Plus Kit and converted to cDNA with the qScript cDNA SuperMix (Quanta Biosciences). qPCR was performed as described previously (Fidalgo et al., 2011). Oligo sequences are listed in Table S1.

**Microarray Profiling, RNA-Seq, and GO Analyses**

RNAs from day 0 (control non-transfected), 2, and 5 samples of the hNAC1 time course rescue experiments from two different null iPSC lines, were analyzed on an Illumina MouseWG-6 v.2.0 Expression BeadChip at the Genomics Core Facility, Icahn School of Medicine at Mount Sinai.

Cellular compartment functional annotation for genes differentially regulated in the Nac1 iPSC microarrays was performed by using David bioinformatics tools (Huang et al., 2009).

For RNA-seq analyses, total RNAs were extracted as above. RNA-seq libraries were prepared at Beijing Genomics Institute, and their quality and yield analyzed by an Agilent 2100 Bioanalyzer and ABI
StepOnePlus Real-Time PCR system, and sequenced on an Illumina HiSeq 2500/4000 instrument. Reads were filtered and then aligned to the reference genome with Bowtie2. Quantitative gene expression was determined by the RSEM software.

Reprogramming and iPSC Rescue Experiments
MEF reprogramming experiments were performed as published previously (Costa et al., 2013; Fidalgo et al., 2012) and described in Supplemental Experimental Procedures.

For rescue experiments (with clonal and/or bulk populations), iPSCs were transfected with piggybac-based expression vectors for indicated proteins and selected with 200 μg/mL hygromycin for a week. Colonies were then photographed and collected for gene expression analyses. For Zeb1 KD rescue assays, cells where infected with pLKO-pim-based lentiviruses with two Zeb1 shRNAs and one empty control, and selected with 1 μg/mL puromycin for a few days. For miR-200 rescues, viruses were prepared and cells infected as described in the Supplemental Experimental Procedures.

ChIP-qPCR Assays
ChIP experiments were performed as in (Lee et al., 2006) with a few modifications described in Supplemental Experimental Procedures.

ACCESSION NUMBERS
The accession number for the microarrays is GEO: GSE100350. The accession numbers for RNA-seq are SRA: SRX2885263, SRX2896775, SRX2896971, SRX2897024, SRX2897194, SRX2899130, SRX2899145, SRX2900609-14.

SUPPLEMENTAL INFORMATION
Supplemental Information includes Supplemental Experimental Procedures, four figures, three tables, and one movie and can be found in the supplementary material.
found with this article online at http://dx.doi.org/10.1016/j.stemcr.2017.07.002.

AUTHOR CONTRIBUTIONS

F.F., M.F., and J.W. designed the research, performed the experiments, and analyzed the data. N.Y. performed the experiments and analyzed the data. X.H. analyzed the microarray data. A.S., J.D., D.G., and B.D. provided technical assistance. E.F. wrote the manuscript draft. J.W. conceived the project and revised and approved the final manuscript.

ACKNOWLEDGMENTS

We thank Drs. Jie Hong and Jing-Yuan Fang for the pGL3-basic-hE-cad1P/hZeb1P constructs, Dr. Takeshi Urano for the pMXs-FHG-NAC1mut-C plasmid, Dr. Guoliang Xu for the pMX-miR-200 plasmids, and Dr. Greg Goodall for the pLenti-miR200 plasmids. This research was funded by grants from the NIH to J.W. (1R01-GM095942), the Empire State Stem Cell Fund through New York State Department of Health (NESTYEM) to J.W. (C028103, C028121), the Chinese Academy of Sciences Strategic Priority Research Program (XDB14040301 to F.F.), the National Natural Science Foundation of China (21577166 to F.F.), the Chinese Academy of Sciences Hundred Talent Program (29[2015]30 to F.F.), and the Key Research Program of Frontier Sciences, CAS (QYZDJ-SSW-DQC017 to F.F.). A.S. is an awardee of the Traineeship of the Irma T. Hirschl and Weill-Caulier Trusts Career Scientist Award. This work was supported by grants from the Irma T. Hirschl and Weill-Caulier Trusts Career Scientist Award.

REFERENCES

Apostolou, E., and Hochdelinger, K. (2013). Chromatin dynamics during cellular reprogramming. Nature 502, 462–471.

Cha, X.Y., Pierce, R.C., Kalivas, P.W., and Mackler, S.A. (1997). NAC-1, a rat brain mRNA, is increased in the nucleus accumbens three weeks after chronic cocaine self-administration. J. Neurosci. 17, 6864–6871.

Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Roberton, M., Vrana, J., Jones, K., Groteveld, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. Nature 450, 1230–1234.

Chen, T., Yuan, D., Wei, B., Jiang, J., Kang, J., Ling, K., Gu, Y., Li, J., Xiao, L., and Pei, G. (2010). E-cadherin-mediated cell-cell contact is critical for induced pluripotent stem cell generation. Stem Cells 28, 1315–1325.

Costa, Y., Ding, J., Theunissen, T.W., Faiola, F., Hore, T.A., Shilha, P.V., Fidalgo, M., Saunders, A., Lawrence, M., Dietmann, S., et al. (2013). NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. Nature 495, 370–374.

Dai, Q., Ren, A., Westholm, J.O., Serganov, A.A., Patel, D.J., and Lai, E.C. (2013). The BEN domain is a novel sequence-specific DNA-binding domain conserved in neural transcriptional repressors. Genes Dev. 27, 602–614.

Ding, J., Xu, H., Faiola, F., Ma’ayan, A., and Wang, J. (2012). Oct4 links multiple epigenetic pathways to the pluripotency network. Cell Res. 22, 155–167.

Ding, J., Huang, X., Shao, N., Zhou, H., Lee, D.F., Faiola, F., Fidalgo, M., Guallar, D., Saunders, A., Shilha, P.V., et al. (2015). Tex10 coordinates epigenetic control of super-enhancer activity in pluripotency and reprogramming. Cell Stem Cell 16, 653–668.

Eger, A., Aigner, K., Sonderegger, S., Dampier, B., Oehler, S., Schreiber, M., Berx, G., Cano, A., Beug, H., and Foisner, R. (2005). DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells. Oncogene 24, 2375–2385.

Esteban, M.A., Bao, X., Zhuang, Q., Zhou, T., Qin, B., and Pei, D. (2012). The mesenchymal-to-epithelial transition in somatic cell reprogramming. Curr. Opin. Genet. Dev. 22, 423–428.

Festuccia, N., Osorno, R., Halbritter, F., Karwacki-Neisius, V., Navarro, P., Colby, D., Wong, F., Yates, A., Tomlinson, S.R., and Chambers, I. (2012). Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. Cell Stem Cell 11, 477–490.

Fidalgo, M., Shekar, P.C., Ang, Y.S., Fujiwara, Y., Orkin, S.H., and Wang, J. (2011). Zip281 functions as a transcriptional repressor for pluripotency of mouse embryonic stem cells. Stem Cells 29, 1705–1716.

Fidalgo, M., Faiola, F., Pereira, C.F., Ding, J.J., Saunders, A., Gingold, J., Schaniel, C., Lemischka, I.R., Silva, J.C.R., and Wang, J.L. (2012). Zip281 mediates Nanog autorepression through recruitment of the NuRD complex and inhibits somatic cell reprogramming. Proc. Natl. Acad. Sci. USA 109, 16202–16207.

Gao, M., Wu, R.C., Herlinger, A.L., Yap, K., Kim, J.W., Wang, T.L., and Shih, Ie.M. (2014). Identification of the NAC1-regulated genes in ovarian cancer. Am. J. Pathol. 184, 133–140.

Golipour, A., David, L., Liu, Y., Jayakumaran, G., Hirsch, C.L., Trcka, D., and Wrana, J.L. (2012). A late transition in somatic cell reprogramming requires regulators distinct from the pluripotency network. Cell Stem Cell 11, 769–782.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4, 44–57.

Isihikawa, M., Nakayama, K., Yeasmin, S., Katagiri, A., Iida, K., Nakayama, N., and Miyazaki, K. (2010). NAC1, a potential stem cell pluripotency factor expression in normal endometrium, endometrial hyperplasia and endometrial carcinoma. Int. J. Oncol. 36, 1097–1103.

Katz, J.P., Peretault, N., Goldstein, B.G., Lee, C.S., Labosky, P.A., Yang, V.W., and Kaestner, K.H. (2002). The zinc-finger transcription factor Klf4 is required for terminal differentiation of goblet cells in the colon. Development 129, 2619–2628.

Kim, J., Chu, J., Shen, X., Wang, J., and Orkin, S.H. (2008). An extended transcriptional network for pluripotency of embryonic stem cells. Cell 132, 1049–1061.

Koopmansch, B., Berx, G., Foidart, J.M., Gilles, C., and Winkler, R. (2013). Interplay between KLF4 and ZEB2/SIP1 in the regulation of
E-cadherin expression. Biochem. Biophys. Res. Commun. 431, 652–657.

Lee, T.I., Johnstone, S.E., and Young, R.A. (2006). Chromatin immunoprecipitation and microarray-based analysis of protein location. Nat. Protoc. 1, 729–748.

Leonardo, T.R., Schultheis, H.L., Loring, J.F., and Laurent, L.C. (2012). The functions of microRNAs in pluripotency and reprogramming. Nat. Cell Biol. 14, 1114–1121.

Li, R., Liang, J., Ni, S., Zhou, T., Qing, X., Li, H., He, W., Chen, J., Li, E., Zhuang, Q., et al. (2010). A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. Cell Stem Cell 7, 51–63.

Mackler, S.A., Kornut, L., Cha, X.Y., Koebbe, M.J., Fournier, K.M., Bowers, M.S., and Kalivas, P.W. (2000). NAC-1 is a brain POZ/BTB factor required for establishing the barrier function of the skin. J. Neurosci. 20, 6210–6217.

Mackler, S.A., Homan, Y.X., Korutla, L., Conti, A.C., and Blendy, J.A. (2003). The mouse nac1 gene, encoding a cocaine-regulated actin-binding protein that can prevent cocaine-induced sensitization in the rat. J. Neurosci. 20, 6210–6217.

Mackler, S.A., Homan, Y.X., Kornut, L., Conti, A.C., and Blendy, J.A. (2003). The mouse nac1 gene, encoding a cocaine-regulated actin-binding protein that can prevent cocaine-induced sensitization in the rat. J. Neurosci. 20, 6210–6217.

Papp, B., and Plath, K. (2013). Epigenetics of reprogramming to induced pluripotency. Cell 152, 1324–1343.

Pieters, T., and van Roy, F. (2014). Role of cell-cell adhesion complexes in embryonic stem cell biology. J. Cell Sci. 127, 2603–2613.

Redmer, T., Diecke, S., Grigoryan, T., Quiroga-Negreir, A., Birchmeier, W., and Besser, D. (2011). Epac-atr is crucial for embryonic stem cell pluripotency and can replace OCT4 during somatic cell reprogramming. Cell Stem Cell 13, 491–504.

Okazaki, K., Nakayama, N., Nariai, Y., Nakayama, K., Miyazaki, K., Maruyama, R., Kato, H., Kusugi, S., Urano, T., and Sakashita, G. (2012). Nuclear localization signal in a cancer-related transcriptional regulator protein NAC1. Carcinogenesis 33, 1854–1862.

Papp, B., and Plath, K. (2013). Epigenetics of reprogramming to induced pluripotency. Cell 152, 1324–1343.

Pieters, T., and van Roy, F. (2014). Role of cell-cell adhesion complexes in embryonic stem cell biology. J. Cell Sci. 127, 2603–2613.

Redmer, T., Diecke, S., Grigoryan, T., Quiroga-Negreir, A., Birchmeier, W., and Besser, D. (2011). Epac-atr is crucial for embryonic stem cell pluripotency and can replace OCT4 during somatic cell reprogramming. Cell Stem Cell 13, 491–504.

Saunders, A., Faiola, F., and Wang, J. (2013). Concise review: pursuing self-renewal and pluripotency with the stem cell factor Nanog. Stem Cells 31, 1227–1236.

Segre, J.A., Bauer, C., and Fuchs, E. (1999). Klf4 is a transcription factor required for establishing the barrier function of the skin. Nat. Genet. 22, 356–360.

Shih, Ie.M., Nakayama, K., Wu, G., Nakayama, N., Zhang, J., and Wang, T.L. (2011). Amplification of the ch19p13.2 NACC1 locus in ovarian high-grade serous carcinoma. Mod. Pathol. 24, 638–645.

Shirakihara, T., Saiko, M., and Miyazono, K. (2007). Differential regulation of epithelial and mesenchymal markers by deltaEF1 proteins in epithelial mesenchymal transition induced by TGF-beta. Mol. Biol. Cell 18, 3533–3544.

Shu, X., and Pei, D. (2014). The function and regulation of mesenchymal-to-epithelial transition in somatic cell reprogramming. Curr. Opin. Genet. Dev. 28, 32–37.

Silva, J., Chambers, I., Pollard, S., and Smith, A. (2006). Nanog promotes transfer of pluripotency after cell fusion. Nature 441, 997–1001.

Silva, J., Barrandon, O., Nichols, J., Kawaguchi, J., Theunissen, T.W., and Smith, A. (2008). Promotion of reprogramming to ground state pluripotency by signal inhibition. PLoS Biol. 6, e253.

Silva, J., Nichols, J., Theunissen, T.W., Guo, G., van Oosten, A.L., Barrandon, O., Wray, J., Yamanaka, S., Chambers, I., and Smith, A. (2009). Nanog is the gateway to the pluripotent ground state. Cell 138, 722–737.

Sridharan, R., Tchieu, J., Mason, M.J., Yachechko, R., Kuoy, E., Horvath, S., Zhou, Q., and Plath, K. (2009). Role of the murine reprogramming factors in the induction of pluripotency. Cell 136, 364–377.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126, 663–676.

Theunissen, T.W., and Jaenisch, R. (2014). Molecular control of induced pluripotency. Cell Stem Cell 14, 720–734.

Tonge, P.D., Corso, A.J., Monetti, C., Hussein, S.M., Puri, M.C., Michael, I.P., Li, M., Lee, D.S., Mar, J.C., Cloonan, N., et al. (2014). Divergent reprogramming routes lead to alternative stem-cell states. Nature 516, 192–197.

Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D.N., Theunissen, T.W., and Orkin, S.H. (2006). A protein interaction network for pluripotency of embryonic stem cells. Nature 444, 364–368.

Watanabe, A., Yamada, Y., and Yamanaka, S. (2013). Epigenetic regulation in pluripotent stem cells: a key to breaking the epigenetic barrier. Philos. Trans. R. Soc. Lond. B Biol. Sci. 368, 20120292.

Winzi, M., Paszkowski-Rogacz, M., and Buchholz, F. (2014). Another brick in the wall: RNAi screens identify new barriers in iPSC reprogramming. Cell Stem Cell 15, 116–118.

Wu, P.H., Hung, S.H., Ren, T., Shih, Ie.M., and Tseng, Y. (2011). Cell cycle-dependent alteration in NAC1 nuclear body dynamics and morphology. Phys. Biol. 8, 015005.

Yap, K.L., Fraley, S.I., Thiaville, M.M., Jinawath, N., Nakayama, K., Wang, J., Wang, T.L., Wirtz, D., and Shih, Ie.M. (2012). NAC1 is an actin-binding protein that is essential for effective cytokinesis in cancer cells. Cancer Res. 72, 4085–4096.

Yap, K.L., Sysa-Shah, P., Bolon, B., Wu, R.C., Gao, M., Herlinger, A.L., Wang, F., Faiola, F., Huso, D., Gabrielson, K., et al. (2013). Loss of NAC1 expression is associated with defective bony patterning in the murine vertebral axis. PLoS One 8, e69099.

Yeom, Y.I., Fuhrmann, G., Ovitt, C.E., Brehn, A., Ohbo, K., Gross, M., Huhner, K., and Scholer, H.R. (1996). Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. Development 122, 881–894.
Supplemental Information

NAC1 Regulates Somatic Cell Reprogramming by Controlling Zeb1 and E-cadherin Expression

Francesco Faiola, Nuoya Yin, Miguel Fidalgo, Xin Huang, Arven Saunders, Junjun Ding, Diana Guallar, Baoyen Dang, and Jianlong Wang
Figure S1
Figure S1. *Nac1* is required for somatic cell reprogramming. Related to Figure 1.

(A) Schematic diagram of the reprogramming process.

(B) Proliferation of MEFs upon scramble (SCR) and *Nac1* shRNAs infection during the first four days of reprogramming.

(C) FACS measurements of total GFP positive iPS cells during control and *Nac1* KD MEF reprogramming.

(D) RT-qPCR analyses showing the efficiency of *Nac1* KD in MEFs (Top) and *Nac1* mRNA expression levels during reprogramming in *Nac1* WT, het, and null lines, as compared to ESCs (Bottom).

(E) *Nac1* WT and null Oct4-GFP MEF derivation and reprogramming scheme (top), and representative pictures of iPS colonies in bright field and GFP fluorescence (bottom).
Figure S2
Figure S2. NAC1 re-introduction rescues the null iPSC phenotype. Related to Figure 2.

(A) Enlarged images of representative colonies for the EV-, hNAC1 WT-, and hNAC1 Cyt-mediated rescues.

(B) WB images showing NAC1 nuclear and cytosolic fractionation in ESCs.

(C) Schematic of the 5-day long experiments of Nac1 null aberrant iPS clones rescued with ectopic hNAC1.

(D) Samples from (C) were taken at days 0, 1, 2, 3, 4, and 5, and the expression of a known NAC1 target and late reprogramming marker, Dppa3, was tested by RT-qPCR.

(E) Expression of selected pluripotency, late reprogramming, and cell adhesion genes in four different Nac1 null abnormal iPS clones upon 2-3 week-long rescues with over-expressed empty vector (EV) or hNAC1 (WT).

(F) Images of Nac1 WT and null neural progenitor cells (NPCs) in bright field.

(G) RT-qPCR analyses showing down-regulation of Oct4 and Nanog, up-regulation of Tet3 and neuro-ectoderm markers, and expression of Brachyury (T) and Myh6 during the differentiation of Nac1 WT and null ESCs into NPCs. Note that the mesoderm marker T was not expressed in NPCs as expected, and the cardiomyocyte gene Myh6 never significantly expressed during the differentiation process. Both genes serve as negative controls.

(H) Image of a typical day-14 iPS colony derived from the reprogramming of Nac1 null NPCs (top). A rare compact colony (bottom) resembling fully reprogrammed iPS colony is also shown for direct comparison. p⁰ stands for passage zero.

(I) Quantification of iPS colonies obtained by reprogramming of Nac1 WT and null NPCs. Colonies were scored based on morphology upon AP staining.
Data in (G) are average values ± SD from representative experiments in triplicates. Day 0 represents ESCs, Day 4 symbolizes Day 4 embryoid bodies, and Day 9 denotes NPCs on monolayer.
Figure S3
Figure S3. E-CADHERIN over-expression rescues the Nac1 WT abnormal iPSC phenotype. Related to Figure 4.

(A) Representative images of Nac1 WT bad morphology iPSC cells after attempted EV, E-CAD and hNAC1 rescues. EV means empty vector and E-CAD stands for E-CADHERIN.

(B-C) RT-qPCR assays showing the over-expression of hNac1 (B) and E-cadherin (C) in the rescue experiments in (A). Data are average ± SD of one representative experiment.

(D) Heat-map of gene expression of several NAC1 WT and null iPSCs, with or without rescuing attempts with either empty vector (EV), or hNAC1, or E-CADHERIN (E-CAD). X represent the relative amount of the reprogramming virus cassette employed.

(E) Hierarchical clustering of the iPSC samples analyzed in (D). X is as in (D).
**Figure S4**

A. Graph showing relative expression levels of various genes under different conditions.

B. Microscope images of samples labeled Mut, 200a, 200b, 200c, 141, 429, 200b-200a-429, 200c-141, and 200b-200a-429 + 200c-141. Each image is labeled with the sample name and has a scale bar of 1000 μm.
Figure S4. Down-regulation of Zeb1 rescues the Nac1 null iPS morphology. Related to Figure 6.

(A) Expression of selected genes by RT-qPCR upon Zeb1 RNAi. Data are plotted as average ± SD from one representative experiment in triplicates. The black bars represent the WT iPSC controls.

(B) Morphology of Nac1 null abnormal iPSCs upon attempted rescue with indicated miR-200 family members. Mut stands for negative mutant control.
**Table S1** List of primers used in RT-qPCR assays

| Primer          | Sequence (5'-3')                   |
|-----------------|------------------------------------|
| Brachyury (F)   | AAGCCTTCTTTGATGCCAAA               |
| Brachyury (R)   | GAGGCTCCAAACACTGAGGGTG             |
| Cdh3 (F)        | CTGATTCAGCTCTGAGGGA                |
| Cdh3 (R)        | CCGCATCTTAAAGGAGACGAA              |
| Cobl (F)        | CTGTTCCTCACTCTCCCGCT               |
| Cobl (R)        | CCTCAGAGCTGAGACCCAAA               |
| Dennd2c (F)     | GCTTGTGCGAGATTTTTGG                |
| Dennd2c (R)     | CTTTCACTGGAAAGGAAGGCC              |
| Dll1 (F)        | CTCCCCCTGGTGGACAGCTCGT             |
| Dll1 (R)        | GGAGAAGGATGTCGACCTC               |
| Dnmt3l (F)      | GGATGACCAAGAGACACTACC              |
| Dnmt3l (R)      | TCTTCAGCCCTGGAAATGTTG              |
| Dppa3 (F)       | CGGGGTATAGGGTACTGTTT               |
| Dppa3 (R)       | GGACCCCTGAAACTCCTCAGA              |
| Dppa4 (F)       | ATTCCACTTCTCTTCCTTCCGC             |
| Dppa4 (R)       | TGAGGAAGACTCTTCTCTCGGA             |
| Dppa5 (F)       | TCATGGATCCTCCAGCTTC                |
| Dppa5 (R)       | CAAGATTGCCGGCTAAATGGA              |
| E-cadherin (F)  | AAAAGAAGGCCGTGTCCTTGGG             |
| E-cadherin (R)  | GAGGTCTACACCTTCCCAGGT              |
| Epcam (F)       | GCTGGCAACAAAGTTGCTCTCTGAA          |
| Epcam (R)       | CGGTGACACTGCTGGGTTTGAAGGA          |
| Fabp3 (F)       | CTTGTGCATGCTAGCCACCT               |
| Fabp3 (R)       | CTCTGTCGATCCCTGAAACG              |
| Fg5 (F)         | AAAACCTGCTGACCCCTAGA               |
| Fg5 (R)         | CATCACATCCCGAATTTAGC               |
| Foxh1 (F)       | CCTGAAGAGGGGGAACACTG               |
| Foxh1 (R)       | GATGCCCTTGGAATTTCAAGG              |
| Gapdh (F)       | ACCAGAAAGACTGTTGGATGG              |
| Gapdh (R)       | CACATTGGGGTTAGGAACAC              |
| Ilf1 (F)        | CCTGTCCTTCAGACCTCACG               |
| Ilf1 (R)        | GACCATGTGATCCTGCTCTCC              |
| Jam2 (F)        | GCCAAAATAGCCCTTTGGA                |
| Jam2 (R)        | ACTACTGATCGTGGCTGCCTG             |
| Klf2 (F)        | GTGGCAGGTGGAGCCAG                 |
| Klf2 (R)        | GTTGCACTACGGGCTTCC                 |
| Klf4 (F)        | AACCTTACACTGTGACTAGG               |
| Klf4 (R)        | AAAAGTGCCCTTTCATGTGT               |
| Lama1 (F)       | GCTCCAAAATCCAGTTTCCA               |
| Lama1 (R)       | CTGTCACCCCTGGACTTACGG              |
| Lefty1 (F)      | TGTTGTGTGCTCTGTTCC                 |
| Lefty1 (R)      | GGAGAGGATGTCGACACTG                |
| Lefty2 (F)      | GGAGATGTACCTGGAGACACTGC            |
| Lefty2 (R)      | CATCTGAGGCCAGCTACAG                |
| Lin28 (F)       | TGTTCTGTATTGGAGTGAGC               |
| Lin28 (R)       | GCTTGCATTCTCTGGCATG                |
Map2 (F) CATCGCCAGCCTCGGAACAAACAG
Map2 (R) TCGCAGAAATGGAAGCTGGAGGCAAC
Myh6 (F) CTTCATCCATGGCCTATTCT
Myh6 (R) GCGCATTGAGTTCAAGAGAGA
Nac1 (F) GTGCAGTAGGGGGGATGTGT
Nac1 (R) AGGGCCTTACCTTCTCAGC
Nanog (F) AGGGTGCTGCTACTGAGATGCTCTG
Nanog (R) CAACCACTGTTTTCTGCGCACC
N-cadherin (F) CAGGGTGGACGTCATTGTAG
N-cadherin (R) AGGGTCTCCACCACTGATTC
Nestin (F) TGGCCAGTACATGCTGCTG
Nestin (R) AGATCGCTCAGATCCTGGAA
Nodal (F) TCTGGCGTACATGTTGAGC
Nodal (R) GGTGGAAAATGTCAATGGTGAG
Notch2 (F) TGTGCGTGTGTTGTTAGGTA
Notch2 (R) TGCTGTGGGCTCTGGCTG
Notch3 (F) GAATCTGGAAGACACCCTGG
Notch3 (R) AAGCGTCTCCTGGATGCTG
Oct4 (F) AGTGAATGGGCGGAGTTATG
Oct4 (R) ACTTGAATGGGCGGAGTTTAG
Pax6 (F) AGTGAATGGGCGGAGTTATG
Pax6 (R) ACTTGAATGGGCGGAGTTTAG
pri-miR-200ab-429 (F) AGGCTAGGGCGGAGACTTAGC
pri-miR-200ab-429 (R) AGTGCGCTGGGTCTGCATAC
pri-miR-200c-141 (F) TAGACAATCCCAAGGCCAAG
pri-miR-200c-141 (R) GCCCACTGGAAGAACACAAT
Rex1 (F) ACGAGTGGCTGTTTCTTGAGGA
Rex1 (R) TATGCCTCATTCCAGGGGCAC
Sall4 (F) AGTGATGTGGCTTGTGACCA
Sall4 (R) AACCAGCTTCTTTCCAAAAT
Scl2a3 (F) ATCGTGGCAGATGGCTGCTC
Scl2a3 (R) CTCTCAGACGCTCTGCGGAT
Scl7a3 (F) CTCTTGCAATTTCCATTCCC
Scl7a3 (R) GACTCTGCTTTGCTCAGTCCC
Snai1 (F) TTTTGCTCAGAGCCCTGTTGA
Snai1 (R) TCTTACATCCGAGTGTTTGG
Snai2 (F) CACATCGAGGAACCACATGGCC
Snai2 (R) TGTCGAGGCTGTTGATGCTG
Sulf2 (F) GTTGCAGGATGGGATTG
Sulf2 (R) GGTTCCCTCAGCGATCTAGC
TceA3 (F) GTTGAGATTTCTCAGGCAAG
TceA3 (R) CTCCACGGAGGAAAGAAAGGA
Tcfcp2ll (F) AGCCGAGGTTTTCTACAGCTG
Tcfcp2ll (R) CGCCCCCTCAGATGTTGGT
Tct (F) CGCAAGATCACCAGTGAATTT
Tct (R) TGTGGAGGAGTGGAGAGCAAG
Tqfl (F) GGTCCAAATTCACAACGC
Tqfl (R) CCTCAAAGGCTTTTACAGAGC
Tet1 (F) ATTTGAAGGAGGAGAAGTGGG
| Primer       | Sequence (5’-3’)                             |
|-------------|---------------------------------------------|
| Gapdh (F)   | AAGCTCATGAGGCACAGAATGGTC                    |
| Gapdh (R)   | TGGGTCATATGGTGACTTCTCTTAGGC                |
| Nanog 1 (F) | GTTTTGACTGCTAAACCACCAGAG                   |
| Nanog 1 (R) | GGCAGGCTTGCTACATTTCTAC                     |
| Nanog 2 (F) | GCCATTATGTAGATAGGGGTAGG                    |
| Nanog 2 (R) | CTTGTTCTGTGCTAGCAGACACTTAG                |
| E-cad -1K (F) | CATGCTGGGCTCAGACAA                      |
| E-cad -1K (R) | TGGGTCATATGGTGACTTCTCTAG                |
| E-cad -0.2K (F) | ACAGCTAGGCTAGATTCTGAC                 |
| E-cad -0.2K (R) | TGGTCTGTAATGGTGCTCAGGAT                |
| E-cad 0.1K (F) | ACCCGAGCTCAGTTTGC                         |
| E-cad 0.1K (R) | GGCAGAAAAACACAGCGAAG                     |
| E-cad 0.2K (F) | TCCGCTCAGCTGCTAGATAGG                    |
| E-cad 0.2K (R) | CTCGGGATCCCCACTTCTTGT                   |
| Zeb1 -2K (F) | ATCCACAGGAAAGGATTTACTGTC                  |
| Zeb1 -2K (R) | ACACACAAACAGCCCGAGAAG                    |
| Zeb1 -1.9K (F) | TCTTCTGGGCTGTTTGTGTC                    |
| Zeb1 -1.9K (R) | TCCGCTCAGCTAGATAGG                       |
| Zeb1 -1K (F) | AGGGTGTTACTATGTTGAGGCTGAC                |
| Zeb1 -1K (R) | CAAATGGAAGGCTCAGATAGG                    |
| Zeb1 -0.7K (F) | AATCCGGTCAGAGGAAGAAG                     |
| Zeb1 -0.7K (R) | CTCGAGCAATTTAGACACAAGAGCG                |
| Zeb1 0K (F) | CTCCGGATCCCCACTTCTTCT                    |
| Zeb1 0K (R) | TCCGACATCTAGACACAGCC                     |
| Zeb1 1K (F) | ATGATTTCAGCGGGATGTC                     |
| Zeb1 1K (R) | AATCCGGATCCCCACTTCTTCT                   |
| miR-141/200c 1 (F) | CTTCCGGTGCCTTCC                          |
| miR-141/200c 1 (R) | GCGCTCAGCTAGATAGG                     |
| miR-141/200c 2 (F) | GGAATGAGGAGCGAGGT                    |
| miR-141/200c 2 (R) | CCCGTAACCCCTGACCA                     |
| miR-200ab/429 1 (F) | CTTGTCAGTTGAGCCATTTC                  |
| miR-200ab/429 1 (R) | CTACAGATCCCAGCCAC                    |
| miR-200ab/429 2 (F) | TGGCCTAGAACTCCACAG                    |
| miR-200ab/429 2 (R) | AGAAACACATTATGCTTC                   |

F: forward primer
R: reverse primer
Supplemental Experimental Procedures

Plasmids

Piggybac-based expression plasmids for hNAC1, KLF4, E-CAD, OCLN, and ESRRB, were generated by PCR amplification of the respective genes with the following oligos.

\(hNAC1: \)
forward 5'-AGGGATCCGCCCAGACACTGCAGATGG,
reverse 5'
TTAATTAATTACTGCAGGGCTTCAGCC;

\(Klf4: \)
forward 5'-GCTAGCAGGCACCTGGGCGAGTCTG,
reverse 5'

TTAATTAATTAAAGTGCTCTTCTTAGTG;

\(E\)-cadherin:
forward 5'-AGCTAGCGGAGCCCGGTGCCGCAGCTTTTC,
reverse 5'
CTTAATTAACTAGTGCTCCTCGCCACCGCC;

\(Ocln: \)
5'-AGCTAGCTCCGTGAGGCCTTGGAAAGT,
reverse 5'
CTTAATTAAGGTTTCGCTCTGCACTAGTCT;

\(Esrrb: \)
forward 5'-GCTAGCGACGTGTCCGAACTCTGCATCC,
reverse 5'
TTAATTAATCACAACCTTGCCCTCCACCATC.

The PCR fragment was then purified from an agarose gel, digested with Nhe I (or BamH I) and Pac I, and ligated into the piggybac vector pPB-3XFLAG described in (Theunissen et al., 2011). pPB-NANOG was previously described (Costa et al., 2013). pPB-hNAC1 Cyt was cloned as the hNAC1 WT plasmid, but the PCR fragment was amplified from the plasmid pMXs-FHG-NAC1mut-C described in (Okazaki et al., 2012).

For the cloning of pLKO-shZeb1-pim, two shRNA duplexes were synthesized (Sigma-Aldrich) and cloned into the lentiviral vector pLKO.1PuroR-IRES-mCherry (pLKO-pim), as described in (Gingold et al., 2014). The Zeb1 shRNA sequences were: shZEB1 I: CCTGTGGATTATGAGTTCAA; shZEB1 II: CCGAAGCTGCAAGACCGGTTT.
To knock-down Nac1 during reprogramming, we employed the retroviral plasmid MSCV-LTR-miR30(Nac1-2704)Pgk-PURO with the following Nac1 RNAi sequence: 5’-CGAATGAAATATTGAAGTA-3’.

As control we used the corresponding pLMP vector with a scramble RNA sequence (Fidalgo et al., 2012).

The luciferase reporter plasmids, pGL3-basic-hCdh1P and pGL3-basic-hZeb1P, are described in (Xiong et al., 2012).

The pMX retroviral plasmids for miR200a/200b/200c/429/141/200cmut7n are described in (Hu et al., 2014)

The plasmids pLenti4.1ExmiR200b-200a-429 and pLenti4.1ExmiR200c-141 were a gift from Greg Goodall (Addgene plasmids # 35533 and 35534) (Gregory et al., 2008)

**Cellular protein fractionations, SDS-PAGE, and Western Blotting**

Non-nuclear extracts were obtained by lysing the cells in Lysis Buffer 1 (20 mM Hepes (pH 7.0, 10 mM KCl, 2 mM MgCl2, 0.5% NP-40) supplemented with 1 mM PMSF and protease inhibitor cocktail (Sigma). Nuclear pellets were first washed with Lysis Buffer 2 (150 mM NaCl, 1 mM EDTA, 20 mM Tris pH 8.0, 0.5% NP-40) without inhibitors, and then lysed by pipetting up and down in Buffer 2 supplemented with PMSF and protease inhibitor cocktail. Polyacrylamide gel electrophoresis and Western Blotting were performed according to standard procedures.

**Transfection, virus production, and cell transduction**

For transfections, iPSCs were detached by trypsinization. Meanwhile, the indicated pPB plasmids and the transposase pBASE with a 1:2 ratio, were pre-incubated with Lipofectamine 2000 or 3000 (Life Technologies) in OptiMEM (Life Technologies) for 10 minutes at room temperature. DNA-lipo complexes were then used to re-suspend the cells at room temperature for 10 min. Transfected cells were then plated in Serum/LIF ES medium. Media were replaced
the following day. For HEK293T cell transfection, jetPEI (Polyplus Transfection) was used according to the manufacturer's instructions.

Lentiviruses were generated in HEK293T cells with the lentiviral pLKO or STEMCCA vectors, and the packaging pCMV-dR8.2 and the pCMV-VSVG envelope plasmids, by transfection with jetPEI or Lipofectamine 3000. Media were changed 4-8 hrs after transfections. Two days later, supernatants were collected and either directly used for infections, or the viruses were concentrated in Amicon Ultra centrifugal filter units (Merck Millipore), following the manufacturer’s suggested procedures. Cells were infected in presence of 8 µg/ml polybrene (Sigma Aldrich), and when appropriate, selected with 1.5 µg/ml puromycin for 3-5 days before RNA extraction.

For retrovirus generation, Plat-E cells were transfected with the pLMP or the pMX plasmids by the calcium chloride and 2X BES (14280, Sigma Aldrich) method. Viruses were collected and concentrated as detailed above. Alternatively, cells were transfected with Lipofectamine 3000 and the supernatants directly collected and used for infections.

**Microarray and RNA-seq analyses**

Microarray data for iPS samples were collected using an Illumina Bead Array Reader confocal scanner. Image raw data were processed with the GenomeStudio software (Illumina) and exported with basic normalization and background correction. Then gene expression data were imported and analyzed using the BRB-ArrayTools (version 4.4.0., http://linus.nci.nih.gov./BRB-ArrayTools.html). Significantly up- or down-regulated genes were obtained by fold change > 2 and P-value < 0.05 by student t-test between Day 0 and Day 5 samples. Hierarchy clustering analysis was performed for the differentially regulated genes with Gene Cluster 3.0 software (http://bonsai.hgc.jp/~mdehoon/software/cluster/) and visualized with the TreeView program (version 1.1.6, http://jtreeview.sourceforge.net).
For RNA-seq analyses, raw reads in fastq format were filtered by removing adaptor sequences and low quality reads. Only samples with Q20 ≥ 95% and Q30 ≥ 90% were considered for further analyses. The genome sequence of mouse ESCs retrieved from the Mouse GRC38/mm10 database (http://genome.uscs.edu), was used as the reference database. All the clean reads were then mapped to the reference genome using Bowtie2 (Langmead et al., 2009). Quantitative gene expression was determined by the RSEM software (Li and Dewey, 2011). Annotation of differentially expressed genes (DEGs) was performed by analyzing syntenic relationships between samples. The significance of gene expression difference was set at Log2 (fold change) ≥ 2 with a false discovery rate (FDR) < 0.001. After normalization, hierarchical clustering was performed with Cluster (de Hoon et al., 2004; Eisen et al., 1998), and visualized with TreeView (Saldanha, 2004). Principal component analysis (from PC1 to PC13) was performed at BGI according to their standard guidelines.

Reprogramming assays

For Nac1 KD assays, fifty thousands Oct4-GFP MEFs were seeded in 6-well plates in DMEM supplemented with 10% serum one day before infection. At day 0, cell were infected simultaneously with shRNA lentiviruses for Nac1 or scramble control, and the hSTEMCCA lentivirus (Merck Millipore) expressing the four human Yamanaka factors. The following day, cells were selected with 5 µg/mL puromycin in ES medium for four days, and then five thousands of them reseeded on feeders in triplicates under 1.5 µg/mL puromycin selection. iPS colonies were scored according to GFP fluorescence or upon alkaline phosphatase staining, according to color intensity.

Nac1 WT, het, and null MEFs were reprogrammed with hSTEMCCA as above, but without puromycin selection.

Neural progenitor cell generation and reprogramming
To obtain neural progenitor cells (NPCs) from mESCs we followed a protocol described in (Bibel et al., 2007; Li et al., 2015). Briefly, day 8 EBs incubated with retinoic acid (RA) for four days, were enzymatically dissociated in single cells and plated on monolayer in dishes coated with laminin/poly-1-lysine. To generate iPSCs, 500 K NPCs were infected with STEMMCA viruses as described above.

ChIP-qPCR assays
For ChIP experiments, 3-5 µg anti-FLAG M2 (F1804, Sigma), and anti-NANOG (A300-397A, Bethyl Laboratories) antibodies, were pre-bound to 30-50 µL protein G Dynabeads magnetic beads (Life Technologies) over night at 4°C, with gentle rotation. For each IP, 150-200 µg chromatin were incubated with the pre-bound antibodies for 6-8 hrs at 4°C, with gentle rotation. IPs were then washed four times with RIPA Buffer (50 mM HEPES pH 7.5, 500 mM LiCl, 1 mM EDTA, 1% NP-40, 0.7% Na-Deoxycholate), once with TE pH 8.0, eluted once, reverse cross-linked for ≥ 6 hrs at 65°C, and purified with QIAquick PCR purification kit (Qiagen). qPCR primer sequences are provided in Table S2.

Luciferase assays
For luciferase detection experiments, Nact1−/− ESCs or HEK293T cells were transfected with Lipofectamine 2000 or jetPEI, respectively, with the indicated plasmids (200-1000 ng each). 48 hrs after transfections, cells were lysed and luciferase activity measured according to the luciferase assay system from Promega. Luciferase values were normalized to protein content.

Statistical analysis
Statistical analyses were performed with an unpaired, two-tailed Student's t-test. Significance values are indicated in the figure legends.
Supplemental references

Bibel, M., Richter, J., Lacroix, E., and Barde, Y.A. (2007). Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells. Nature protocols 2, 1034-1043.

de Hoon, M.J., Imoto, S., Nolan, J., and Miyano, S. (2004). Open source clustering software. Bioinformatics 20, 1453-1454.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proceedings of the National Academy of Sciences of the United States of America 95, 14863-14868.

Gingold, J.A., Fidalgo, M., Guallar, D., Lau, Z., Sun, Z., Zhou, H., Faiola, F., Huang, X., Lee, D.F., Waghrey, A., et al. (2014). A genome-wide RNAi screen identifies opposing functions of Snai1 and Snai2 on the Nanog dependency in reprogramming. Mol Cell 56, 140-152.

Gregory, P.A., Bert, A.G., Paterson, E.L., Barry, S.C., Tsykin, A., Farshid, G., Vadas, M.A., Khew-Goodall, Y., and Goodall, G.J. (2008). The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. Nat Cell Biol 10, 593-601.

Hu, X., Zhang, L., Mao, S.Q., Li, Z., Chen, J., Zhang, R.R., Wu, H.P., Gao, J., Guo, F., Liu, W., et al. (2014). Tet and TDG mediate DNA demethylation essential for mesenchymal-to-epithelial transition in somatic cell reprogramming. Cell stem cell 14, 512-522.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10, R25.

Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12, 323.

Li, T., Yang, D., Li, J., Tang, Y., Yang, J., and Le, W. (2015). Critical role of Tet3 in neural progenitor cell maintenance and terminal differentiation. Molecular neurobiology 51, 142-154.
Saldanha, A.J. (2004). Java Treeview--extensible visualization of microarray data. Bioinformatics 20, 3246-3248.

Theunissen, T.W., Costa, Y., Radzisheuskaya, A., van Oosten, A.L., Lavial, F., Pain, B., Castro, L.F., and Silva, J.C. (2011). Reprogramming capacity of Nanog is functionally conserved in vertebrates and resides in a unique homeodomain. Development 138, 4853-4865.

Xiong, H., Hong, J., Du, W., Lin, Y.W., Ren, L.L., Wang, Y.C., Su, W.Y., Wang, J.L., Cui, Y., Wang, Z.H., et al. (2012). Roles of STAT3 and ZEB1 proteins in E-cadherin down-regulation and human colorectal cancer epithelial-mesenchymal transition. The Journal of biological chemistry 287, 5819-5832.