

# A core erythroid transcriptional network is repressed by a master regulator of myelo-lymphoid differentiation

Sandeep N. Wontakal<sup>a,1</sup>, Xingyi Guo<sup>b,1</sup>, Cameron Smith<sup>c,2</sup>, Thomas MacCarthy<sup>c,d,2</sup>, Emery H. Bresnick<sup>e</sup>, Aviv Bergman<sup>c</sup>, Michael P. Snyder<sup>f,g</sup>, Sherman M. Weissman<sup>h,3</sup>, Deyou Zheng<sup>b,i,3</sup>, and Arthur I. Skoultschi<sup>a,3</sup>

Departments of <sup>a</sup>Cell Biology, <sup>b</sup>Neurology, <sup>c</sup>Systems and Computational Biology, and <sup>i</sup>Genetics and Neuroscience, Albert Einstein College of Medicine, Bronx, NY 10461; <sup>d</sup>Department of Applied Mathematics and Statistics, State University of New York, Stony Brook, NY 11794; <sup>e</sup>Wisconsin Institutes for Medical Research, University of Wisconsin School of Medicine and Public Health, Madison, WI 53705; <sup>f</sup>Department of Genetics, Yale University School of Medicine, New Haven, CT 06520; <sup>g</sup>Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, CT 06520; and <sup>h</sup>Department of Genetics, Stanford University, Stanford, CA 94305

Contributed by Sherman M. Weissman, December 29, 2011 (sent for review October 10, 2011)

Two mechanisms that play important roles in cell fate decisions are control of a “core transcriptional network” and repression of alternative transcriptional programs by antagonizing transcription factors. Whether these two mechanisms operate together is not known. Here we report that GATA-1, SCL, and Klf1 form an erythroid core transcriptional network by co-occupying >300 genes. Importantly, we find that PU.1, a negative regulator of terminal erythroid differentiation, is a highly integrated component of this network. GATA-1, SCL, and Klf1 act to promote, whereas PU.1 represses expression of many of the core network genes. PU.1 also represses the genes encoding GATA-1, SCL, Klf1, and important GATA-1 cofactors. Conversely, in addition to repressing PU.1 expression, GATA-1 also binds to and represses >100 PU.1 myelo-lymphoid gene targets in erythroid progenitors. Mathematical modeling further supports that this dual mechanism of repressing both the opposing upstream activator and its downstream targets provides a synergistic, robust mechanism for lineage specification. Taken together, these results amalgamate two key developmental principles, namely, regulation of a core transcriptional network and repression of an alternative transcriptional program, thereby enhancing our understanding of the mechanisms that establish cellular identity.

ChIP sequencing | erythropoiesis | cross antagonism

Although cells have hundreds of transcriptional regulators, the function of only a few key factors has been proposed to be critical for establishing and/or maintaining cellular identity (1). Studies in just a few cell types, particularly embryonic stem (ES) cells, support this concept (2, 3). In ES cells, the “core pluripotency factors” Oct4, Nanog, and Sox2 co-occupy ~300 genes that are enriched for developmental regulators and genes involved in self-renewal (2). The gene network formed by these three core ES cell factors exhibits several types of regulatory circuitry including a multi-input motif and feed-forward loops (2, 3). However, whether such a “core transcriptional network” exists in hematopoietic cells is not known.

GATA-1, SCL, and Klf1 are three essential erythroid promoting transcription factors that play critical roles in establishing erythroid identity through the up-regulation of erythroid-specific genes (4, 5). GATA-1 regulates expression of some erythroid-specific genes, such as globin genes, in association with SCL (6–8) and Klf1 (9–11). Recent studies of transcription factor occupancy in erythroid progenitors by chromatin immunoprecipitation and high-throughput sequencing (ChIP-Seq) revealed that GATA-1 bound regions are enriched for SCL binding elements (12, 13), SCL bound regions are enriched for potential GATA-1 binding sites (14), and Klf1 occupied regions are enriched for putative GATA-1 and SCL binding motifs (15). Although there is evidence that these three factors cooperatively regulate certain erythroid-specific genes, whether they form a network with features similar to the ES cell core transcriptional network is not known.

Whereas GATA-1, SCL, and Klf1 are essential for erythroid development, the myelo-lymphoid promoting transcription factor,

PU.1, is a negative regulator of terminal erythroid differentiation (16–19). Surprisingly, PU.1 was found to occupy more genes in erythroid progenitors than the three erythroid-promoting factors (20). However, the extent of overlap between the genes bound by PU.1 and the three erythroid factors is not known.

In this study, we provide genomic evidence for the existence of a core erythroid network of >300 genes that are co-occupied and regulated by GATA-1, SCL, and Klf1. This network has characteristic features of core transcriptional networks, including a multi-input motif and feed-forward loops. Furthermore, we also find that PU.1 binds to and represses most of the genes in this network, indicating that PU.1 is a highly integrated negative regulator of the core erythroid network. Conversely, we also find that GATA-1 binds to and represses >100 PU.1 myelo-lymphoid gene targets in erythroid progenitors. Finally, mathematical modeling reveals that the dual mechanism used by both GATA-1 and PU.1 to repress an alternative lineage-specific transcriptional program provides a robust mechanism for lineage specification.

## Results

**GATA-1 Preferentially Binds Distal to Genes.** To begin to investigate the possible existence of a “core erythroid network,” we carried out ChIP-Seq experiments on endogenous GATA-1 in normal, murine ES cell-derived erythroid progenitors (ES-EP) (21) in both proliferating and differentiating conditions. The following results indicate that our ChIP-Seq data are of high quality. Several well-established GATA-1 binding sites are found in the dataset, including binding sites in the  $\beta$ -globin locus control region (LCR) (Fig. S1C). GATA-1 bound regions in both proliferating and differentiating conditions are highly enriched with a characteristic GATA-1 binding motif (Fig. S1D), similar to that observed in recent studies (12, 13). Using quantitative chromatin immunoprecipitation (qChIP) for GATA-1, we validated 100% (18/18) of the sites bound in both proliferating and differentiating conditions (Fig. S2).

Several recent reports describe genome-wide occupancy of GATA-1 in several murine and human erythroid cell lines (12, 13, 22, 23). Similar to these other reports, we find that GATA-1

Author contributions: S.N.W., S.M.W., and A.I.S. designed research; S.N.W., C.S., and T.M. performed research; S.N.W., X.G., C.S., T.M., E.H.B., S.M.W., M.P.S., and D.Z. contributed new reagents/analytic tools; S.N.W., X.G., C.S., T.M., A.B., S.M.W., D.Z., and A.I.S. analyzed data; and X.G., C.S., T.M., S.N.W., D.Z., and A.I.S. wrote the paper.

The authors declare no conflict of interest.

Database deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, [www.ncbi.nlm.nih.gov/geo](http://www.ncbi.nlm.nih.gov/geo) (accession no. GSE35385).

<sup>1</sup>S.N.W. and X.G. contributed equally to this work.

<sup>2</sup>C.S. and T.M. contributed equally to this work.

<sup>3</sup>To whom correspondence may be addressed. E-mail: [arthur.skoultschi@einstein.yu.edu](mailto:arthur.skoultschi@einstein.yu.edu), [sherman.weissman@yale.edu](mailto:sherman.weissman@yale.edu), or [deyou.zheng@einstein.yu.edu](mailto:deyou.zheng@einstein.yu.edu).

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1121019109/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1121019109/-DCSupplemental).



bound by GATA-1, SCL, and Klf1. The analysis showed that these genes are indeed enriched for genes involved in development ( $P$  value =  $3.24 \times 10^{-7}$ ) and hematopoiesis ( $P$  value =  $1.52 \times 10^{-6}$ ) (Fig. 2A). Moreover, comparison of the gene list with the Gene Ontology (GO) term Erythrocyte Differentiation (GO term 0030218) showed that the genes occupied by the three factors are enriched for erythroid-specific genes (binomial  $P$  value =  $1.0 \times 10^{-5}$ ). Consistent with these findings, we also find that these triply occupied genes are overrepresented (binomial  $P$  value  $< 2.2 \times 10^{-16}$ ) in the group of genes that were previously found to be highly expressed in the erythroid lineage (27). Three such genes are band 4.1 (Epb4.1), glycophorin C (Gypc), and aminolevulinic acid synthase 2 (Alas2) (Fig. 2B–D). These results indicate that the genes bound by GATA-1, SCL, and Klf1 constitute an erythroid core network.

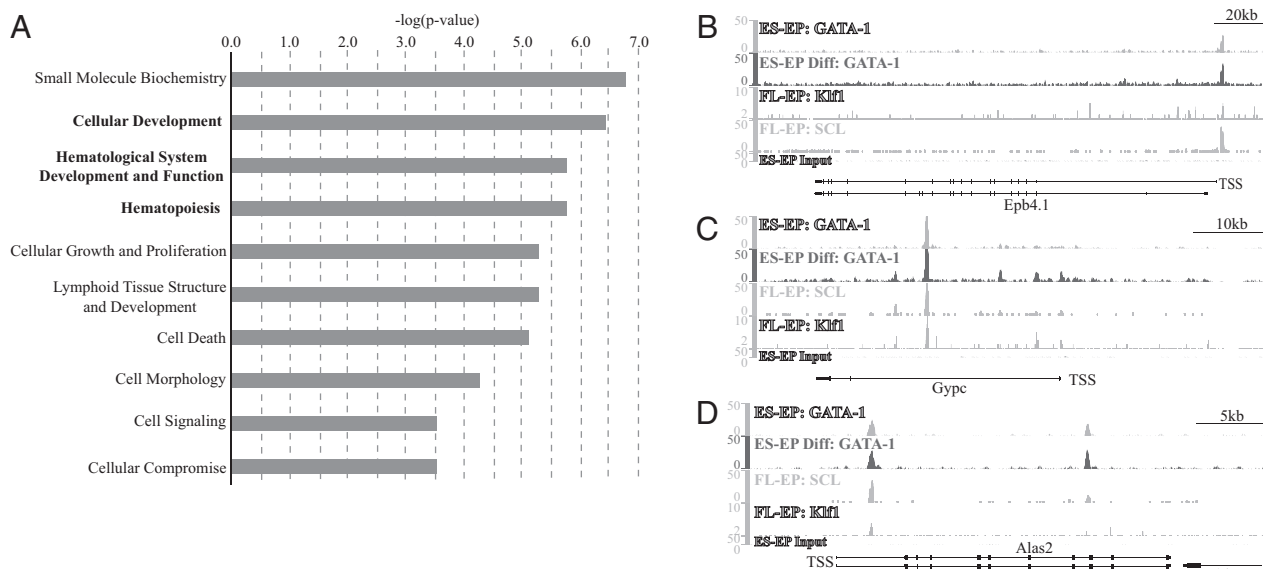
As in the genes encoding the core ES cell factors, expression of the GATA-1 and SCL genes is subject to autoregulation (28, 29). Although the GATA-1 and SCL proteins exhibit very similar binding patterns near the *GATA-1*, *SCL*, and *Klf1* genes, only GATA-1 up-regulates expression of all three of these erythroid promoting factors, indicating that GATA-1 is upstream of SCL and Klf1 (Fig. S4). This relationship is consistent with GATA-1 forming a coherent type I feed-forward loop with both SCL and Klf1 (Fig. S4F), similar to that reported in the other core transcriptional networks (2, 3). These findings may also help to explain the observation that overexpression of GATA-1 is sufficient to reprogram nonerythroid cells (30–32), whereas, as far as we are aware, this property has not been attributed to SCL or Klf1.

**PU.1 Is a Highly Integrated, Negative Regulator of the Core Erythroid Network.** In addition to interacting with SCL and Klf1, GATA-1 directly interacts with the myelo-lymphoid promoting factor PU.1 (16, 33, 34). The interplay between PU.1 and GATA-1 has served as an important model for understanding the mechanisms underlying lineage specification (35). Although PU.1 is essential for myeloid and B-cell development (36, 37), it is also expressed in erythroid progenitors, where it plays an important role in regulating the terminal differentiation decision (18, 38). PU.1

blocks erythroid differentiation by inhibiting GATA-1 transcriptional activity (39, 40), as well as by directly regulating many genes in immature erythroid cells (20). Unexpectedly, ChIP-Seq studies of PU.1 in erythroid progenitors revealed that PU.1 occupies many more sites in these cells than any of the three erythroid-promoting factors (20). However, whether PU.1 affects erythroid differentiation by regulating the core erythroid network formed by the three erythroid factors is not known.

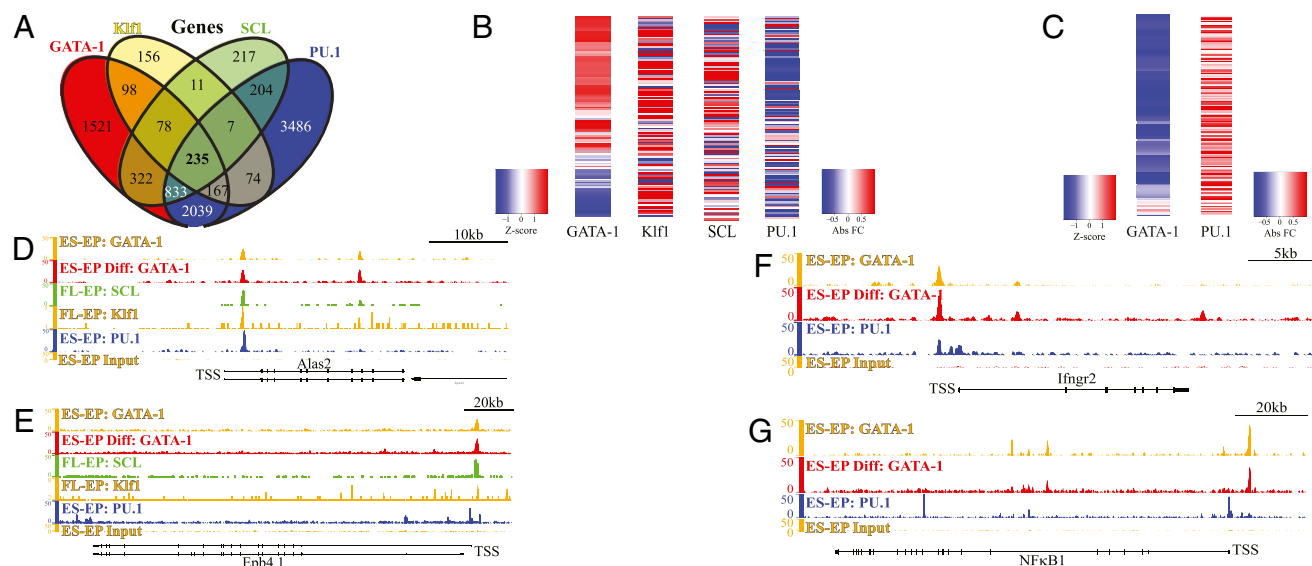
To test this possibility, we used our recently published ChIP-Seq data for PU.1 in ES-EP (20) to interrogate PU.1 occupancy near the 313 genes bound by the three erythroid factors. Strikingly, we find that 75% (235/313) of these genes are occupied by PU.1 (Fig. 3A). Using recently published data for GATA-1- (13, 23), SCL- (14), Klf1- (41), and PU.1-dependent (20) gene expression changes in erythroid progenitors, we investigated the response of the 235 genes to the four factors. Remarkably, we find that GATA-1, SCL, and Klf1 act to positively regulate expression of 68%, 67%, and 52% of these genes, respectively, whereas PU.1 represses expression of 72% of the genes (Fig. 3B). For example, *Alas2* and *Epb4.1* are two erythroid-specific genes that are activated by the three erythroid factors and repressed by PU.1 (Fig. 3C and D). These results reveal a unique feature of a core transcriptional network, namely that it can be subject to negative regulation by a transcription factor that promotes other closely related lineages.

**PU.1 Directly Represses Expression of GATA-1, SCL, Klf1, and Important GATA-1 Cofactors.** As mentioned, two features of the core erythroid network formed by GATA-1, SCL, and Klf1 are regulation of SCL and Klf1 by GATA-1 (Fig. S4) and autoregulation of the GATA-1 and SCL genes (28, 29). Therefore, it was of interest to determine whether PU.1, in addition to repressing many downstream targets of the three erythroid factors, also negatively regulates expression of the factors themselves. Indeed, we find that PU.1 binds in close proximity to the *GATA-1*, *SCL*, and *Klf1* genes (Fig. S5). PU.1 represses expression of the three genes, albeit to a greater degree for GATA-1 and Klf1 than SCL (Fig. S5). Moreover, we find that PU.1 binds near the TSS of genes encoding two critical GATA-1 cofactors, FOG1 (*Zfpml*), and



**Fig. 2.** Genes bound by GATA-1, SCL, and Klf1 are enriched for erythroid genes. (A) Ingenuity pathway analysis (IPA) was performed on genes co-occupied by all three erythroid factors. The 10 most significantly enriched molecular and cellular functions/physiological system development and functions within this set of genes are displayed. (B) Occupancy maps are shown for GATA-1, SCL, and Klf1 in the vicinity of the genes that encode for band 4.1 (*Epb4.1*), glycophorin C (*Gypc*), and aminolevulinic acid synthase 2 (*Alas2*), which are three erythroid-specific genes.





**Fig. 3.** PU.1 is an integral repressor of a core erythroid network, and conversely, GATA-1 represses numerous PU.1 myelo-lymphoid gene targets in erythroid progenitors. (A) PU.1 ChIP-Seq analysis in ES-EP (20) was used to determine the genes occupied by PU.1 in erythroid progenitor cells. Using the same criterion for associating bound regions with genes as in Fig. 1A, a four-way Venn diagram displays the overlap of genes occupied by PU.1, GATA-1 (in proliferating and/or differentiating ES-EP cells), SCL, and Klf1. (B) GATA-1- (13, 23), Klf1- (41), SCL- (14), and PU.1-dependent (20) gene expression datasets were used to determine how each factor regulates the expression of genes co-occupied by all four factors in erythroid progenitor cells. GATA-1-dependent gene expression was generated by performing expression array analysis at multiple time points after the introduction of GATA-1 into GATA-1 null erythroblasts (13, 23). The heatmap displayed for GATA-1 represents the final time point of this induction series represented as a Z-score relative to all time points. Heatmaps for Klf1, SCL, and PU.1 represent the absolute fold change in gene expression in fetal-liver-derived erythroid progenitors from wild-type animals relative to those from mutant animals. (C and D) Occupancy maps from ChIP-Seq data for GATA-1, SCL, Klf1, and PU.1 in the vicinity of the genes encoding (C) aminolevulinic acid synthase 2 (*Alas2*) and (D) band 4.1 (*Epb4.1*). (E) Heatmaps displayed show GATA-1- (13, 23) and PU.1-dependent (20) gene regulation of 151 myelo-lymphoid genes bound by both proteins in erythroid progenitors. Myelo-lymphoid genes were identified as genes that are highly expressed in granulocytic, monocytic, T-cell, or B-cell lineages compared with other hematopoietic cells (27). The heatmap displayed for GATA-1 represents the 21-h time point of this induction series represented as a Z-score relative to all time points, whereas the PU.1 heatmap represents the absolute fold change in gene expression in fetal-liver-derived early erythroid progenitor cells from wild-type animals compared with that in mice that have ~70% reduction in PU.1 levels. (F and G) Occupancy maps of GATA-1 and PU.1 in ES-EP cells are displayed in the vicinity of the myelo-lymphoid genes that encode for (F) *Ifngr2* and (G) *Nfkb1*.

Gfi-1b, and represses their expression (Fig. S6). Thus, the negative effect of PU.1 on the erythroid core network extends to a group of factors that help drive the network.

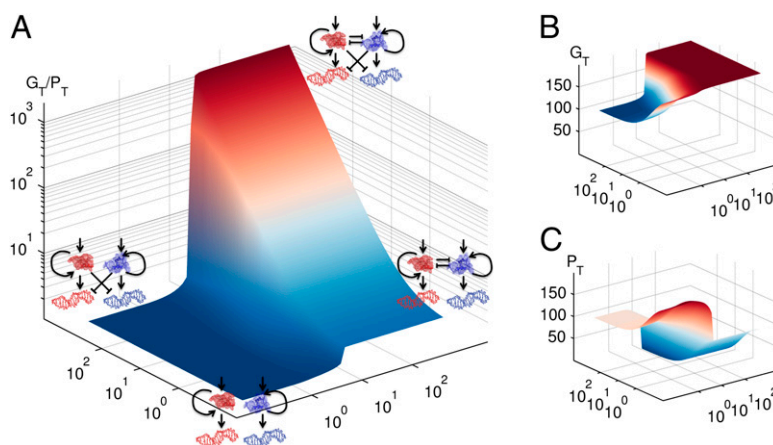
**GATA-1 Binds to and Represses Many PU.1 Myeloid-Lymphoid Gene Targets in Erythroid Progenitors.** We previously found that PU.1 occupies several myelo-lymphoid-specific genes in erythroid progenitors (20). Given the mutual antagonism between PU.1 and GATA-1 (16, 33, 34), one might predict that GATA-1 represses these gene targets in erythroid cells. Indeed, we find 151 myelo-lymphoid genes that are occupied by GATA-1 and PU.1 and that are positively regulated by PU.1 and repressed by GATA-1 (Fig. 3E). For example, we find that GATA-1 binds in close proximity to and represses the myelo-lymphoid genes *NFkB1* and the IFN- $\gamma$  receptor 2 (*Ifngr2*), which are essential for immune cell function (Fig. 3F and G). Interestingly, a recent report demonstrated that during erythroid differentiation GATA-1 represses expression of PU.1 itself (42). Taken together with all of the aforementioned effects of PU.1 on the erythroid core network, these results suggest that an important aspect of lineage specification is negative cross-regulation directed at both the downstream gene targets of the factors and the genes encoding the factors themselves.

**Mathematical Modeling of the GATA-1 and PU.1 Transcriptional Interaction.** To better understand the consequences of this dual mechanism of repressing both the opposing upstream activator and its downstream targets, we developed a mathematical model (*SI Materials and Methods*) to describe the dynamics and steady-state expression profiles of GATA-1 and PU.1 target genes. Fig.

4A shows the equilibrium expression-level ratio [GATA-1 targets/PU.1 targets ( $G_T/P_T$ ); Fig. 4B and C], following a transient stimulus favoring GATA-1, over the plane representing the 2D continuum of possible antagonistic effects: (i) between the factors themselves and (ii) on the downstream targets of the opposing factor (Figs. S7–S11 and Tables S1–S2). Whereas mutual inhibition between GATA-1 and PU.1 alone increases the  $G_T/P_T$  ratio (right corner of Fig. 4A), the model behavior illustrates that mutual inhibition and repression of opposing downstream targets act synergistically (center top corner of Fig. 4A) to maximize the  $G_T/P_T$  ratio. These results suggest that the dual mechanism identified here provides, in comparison with either cross-inhibition or target inhibition alone, more robust suppression of an alternative gene expression program during lineage specification.

## Discussion

Work in a limited number of cell types suggests that one important aspect of cellular identity is determined by the concerted actions of a few key transcriptional regulators controlling a subset of genes referred to as a core transcriptional network (1). In this study, we find that GATA-1, SCL, and Klf1, three essential erythroid transcription factors, form such a network in erythroid cells (Fig. 1). Several characteristics of this network are highly reminiscent of the ES cell core network (2), including the existence of a multi-input motif formed by the three factors (Fig. 1B), as well as coherent type I feed-forward loops (Fig. S4). Because both inputs are required for optimal expression of the output signal, coherent type I feed-forward loops may be important for combinatorial transcriptional regulation (43). Indeed, absence of SCL leads to a reduction in GATA-1 binding at some erythroid-specific



**Fig. 4.** Mathematical model demonstrating the synergistic effect of repressing both the upstream activator and its downstream targets during lineage specification. (A) We determined the equilibrium expression-level ratio of GATA-1 targets/PU.1 targets ( $G_T/P_T$ , on the z axis using logarithmic scale) following a transient stimulation of GATA-1 for different combinations of the parameter values that independently modulate the mutual inhibition between GATA-1 and PU.1 (right corner, x axis) and inhibition of the other's downstream targets (left corner, y axis) from the base network topology represented in the bottom center corner (all other parameter values are constant and tabulated in *SI Materials and Methods*). The top center corner represents simultaneous high-level mutual inhibition and repression of the downstream targets. (B and C) GATA-1 target ( $G_T$ ) and PU.1 target ( $P_T$ ) gene expression levels, respectively, used to compute the  $G_T/P_T$  ratio. The highest  $G_T/P_T$  ratio observed in the top center corner of A is due to sustained elevated  $G_T$  concentration with decreased  $P_T$  concentration relative to all other corners.

genes that results in suboptimal expression of these genes (14). Interestingly, in all cases studied thus far, the critical factors co-occupy ~300 genes (2, 3) (Fig. 1A). The fact that these factors often bind in close proximity to one another (2) (Fig. 1B) suggests that the *cis*-binding elements may have coevolved together. In addition to binding in close proximity, another feature shared by the core ES factors and the three erythroid factors is the ability to physically interact. Nanog and Oct4 interact with one another (44), whereas GATA-1 interacts with both SCL (6) and Klf1 (10). This observation raises an interesting question. Did the ability of these transcription factors to interact favor the evolution of *cis*-elements in close proximity or did the evolution of *cis*-binding sites facilitate the ability of these factors to interact? Phylogenetic analysis of the binding elements, along with studies of how the protein interaction interfaces evolved, could provide novel insights into the evolution of these core transcriptional networks.

Interestingly, we also find that PU.1, a negative regulator of terminal erythroid differentiation, is a highly integrated component of the erythroid core network (Fig. 3A). Furthermore, we show that, in addition to negatively regulating the expression of many erythroid core network genes, PU.1 also represses expression of GATA-1, SCL, and Klf1 themselves (Fig. S5), as well as some key GATA-1 cofactors (Fig. S6). In this way, PU.1 appears to antagonize both the erythroid-specific transcriptional and proteomic networks in erythroid progenitors. Importantly, we also find that GATA-1 represses many of the myelo-lymphoid downstream gene targets of PU.1 (Fig. 3C), as well as the PU.1 gene itself (42). Mathematical modeling reveals that this dual mechanism of repressing both the opposing upstream activator and its downstream targets provides for a robust method of silencing alternative gene expression programs (Fig. 4). This result suggests that such a mechanism may be used by antagonizing transcription factors in other lineages.

The work reported in this study unifies two key concepts that are important for establishing cellular identity, namely the core transcriptional network and the mutual antagonism between master transcriptional regulators. Our findings demonstrate that a “core transcriptional network” can be subject to negative regulation by a master regulatory transcription factor from a closely related lineage. In the future, it will be important to determine whether the types of positive and negative effects on a core

transcriptional network found here in the erythroid lineage are also present in other developmental systems.

## Materials and Methods

**Cell Culture Conditions.** ES-EP were cultured as previously described (21). Briefly, cells were grown in StemPro34 medium (Invitrogen) supplemented with 2 units/mL Epogen (Amgen), 40 ng/mL hGF-1 (Sigma), 1  $\mu$ M dexamethasone (Sigma), 100 ng/mL murine SCF (R&D Systems/Invitrogen), and 0.1%  $\beta$ -mercaptoethanol (Gibco) at 37  $^{\circ}$ C in a humidified 10%  $\text{CO}_2$  atmosphere. The cell concentration was maintained between  $2 \times 10^6$  and  $6 \times 10^6$  cells/mL by daily medium changes. ES-EP cells were differentiated for 24 h by culturing in StemPro34 media supplemented with 10 units/mL Epogen, 10  $\mu$ g/mL insulin (Sigma), 3  $\mu$ M mifepristone (Sigma), and 0.1%  $\beta$ -mercaptoethanol.

**qChIP and ChIP-Seq.** qChIP was performed as previously described (45). Briefly, cross-linked chromatin from  $2.5 \times 10^6$  cells was immunoprecipitated with Protein A agarose beads (Roche), using 2  $\mu$ g of HA (Y-11; Santa Cruz) or GATA-1 (46) antisera. Chromatin bound to beads was eluted with 1% SDS and 0.1 M  $\text{NaCHO}_3$  and the eluate was incubated at 65  $^{\circ}$ C overnight. Protein and RNA were digested with Proteinase K (Invitrogen) and RNase A (Roche) following the manufacturers' instructions and DNA was isolated using a PCR purification column (Qiagen). qPCR was performed using the primers indicated in Table S3. All qChIP experiments were performed using two independent chromatin preparations.

ChIP-Seq samples were prepared similarly, using chromatin from  $5 \times 10^7$  cells and immunoprecipitation with 40  $\mu$ g of GATA-1 antiserum. ChIP-Seq was performed in duplicate, using two independent chromatin preparations. DNA was isolated before immunoprecipitation and used as an input control sample. DNA was isolated and libraries were prepared for sequencing as described previously (47). Libraries were sequenced using an Illumina Analyzer GAI and processed with the Illumina ELAND pipeline. Uniquely mapped reads were aligned to the mouse genome (mm9). Further details on processing of ChIP-Seq samples can be found in ref. 20.

**Data Analysis and Mathematical Modeling.** ChIP-Seq data analysis; assignment of GATA-1, SCL, Klf1, and PU.1 binding sites to genes; integrated analysis of gene expression data and e-4C interaction data; statistical analysis; and mathematical modeling are described in *SI Materials and Methods*.

**ACKNOWLEDGMENTS.** We thank Dr. Andrew Perkins for kindly providing access to Klf1-dependent gene expression data. We thank Drs. Ghia Euskirchen, Hannah Monahan, Debasish Raha, and Minyi Shi for their help with Illumina sequencing. In addition, we thank Drs. Matthew Scharff and Ulrich Steidl for helpful comments on the manuscript. Work presented in this report was supported by National Institutes of Health Grants HL078381/

CA154239 (to A.I.S.) and DK68634 (to E.H.B.) and National Eye Institute Grant EY012200 (to D.Z.). S.N.W. was supported by National Institutes of Health/Medical Scientist Training Program Grant 5T32GM07288. A.I.S. also

receives support from National Cancer Institute Cancer Center Grant 2P30CA13330. This work was also partially supported by startup funds from Albert Einstein College of Medicine of Yeshiva University (to D.Z.).

- Cole MF, Young RA (2008) Mapping key features of transcriptional regulatory circuitry in embryonic stem cells. *Cold Spring Harb Symp Quant Biol* 73:183–193.
- Boyer LA, et al. (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122:947–956.
- Odom DT, et al. (2004) Control of pancreas and liver gene expression by HNF transcription factors. *Science* 303:1378–1381.
- Cantor AB, Orkin SH (2002) Transcriptional regulation of erythropoiesis: An affair involving multiple partners. *Oncogene* 21:3368–3376.
- Tsiftoglou AS, Vizirianakis IS, Strouboulis J (2009) Erythropoiesis: Model systems, molecular regulators, and developmental programs. *IUBMB Life* 61:800–830.
- Wadman IA, et al. (1997) The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. *EMBO J* 16:3145–3157.
- Anguita E, et al. (2004) Globin gene activation during haemopoiesis is driven by protein complexes nucleated by GATA-1 and GATA-2. *EMBO J* 23:2841–2852.
- Lahlil R, Lécuyer E, Herblot S, Hoang T (2004) SCL assembles a multifactorial complex that determines glycophorin A expression. *Mol Cell Biol* 24:1439–1452.
- Rodriguez P, et al. (2005) GATA-1 forms distinct activating and repressive complexes in erythroid cells. *EMBO J* 24:2354–2366.
- Gregory RC, et al. (1996) Functional interaction of GATA1 with erythroid Krüppel-like factor and Sp1 at defined erythroid promoters. *Blood* 87:1793–1801.
- Merika M, Orkin SH (1995) Functional synergy and physical interactions of the erythroid transcription factor GATA-1 with the Krüppel family proteins Sp1 and EKLF. *Mol Cell Biol* 15:2437–2447.
- Fujiwara T, et al. (2009) Discovering hematopoietic mechanisms through genome-wide analysis of GATA factor chromatin occupancy. *Mol Cell* 36:667–681.
- Yu M, et al. (2009) Insights into GATA-1-mediated gene activation versus repression via genome-wide chromatin occupancy analysis. *Mol Cell* 36:682–695.
- Kassouf MT, et al. (2010) Genome-wide identification of TAL1's functional targets: Insights into its mechanisms of action in primary erythroid cells. *Genome Res* 20:1064–1083.
- Tallack MR, et al. (2010) A global role for KLF1 in erythropoiesis revealed by ChIP-seq in primary erythroid cells. *Genome Res* 20:1052–1063.
- Rekhtman N, Radparvar F, Evans T, Skoultschi AI (1999) Direct interaction of hematopoietic transcription factors PU.1 and GATA-1: Functional antagonism in erythroid cells. *Genes Dev* 13:1398–1411.
- Rao G, Rekhtman N, Cheng G, Krasikov T, Skoultschi AI (1997) Deregulated expression of the PU.1 transcription factor blocks murine erythroleukemia cell terminal differentiation. *Oncogene* 14:123–131.
- Pop R, et al. (2010) A key commitment step in erythropoiesis is synchronized with the cell cycle clock through mutual inhibition between PU.1 and S-phase progression. *PLoS Biol* 8:e1000484.
- Rhodes J, et al. (2005) Interplay of pu.1 and gata1 determines myelo-erythroid progenitor cell fate in zebrafish. *Dev Cell* 8:97–108.
- Wontakal SN, et al. (2011) A large gene network in immature erythroid cells is controlled by the myeloid and B cell transcriptional regulator PU.1. *PLoS Genet* 7:e1001392.
- Dolznig H, et al. (2005) Expansion and differentiation of immature mouse and human hematopoietic progenitors. *Methods Mol Med* 105:323–344.
- Soler E, et al. (2010) The genome-wide dynamics of the binding of Ldb1 complexes during erythroid differentiation. *Genes Dev* 24:277–289.
- Cheng Y, et al. (2009) Erythroid GATA1 function revealed by genome-wide analysis of transcription factor occupancy, histone modifications, and mRNA expression. *Genome Res* 19:2172–2184.
- Kerenyi MA, Orkin SH (2010) Networking erythropoiesis. *J Exp Med* 207:2537–2541.
- Macquarrie KL, Fong AP, Morse RH, Tapscott SJ (2011) Genome-wide transcription factor binding: Beyond direct target regulation. *Trends Genet* 27(4):141–148.
- Schoenfelder S, et al. (2010) Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nat Genet* 42:53–61.
- Novershtern N, et al. (2011) Densely interconnected transcriptional circuits control cell states in human hematopoiesis. *Cell* 144:296–309.
- Tsai SF, Strauss E, Orkin SH (1991) Functional analysis and in vivo footprinting implicate the erythroid transcription factor GATA-1 as a positive regulator of its own promoter. *Genes Dev* 5:919–931.
- Ogilvy S, et al. (2007) The SCL +40 enhancer targets the midbrain together with primitive and definitive hematopoiesis and is regulated by SCL and GATA proteins. *Mol Cell Biol* 27:7206–7219.
- Kulesa H, Frampton J, Graf T (1995) GATA-1 reprograms avian myelomonocytic cell lines into eosinophils, thrombocytes, and erythroblasts. *Genes Dev* 9:1250–1262.
- Visvader JE, Elefanty AG, Strasser A, Adams JM (1992) GATA-1 but not SCL induces megakaryocytic differentiation in an early myeloid line. *EMBO J* 11:4557–4564.
- Iwasaki H, et al. (2003) GATA-1 converts lymphoid and myelomonocytic progenitors into the megakaryocyte/erythrocyte lineages. *Immunity* 19:451–462.
- Zhang P, et al. (1999) Negative cross-talk between hematopoietic regulators: GATA proteins repress PU.1. *Proc Natl Acad Sci USA* 96:8705–8710.
- Nerlov C, Querfurth E, Kulesa H, Graf T (2000) GATA-1 interacts with the myeloid PU.1 transcription factor and represses PU.1-dependent transcription. *Blood* 95:2543–2551.
- Graf T, Enver T (2009) Forcing cells to change lineages. *Nature* 462:587–594.
- Scott EW, Simon MC, Anastasi J, Singh H (1994) Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science* 265:1573–1577.
- McKercher SR, et al. (1996) Targeted disruption of the PU.1 gene results in multiple hematopoietic abnormalities. *EMBO J* 15:5647–5658.
- Back J, Dierich A, Bronn C, Kastner P, Chan S (2004) PU.1 determines the self-renewal capacity of erythroid progenitor cells. *Blood* 103:3615–3623.
- Rekhtman N, et al. (2003) PU.1 and pRB interact and cooperate to repress GATA-1 and block erythroid differentiation. *Mol Cell Biol* 23:7460–7474.
- Stopka T, Amanatullah DF, Papetti M, Skoultschi AI (2005) PU.1 inhibits the erythroid program by binding to GATA-1 on DNA and creating a repressive chromatin structure. *EMBO J* 24:3712–3723.
- Hodge D, et al. (2006) A global role for EKLF in definitive and primitive erythropoiesis. *Blood* 107:3359–3370.
- Chou ST, et al. (2009) Graded repression of PU.1/Sfpi1 gene transcription by GATA factors regulates hematopoietic cell fate. *Blood* 114:983–994.
- Alon U (2007) Network motifs: Theory and experimental approaches. *Nat Rev Genet* 8:450–461.
- Wang J, et al. (2006) A protein interaction network for pluripotency of embryonic stem cells. *Nature* 444:364–368.
- Choe KS, Ujhelly O, Wontakal SN, Skoultschi AI (2010) PU.1 directly regulates cdk6 gene expression, linking the cell proliferation and differentiation programs in erythroid cells. *J Biol Chem* 285:3044–3052.
- Im H, et al. (2005) Chromatin domain activation via GATA-1 utilization of a small subset of dispersed GATA motifs within a broad chromosomal region. *Proc Natl Acad Sci USA* 102:17065–17070.
- Robertson G, et al. (2007) Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat Methods* 4:651–657.