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
6Correspondence: faiola@rcees.ac.cn (F.F.), jianlong.wang@mssm.edu (J.W.)

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 Jänecke, 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 tramtrack 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 iPS 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 nullcolonies 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<sup>−/−</sup> 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 iPS colonies and kill partially reprogrammed cells (Silva et al., 2008). As depicted in Figure 1I, about 50% of Nac1 WT iPS colonies 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 iPSC, and MET/cell adhesion, by qRT-PCR analyses. We found incomplete upregulation of a number of late-acting pluripotency genes including Nanog, Lin28, Tcl1, Dmnt3I, 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...
Nac1 in (Sridharan et al., 2009) were also not properly regulated other late reprogramming markers previously defined morphology. As shown in Figure 2A, re-introduction of passaged extensively but yet retained their aberrant iPS NAC1 (hNAC1) in null atypical iPS clones that had been and not other unknown reasons, we overexpressed human obtained was due to lack 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 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

**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 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 iPSC morphology, we overexpressed a form of hNAC1 mutated in its nuclear import sequence, which has been reported to be exclusively cytosolic (Okazaki et al., 2012). As depicted in Figures 2A and 2C, the cytosolic version of hNAC1 (hNAC1 Cyt), based on overall colony morphology, seemed to only slightly rescue the phenotype, despite its expression being higher than the WT

*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 in *Nac1* null iPSCs (Figure 1J). Taken together, these results suggest that reprogramming of *Nac1*-depleted cells is halted at an either intermediate or atypical reprogramming stage.

**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.
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, Rhnase4, Creb3, and Ifitm3), known to be downregulated during the pre-iPS to iPS transition (Sridharan et al., 2009), are all repressed upon the rescue of 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 iPS colonies 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 iPS 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

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 validate the microarray results, we also investigated the expression of several genes during the rescue by qRT-PCR. Besides genes from the heatmap of Figure 2D, we selected a few additional ones involved in MET/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, Sall4, Dppa3, E-cadherin, and Occludin (Occl), 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 iPS 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 defects 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 iPS colonies still retained the Nac1 null atypical phenotype observed during MEF reprogramming (see Figure S2H for morphology, and Figure S2I for quantification).
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/C0 aberrant iPSCs are distinct from Nanog/C0 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 iPSC 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 iPS colonies (Figure 3A), despite its ectopic expression being much lower than hNAC1 overexpression (Figure 3C). In contrast, OCLN did not rescue the Nac1 null iPS states (Figure 3D). Additional pluripotency and adhesion factors, i.e., Dppa3, Cdh3, 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 iPS 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 iPS colonies with abnormal morphology. To assess whether those rare WT colonies were pre-iPS or colonies with an atypical pluripotent state similar to the majority of Nac1 null colonies, we harvested two of them (together with two iPS 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 iPS cells. Noteworthy, the NAC1 overexpression alone phenocopied hNAC1 overexpression to generate typical dome-shaped compact iPS colonies (Figure 3D). Additional pluripotency and adhesion factors, i.e., Dppa3, Cdh3, 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.

**Figure 4. Nac1 Protects Reprogrammed Cells from Acquiring Abnormal iPS States upon High Expression of the Reprogramming Factors**

(A) Heatmap showing differential gene expression between two Nac1 WT iPS colonies with normal morphology (G1 and G2) and two with aberrant morphology (B1 and B2).

(B) GO analyses of the differentially expressed genes in (A).

(C) qRT-PCR analyses of selected pluripotency, late reprogramming, and MET/adhesion markers in G1, G2, B1, and B2 WT iPS cells.

(D) Principal-component (PC) analyses of various Nac1 WT and null iPS samples before and after indicated rescues. Note how the position in the 2D space is mostly determined by the amount of the reprogramming STEMCCA virus used. Color code: red, bad morphology; green, good morphology; blue, differentiated morphology. See also Figure S3 and Movie S1.
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

Moreover, we analyzed the expression of a few pluripotency, Nac1

bad colonies was strikingly similar to the one between 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 (Tonge 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 cell 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 (Figer 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
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) promoted 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 assays. 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 iPSC 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 (Chen et al., 2010; Li et al., 2010; Redmer et al., 2011; Sama-varchi-Tehrani et al., 2010). However, how E-cadherin is regulated during reprogramming is incompletely understood. Our study demonstrates a previously unappreciated role of the BTB-POZ transcriptional regulator NAC1 in direct transcriptional and indirect, via miR-200 and ZEB1, post-transcriptional control of E-cadherin expression during the reprogramming process.

We have described previously how NAC1 participates in the intricate protein interaction and transcription regulatory networks that regulate stem cell maintenance and pluripotency, suggesting a potentially significant role for NAC1 in self-renewal and/or pluripotency (Costa et al., 2013; 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 theOct4 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 (Kooptmansch 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 againstNac1null 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 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 (NYSTEM) 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.

REFERENCES


Koopmansch, B., Berx, G., Foidart, J.M., Gilles, C., and Winkler, R. (2013). Interplay between KLF4 and ZEB2/SIP1 in the regulation of