Cell

# A Myc Network Accounts for Similarities between Embryonic Stem and Cancer Cell Transcription Programs

Jonghwan Kim,[1,2,3] Andrew J. Woo,[1,3] Jianlin Chu,[1,2,3] Jonathan W. Snow,[1,2,3] Yuko Fujiwara,[1,2,3,4] Chul Geun Kim,[5] Alan B. Cantor,[1,3] and Stuart H. Orkin[1,2,3,4,*]
[1]Department of Pediatric Oncology, Children's Hospital and Dana Farber Cancer Institute
[2]Harvard Stem Cell Institute
[3]Harvard Medical School
[4]Howard Hughes Medical Institute
Boston, MA 02115, USA
[5]Department of Life Science, Hanyang University, Seoul 133-791, Republic of Korea
*Correspondence: stuart_orkin@dfci.harvard.edu
DOI 10.1016/j.cell.2010.09.010

## SUMMARY

c-Myc (Myc) is an important transcriptional regulator in embryonic stem (ES) cells, somatic cell reprogramming, and cancer. Here, we identify a Myc-centered regulatory network in ES cells by combining protein-protein and protein-DNA interaction studies and show that Myc interacts with the NuA4 complex, a regulator of ES cell identity. In combination with regulatory network information, we define three ES cell modules (Core, Polycomb, and Myc) and show that the modules are functionally separable, illustrating that the overall ES cell transcription program is composed of distinct units. With these modules as an analytical tool, we have reassessed the hypothesis linking an ES cell signature with cancer or cancer stem cells. We find that the Myc module, independent of the Core module, is active in various cancers and predicts cancer outcome. The apparent similarity of cancer and ES cell signatures reflects, in large part, the pervasive nature of Myc regulatory networks.

## INTRODUCTION

The pluripotent state of embryonic stem (ES) cells is maintained through the combinatorial actions of core transcription factors, including Oct4, Sox2, and Nanog (Boyer et al., 2005; Chen et al., 2008; Kim et al., 2008; Loh et al., 2006), in addition to other regulatory mechanisms encompassing epigenetic regulation (Boyer et al., 2006; Lee et al., 2006), microRNAs (Marson et al., 2008; Melton et al., 2010), and signaling pathways (Niwa et al., 1998; Sato et al., 2004). The discovery that cocktails of core pluripotency factors and selected widely expressed factors, such as Myc and Lin28, reprogram differentiated cells to an ES-like state (Park et al., 2008; Takahashi and Yamanaka, 2006; Yu et al., 2007) underscores the central role of transcription factors in cell fate decisions (Graf and Enver, 2009). Comprehensive protein interaction and target gene assessment of core pluripotency factors has provided a framework for conceptualizing the regulatory network that supports the ES cell state. Striking among the features of this network is the extent to which the core factors physically associate within protein complexes, co-occupy target genes, and cross-regulate each other (Boyer et al., 2005; Chen et al., 2008; Kim et al., 2008; Loh et al., 2006; Wang et al., 2006).

Although its expression dramatically enhances induced pluripotent (iPS) cell formation, Myc is not an integral member of the core pluripotency network (Chen et al., 2008; Hu et al., 2009; Kim et al., 2008). Myc occupies considerably more genomic target genes than the core factors, and Myc targets are involved predominantly in cellular metabolism, cell cycle, and protein synthesis pathways, whereas the targets of core factors relate more toward developmental and transcription-associated processes (Kim et al., 2008). Interestingly, promoters occupied by Myc show a strong correlation with a histone H3 lysine 4 trimethylation (H3K4me3) signature and a reverse correlation with histone H3 lysine 27 trimethylation (H3K27me3), suggesting a connection between Myc and epigenetic regulation (Kim et al., 2008). It is notable that the H3K4me3 signature has a positive correlation with active genes, and an open chromosomal structure, a distinctive feature of ES cells (Meshorer et al., 2006). Studies in non-ES cells have also revealed that Myc interacts with histone acetyltransferases (HATs) (Doyon and Côté, 2004; Frank et al., 2003). Improved iPS cell generation by addition of histone deacetylase inhibitors implies that global changes in epigenetic signatures are critical to efficient somatic cell reprogramming (Huangfu et al., 2008).

Although they remain pluripotent, ES cells are capable of indefinite self-renewal. Both blocked differentiation and the capacity for self-renewal, hallmarks of ES cells and adult stem cells, are shared in part by cancer cells (Clarke and Fuller, 2006; Reya et al., 2001). Although contested in the literature, expression of pluripotency factors, such as Oct4 and Nanog, has been

described in some cancers (Kang et al., 2009; Schoenhals et al., 2009). The involvement of Myc in many cancers (Cole and Henriksson, 2006) and its effects in iPS cell generation raise important issues regarding the relationship between cancer and embryonic stem cell states. Moreover, renewed focus on tumor subpopulations that initiate tumor formation on transfer to a suitable host (cancer stem cells) has contributed to the comparison of cancers and stem cells and to the potential resemblance of metastatic cancer cells to stem cells.

These relationships have been reinforced by reports of ''stem cell'' or ''embryonic stem cell'' (ESC)–like signatures in human and mouse cancers (Ben-Porath et al., 2008; Wong et al., 2008a; Wong et al., 2008b). The properties of such ESC-like signatures have thus far not been clearly defined, leaving open the possibility that they are composed of multiple gene expression signatures that are the outcomes of functionally independent transcriptional regulatory networks. Cancer cells may share only one or few of these subdivided signatures observed in ES cells, and thus have relatively less in common with the ''embryonic state'' than recently suggested.

In the present study, we sought to define how the regulatory network controlled by Myc relates to the previously defined core pluripotency network (Boyer et al., 2005; Chen et al., 2008; Kim et al., 2008; Loh et al., 2006). We first identified a Myc-centered regulatory network in ES cells and revealed that this Myc-centered network is largely independent of the core ES cell pluripotency network. On the basis of these findings, we propose that the overall ES cell specific gene expression signature is composed of smaller sets of subsignatures, which are represented as ''modules''—modules for the core pluripotency factors (Core module), the Polycomb complex factors (PRC module), and the Myc-related factors (Myc module). We provide evidence that these modules are functionally independent in ES cells, as well as during somatic cell reprogramming. With these modules as analytical tools, we observe that ES cells and cancer cells share Myc module activity, but generally do not share Core module activity. These findings argue against the hypothesis that cancer cells often reactivate an embryonic stem cell gene signature, even as they progress to a more highly invasive or metastatic state. Instead, the common features of ES cells and cancer cells reflect in large part the pervasive nature of the Myc regulatory network.

## RESULTS

### Construction of a Myc-Centered Protein-Protein Interaction Network in ES Cells

Previous protein-DNA interaction studies in ES cells indicated that targets occupied by the core pluripotency factors differ from genes bound by Myc (Chen et al., 2008; Kim et al., 2008). A recent RNA interference–based functional screen additionally suggested the existence of a second network linked functionally with Myc (Hu et al., 2009). Because coregulators that function with Myc have not been characterized previously in ES cells, we first sought to identify protein complexes that contain Myc with Myc-associated factors in ES cells. Using the in vivo metabolic biotin tagging method (de Boer et al., 2003; Wang et al., 2006), protein complexes containing tagged Myc in ES cells were affinity purified and analyzed by mass-spectrometry. We identified several proteins known to interact with Myc in other cell types, including Max, Ep400, Dmap1, and Trrap (Figure 1A) (Cai et al., 2003; Fuchs et al., 2001; McMahon et al., 1998). To expand and validate the protein-protein interaction network encompassing Myc, we subsequently generated ES cell lines expressing tagged Max and tagged Dmap1. ES cells expressing tagged Tip60 and tagged Gcn5 were also generated because they are HATs and known interacting partners of Trrap (Ikura et al., 2000; McMahon et al., 2000). We also generated tagged E2F4 ES cells, because another E2F family member E2F1 shares many common targets with Myc (Chen et al., 2008). E2F1 and E2F4 have many common targets and interchangeable roles in normal and tumor cells (Xu et al., 2007). Among E2F family proteins, E2F4 shows strongest expression in ES cells. In summary, we established ES cell lines expressing tagged Myc, Max, Dmap1, Tip60, Gcn5, and E2F4 (Figure 1A and Figure S1A available online) and identified their interacting partner proteins (summarized in Table S1). Figure 1A shows lists of high confidence interacting partner proteins of each factor tested. Interactions were independently validated by coimmunoprecipitation (Figure 1C and Figure S1B).

### Myc Interacts with the NuA4 HAT Complex in ES Cells

We did not observe overlap of proteins existing between the core protein interaction network (Wang et al., 2006) and the Myc-centered protein interaction network (Figure S1C). Although this may be due to the stringency of our conditions for recovery of protein complexes, within each network we observed a high degree of interactions, strongly suggesting that these two networks, and their protein complexes, are physically separate. Interestingly, we observed that Myc interacts with many proteins in a recognized conserved protein complex known as NuA4 HAT (or the Tip60-Ep400 complex) (Doyon and Côté, 2004) as shown in Figure 1A (pink cells) and Figure 1B (proteins in a pink circle). Myc, Max, Dmap1, Tip60, Trrap, and Ep400 are tightly interconnected within the network; however, Gcn5 and E2F4 show a lower degree of association, suggesting their weak or indirect interaction with Myc/NuA4. It has been suggested that transcription factors, such as Myc, p53, and E2Fs, require the NuA4 complex to activate downstream targets in non-ES cell contexts (Ard et al., 2002; McMahon et al., 1998). Our data (Figure 1 and Table S1) strongly support the view that Myc interacts with an intact NuA4 HAT complex in ES cells, also implying that histone 3 and 4 acetylation (AcH3 and AcH4, respectively) signatures may also be generated in part by the Myc/NuA4 complex via Tip60 in ES cells. Previous RNAi-based phenotypic analyses in ES cells revealed that factors in the NuA4 HAT complex, including Ep400, Dmap1, Tip60, Trrap, Ruvb1, and Ruvb2, are critical to ES cell identity (Fazzio et al., 2008) (also our observation, Figures S1D and S1E). These findings imply a crucial role for the Myc/NuA4 complex in ES cells.

### Construction of a Myc-Centered Protein-DNA Interaction Network in ES Cells

To identify genomic targets of Myc and its associated factors tested in Figure 1, we performed bioChIP-chip (Kim et al., 2008). Because Tip60 and Gcn5 generate AcH3 and AcH4