Published in final edited form as:

Curr Top Dev Biol. 2016 ; 118: 45–76. doi:10.1016/bs.ctdb.2016.01.002.

The Hematopoietic Stem and Progenitor Cell Cistrome: GATA Factor-Dependent *cis*-Regulatory Mechanisms

Kyle J. Hewitt^{1,2}, Kirby D. Johnson^{1,2}, Xin Gao^{1,2}, Sunduz Keles³, Emery H. Bresnick^{1,2,*}

¹University of Wisconsin School of Medicine and Public Health, Department of Cell and Regenerative Biology, Carbone Cancer Center, Madison, WI 53705

²UW-Madison Blood Research Program

³Department of Biostatistics and Medical Informatics, University of Wisconsin School of Medicine and Public Health

Abstract

Transcriptional regulators mediate the genesis and function of the hematopoietic system by binding complex ensembles of *cis*-regulatory elements to establish genetic networks. While thousands to millions of any given *cis*-element reside in a genome, how transcriptional regulators select these sites and how site attributes dictate the functional output is not well understood. An instructive system to address this problem involves the GATA family of transcription factors that control vital developmental and physiological processes and are linked to multiple human pathologies. Although GATA factors bind DNA motifs harboring the sequence GATA, only a very small subset of these abundant motifs are occupied in genomes. Mechanistic studies revealed a unique configuration of GATA factor-regulated *cis*-element consisting of an E-box and a downstream GATA motif separated by a short DNA spacer. GATA-1- or GATA-2-containing multi-protein complexes at these composite elements control transcription of genes critical for hematopoietic stem cell emergence in the mammalian embryo, hematopoietic progenitor cell regulation, and erythroid cell maturation. Other constituents of the complex include the basic helix-loop-loop transcription factor Scl/TAL1, its heterodimeric partner E2A, and the Lim domain proteins LMO2 and LDB1. This article reviews the structure/function of E-box-GATA composite *cis*-elements and evidence implicating these elements as critical constituents of a genetic circuit we term the hematopoietic stem and progenitor cell cistrome.

I. *Cis*-regulatory mechanisms

Many proteins and protein networks are implicated in controlling the generation and function of hematopoietic stem and progenitor cells (HSPCs) and their hematopoietic cell progeny (Rossi et al., 2012). However, only recently have DNA regulatory motifs (*cis*-elements) in the genome been described that non-redundantly control cell fate decisions during hematopoiesis (Gao et al., 2013; Johnson et al., 2012; Johnson et al., 2015). As a single *cis*-element integrates activities of a limited ensemble of regulatory proteins and signals, it provides a unique path to discover important pathways and regulatory networks.

*Corresponding author: ehbresni@wisc.edu.

While *cis*-elements reside at distinct gene locations, including promoters, enhancers, and insulators, and there can be both similarities and differences in their functions at these locations, the focus of this review is on enhancer mechanisms.

Enhancers are short (~50-500 bp) DNA sequences residing distal to genes or in introns that function in *cis* to increase transcription (Plank and Dean, 2014). Typical enhancers contain clustered *cis*-elements, each binding one or more transcription factor (TF). Mechanistic steps in enhancer function include TF recruitment of coregulators that modify or remodel chromatin, thereby regulating chromatin template access to additional regulators and RNA Polymerase II (Pol II). Whereas Pol II recruitment to promoters is commonly stimulated by enhancers, enhancers also recruit Pol II directly to the enhancer (Johnson et al., 2001; Johnson et al., 2003; Szutorisz et al., 2005). Consequences of Pol II recruitment to an enhancer include generation of enhancer-derived RNA transcripts (Lai and Shiekhattar, 2014), the exact function of which remains controversial, and Pol II transfer to the promoter, a poorly understood reaction (Johnson et al., 2001; Vieira et al., 2004). In principle, transfer may involve processive Pol II tracking along the template (Ling et al., 2004a) or DNA/chromatin looping in which the enhancer complex physically associates with promoter components (Cullen et al., 1993; Dekker et al., 2002; Schleif, 1992; Willis and Seyfred, 1996). Finally, enhancers control locus subnuclear localization, which can expel a locus from a nuclear microenvironment unfavorable for high-level transcription, while positioning it in a microenvironment permissive for high-level transcription (Fraser and Bickmore, 2007; Jackson and Cook, 1995; Kosak et al., 2002; Lee et al., 2009; Lee et al., 2011; Rajapakse and Groudine, 2011; Schubeler et al., 2000; Zullo et al., 2012). Enhancer mechanisms at large cohorts of enhancer-regulated genes are vital determinants of cell phenotype in a wide swath of biological and pathological contexts.

II. GATA factor interactions with DNA and chromatin

Enhancers involved in the development and function of the hematopoietic system are often activated by members of the dual zinc finger GATA transcription factor family (GATA-1 to GATA-6) (Bresnick et al., 2012; Charron and Nemer, 1999; Molkenin, 2000), which recognize naked DNA containing the consensus (A/T)GATA(A/G) (Evans et al., 1988; Ko and Engel, 1993; Martin and Orkin, 1989; Merika and Orkin, 1993). While certain transcription factors (TFs) commonly localize to promoter complexes, GATA factors have a considerably higher propensity to occupy enhancers (Cheng et al., 2009; Fujiwara et al., 2009; Kang et al., 2012; Linnemann et al., 2011; Yu et al., 2009). GATA-1, -2, and -3 are related GATA factors that control stem and progenitor cell transitions that give rise to the hematopoietic system (Bresnick et al., 2012). These proteins also control important functions of differentiated hematopoietic cells.

This review will focus on GATA-1 (Evans and Felsenfeld, 1989; Tsai et al., 1989) and GATA-2 (Dorfman et al., 1992; Tsai et al., 1994), key regulators of HSPCs and myeloid-erythroid lineage cells in which considerable mechanistic/biological insights and principles have accrued. We will not discuss GATA-3, an important regulator of lymphopoiesis (Ho et al., 1991; Ko et al., 1991; Scripture-Adams et al., 2014; Ting et al., 1996). GATA-2 is expressed in HSPCs and promotes their generation and function (de Pater et al., 2013; Gao

et al., 2013; Johnson et al., 2012; Ling et al., 2004b; Rodrigues et al., 2005; Tsai et al., 1994). The targeted deletion of *Gata2* in the mouse is embryonic lethal at E10.5 due to lack of hematopoietic development (Tsai et al., 1994). GATA-2 is also expressed in early erythroid precursors, megakaryocytes, and mast cells (Huang et al., 2009; Mouthon et al., 1993; Tsai and Orkin, 1997; Weiss et al., 1994; Zon et al., 1991). GATA-2 expression in HSPCs requires an intronic enhancer (+9.5) containing an E-box-GATA composite element (Gao et al., 2013; Grass et al., 2006; Johnson et al., 2012; Sanalkumar et al., 2014). The +9.5 site resides 9.5 kb downstream of the *Gata2* transcription start site, and its deletion abrogates hematopoietic stem cell (HSC) generation in the aorta-gonad-mesonephros (AGM) during embryogenesis (Gao et al., 2013). The GATA-2 concentration must be maintained within a critical physiological window, as ectopically elevated GATA-2 blocks erythroid cell differentiation (Briegel et al., 1993; Persons et al., 1999), and HSC expansion and function are defective in *Gata2*^{+/-} mice (Ling et al., 2004b; Rodrigues et al., 2005). Consistent with the tight control of *Gata2* transcription, GATA-2 protein has a short half-life of <1 hour, which provides opportunities to rapidly change its cellular concentration (Lurie et al., 2008). GATA-2 deficiency induces a wide spectrum of human disorders (Dickinson et al., 2014; Spinner et al., 2013). Familial and sporadic heterozygous mutations in *GATA2* cause a complex immunodeficiency syndrome (MonoMAC) that often progresses to myelodysplastic syndrome and acute myeloid leukemia (AML) (Dickinson et al., 2011; Hahn et al., 2011; Hsu et al., 2011). Mutation of the *GATA2* +9.5 *cis*-element also yields MonoMAC syndrome (Hsu et al., 2013; Johnson et al., 2012). Furthermore, high GATA-2 correlates with poor prognosis of pediatric and adult AML (Luesink et al., 2012; Vicente et al., 2012).

Unlike GATA-2, GATA-1 is not required for HSPC genesis and function. GATA-1 is expressed predominantly in megakaryocyte and erythroid lineage cells and regulates red blood cell, platelet, mast cell and eosinophil differentiation (Fujiwara et al., 1996; Pevny et al., 1995; Pevny et al., 1991; Shivdasani et al., 1997; Simon et al., 1992; Weiss et al., 1994). Mutational disruption of *GATA1* yields familial dyserythropoietic anemia and thrombocytopenia (Nichols et al., 2000) and megakaryoblastic leukemia (Wechsler et al., 2002). In many contexts, GATA-1 requires a multi-zinc finger coregulator Friend of GATA-1 (FOG-1) (Tsang et al., 1998; Tsang et al., 1997) for both activation and repression of target gene transcription (Crispino et al., 1999). Whereas FOG-1 lacks DNA binding activity, it promotes GATA-1 chromatin occupancy (Letting et al., 2004; Pal et al., 2004) and recruits the chromatin remodeling NuRD complex to GATA-1-bound sites (Gao et al., 2010; Hong et al., 2005; Miccio et al., 2010; Ross et al., 2012). During erythropoiesis, GATA-1 replaces GATA-2 at many of its chromatin sites (Bresnick et al., 2010; Grass et al., 2003b). This “GATA switch” mechanism has been proposed to control the transition of early erythroid precursors to terminally-differentiating erythroid cells (Bresnick et al., 2012; Bresnick et al., 2010; Bresnick et al., 2005; DeVilbiss et al., 2014; Grass et al., 2003a). However, GATA switches do not occur at all GATA-2-occupied loci (May et al., 2013). The often different GATA factor expression patterns in various cell types reflect their qualitatively distinct biological roles, and GATA switching illustrates one example by which different GATA factors function differentially through a common chromatin site. In addition, in certain contexts, GATA factors appear to function redundantly. Yolk sac-derived primitive

erythroblasts express both GATA-1 and GATA-2. In these cells, targeted deletion of both factors is required to ablate a survival mechanism (Fujiwara et al., 2004).

GATA factors co-localize on chromatin with the E-box-binding basic helix-loop-helix (bHLH) protein Scl/TAL1 (Aplan et al., 1992b; Begley et al., 1989; Kallianpur et al., 1994). Scl/TAL1 heterodimerizes with a class A bHLH E-box protein, typically E12 or E47, and the heterodimer binds with high affinity to the E-box motif (Hsu et al., 1991; Hsu et al., 1994b). The Scl/TAL1-GATA complex occupies chromosomal sites containing a GATA motif, with or without a neighboring E-box (Cheng et al., 2009; Kassouf et al., 2010; Tripic et al., 2008; Wozniak et al., 2008; Yu et al., 2009). In ChIP-seq studies, approximately 20% of Scl/TAL1 occupancy does not coincide with an E-box motif (Kassouf et al., 2010). At an E-box-GATA element, CRISPR/Cas9-mediated deletion of only the E-box motif abrogates GATA-2 occupancy and reduces transcription of the endogenous loci (Hewitt et al., 2015), indicating that the E-box is required for complex formation at certain E-box-GATA elements. ChIP studies indicate that Scl/TAL1 has a higher propensity for occupying E-box-GATA elements in a transcriptionally active state in comparison with a repressed locus (Tripic et al., 2008; Yu et al., 2009).

Resembling GATA-2, Scl/TAL1 is crucial for definitive hematopoiesis, since targeted deletion of *Scl/Tal1* in mice severely impairs hematopoiesis (Porcher et al., 1996; Shivdasani et al., 1995). The conditional ablation of *Scl/Tal1* in mice provided evidence for an Scl/TAL1 function to generate HSCs, but not for their function (Mikkola et al., 2003). As expected for such a critical hematopoietic regulator, dysregulated Scl/TAL1 expression is linked to leukemogenesis, specifically T-cell acute lymphocytic leukemia (Aplan et al., 1992a; Aplan et al., 1992c; Bash et al., 1993; Bash et al., 1995; Ferrando et al., 2004; Gerby et al., 2014; Sanda et al., 2012; Tremblay et al., 2010). A non-DNA-binding Scl/TAL1 mutant delays embryonic lethality in a *Scl/Tal1*^{-/-} context, suggesting that intrinsic DNA binding activity is not essential for hematopoietic development (Kassouf et al., 2008). However, since these partially rescued mice have severe survival defects and are grossly anemic, the DNA binding function likely contributes to Scl/TAL1 activity. Since Scl/TAL1 occupies certain chromatin sites lacking an E-box, the partial rescue may also reflect activity via tethering to a GATA factor at these target sites.

Sequence requirements

Analysis of GATA-2 chromatin occupancy in erythroid cells reiterated the importance of the naked DNA consensus and revealed a more complex consensus (C/G)(A/T)GATAA(G/A/C)(G/A/C) enriched at chromatin occupancy sites (Fujiwara et al., 2009). While sequences conforming to this consensus (297,124 copies in the human genome) mediate GATA factor chromatin occupancy, these sequences are insufficient for chromatin occupancy and enhancer activity. Based on chromatin immunoprecipitation (ChIP) analysis, the vast majority of the sequences (>99%) are not GATA factor-occupied in cells (Fujiwara et al., 2009; Kang et al., 2012; Linnemann et al., 2011), presumably due to their occlusion by inaccessible chromatin and/or the absence of neighboring *cis*-elements that confer stable complex formation on chromatin. GATA factor occupancy is not restricted to enhancers,

however, and occur at certain promoters correlates with transcriptional regulation (Johnson et al., 2007).

A pioneering study discovered an E-box (CANNTG) followed by a GATA motif, termed an E-box-GATA composite element, that mediates multimeric protein complex assembly in GATA-1-expressing erythroid cells (Wadman et al., 1997). The precise orientation (CANNTG upstream of GATA) and spacing (8-10 base pairs) support maximal protein complex assembly on naked DNA templates and enhancer activity in transfection assays (Wadman et al., 1997; Wozniak et al., 2007). Either E-box or GATA motif mutations prevent multimeric protein complex formation *in vitro* (Wadman et al., 1997). Subsequent work demonstrated importance of E-box-GATA elements as GATA-2-regulated elements in GATA-2-expressing HSPCs and erythroid precursor cells. GATA-1 and GATA-2 occupy several hundred E-box-GATA elements (Fujiwara et al., 2009). More than 300 E-box-GATA elements are GATA-2-occupied in human K562 erythroleukemia cells. Using a more restricted sequence, CATCTG-8bp-AGATAA (corresponding to the *Gata2*+9.5 site), 62 out of 797 sites are GATA-2-occupied (Hewitt et al., 2015). Many of these sites are GATA-1 or GATA-2-occupied in a context-dependent manner, i.e. multimeric complexes are GATA-2-occupied in HSPCs and GATA-1-occupied in erythroid cells (Fujiwara et al., 2009; Li et al., 2011; Wilson et al., 2010).

How widespread are E-box-GATA motifs in the human genome? Over 1×10^5 human motifs contain some permutation of the CANNTG-N8-N14-AGATAA sequence. Using the E-box motif specific to the *Gata2*+9.5 site, CATCTG-N8-N14-AGATAA, close to 9,000 E-box-GATA motifs exist, and 25% are evolutionarily conserved in mouse by location analysis. Further narrowing the spacer sequence constraints yields nearly 800 mouse sequences conform to CATCTG-8bp-AGATAA, and ~5% of these sequences are GATA-2 and Scl/TAL1-occupied, based on ChIP-seq analysis (Hewitt et al., 2015).

It is instructive to consider how GATA factors function through E-box-GATA elements, in comparison with GATA motifs lacking E-boxes. Whereas this relationship is not understood, E-box-GATA composite elements have been functionally analyzed at several hematopoietic loci. At *Epb4.2* encoding an erythroid cytoskeletal component, tandem E-box-GATA motifs are each required for *Epb4.2* transcriptional activation in a transient transfection assay (Xu et al., 2003). At the *Klf1* locus, an E-box-GATA motif is required for transcriptional activation in a transient transfection assay (Anderson et al., 1998) and with a reporter-containing transgene (Anderson et al., 2000). At the *Gata2*+9.5 E-box-GATA composite element, the 7-9 bp distance between the E-box and GATA motifs is an important determinant of enhancer activity in a transient transfection assay (Wozniak et al., 2007). The specific spacer sequence can be altered without functional consequences, at least in this transient transfection context (Wozniak et al., 2007). The spacer length of E-box-GATA elements varies at different loci. E-box-GATA elements consisting of longer spacer sequences at the *Kit* promoter (Lecuyer et al., 2002), *Runx1* intron (Bee et al., 2009), *Smad1* intron, and *Smad5* promoter (Marks-Bluth et al., 2015) exhibit E-box- and GATA motif-dependent enhancer activity in transient transfection and transgenic mouse assays.

The principle of a restricted spacer length segregating the E-box and GATA motif has been reinforced by chromatin occupancy studies. GATA-1 and GATA-2 occupy E-box-GATA elements with a spacer length of 9 bp (Fujiwara et al., 2009; Hewitt et al., 2015; Li et al., 2013). Naked DNA binding studies indicate that Scl/TAL1, a GATA-2 target gene and essential determinant of hematopoiesis, prefers E-box sequences CAGGTG (Wadman et al., 1997) or CAGATG (Hsu et al., 1994a) at composite elements. In erythroid cells, however, Scl/TAL1 occupies E-boxes with a CATCTG consensus sequence at composite elements (Hewitt et al., 2015). GATA-1 and GATA-2 occupy E-box-GATA elements in K562 erythroleukemia cells with a more degenerate E-box sequence (CANNTG) (Fujiwara et al., 2009). Scl/TAL1 chromatin occupancy has been detected at an E-box-GATA motif with a 13-bp spacer, suggesting potential flexibility in the spacer length in certain contexts (Cohen-Kaminsky et al., 1998). In certain cases, a degenerate half E-box motif (CTG) can supplant the requisite E-box motif at hematopoietic enhancers with a spacing upstream of the GATA motif identical to that of the full E-box-GATA element (Bauer et al., 2013; Kassouf et al., 2010; Tallack et al., 2010). E-twenty six (ETS) motifs (TTCC), mediating ETS transcription factor DNA binding, are often associated with GATA-factor-occupied E-box-GATA motifs (Dore et al., 2012; Linnemann et al., 2011), revealing another layer of sequence complexity at composite enhancers. Furthermore, mutations of an ETS motif immediately downstream of the *Gata2*+9.5 E-box-GATA element is associated with MonoMAC syndrome, and mutation of the ETS motif reduces transactivation mediated by the E-box-GATA element in a transient transfection assay (Hsu et al., 2013).

Despite high-affinity GATA factor binding to naked DNA, most E-box-GATA elements are not GATA factor-occupied in mouse and human hematopoietic cell genomes (Fujiwara et al., 2009; Hewitt et al., 2015; Linnemann et al., 2011; Wozniak et al., 2007). Although multi-factorial mechanisms governing *cis*-element occupancy in chromatin are complex, a rudimentary principle is that contextual or “epigenetic” parameters irrespective of DNA sequence impact DNA structure in chromatin and *cis*-element accessibility. These parameters include histone modifications (DeVilbiss et al., 2013), DNA methylation (Jeong and Goodell, 2014; Jones, 2012), and DNA hydroxymethylation (Koh et al., 2011; Madzo et al., 2014; Tahiliani et al., 2009). In addition, mechanisms that expel or attract loci to nuclear microenvironments favor or preclude regulatory modes (Kosak et al., 2002; Lee et al., 2009; Lee et al., 2011; Ragoczy et al., 2006; Ragoczy et al., 2003; Rajapakse and Groudine, 2011; Schubeler et al., 2000; Zullo et al., 2012). Further complicating the matter, while GATA factor occupancy is required for transcriptional regulation of endogenous loci, occupancy does not suffice to generate a measurable functional output. Many occupancy sites are not physically linked to GATA factor-regulated genes. In principle, occupied *cis*-elements might control neighboring genes in three-dimensional space or might generate regulatory RNAs that act in *trans* to control distinct loci. Deciphering parameters dictating TF binding at endogenous loci and endowing non-coding sequences with *cis*-element activity remains challenging, and there is no universally accepted methodology to achieve this goal.

Is evolutionary conservation the ultimate parameter for predicting the functional importance of GATA factor-regulated enhancers? The current evidence suggests that while conservation has utility to winnow and sift through large ensembles of GATA-bound sequences, as a sole parameter, it lacks predictive power. Though essential *cis*-elements are likely to be

retained through speciation events (Pennacchio et al., 2006), human to mouse conservation of E-box-GATA sequences is well below 1% (Hewitt et al., 2015), far lower than the ~85% conservation of protein-coding sequences. Furthermore, there is little to no evidence that positional conservation of enhancers universally equates to functional importance. Orthologous alignments of candidate enhancers reveal that most GATA-occupied sites are poorly conserved, and only a small fraction of conserved elements are GATA factor-occupied in mouse and human (Cheng et al., 2014; Hewitt et al., 2015). Thus, mouse and human genomes might be regulated by distinct enhancer mechanisms, or similar enhancer mechanisms in these species might exhibit considerable fluidity in genomic positioning of the *cis*-elements. Of course, these two possibilities are not mutually exclusive. By using less-stringent location analysis to identify similar sequence motifs at orthologous gene loci, more than 25% of human E-box-GATA sequences are detected in mouse (Hewitt et al., 2015). While it is not known whether all of these elements are functionally conserved, there is greater conservation of E-box-GATA composite elements at loci than defined by strict positional alignments. A noteworthy example of enhancer conservation at distinct locations involves *Gata1*, where an upstream GATA-occupied mouse enhancer corresponds to an intronic enhancer in human (Valverde-Garduno et al., 2004). Enhancer diversity and locus organization is a feature that almost certainly contributes to mammalian phenotypic diversity and inter-species trait variability.

Genomic and bioinformatics analysis

ChIP coupled with next-generation sequencing (ChIP-seq) permits detailed mapping of TF interactions and chromatin attributes genome-wide. Enormous efforts have established TF and histone occupancy profiles genome-wide (Giresi et al., 2007; Hannah et al., 2011; Heintzman et al., 2009; Li et al., 2013; May et al., 2013; Shen et al., 2012). These studies suggest the following parameters to predict enhancers: chromatin occupancy, chromatin accessibility, histone modifications (which can regulate accessibility and occupancy), and DNA sequence attributes (including disease mutations and polymorphisms). Statistical tools to integrate multiple ChIP-seq datasets in parallel have revealed combinatorial patterns of protein-DNA interactions genome-wide in diverse cell types (Zeng et al., 2013). These efforts have led to the annotation of thousands of putative enhancers and have devised innovative and powerful methods when combined with direct functional analysis.

While chromatin occupancy and accessibility have merit for localizing enhancers in complex genomes (Kellis et al., 2014; Yue et al., 2014), these parameters are insufficient to predict enhancers at endogenous loci. Prospective enhancers with promising attributes can be deleted from the genome with little to no impact on target gene expression (Sanalkumar et al., 2014; Snow et al., 2010; Snow et al., 2011). By contrast, enhancers with essential non-redundant functions have been identified (Johnson et al., 2012; Johnson et al., 2015). Recently, “super-enhancers” have been described as a subclass of large, clustered enhancers with particularly high ChIP-seq signals that can be several thousand base pairs in length (Hnisz et al., 2013). The attributes of these complex sequences recapitulate those of conventional enhancers, albeit super-enhancers are considerably longer and therefore contain many more prospective *cis*-elements. The mechanistic relationship of super-enhancers to locus control regions (LCRs) (Forrester et al., 1987; Grosveld. et al., 1987) and conventional

enhancers is unclear. The β -globin LCR confers high-level transcription to all of the β -like globin genes at different stages of development (Bender et al., 2000; Bender et al., 2001; Fiering et al., 1995; Forrester et al., 1987; Grosveld. et al., 1987), and this activity involves mechanistic steps shared with enhancers, namely chromatin looping, coregulator recruitment, Pol II recruitment to the enhancer and promoter, and control of target loci subnuclear localization (Bresnick et al., 2006; Katsumura et al., 2013).

Despite limitations to deriving definitive functional insights, chromatin attribute datasets permit comparisons of multiple parameters at cohorts of predicted enhancers. Enhancer sequences can be stratified genome-wide based on their contextual parameters at endogenous loci (Hannah et al., 2011). Enhancer activity in transient transfection assays with reporter genes in plasmids and/or in transgenic mouse harboring reporter genes integrated at ectopic chromosomal loci can be utilized as training sets for algorithms that may improve upon existing approaches to predict enhancers. Comparison of the attributes (e.g., chromatin occupancy, accessibility, and function) of sequence motifs, such as E-box-GATA elements, genome-wide allows one to classify these elements based on shared parameters, which may infer functional similarity (Hewitt et al., 2015). Methodological development continues to improve with the goal to predict the full complement of genomic enhancers with activity at endogenous loci. As functional validation at endogenous loci has lagged behind the development of enhancer predictive utilities, the picture remains murky regarding the extent to which the thousands of predicted enhancers increase promoter activity at their endogenous loci. The advent of facile gene editing technologies has revolutionized validation strategies, which can now rigorously test function at endogenous loci in cultured cells and in mice.

Discovery and validation of enhancers controlling hematopoietic stem and progenitor cells

Historically, enhancers were defined based on activity to increase reporter gene activity in a position- and orientation-independent manner in transiently transfected plasmids (Khoury and Gruss, 1983; Wasylyk et al., 1984). Electrophoretic mobility shift assays (EMSA) (Garner and Revzin, 1981) were used to discover proteins that bind with sequence-specificity to the functionally important sequences. This conceptually straightforward approach was extended to transgenic mice in which one or more copies of reporter genes are integrated at ectopic chromosomal sites (Pennacchio et al., 2006). While these approaches can identify enhancers, enhancer activity at endogenous loci almost invariably differs from that at ectopic chromosomal sites. In part, this reflects the differential chromatin environments at endogenous loci that create unique mechanistic requirements for conferring activation or repression. DNA sequences may elicit promiscuous enhancer activity at ectopic chromosomal sites due to the absence of sequences at the endogenous locus that occlude the respective *cis*-elements. Distinct ensembles of neighboring *cis*-elements, may be permissive, stimulatory or antagonistic for enhancer function.

Technologies to discover and analyze enhancers fall into two groups. Correlative methods infer function from DNA-protein interaction and chromatin attribute patterns. Causative methods involve disrupting or deleting the endogenous sequence, which can provide a definitive assessment of functional importance, at least in the specific context analyzed.

However, as enhancers often have cell type- and developmental-specific activities, deletion analysis in any single cell would not be expected to identify all enhancers.

Multiple technologies have emerged to establish the chromatin accessibility state and TF composition of enhancers. DNase hypersensitivity, formaldehyde assisted isolation of regulatory elements (FAIRE) (Giresi et al., 2007), or ATAC-seq (Buenrostro et al., 2015) are used to identify accessible or “open” chromatin regions that often reflect sites lacking nucleosomes or containing altered nucleosomes. Such sites often harbor clusters of *cis*-regulatory elements bound by TFs and can demarcate enhancers. ChIP relies on formaldehyde crosslinking of proteins to DNA on chromosomes to measure TF, coregulator, and modified histone occupancy of specific chromatin sites (DeVilbiss et al., 2014). A variation of ChIP, chromatin conformation capture (3C) (Dekker et al., 2002), is used to detect higher-order chromatin structure, specifically long range interactions between chromosomal segments, and powerful high-throughput permutations of 3C have been developed (Dekker et al., 2013). This is important, as enhancer localization near or in a gene does not definitively establish that the enhancer is committed to this particular gene. The enhancer could regulate a gene residing in close proximity in three-dimensional space, which might be distant along the DNA template. Alternatively, the target gene might reside on another chromosome. Proteins bound at distal sites that interact with promoters via chromatin looping might crosslink to the promoter, even though the factors are bound directly to distal DNA sequences. Whereas the aggregate data has painted detailed pictures of chromatin landscapes and pinpointed many prospective enhancers, these assays need to be used in concert with direct functional analyses to identify enhancers.

Enhancer deletions at endogenous loci represent the acid-test of functional significance, and this direct analysis can yield results contrasting with those inferred from correlations. DNase hypersensitivity sites 2 and 3 (HS3) upstream of *Hbb* locus have all the hallmarks of a classic enhancer: DNase hypersensitivity, reporter activity and multi-TF occupancy including GATA-1. However, the targeted deletion of either HS2 or HS3 individually in mice only slightly decreases globin gene expression (Bender et al., 2001; Fiering et al., 1995; Hug et al., 1996). The targeted deletion of a major DNase hypersensitivity site at *Gata1* does not influence GATA-1 expression (Guyot et al., 2006). Knockout of a candidate enhancer in the *Scf/TAL1* locus, exhibiting DNase hypersensitivity, robust reporter activity, and transcription factor occupancy, does not impact *Scf/TAL1* expression (Gottgens et al., 2004). As enhancers can be cell type- and developmental-stage specific, minimal or no activity might reflect an inappropriate context for the analysis. Alternatively, certain enhancers may modulate promoter activity, without inducing quantitatively large changes. In principle, incremental regulation may have critical physiological consequences. Finally, certain enhancers may not be particularly crucial in homeostasis, but might function in contexts of stress or other pathological insults. It can therefore be difficult to comprehend negative or incremental results obtained from endogenous locus gene editing analysis.

The issues discussed above are well illustrated by consideration of enhancers controlling *Gata2* expression. Five GATA-1/GATA-2-occupied sites at *Gata2* (named for their distance in kilobases from the *Gata2* transcription start site) are referred to as GATA-switch sites due to the replacement of GATA-2 by GATA-1 at these sites during erythroid

differentiation (Bresnick et al., 2010). Mutant mouse strains lacking individual sites from the endogenous *Gata2* locus have been generated. As the targeted deletion of *Gata2* yields a severe block in definitive hematopoiesis and embryonic lethality, one expectation was that deletion of enhancers essential for conferring *Gata2* expression in HSPCs would also block hematopoiesis and be embryonic lethal. However, -1.8 , -2.8 and -3.9 sites were not individually required for hematopoietic development, despite GATA factor occupancy and capacity to increase reporter gene expression in a transient transfection assay (Grass et al., 2006; Martowicz et al., 2005). The -1.8 site has a unique activity to maintain *Gata2* repression in late-stage erythroblasts, but its deletion has little to no influence on hematopoiesis (Snow et al., 2010). The -2.8 site confers maximal *Gata2* expression, but is not required for *Gata2* expression or repression (Snow et al., 2011). The -3.9 does not influence *Gata2* expression or repression (Sanalkumar et al., 2014). In stark contrast to these results, the intronic $+9.5$ enhancer, containing an E-box-GATA element, is critically required for HSC genesis (Johnson et al., 2012), establishment of HSPCs in the fetal liver, and embryonic development. A distal -77 *cis*-element (Grass et al., 2006) is also required for embryonic development, but operates distinctly from the $+9.5$ site (Johnson et al., 2015). Whereas HSCs are generated in $-77^{-/-}$ mouse embryos, myeloid progenitors accumulate and are defective in their activity to generate diverse myelo-erythroid progeny. Consistent with this impairment of myeloid progenitor cell function, a sector of the myeloid progenitor transcriptome is altered. Thus, multiple *Gata2* sites bearing enhancer attributes exhibit remarkably distinct functions at the endogenous locus *in vivo*. Elucidating mechanisms that endow these distinct functions may permit improved predictions of enhancer function.

Simple gene editing technologies, such as TALENs and CRISPR/Cas9, permit enhancer deletions in cultured cells (Cho et al., 2013; Kim et al., 2013). Thus, it has become relatively straightforward to validate enhancer function at endogenous loci. Moreover, both TALENs and CRISPR/Cas9 have greatly reduced the time required to make mutant mouse strains (Kim and Kim, 2014; Yang et al., 2013). Genome-wide screens using CRISPR libraries can be used to elucidate details of complex genetic networks (Zhou et al., 2014), and this technology is readily adaptable to interrogate candidate enhancers. Despite lacking enhancer activity in a transient transfection assay in a biologically relevant cell type, CRISPR/Cas9-mediated deletion of an E-box-GATA element at the *Bcl2l1* locus indicated its importance at the endogenous locus. This result further illustrates a lack of concordance between correlative and causative data (Hewitt et al., 2015). Predicted disease-associated loci can be deleted in cell lines to elucidate their molecular mechanisms and to inform disease mechanisms (Bauer et al., 2013). These gene editing technologies have transformed efforts to discover functionally critical *cis*-elements, to dissect their sequence requirements, and to establish their roles in cell function and phenotype.

III. Components of Scl/TAL1 and GATA factor complexes at E-box-GATA elements

Scores of TFs that function directly or indirectly through enhancers are implicated in development and function of the hematopoietic system. The scope of this review is restricted to TFs that function through E-box-GATA elements. Whereas it has been known for more

than a decade that multimeric protein complexes assemble at these elements to generate an enhancer (Wadman et al., 1997), many questions remain unanswered regarding the exact nature of the constituents involved and the underlying mechanisms. Given the critical functions of the established components, e.g. GATA-1, GATA-2, Scl/TAL1, LDB1 and LMO2, the future development of work on this problem promises to yield vital biological, mechanistic, and pathological insights.

LMO2

The Scl/TAL1-GATA complex contains the non-DNA binding components LIM-only (LMO) proteins 1 and 2 (Love et al., 2014; Osada et al., 1995). The LIM domain is a cysteine rich zinc-finger-like motif that acts as a docking site for proteins and mediates multi-protein complex assembly. As with GATA-1, GATA-2, and Scl/TAL1, LMO2 promotes erythropoiesis (Warren et al., 1994). LMO2 has little to no transcriptional activation activity with reporter genes in transient transfection assays. LMO2 bridges the GATA factor and Scl/TAL1 at E-box-GATA elements and functions as a scaffold for additional components of the complex (Lecuyer et al., 2007; Wadman et al., 1997). Disruption of *LMO2*, e.g. due to retroviral insertion in a gene therapy context, can cause human myeloid leukemia (Hacein-Bey-Abina et al., 2003). Conditional overexpression of LMO2 in T-cells results in the development of T-cell acute lymphoblastic lymphoma (T-ALL) (Fisch et al., 1992). The LMO2-containing DNA-binding complex in this mouse model of T-ALL forms at a bipartite E-box-E-box sequence, rather than E-box-GATA sequences (Grutz et al., 1998), which highlights a potential new configuration of Scl/TAL1-responsive *cis*-elements lacking a GATA motif.

LDB1

Another well-studied component of the Scl/TAL1-GATA complex at E-box-GATA elements is the non-DNA binding protein LIM-Domain-Binding protein 1 (LDB1). LDB1 binds the LMO2 LIM domain and contributes to the assembly of the Scl/TAL1-GATA complex. LDB1 occupies chromatin at many erythroid-specific genes (Li et al., 2013). *Ldb1* knockout embryos are severely anemic and die at ~E9.5 (Mukhopadhyay et al., 2003). LDB1 promotes formation of higher-order chromatin loops, e.g. at *Hbb* (Song et al., 2007), and these long-range interactions are believed to promote transcriptional activation. LDB1 is also required for subnuclear localization of *Hbb* to sites of active transcription (Song et al., 2010), which may or may not reflect its capacity to induce chromatin looping; the relationship between chromatin looping and alteration of subnuclear localization is not known. The mechanism for the LDB1-induced long-range looping involves LDB1 interactions at promoters and distal enhancers and its capacity to homodimerize, thus tethering together LDB1-bound sites (Soler et al., 2010). While LMO2 provides a scaffold connecting Scl/TAL1 to LDB1, LDB1 self-association appears to underlie promoter-enhancer interactions in certain contexts. This self-association capacity has been exploited to engineer fusion proteins that dock at specific chromosomal sites and have the capacity to interact. These proteins represent innovative tools to study long-range chromatin interactions (Deng et al., 2012).

KLF1

Erythroid Kruppel-like factor 1/Kruppel-like factor 1 (EKLF or KLF1) (Miller and Bieker, 1993) promotes erythropoiesis (Perkins et al., 1995) by regulating a large cohort of erythroid genes (Tallack et al., 2012). KLF1^{-/-} mice die of anemia in late embryonic development, primarily due to defective terminal erythroid maturation (Nuez et al., 1995; Perkins et al., 1995). Nearly 50% of KLF1-occupied sites have neighboring GATA-1-occupancy in erythroid cells (Tallack et al., 2010). A half E-box-GATA motif, with identical spacing to the canonical E-box-GATA element, is frequently associated with KLF1-occupied chromatin sites (Tallack et al., 2010). At LDB1-occupied chromatin, the KLF1 binding motif is also enriched (Li et al., 2013). While this proximity suggests that KLF1 may interact or functionally interface with Scl/TAL1-GATA-1 complexes to coordinate gene expression, understanding the functional relationship between Scl/TAL1-GATA-1 complexes and KLF1 requires additional investigation. Whereas KLF1 is undoubtedly crucial for establishment/maintenance of the erythroid cell transcriptome (Nilson et al., 2006; Pilon et al., 2008; Tallack et al., 2012; Tallack et al., 2010), many questions remain regarding how it functions with GATA-1 and Scl/TAL1, whether it functions via qualitatively distinct mechanisms, or whether it predominantly amplifies GATA-1- and Scl/TAL1-dependent transcriptional outputs.

Chromatin Remodeling SWI/SNF Complex

Bhrama-related gene 1 (BRG1) is a catalytic subunit of the SWI/SNF complex, which has a broad function in diverse cell types to remodel nucleosome structure (Bevilacqua et al., 2014; Hargreaves and Crabtree, 2011). By altering nucleosome structure and/or positioning, which may expose or occlude *cis*-elements and/or alter higher-order chromatin transitions, the SWI/SNF complex can promote or inhibit transcriptional activation. In hematopoietic cells, BRG1 confers chromatin accessibility at enhancers (Hu et al., 2011; Kim et al., 2009a; Kim et al., 2009b). Studies using *Brg1*^{-/-} cells demonstrated that BRG1 regulates chromatin looping at the α - and β -globin loci (Kim et al., 2009a; Kim et al., 2009b). While BRG1 lacks intrinsic sequence specificity for chromosomal site occupancy, it is recruited to chromatin via protein-protein interactions with DNA-bound factors. For example, it is recruited by the Scl/TAL1-GATA-1 complex at an E-box-GATA composite element in the *Epb4.2* promoter (Xu et al., 2006). In erythroid cells, BRG1 also occupies the E-box-GATA element at *Gata2*+9.5 (Sanalkumar et al., 2014). In leukemic cells, BRG1 co-localizes with LMO2 at an enhancer regulating *Myc* expression (Shi et al., 2013). Given the fundamental and broad BRG1 functions, BRG1-dependent mechanisms at E-box-GATA elements are likely to be crucial in both physiological and pathological contexts. However, it would not be surprising if additional chromatin remodelers and modifiers are involved, and the full complement of coregulators that trigger or suppress E-box-GATA element activity is unknown.

Single-stranded DNA binding proteins

Single-stranded DNA binding proteins (SSBPs) bind LDB1 and regulate its function (Chen et al., 2002). This family of proteins exert important roles in embryonic developmental patterning, and hypomorphic alleles of SSBPs have developmental phenotypes resembling

that of LDB1 and LMO2 mutants (van Meyel et al., 2003). SSBP proteins Ssbp2 and Ssbp3 occupy chromatin at the Ebp4.2 E-box-GATA site and modestly increase its transcriptional activity in transfection assays (Xu et al., 2007). Their association with both LDB1 and LMO2 is essential for maintaining the abundance of both proteins in erythroid precursor cells, presumably by inhibiting ubiquitin-mediated protein degradation (Xu et al., 2007). Thus, this mechanism sustains LDB1 levels and promotes multimeric complex assembly at E-box-GATA sites. The human *SSBP2* locus is disrupted in a subset of therapy-induced AML patients, and *Ssbp2*^{-/-} mice are susceptible to lymphoma and other carcinomas, supporting its role as a tumor suppressor (Wang et al., 2010). Ssbp2 also plays a critical role in the maintenance and stress response of HSPC (Li et al., 2014).

IV. Hematopoietic stem/progenitor cell cistrome

A common pattern emerges when examining each factor of the multiprotein complex at E-box-GATA elements. Each TF is an obligate regulator of hematopoiesis during embryonic development, and all are implicated in hematopoietic pathophysiology. Presumably, these factors synergize through E-box-GATA elements to control hematopoietic development and homeostasis. Thus, it is reasonable to assume that alterations in any of the mechanistic steps involving these factors will dysregulate E-box-GATA element function, thereby impairing regulation of the important cohort of genes harboring this element.

E-box-GATA elements regulate gene expression at hematopoietic loci involved in critical cellular processes (Figure 1). An E-box-GATA element controls transcription of *Gata1* (Vyas et al., 1999). As described earlier, the *Gata2*+9.5 E-box-GATA element controls *Gata2* transcription during HSPC genesis/function and is essential for HSC genesis and function (Gao et al., 2013; Johnson et al., 2012; Sanalkumar et al., 2014). At *Epb4.2*, an E-box-GATA element controls expression of a key membrane protein in erythrocytes critical for maintaining red blood cell size and cation transport (Peters et al., 1999; Xu et al., 2003). A *Klf1* E-box GATA element is implicated in *Klf1* expression (Anderson et al., 2000). E-box-GATA elements have been functionally validated using gene editing techniques at the *Bcl2l1*, *Dapp1*, and *Samd14* loci (Hewitt et al., 2015). *Bcl2l1*+12.2 controls expression of Bcl-xL, an essential constituent of apoptotic pathways and survival determinant in HSPCs and other hematopoietic cells (Chao and Korsmeyer, 1998; Gregory et al., 1999). *Dapp1*+23.5 controls expression of Bam32, which regulates germinal center progression, mast cell activation, and B- and T-cell receptor signaling (Han et al., 2003).

Samd14+2.5 contains an E-box-GATA element conserved in mouse and humans. GATA-2 occupies the +2.5 site and regulates *Samd14* transcription in HSPCs (Hewitt et al., 2015). While E-box sequences are not believed to be essential for multimeric complex assembly in all contexts, deletion of E-box sequences from the *Samd14* E-box-GATA composite element severely impairs *Samd14*+2.5 function. +2.5 site function requires LDB1, since LDB1 knockdown in *Samd14*+2.5^{+/-} cells only affects the wild type *Samd14* allele. *Samd14* promotes c-Kit signaling in fetal liver HSPCs, and shRNA-mediated reduction of *Samd14* expression reduces fetal liver hematopoietic progenitor levels/activity. As GATA-2 also activates *Kit* expression (Jing et al., 2008; Munugalavadla et al., 2005), *Gata2*, *Samd14* and *Kit* constitute a type I coherent feedforward loop that controls the function of hematopoietic

progenitors (Hewitt et al., 2015). One would assume that this axis would also regulate HSCs, but this has not yet been described.

V. Perspective

Dissecting mammalian genome function is complicated by the extraordinary diverse, large ensemble of *cis*-elements, a subset of which comprise enhancers. E-box-GATA elements nucleate chromatin assembly of a multiprotein complex that regulates expression of key hematopoietic genes. Given the thousands of E-box-GATA elements in the genome, identifying those with critical functions at endogenous loci could be considered to be formidable. However, recent efforts utilizing sequence and chromatin attributes of essential enhancer elements have identified essential components of the HSPC cistrome (Hewitt et al., 2015). By linking these enhancers to the cognate genes, this approach has led to the discovery of unknown regulators of HSPCs, including a facilitator of SCF-c-Kit signaling (Hewitt et al., 2015). Ultimately, one can assemble a systems-wide network of functionally critical enhancers and their respective genes, rather than factor-occupied sites that might or might not reflect functionally critical sites. This global network derived from the HSPC cistrome will almost certainly inform fundamental processes including, but not limited to, stem cell genesis, self-renewal, and differentiation. It will be extremely interesting to establish the full complement of enhancers resembling the *Gata2*+9.5 and -77 sites in controlling hematopoiesis nonredundantly. Do these enhancers parse into subgroups based on their functions to control distinct sectors of the hematopoietic cellular hierarchy? Do they parse into subgroups based on distinct molecular mechanisms? Clearly, major mechanistic questions remain unanswered, but given existing systems and tools, it has become tractable to sift through genomes with considerably greater ease than ever before possible. Obviously, similar approaches can be deployed to analyze enhancers at loci encoding master regulators that control distinct biological processes.

References

- Anderson KP, Crable SC, Lingrel JB, 1998. Multiple proteins binding to a GATA-E box-GATA motif regulate the erythroid Kruppel-like factor (EKLF) gene. *J. Biol. Chem* 273, 14347–14354. [PubMed: 9603943]
- Anderson KP, Crable SC, Lingrel JB, 2000. The GATA-E-box motif in the EKLF promoter is required for in vivo expression. *Blood* 95, 1652–1655. [PubMed: 10688820]
- Aplan PD, Lombardi DP, Reaman GH, Sather HN, Hammond GD, Kirsch IR, 1992a. Involvement of the putative hematopoietic transcription factor SCL in T-cell acute lymphoblastic leukemia. *Blood* 79, 1327–1333. [PubMed: 1311214]
- Aplan PD, Nakahara K, Orkin SH, Kirsch IR, 1992b. The SCL gene product: a positive regulator of erythroid differentiation. *EMBO J.* 11, 4073–4081. [PubMed: 1396592]
- Aplan PD, Raimondi SC, Kirsch IR, 1992c. Disruption of the SCL gene by a t(1;3) translocation in a patient with T cell acute lymphoblastic leukemia. *J. Exp. Med* 176, 1303–1310. [PubMed: 1402676]
- Bash RO, Crist WM, Shuster JJ, Link MP, Amylon M, Pullen J, Carroll AJ, Buchanan GR, Smith RG, Baer R, 1993. Clinical features and outcome of T-cell acute lymphoblastic leukemia in childhood with respect to alterations at the TAL1 locus: a Pediatric Oncology Group study. *Blood* 81, 2110–2117. [PubMed: 8471769]
- Bash RO, Hall S, Timmons CF, Crist WM, Amylon M, Smith RG, Baer R, 1995. Does activation of the TAL1 gene occur in a majority of patients with T-cell acute lymphoblastic leukemia? A pediatric oncology group study. *Blood* 86, 666–676. [PubMed: 7605997]

- Bauer DE, Kamran SC, Lessard S, Xu J, Fujiwara Y, Lin C, Shao Z, Canver MC, Smith EC, Pinello L, Sabo PJ, Vierstra J, Voit RA, Yuan GC, Porteus MH, Stamatoyannopoulos JA, Lettre G, Orkin SH, 2013. An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. *Science* 342, 253–257. [PubMed: 24115442]
- Bee T, Liddiard K, Swiers G, Bickley SR, Vink CS, Jarratt A, Hughes JR, Medvinsky A, de Bruijn MF, 2009. Alternative Runx1 promoter usage in mouse developmental hematopoiesis. *Blood cells, molecules & diseases* 43, 35–42.
- Begley CG, Aplan PD, Denning SM, Haynes BF, Waldmann TA, Kirsch IR, 1989. The gene SCL is expressed during early hematopoiesis and encodes a differentiation-related DNA-binding motif. *Proc. Natl. Acad. Sci. U. S. A* 86, 10128–10132. [PubMed: 2602361]
- Bender MA, Bulger M, Close J, Groudine M, 2000. Beta-globin gene switching and DNaseI sensitivity of the endogenous beta-globin locus in mice do not require the locus control region. *Mol. Cell* 5, 387–393. [PubMed: 10882079]
- Bender MA, Roach JN, Halow J, Close J, Alami R, Bouhassira EE, Groudine M, Fiering SN, 2001. Targeted deletion of 5'HS1 and 5'HS4 of the beta-globin locus control region reveals additive activity of the DNaseI hypersensitive sites. *Blood* 98, 2022–2027. [PubMed: 11567985]
- Bevilacqua A, Willis MS, Bultman SJ, 2014. SWI/SNF chromatin-remodeling complexes in cardiovascular development and disease. *Cardiovascular pathology : the official journal of the Society for Cardiovascular Pathology* 23, 85–91. [PubMed: 24183004]
- Bresnick EH, Johnson KD, Kim S-I, Im H, 2006. Establishment and regulation of chromatin domains: mechanistic insights from studies of hemoglobin synthesis. *Prog. Nucl. Acids Res. Mol. Biol* 81, 435–471.
- Bresnick EH, Katsumura KR, Lee HY, Johnson KD, Perkins AS, 2012. Master regulatory GATA transcription factors: mechanistic principles and emerging links to hematologic malignancies. *Nucleic Acids Res* 40, 5819–5831. [PubMed: 22492510]
- Bresnick EH, Lee HY, Fujiwara T, Johnson KD, Keles S, 2010. GATA switches as developmental drivers. *J Biol Chem* 285, 31087–31093. [PubMed: 20670937]
- Bresnick EH, Martowicz ML, Pal S, Johnson KD, 2005. Developmental control via GATA factor interplay at chromatin domains. *J Cell Physiol* 205, 1–9. [PubMed: 15887235]
- Briegel K, Lim KC, Plank C, Beug H, Engel JD, Zenke M, 1993. Ectopic expression of a conditional GATA-2/estrogen receptor chimera arrests erythroid differentiation in a hormone-dependent manner. *Genes Dev* 7, 1097–1109. [PubMed: 8504932]
- Buenrostro JD, Wu B, Chang HY, Greenleaf WJ, 2015. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Current protocols in molecular biology*/edited by Frederick M. Ausubel ... [et al.] 109, 21 29 21–29.
- Chao DT, Korsmeyer SJ, 1998. BCL-2 family: regulators of cell death. *Annual review of immunology* 16, 395–419.
- Charron F, Nemer M, 1999. GATA transcription factors and cardiac development. *Semin Cell Dev Biol* 10, 85–91. [PubMed: 10355032]
- Chen L, Segal D, Hukriede NA, Podtelejnikov AV, Bayarsaihan D, Kennison JA, Ogryzko VV, Dawid IB, Westphal H, 2002. Ssdp proteins interact with the LIM-domain-binding protein Ldb1 to regulate development. *Proc. Natl. Acad. Sci. U. S. A* 99, 14320–14325. [PubMed: 12381786]
- Cheng Y, Ma Z, Kim BH, Wu W, Cayting P, Boyle AP, Sundaram V, Xing X, Dogan N, Li J, Euskirchen G, Lin S, Lin Y, Visel A, Kawli T, Yang X, Patacsil D, Keller CA, Giardine B, Mouse EC, Kundaje A, Wang T, Pennacchio LA, Weng Z, Hardison RC, Snyder MP, 2014. Principles of regulatory information conservation between mouse and human. *Nature* 515, 371–375. [PubMed: 25409826]
- Cheng Y, Wu W, Kumar SA, Yu D, Deng W, Tripic T, King DC, Chen K-B, Zhang Y, Drautz D, Giardine B, Schuster SC, Miller W, Chiaromonte F, Zhang Y, Blobel GA, Weiss MJ, Hardison RC, 2009. Erythroid GATA1 function revealed by genome-wide analysis of transcription factor occupancy, histone modifications, and mRNA expression. *Genome Res.* 19, 2172–2184. [PubMed: 19887574]
- Cho SW, Kim S, Kim JM, Kim JS, 2013. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nature biotechnology* 31, 230–232.

- Cohen-Kaminsky S, Maouche-Chretien L, Vitelli L, Vinit MA, Blanchard I, Yamamoto M, Peschle C, Romeo PH, 1998. Chromatin immunoselection defines a TAL-1 target gene. *Embo J* 17, 5151–5160. [PubMed: 9724651]
- Crispino JD, Lodish MB, MacKay JP, Orkin SH, 1999. Use of altered specificity mutants to probe a specific protein-protein interaction in differentiation: the GATA-1:FOG complex. *Mol Cell* 3, 219–228. [PubMed: 10078204]
- Cullen KE, Kladde MP, Seyfred MA, 1993. Interaction between transcriptional regulatory elements of prolactin chromatin. *Science* 261, 203–206. [PubMed: 8327891]
- de Pater E, Kaimakis P, Vink CS, Yokomizo T, Yamada-Inagawa T, van der Linden R, Kartalaei PS, Camper SA, Speck N, Dzierzak E, 2013. Gata2 is required for HSC generation and survival. *J Exp Med* 210, 2843–2850. [PubMed: 24297996]
- Dekker J, Marti-Renom MA, Mirny LA, 2013. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat Rev Genet* 14, 390–403. [PubMed: 23657480]
- Dekker J, Rippe K, Dekker M, Kleckner N, 2002. Capturing chromosome conformation. *Science* 295, 1306–1311. [PubMed: 11847345]
- Deng W, Lee J, Wang H, Miller J, Reik A, Gregory PD, Dean A, Blobel GA, 2012. Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* 149, 1233–1244. [PubMed: 22682246]
- DeVilbiss AW, Boyer ME, Bresnick EH, 2013. Establishing a hematopoietic genetic network through locus-specific integration of chromatin regulators. *Proc Natl Acad Sci U S A* 110, E3398–3407. [PubMed: 23959865]
- DeVilbiss AW, Sanalkumar R, Johnson KD, Keles S, Bresnick EH, 2014. Hematopoietic transcriptional mechanisms: from locus-specific to genome-wide vantage points. *Exp Hematol* 42, 618–629. [PubMed: 24816274]
- Dickinson RE, Griffin H, Bigley V, Reynard LN, Hussain R, Haniffa M, Lakey JH, Rahman T, Wang XN, McGovern N, Pagan S, Cookson S, McDonald D, Chua I, Wallis J, Cant A, Wright M, Keavney B, Chinnery PF, Loughlin J, Hambleton S, Santibanez-Koref M, Collin M, 2011. Exome sequencing identifies GATA-2 mutation as the cause of dendritic cell, monocyte, B and NK lymphoid deficiency. *Blood* 118, 2656–2658. [PubMed: 21765025]
- Dickinson RE, Milne P, Jardine L, Zandi S, Swierczek SI, McGovern N, Cookson S, Ferozepurwalla Z, Langridge A, Pagan S, Gennery A, Heiskanen-Kosma T, Hamalainen S, Seppanen M, Helbert M, Tholouli E, Gambineri E, Reykdal S, Gottfrethsson M, Thaventhiran JE, Morris E, Hirschfield G, Richter AG, Jolles S, Bacon CM, Hambleton S, Haniffa M, Bryceson Y, Allen C, Prchal JT, Dick JE, Bigley V, Collin M, 2014. The evolution of cellular deficiency in GATA2 mutation. *Blood* 123, 863–874. [PubMed: 24345756]
- Dore LC, Chlon TM, Brown CD, White KP, Crispino JD, 2012. Chromatin occupancy analysis reveals genome-wide GATA factor switching during hematopoiesis. *Blood* 119, 3724–3733. [PubMed: 22383799]
- Dorfman DM, Wilson DB, Bruns GA, Orkin SH, 1992. Human transcription factor GATA-2. Evidence for regulation of preproendothelin-1 gene expression in endothelial cells. *J. Biol. Chem* 267, 1279–1285. [PubMed: 1370462]
- Evans T, Felsenfeld G, 1989. The erythroid-specific transcription factor Eryf1: a new finger protein. *Cell* 58, 877–885. [PubMed: 2776214]
- Evans T, Reitman M, Felsenfeld G, 1988. An erythrocyte-specific DNA-binding factor recognizes a regulatory sequence common to all chicken globin genes. *Proc. Natl. Acad. Sci. U. S. A* 85, 5976–5980. [PubMed: 3413070]
- Ferrando AA, Herblot S, Palomero T, Hansen M, Hoang T, Fox EA, Look AT, 2004. Biallelic transcriptional activation of oncogenic transcription factors in T-cell acute lymphoblastic leukemia. *Blood* 103, 1909–1911. [PubMed: 14604958]
- Fiering S, Epner E, Robinson K, Zhuang Y, Telling A, Hu M, Martin DI, Enver T, Ley TJ, Groudine M, 1995. Targeted deletion of 5'HS2 of the murine beta-globin LCR reveals that it is not essential for proper regulation of the beta-globin locus. *Genes Dev.* 9, 2203–2213. [PubMed: 7557375]

- Fisch P, Boehm T, Lavenir I, Larson T, Arno J, Forster A, Rabbitts TH, 1992. T-cell acute lymphoblastic lymphoma induced in transgenic mice by the RBTN1 and RBTN2 LIM-domain genes. *Oncogene* 7, 2389–2397. [PubMed: 1461647]
- Forrester WC, Takegawa S, Papayannopoulou T, Stamatoyannopoulos G, Groudine M, 1987. Evidence for a locus activation region: the formation of developmentally stable hypersensitive sites in globin-expressing hybrids. *Nucleic Acids Res.* 15, 10159–10177. [PubMed: 3480506]
- Fraser P, Bickmore WA, 2007. Nuclear organization of the genome and the potential for gene regulation. *Nature* 447, 413–417. [PubMed: 17522674]
- Fujiwara T, O'Geen H, Keles S, Blahnik K, Linnemann AK, Kang YA, Choi K, Farnham PJ, Bresnick EH, 2009. Discovering hematopoietic mechanisms through genome-wide analysis of GATA factor chromatin occupancy. *Mol Cell* 36, 667–681. [PubMed: 19941826]
- Fujiwara Y, Browne CP, Cunniff K, Goff SC, Orkin SH, 1996. Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. *Proc. Natl. Acad. Sci. U. S. A* 93, 12355–12358. [PubMed: 8901585]
- Fujiwara Y, Chang AN, Williams AM, Orkin SH, 2004. Functional overlap of GATA-1 and GATA-2 in primitive hematopoietic development. *Blood* 103, 583–585. [PubMed: 14504093]
- Gao X, Johnson KD, Chang YI, Boyer ME, Dewey CN, Zhang J, Bresnick EH, 2013. Gata2 cis-element is required for hematopoietic stem cell generation in the mammalian embryo. *J Exp Med* 210, 2833–2842. [PubMed: 24297994]
- Gao Z, Huang Z, Olivey HE, Gurbuxani S, Crispino JD, Svensson EC, 2010. FOG-1-mediated recruitment of NuRD is required for cell lineage re-enforcement during haematopoiesis. *EMBO J* 29, 457–468. [PubMed: 20010697]
- Garner MM, Revzin A, 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia coli* lactose operon regulatory system. *Nucleic Acids Res.* 9, 3047–3060. [PubMed: 6269071]
- Gerby B, Tremblay CS, Tremblay M, Rojas-Sutterlin S, Herblot S, Hebert J, Sauvageau G, Lemieux S, Lecuyer E, Veiga DF, Hoang T, 2014. SCL, LMO1 and Notch1 reprogram thymocytes into self-renewing cells. *PLoS genetics* 10, e1004768. [PubMed: 25522233]
- Giresi PG, Kim J, McDaniell RM, Iyer VR, Lieb JD, 2007. FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome Res* 17, 877–885. [PubMed: 17179217]
- Gottgens B, Broccardo C, Sanchez MJ, Deveau S, Murphy G, Gothert JR, Kotsopoulou E, Kinston S, Delaney L, Piltz S, Barton LM, Knezevic K, Erber WN, Begley CG, Frampton J, Green AR, 2004. The scl +18/19 stem cell enhancer is not required for hematopoiesis: identification of a 5' bifunctional hematopoietic-endothelial enhancer bound by Fli-1 and Elf-1. *Mol. Cell. Biol* 24, 1870–1883. [PubMed: 14966269]
- Grass JA, Boyer ME, Pal S, Wu J, Weiss MJ, Bresnick EH, 2003a. GATA-1-dependent transcriptional repression of GATA-2 via disruption of positive autoregulation and domain-wide chromatin remodeling. *Proc. Natl. Acad. Sci. U. S. A* 100, 8811–8816. [PubMed: 12857954]
- Grass JA, Boyer ME, Pal S, Wu J, Weiss MJ, Bresnick EH, 2003b. GATA-1-dependent transcriptional repression of GATA-2 via disruption of positive autoregulation and domain-wide chromatin remodeling. *Proc Natl Acad Sci U S A* 100, 8811–8816. [PubMed: 12857954]
- Grass JA, Jing H, Kim S-I, Martowicz ML, Pal S, Blobel GA, Bresnick EH, 2006. Distinct functions of dispersed GATA factor complexes at an endogenous gene locus. *Mol. Cell. Biol* 26, 7056–7067. [PubMed: 16980610]
- Gregory T, Yu C, Ma A, Orkin SH, Blobel GA, Weiss MJ, 1999. GATA-1 and erythropoietin cooperate to promoter erythroid cell survival by regulating bcl-xl expression. *Blood* 94, 87–96. [PubMed: 10381501]
- Grosveld F, van Assendelft GB, Greaves DR, Kollias G, 1987. Position-independent high-level expression of the human b-globin gene in transgenic mice. *Cell* 51, 975–985. [PubMed: 3690667]
- Grutz G, Forster A, Rabbitts TH, 1998. Identification of the LMO4 gene encoding an interaction partner of the LIM-binding protein LDB1/NLI1: a candidate for displacement by LMO proteins in T cell acute leukaemia. *Oncogene* 17, 2799–2803. [PubMed: 9840944]

- Guyot B, Murai K, Fujiwara Y, Valverde-Garduno V, Hammett M, Wells S, Dear N, Orkin SH, Porcher C, Vyas P, 2006. Characterization of a megakaryocyte-specific enhancer of the key hemopoietic transcription factor GATA1. *J Biol Chem* 281, 13733–13742. [PubMed: 16551635]
- Hacein-Bey-Abina S, Von Kalle C, Schmidt M, McCormack MP, Wulffraat N, Leboulch P, Lim A, Osborne CS, Pawliuk R, Morillon E, Sorensen R, Forster A, Fraser P, Cohen JI, de Saint Basile G, Alexander I, Wintergerst U, Frebourg T, Aurias A, Stoppa-Lyonnet D, Romana S, Radford-Weiss I, Gross F, Valensi F, Delabesse E, Macintyre E, Sigaux F, Soulier J, Leiva LE, Wissler M, Prinz C, Rabbitts TH, Le Deist F, Fischer A, Cavazzana-Calvo M, 2003. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* 302, 415–419. [PubMed: 14564000]
- Hahn CN, Chong CE, Carmichael CL, Wilkins EJ, Brautigan PJ, Li XC, Babic M, Lin M, Carmagnac A, Lee YK, Kok CH, Gagliardi L, Friend KL, Ekert PG, Butcher CM, Brown AL, Lewis ID, To LB, Timms AE, Storek J, Moore S, Altree M, Escher R, Bardy PG, Suthers GK, D'Andrea RJ, Horwitz MS, Scott HS, 2011. Heritable GATA2 mutations associated with familial myelodysplastic syndrome and acute myeloid leukemia. *Nat Genet* 43, 1012–1017. [PubMed: 21892162]
- Han A, Saijo K, Mecklenbrauker I, Tarakhovsky A, Nussenzweig MC, 2003. Bam32 links the B cell receptor to ERK and JNK and mediates B cell proliferation but not survival. *Immunity* 19, 621–632. [PubMed: 14563325]
- Hannah R, Joshi A, Wilson NK, Kinston S, Gottgens B, 2011. A compendium of genome-wide hematopoietic transcription factor maps supports the identification of gene regulatory control mechanisms. *Exp Hematol* 39, 531–541. [PubMed: 21338655]
- Hargreaves DC, Crabtree GR, 2011. ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. *Cell research* 21, 396–420. [PubMed: 21358755]
- Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, Ye Z, Lee LK, Stuart RK, Ching CW, Ching KA, Antosiewicz-Bourget JE, Liu H, Zhang X, Green RD, Lobanenko VV, Stewart R, Thomson JA, Crawford GE, Kellis M, Ren B, 2009. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* 459, 108–112. [PubMed: 19295514]
- Hewitt KJ, Kim DH, Devadas P, Prathibha R, Zuo C, Sanalkumar R, Johnson KD, Kang YA, Kim JS, Dewey CN, Keles S, Bresnick EH, 2015. Hematopoietic Signaling Mechanism Revealed from a Stem/Progenitor Cell Cistrome. *Mol Cell*.
- Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, Sigova AA, Hoke HA, Young RA, 2013. Super-enhancers in the control of cell identity and disease. *Cell* 155, 934–947. [PubMed: 24119843]
- Ho IC, Vorhees P, Marin N, Oakley BK, Tsai SF, Orkin SH, Leiden JM, 1991. Human GATA-3: a lineage-restricted transcription factor that regulates the expression of the T cell receptor alpha gene. *EMBO J* 10, 1187–1192. [PubMed: 1827068]
- Hong W, Nakazawa M, Chen YY, Kori R, Vakoc CR, Rakowski C, Blobel GA, 2005. FOG-1 recruits the NuRD repressor complex to mediate transcriptional repression by GATA-1. *EMBO J* 24, 67–78.
- Hsu AP, Johnson KD, Falcone EL, Sanalkumar R, Sanchez L, Hickstein DD, Cuellar-Rodriguez J, Lemieux JE, Zerbe CS, Bresnick EH, Holland SM, 2013. GATA2 haploinsufficiency caused by mutations in a conserved intronic element leads to MonoMAC syndrome. *Blood* 121, 3830–3837, S3831–3837. [PubMed: 23502222]
- Hsu AP, Sampaio EP, Khan J, Calvo KR, Lemieux JE, Patel SY, Frucht DM, Vinh DC, Auth RD, Freeman AF, Olivier KN, Uzel G, Zerbe CS, Spalding C, Pittaluga S, Raffeld M, Kuhns DB, Ding L, Paulson ML, Marciano BE, Gea-Banacloche JC, Orange JS, Cuellar-Rodriguez J, Hickstein DD, Holland SM, 2011. Mutations in GATA2 are associated with the autosomal dominant and sporadic monocytopenia and mycobacterial infection (MonoMAC) syndrome. *Blood* 118, 2653–2655. [PubMed: 21670465]
- Hsu HL, Cheng JT, Chen Q, Baer R, 1991. Enhancer-binding activity of the tal-1 oncoprotein in association with the E47/E12 helix-loop-helix proteins. *Mol Cell Biol* 11, 3037–3042. [PubMed: 2038315]

- Hsu HL, Huang L, Tsan JT, Funk W, Wright WE, Hu JS, Kingston RE, Baer R, 1994a. Preferred sequences for DNA recognition by the TAL1 helix-loop-helix proteins. *Mol. Cell. Biol* 14, 1256–1265. [PubMed: 8289805]
- Hsu HL, Wadman I, Baer R, 1994b. Formation of in vivo complexes between the TAL1 and E2A polypeptides of leukemic T-cells. *Proc. Natl. Acad. Sci. U.S.A* 91, 3181–3185. [PubMed: 8159721]
- Hu G, Schones DE, Cui K, Ybarra R, Northrup D, Tang Q, Gattinoni L, Restifo NP, Huang S, Zhao K, 2011. Regulation of nucleosome landscape and transcription factor targeting at tissue-specific enhancers by BRG1. *Genome Res* 21, 1650–1658. [PubMed: 21795385]
- Huang Z, Dore LC, Li Z, Orkin SH, Feng G, Lin S, Crispino JD, 2009. GATA-2 reinforces megakaryocyte development in the absence of GATA-1. *Mol. Cell. Biol* 29, 5168–5180. [PubMed: 19620289]
- Hug BA, Wesselschmidt RL, Fiering S, Bender MA, Epner E, Groudine M, Ley TJ, 1996. Analysis of mice containing a targeted deletion of beta-globin locus control region 5' hypersensitive site 3. *Mol. Cell. Biol* 16, 2906–2912. [PubMed: 8649401]
- Jackson DA, Cook PR, 1995. The structural basis of nuclear function. *Int. Rev. Cytol* 162A, 125–149. [PubMed: 8575879]
- Jeong M, Goodell MA, 2014. New answers to old questions from genome-wide maps of DNA methylation in hematopoietic cells. *Exp Hematol* 42, 609–617. [PubMed: 24993071]
- Jing H, Vakoc CR, Ying L, Mandat S, Wang H, Zheng X, Blobel GA, 2008. Exchange of GATA factors mediates transitions in looped chromatin organization at a developmentally regulated gene locus. *Mol. Cell* 29, 232–242. [PubMed: 18243117]
- Johnson KD, Boyer ME, Kang JA, Wickrema A, Cantor AB, Bresnick EH, 2007. Friend of GATA-1-independent transcriptional repression: a novel mode of GATA-1 function. *Blood* 109, 5230–5233. [PubMed: 17339418]
- Johnson KD, Christensen HM, Zhao B, Bresnick EH, 2001. Distinct mechanisms control RNA polymerase II recruitment to a tissue-specific locus control region and a downstream promoter. *Mol. Cell* 8, 465–471. [PubMed: 11545748]
- Johnson KD, Grass JA, Im H, Park C, Choi K, Bresnick EH, 2003. Highly restricted localization of RNA polymerase II to the hypersensitive site cores of a tissue-specific locus control region. *Mol. Cell. Biol* 23, 6468–6493.
- Johnson KD, Hsu AP, Ryu MJ, Wang J, Gao X, Boyer ME, Liu Y, Lee Y, Calvo KR, Keles S, Zhang J, Holland SM, Bresnick EH, 2012. Cis-element mutated in GATA2-dependent immunodeficiency governs hematopoiesis and vascular integrity. *J Clin Invest* 122, 3692–3704. [PubMed: 22996659]
- Johnson KD, Kong G, Gao X, Chang Y-I, Hewitt KJ, Sanalkumar R, Prathibha R, Ranheim EA, Dewey CN, Zhang J, Bresnick EH, 2015. Cis-regulatory mechanisms governing stem and progenitor cell transitions. *Science Advances*.
- Jones PA, 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* 13, 484–492. [PubMed: 22641018]
- Kallianpur AR, Jordan JE, Brandt SJ, 1994. The SCL/TAL-1 gene is expressed in progenitors of both the hematopoietic and vascular systems during embryogenesis. *Blood* 83, 1200–1208. [PubMed: 8118024]
- Kang YA, Sanalkumar R, O'Geen H, Linnemann AK, Chang CJ, Bouhassira EE, Farnham PJ, Keles S, Bresnick EH, 2012. Autophagy driven by a master regulator of hematopoiesis. *Mol Cell Biol* 32, 226–239. [PubMed: 22025678]
- Kassouf MT, Chagraoui H, Vyas P, Porcher C, 2008. Differential use of SCL/TAL-1 DNA-binding domain in developmental hematopoiesis. *Blood* 112, 1056–1067. [PubMed: 18550854]
- Kassouf MT, Hughes JR, Taylor S, McGowan SJ, Soneji S, Green AL, Vyas P, Porcher C, 2010. Genome-wide identification of TAL1's functional targets: insights into its mechanisms of action in primary erythroid cells. *Genome Res* 20, 1064–1083. [PubMed: 20566737]
- Katsumura KR, DeVilbiss AW, Pope NJ, Johnson KD, Bresnick EH, 2013. Transcriptional mechanisms underlying hemoglobin synthesis. *Cold Spring Harbor perspectives in medicine* 3, a015412. [PubMed: 23838521]

- Kellis M, Wold B, Snyder MP, Bernstein BE, Kundaje A, Marinov GK, Ward LD, Birney E, Crawford GE, Dekker J, Dunham I, Elnitski LL, Farnham PJ, Feingold EA, Gerstein M, Giddings MC, Gilbert DM, Gingeras TR, Green ED, Guigo R, Hubbard T, Kent J, Lieb JD, Myers RM, Pazin MJ, Ren B, Stamatoyannopoulos JA, Weng Z, White KP, Hardison RC, 2014. Defining functional DNA elements in the human genome. *Proc Natl Acad Sci U S A* 111, 6131–6138. [PubMed: 24753594]
- Khoury G, Gruss P, 1983. Enhancer elements. *Cell* 33, 313–314. [PubMed: 6305503]
- Kim H, Kim JS, 2014. A guide to genome engineering with programmable nucleases. *Nat Rev Genet* 15, 321–334. [PubMed: 24690881]
- Kim S-I, Bultman SJ, Kiefer CM, Dean A, Bresnick EH, 2009a. BRG1 requirement for long-range interaction of a locus control region with a downstream promoter. *Proc. Natl. Acad. Sci. U. S. A* 106, 2259–2264. [PubMed: 19171905]
- Kim SI, Bresnick EH, Bultman SJ, 2009b. BRG1 directly regulates nucleosome structure and chromatin looping of the alpha globin locus to activate transcription. *Nucleic Acids Res* 37, 6019–6027. [PubMed: 19696073]
- Kim Y, Kweon J, Kim A, Chon JK, Yoo JY, Kim HJ, Kim S, Lee C, Jeong E, Chung E, Kim D, Lee MS, Go EM, Song HJ, Kim H, Cho N, Bang D, Kim S, Kim JS, 2013. A library of TAL effector nucleases spanning the human genome. *Nature biotechnology* 31, 251–258.
- Ko LJ, Engel JD, 1993. DNA-binding specificities of the GATA transcription factor family. *Mol. Cell. Biol* 13, 4011–4022. [PubMed: 8321208]
- Ko LJ, Yamamoto M, Leonard MW, George KM, Ting P, Engel JD, 1991. Murine and human T-lymphocyte GATA-3 factors mediate transcription through a cis-regulatory element within the human T-cell receptor delta gene enhancer. *Mol Cell Biol* 11, 2778–2784. [PubMed: 2017177]
- Koh KP, Yabuuchi A, Rao S, Huang Y, Cunniff K, Nardone J, Laiho A, Tahiliani M, Sommer CA, Mostoslavsky G, Lahesmaa R, Orkin SH, Rodig SJ, Daley GQ, Rao A, 2011. Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* 8, 200–213. [PubMed: 21295276]
- Kosak ST, Skok JA, Medina KL, Riblet R, Le Beau MM, Fisher AG, Singh H, 2002. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science* 296, 158–162. [PubMed: 11935030]
- Lai F, Shiekhattar R, 2014. Enhancer RNAs: the new molecules of transcription. *Curr Opin Genet Dev* 25, 38–42. [PubMed: 24480293]
- Lecuyer E, Herblot S, Saint-Denis M, Martin R, Begley CG, Porcher C, Orkin SH, Hoang T, 2002. The SCL complex regulates c-kit expression in hematopoietic cells through functional interaction with Sp1. *Blood* 100, 2430–2440. [PubMed: 12239153]
- Lecuyer E, Lariviere S, Sincennes MC, Haman A, Lahlil R, Todorova M, Tremblay M, Wilkes BC, Hoang T, 2007. Protein stability and transcription factor complex assembly determined by the SCL-LMO2 interaction. *J Biol Chem* 282, 33649–33658. [PubMed: 17878155]
- Lee H-Y, Johnson KD, Fujiwara T, Boyer ME, Kim S-I, Bresnick EH, 2009. Controlling hematopoiesis through sumoylation-dependent regulation of a GATA factor. *Mol. Cell* 36, 984–995. [PubMed: 20064464]
- Lee HY, Johnson KD, Boyer ME, Bresnick EH, 2011. Relocalizing genetic loci into specific subnuclear neighborhoods. *J Biol Chem* 286, 18834–18844. [PubMed: 21398517]
- Letting DL, Chen YY, Rakowski C, Reedy S, Blobel GA, 2004. Context-dependent regulation of GATA-1 by friend of GATA-1. *Proc. Natl. Acad. Sci. U. S. A* 101, 476–481. [PubMed: 14695898]
- Li J, Kurasawa Y, Wang Y, Clise-Dwyer K, Klumpp SA, Liang H, Taylor RC, Raymond AC, Estrov Z, Brandt SJ, Davis RE, Zweidler-McKay P, Amin HM, Nagarajan L, 2014. Requirement for ssbp2 in hematopoietic stem cell maintenance and stress response. *J Immunol* 193, 4654–4662. [PubMed: 25238756]
- Li L, Freudenberg J, Cui K, Dale R, Song SH, Dean A, Zhao K, Jothi R, Love PE, 2013. Ldb1-nucleated transcription complexes function as primary mediators of global erythroid gene activation. *Blood* 121, 4575–4585. [PubMed: 23610375]

- Li L, Jothi R, Cui K, Lee JY, Cohen T, Gorivodsky M, Tzchori I, Zhao Y, Hayes SM, Bresnick EH, Zhao K, Westphal H, Love PE, 2011. Nuclear adaptor Ldb1 regulates a transcriptional program essential for the maintenance of hematopoietic stem cells. *Nat Immunol* 12, 129–136. [PubMed: 21186366]
- Ling J, Ainol L, Zhang L, Yu X, Pi W, Tuan D, 2004a. HS2 enhancer function is blocked by a transcriptional terminator inserted between the enhancer and the promoter. *J. Biol. Chem* 279, 51704–51713. [PubMed: 15465832]
- Ling KW, Ottersbach K, van Hamburg JP, Oziemlak A, Tsai FY, Orkin SH, Ploemacher R, Hendriks RW, Dzierzak E, 2004b. GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells. *J. Exp. Med* 200, 871–882. [PubMed: 15466621]
- Linnemann AK, O'Geen H, Keles S, Farnham PJ, Bresnick EH, 2011. Genetic framework for GATA factor function in vascular biology. *Proc Natl Acad Sci U S A* 108, 13641–13646. [PubMed: 21808000]
- Love PE, Warzecha C, Li L, 2014. Ldb1 complexes: the new master regulators of erythroid gene transcription. *Trends in genetics : TIG* 30, 1–9. [PubMed: 24290192]
- Luesink M, Hollink IH, van der Velden VH, Knops RH, Boezeman JB, de Haas V, Trka J, Baruchel A, Reinhardt D, van der Reijden BA, van den Heuvel-Eibrink MM, Zwaan CM, Jansen JH, 2012. High GATA2 expression is a poor prognostic marker in pediatric acute myeloid leukemia. *Blood* 120, 2064–2075. [PubMed: 22786876]
- Lurie LJ, Boyer ME, Grass JA, Bresnick EH, 2008. Differential GATA factor stabilities: implications for chromatin occupancy by structurally similar transcription factors. *Biochem.* 47, 859–869. [PubMed: 18154321]
- Madzo J, Liu H, Rodriguez A, Vasanthakumar A, Sundaravel S, Caces DB, Looney TJ, Zhang L, Lepore JB, Macrae T, Duszynski R, Shih AH, Song CX, Yu M, Yu Y, Grossman R, Raumann B, Verma A, He C, Levine RL, Lavelle D, Lahn BT, Wickrema A, Godley LA, 2014. Hydroxymethylation at gene regulatory regions directs stem/early progenitor cell commitment during erythropoiesis. *Cell reports* 6, 231–244. [PubMed: 24373966]
- Marks-Bluth J, Khanna A, Chandrakanthan V, Thoms J, Bee T, Eich C, Kang YC, Knezevic K, Qiao Q, Fitch S, Oxburgh L, Ottersbach K, Dzierzak E, de Bruijn MF, Pimanda JE, 2015. SMAD1 and SMAD5 Expression Is Coordinately Regulated by FLI1 and GATA2 during Endothelial Development. *Mol Cell Biol* 35, 2165–2172. [PubMed: 25870111]
- Martin DI, Orkin SH, 1989. An erythroid-specific DNA binding factor mediates increased gamma-globin expression in hereditary persistence of fetal hemoglobin (HPFH). *Progress in clinical and biological research* 316A, 217–228. [PubMed: 2480608]
- Martowicz ML, Grass JA, Boyer ME, Guend H, Bresnick EH, 2005. Dynamic GATA factor interplay at a multicomponent regulatory region of the GATA-2 locus. *J Biol Chem* 280, 1724–1732. [PubMed: 15494394]
- May G, Soneji S, Tipping AJ, Teles J, McGowan SJ, Wu M, Guo Y, Fugazza C, Brown J, Karlsson G, Pina C, Olariu V, Taylor S, Tenen DG, Peterson C, Enver T, 2013. Dynamic analysis of gene expression and genome-wide transcription factor binding during lineage specification of multipotent progenitors. *Cell Stem Cell* 13, 754–768. [PubMed: 24120743]
- Merika M, Orkin SH, 1993. DNA-binding specificity of GATA family transcription factors. *Mol. Cell. Biol* 13, 3999–4010. [PubMed: 8321207]
- Miccio A, Wang Y, Hong W, Gregory GD, Wang H, Yu X, Choi JK, Shelat S, Tong W, Poncz M, Blobel GA, 2010. NuRD mediates activating and repressive functions of GATA-1 and FOG-1 during blood development. *EMBO J* 29, 442–456. [PubMed: 19927129]
- Mikkola HK, Klintman J, Yang H, Hock H, Schlaeger TM, Fujiwara Y, Orkin SH, 2003. Hematopoietic stem cells retain long-term repopulating activity and multipotency in the absence of stem-cell leukemia SCL/tal-1 gene. *Nature* 421, 547–551. [PubMed: 12540851]
- Miller IJ, Bieker JJ, 1993. A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Kruppel family of nuclear proteins. *Mol. Cell. Biol* 13, 2776–2786. [PubMed: 7682653]

- Molkentin JD, 2000. The zinc finger-containing transcription factors GATA-4, -5, and -6. Ubiquitously expressed regulators of tissue-specific gene expression. *J Biol Chem* 275, 38949–38952. [PubMed: 11042222]
- Mouthon MA, Bernard O, Mitjavila MT, Romeo PH, Vainchenker W, Mathieu-Mahul D, 1993. Expression of Tal-1 and GATA-binding proteins during human hematopoiesis. *Blood* 81, 647–655. [PubMed: 7678994]
- Mukhopadhyay M, Teufel A, Yamashita T, Agulnick AD, Chen L, Downs KM, Schindler A, Grinberg A, Huang SP, Dorward D, Westphal H, 2003. Functional ablation of the mouse *Ldb1* gene results in severe patterning defects during gastrulation. *Development* 130, 495–505. [PubMed: 12490556]
- Munugalavadda V, Dore LC, Tan BL, Hong L, Vishnu M, Weiss MJ, Kapur R, 2005. Repression of c-kit and its downstream substrates by GATA-1 inhibits cell proliferation during erythroid maturation. *Mol Cell Biol* 25, 6747–6759. [PubMed: 16024808]
- Nichols KE, Crispino JD, Poncz M, White JG, Orkin SH, Maris JM, Weiss MJ, 2000. Familial dyserythropoietic anaemia and thrombocytopenia due to an inherited mutation in GATA-1. *Nat. Genet* 24, 266–270. [PubMed: 10700180]
- Nilson DG, Sabatino DE, Bodine DM, Gallagher PG, 2006. Major erythrocyte membrane protein genes in EKLF-deficient mice. *Exp. Hematol* 34, 705–712. [PubMed: 16728274]
- Nuez B, Michalovich D, Bygrave A, Ploemacher R, Grosveld F, 1995. Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. *Nature* 375, 316–318. [PubMed: 7753194]
- Osada H, Grutz G, Axelson H, Forster A, Rabbitts TH, 1995. Association of erythroid transcription factors: complexes involving the LIM protein RBTN2 and the zinc-finger protein GATA1. *Proc Natl Acad Sci U S A* 92, 9585–9589. [PubMed: 7568177]
- Pal S, Cantor AB, Johnson KD, Moran T, Boyer ME, Orkin SH, Bresnick EH, 2004. Coregulator-dependent facilitation of chromatin occupancy by GATA-1. *Proc. Natl. Acad. Sci. U. S. A* 101, 980–985. [PubMed: 14715908]
- Pennacchio LA, Ahituv N, Moses AM, Prabhakar S, Nobrega MA, Shoukry M, Minovitsky S, Dubchak I, Holt A, Lewis KD, Plajzer-Frick I, Akiyama J, De Val S, Afzal V, Black BL, Couronne O, Eisen MB, Visel A, Rubin EM, 2006. In vivo enhancer analysis of human conserved non-coding sequences. *Nature* 444, 499–502. [PubMed: 17086198]
- Perkins AC, Sharpe AH, Orkin SH, 1995. Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. *Nature* 375, 318–322. [PubMed: 7753195]
- Persons DA, Allay JA, Allay ER, Ashmun RA, Orlic D, Jane SM, Cunningham JM, Nienhuis AW, 1999. Enforced expression of the GATA-2 transcription factor blocks normal hematopoiesis. *Blood* 93, 488–499. [PubMed: 9885210]
- Peters LL, Jindel HK, Gwynn B, Korsgren C, John KM, Lux SE, Mohandas N, Cohen CM, Cho MR, Golan DE, Brugnara C, 1999. Mild spherocytosis and altered red cell ion transport in protein 4.2-null mice. *J Clin Invest* 103, 1527–1537. [PubMed: 10359562]
- Pevny L, Lin CS, D'Agati V, Simon MC, Orkin SH, Costantini F, 1995. Development of hematopoietic cells lacking transcription factor GATA-1. *Development* 121, 163–172. [PubMed: 7867497]
- Pevny L, Simon MC, Robertson E, Klein WH, Tsai SF, D'Agati V, Orkin SH, Costantini F, 1991. Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature* 349, 257–260. [PubMed: 1987478]
- Pilon AM, Arcasoy MO, Dressman HK, Vayda SE, Maksimova YD, Sangerman JI, Gallagher PG, Bodine DM, 2008. Failure of terminal erythroid differentiation in EKLF-deficient mice is associated with cell cycle perturbation and reduced expression of E2F2. *Mol Cell Biol* 28, 7394–7401. [PubMed: 18852285]
- Plank JL, Dean A, 2014. Enhancer function: mechanistic and genome-wide insights come together. *Mol Cell* 55, 5–14. [PubMed: 24996062]
- Porcher C, Swat W, Rockwell K, Fujiwara Y, Alt FW, Orkin SH, 1996. The T cell leukemia oncoprotein SCL/tal-1 is essential for development of all hematopoietic lineages. *Cell* 86, 47–57. [PubMed: 8689686]

- Ragoczy T, Bender MA, Telling A, Byron R, Groudine M, 2006. The locus control region is required for association of the murine beta-globin locus with engaged transcription factories during erythroid maturation. *Genes Dev.* 20, 1447–1457. [PubMed: 16705039]
- Ragoczy T, Telling A, Sawado T, Groudine M, Kosak ST, 2003. A genetic analysis of chromosome territory looping: diverse roles for distal regulatory elements. *Chromosome Res.* 11, 513–525.
- Rajapakse I, Groudine M, 2011. On emerging nuclear order. *J Cell Biol* 192, 711–721. [PubMed: 21383074]
- Rodrigues NP, Janzen V, Forkert R, Dombkowski DM, Boyd AS, Orkin SH, Enver T, Vyas P, Scadden DT, 2005. Haploinsufficiency of GATA-2 perturbs adult hematopoietic stem cell homeostasis. *Blood* 106, 477–484. [PubMed: 15811962]
- Ross J, Mavoungou L, Bresnick EH, Milot E, 2012. GATA-1 utilizes Ikaros and polycomb repressive complex 2 to suppress Hes1 and to promote erythropoiesis. *Mol Cell Biol* 32, 3624–3638. [PubMed: 22778136]
- Rossi L, Lin KK, Boles NC, Yang L, King KY, Jeong M, Mayle A, Goodell MA, 2012. Less is more: unveiling the functional core of hematopoietic stem cells through knockout mice. *Cell Stem Cell* 11, 302–317. [PubMed: 22958929]
- Sanalkumar R, Johnson KD, Gao X, Boyer ME, Chang YI, Hewitt KJ, Zhang J, Bresnick EH, 2014. Mechanism governing a stem cell-generating cis-regulatory element. *Proc Natl Acad Sci U S A.*
- Sanda T, Lawton LN, Barrasa MI, Fan ZP, Kohlhammer H, Gutierrez A, Ma W, Tatarek J, Ahn Y, Kelliher MA, Jamieson CH, Staudt LM, Young RA, Look AT, 2012. Core transcriptional regulatory circuit controlled by the TAL1 complex in human T cell acute lymphoblastic leukemia. *Cancer Cell* 22, 209–221. [PubMed: 22897851]
- Schleif R, 1992. DNA looping. *Annu. Rev. Biochem* 61, 199–223. [PubMed: 1497310]
- Schubeler D, Francastel C, Cimbora DM, Reik A, Martin DI, Groudine M, 2000. Nuclear localization and histone acetylation: a pathway for chromatin opening and transcriptional activation of the human beta-globin locus. *Genes Dev.* 14, 940–950. [PubMed: 10783166]
- Scripture-Adams DD, Damle SS, Li L, Elihu KJ, Qin S, Arias AM, Butler RR 3rd, Champhekar A, Zhang JA, Rothenberg EV, 2014. GATA-3 dose-dependent checkpoints in early T cell commitment. *J Immunol* 193, 3470–3491. [PubMed: 25172496]
- Shen Y, Yue F, McCleary DF, Ye Z, Edsall L, Kuan S, Wagner U, Dixon J, Lee L, Lobanenkov VV, Ren B, 2012. A map of the cis-regulatory sequences in the mouse genome. *Nature* 488, 116–120. [PubMed: 22763441]
- Shi J, Whyte WA, Zepeda-Mendoza CJ, Milazzo JP, Shen C, Roe JS, Minder JL, Mercan F, Wang E, Eckersley-Maslin MA, Campbell AE, Kawaoka S, Shareef S, Zhu Z, Kendall J, Muhar M, Haslinger C, Yu M, Roeder RG, Wigler MH, Blobel GA, Zuber J, Spector DL, Young RA, Vakoc CR, 2013. Role of SWI/SNF in acute leukemia maintenance and enhancer-mediated Myc regulation. *Genes Dev* 27, 2648–2662. [PubMed: 24285714]
- Shivdasani RA, Fujiwara Y, McDevitt MA, Orkin SH, 1997. A lineage-selective knockout establishes the critical role of transcription factor GATA-1 in megakaryocyte growth and platelet development. *EMBO J.* 16, 3965–3973. [PubMed: 9233806]
- Shivdasani RA, Mayer EL, Orkin SH, 1995. Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. *Nature* 373, 432–434. [PubMed: 7830794]
- Simon MC, Pevny L, Wiles MV, Keller G, Costantini F, Orkin SH, 1992. Rescue of erythroid development in gene targeted GATA-1-mouse embryonic stem cells. *Nat. Genet* 1, 92–98. [PubMed: 1302015]
- Snow JW, Trowbridge JJ, Fujiwara T, Emambokus NE, Grass JA, Orkin SH, Bresnick EH, 2010. A single cis element maintains repression of the key developmental regulator Gata2. *PLoS genetics* 6, e1001103. [PubMed: 20838598]
- Snow JW, Trowbridge JJ, Johnson KD, Fujiwara T, Emambokus NE, Grass JA, Orkin SH, Bresnick EH, 2011. Context-dependent function of “GATA switch” sites in vivo. *Blood* 117, 4769–4772. [PubMed: 21398579]
- Soler E, Andrieu-Soler C, de Boer E, Bryne JC, Thongjuea S, Stadhouders R, Palstra RJ, Stevens M, Kockx C, van Ijcken W, Hou J, Steinhoff C, Rijkers E, Lenhard B, Grosveld F, 2010. The

genome-wide dynamics of the binding of Ldb1 complexes during erythroid differentiation. *Genes Dev* 24, 277–289. [PubMed: 20123907]

Song SH, Hou C, Dean A, 2007. A positive role for NL1/Ldb1 in long-range beta-globin locus control region function. *Mol. Cell* 28, 810–822. [PubMed: 18082606]

Song SH, Kim A, Ragoczy T, Bender MA, Groudine M, Dean A, 2010. Multiple functions of Ldb1 required for beta-globin activation during erythroid differentiation. *Blood* 116, 2356–2364. [PubMed: 20570862]

Spinner MA, Sanchez LA, Hsu AP, Shaw PA, Zerbe CS, Calvo KR, Arthur DC, Gu W, Gould CM, Brewer CC, Cowen EW, Freeman AF, Olivier KN, Uzel G, Zelazny AM, Daub JR, Spalding CD, Claypool RJ, Giri NK, Alter BP, Mace EM, Orange JS, Cuellar-Rodriguez J, Hickstein DD, Holland SM, 2013. GATA2 deficiency: a protean disorder of hematopoiesis, lymphatics and immunity. *Blood*.

Szutorisz H, Dillon N, Tora L, 2005. The role of enhancers as centres for general transcription factor recruitment. *Trends Biochem. Sci* 30, 593–599. [PubMed: 16126390]

Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A, 2009. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324, 930–935. [PubMed: 19372391]

Tallack MR, Magor GW, Dartigues B, Sun L, Huang S, Fittock JM, Fry SV, Glazov EA, Bailey TL, Perkins AC, 2012. Novel roles for KLF1 in erythropoiesis revealed by mRNA-seq. *Genome Res* 22, 2385–2398. [PubMed: 22835905]

Tallack MR, Whittington T, Yuen WS, Wainwright EN, Keys JR, Gardiner BB, Nourbakhsh E, Cloonan N, Grimmond SM, Bailey TL, Perkins AC, 2010. A global role for KLF1 in erythropoiesis revealed by ChIP-seq in primary erythroid cells. *Genome Res* 20, 1052–1063. [PubMed: 20508144]

Ting CN, Olson MC, Barton KP, Leiden JM, 1996. Transcription factor GATA-3 is required for development of the T-cell lineage. *Nature* 384, 474–478. [PubMed: 8945476]

Tremblay M, Tremblay CS, Herblot S, Aplan PD, Hebert J, Perreault C, Hoang T, 2010. Modeling T-cell acute lymphoblastic leukemia induced by the SCL and LMO1 oncogenes. *Genes Dev* 24, 1093–1105. [PubMed: 20516195]

Tripic T, Deng W, Cheng Y, Vakoc CR, Gregory GD, Hardison RC, Blobel GA, 2008. SCL and associated protein distinguish active from repressive GATA transcription factor complexes. *Blood* 113, 2191–2201. [PubMed: 19011221]

Tsai F-Y, Orkin SH, 1997. Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation. *Blood* 89, 3636–3643. [PubMed: 9160668]

Tsai FY, Keller G, Kuo FC, Weiss M, Chen J, Rosenblatt M, Alt FW, Orkin SH, 1994. An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature* 371, 221–226. [PubMed: 8078582]

Tsai SF, Martin DI, Zon LI, D'Andrea AD, Wong GG, Orkin SH, 1989. Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression in mammalian cells. *Nature* 339, 446–451. [PubMed: 2725678]

Tsang AP, Fujiwara Y, Hom DB, Orkin SH, 1998. Failure of megakaryopoiesis and arrested erythropoiesis in mice lacking the GATA-1 transcriptional cofactor FOG. *Genes Dev.* 12, 1176–1188. [PubMed: 9553047]

Tsang AP, Visvader JE, Turner CA, Fujiwara Y, Yu C, Weiss MJ, Crossley M, Orkin SH, 1997. FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. *Cell* 90, 109–119. [PubMed: 9230307]

Valverde-Garduno V, Guyot B, Anguita E, Hamlett I, Porcher C, Vyas P, 2004. Differences in the chromatin structure and cis-element organization of the human and mouse GATA1 loci: implications for cis-element identification. *Blood* 104, 3106–3116. [PubMed: 15265794]

van Meyel DJ, Thomas JB, Agulnick AD, 2003. SsdP proteins bind to LIM-interacting co-factors and regulate the activity of LIM-homeodomain protein complexes in vivo. *Development* 130, 1915–1925. [PubMed: 12642495]

- Vicente C, Vazquez I, Conchillo A, Garcia-Sanchez MA, Marcotegui N, Fuster O, Gonzalez M, Calasanz MJ, Lahortiga I, Odero MD, 2012. Overexpression of GATA2 predicts an adverse prognosis for patients with acute myeloid leukemia and it is associated with distinct molecular abnormalities. *Leukemia* 26, 550–554. [PubMed: 21904383]
- Vieira KF, Levings PP, Hilll MA, Cruselle VJ, Kang SH, Engel JD, Bungert J, 2004. Recruitment of transcription complexes to the beta-globin gene locus in vivo and in vitro. *J. Biol. Chem* 279, 50350–50357. [PubMed: 15385559]
- Vyas P, McDevitt MA, Cantor AB, Katz SG, Fujiwara Y, Orkin SH, 1999. Different sequence requirements for expression in erythroid and megakaryocytic cells within a regulatory element upstream of the GATA-1 gene. *Development* 126, 2799–2811. [PubMed: 10331989]
- Wadman IA, Osada H, Grutz GG, Agulnick AD, Westphal H, Forster A, Rabbitts TH, 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. [PubMed: 9214632]
- Wang Y, Klumpp S, Amin HM, Liang H, Li J, Estrov Z, Zweidler-McKay P, Brandt SJ, Agulnick A, Nagarajan L, 2010. SSBP2 is an in vivo tumor suppressor and regulator of LDB1 stability. *Oncogene* 29, 3044–3053. [PubMed: 20348955]
- Warren AJ, Colledge WH, Carlton MB, Evans MJ, Smith AJ, Rabbitts TH, 1994. The oncogenic cysteine-rich LIM domain protein rbt2 is essential for erythroid development. *Cell* 78, 45–57. [PubMed: 8033210]
- Wasylyk B, Wasylyk C, Chambon P, 1984. Short and long range activation by the SV40 enhancer. *Nucleic Acids Res* 12, 5589–5608. [PubMed: 6087293]
- Wechsler J, Greene M, McDevitt MA, Anastasi J, J.E. K, LeBeau MM, Crispino JD, 2002. Acquired mutations in GATA-1 in the megakaryoblastic leukemia of Down syndrome. *Nat. Genet* 32, 148–152. [PubMed: 12172547]
- Weiss MJ, Keller G, Orkin SH, 1994. Novel insights into erythroid development revealed through in vitro differentiation of GATA-1 embryonic stem cells. *Genes Dev.* 8, 1184–1197. [PubMed: 7926723]
- Willis SD, Seyfred MA, 1996. Pituitary-specific chromatin structure of the rat prolactin distal enhancer element. *Nucleic Acids Res* 24, 1065–1072. [PubMed: 8604340]
- Wilson NK, Foster SD, Wang X, Knezevic K, Schutte J, Kaimakis P, Chilarska PM, Kinston S, Ouwehand WH, Dzierzak E, Pimanda JE, de Bruijn MF, Gottgens B, 2010. Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell* 7, 532–544. [PubMed: 20887958]
- Wozniak RJ, Boyer ME, Grass JA, Lee Y, Bresnick EH, 2007. Context-dependent GATA factor function: combinatorial requirements for transcriptional control in hematopoietic and endothelial cells. *J Biol Chem* 282, 14665–14674. [PubMed: 17347142]
- Wozniak RJ, Keles S, Lugus JJ, Young K, Boyer ME, Tran TT, Choi K, Bresnick EH, 2008. Molecular hallmarks of endogenous chromatin complexes containing master regulators of hematopoiesis. *Mol. Cell. Biol* 28, 6681–6694. [PubMed: 18779319]
- Xu Z, Huang S, Chang LS, Agulnick AD, Brandt SJ, 2003. Identification of a TAL1 target gene reveals a positive role for the LIM domain-binding protein Ldb1 in erythroid gene expression and differentiation. *Mol Cell Biol* 23, 7585–7599. [PubMed: 14560005]
- Xu Z, Meng X, Cai Y, Koury MJ, Brandt SJ, 2006. Recruitment of the SWI/SNF protein Brg1 by a multiprotein complex effects transcriptional repression in murine erythroid progenitors. *Biochem J.* 399, 297–304. [PubMed: 16800816]
- Xu Z, Meng X, Cai Y, Liang H, Nagarajan L, Brandt SJ, 2007. Single-stranded DNA-binding proteins regulate the abundance of LIM domain and LIM domain-binding proteins. *Genes Dev.* 21, 942–955. [PubMed: 17437998]
- Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R, 2013. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* 154, 1370–1379. [PubMed: 23992847]
- Yu M, Riva L, Xie H, Schindler Y, Moran TB, Cheng Y, Yu D, Hardison R, Weiss MJ, Orkin SH, Bernstein BE, Fraenkel E, Cantor AB, 2009. Insights into GATA-1-mediated gene activation

versus repression via genome-wide chromatin occupancy analysis. *Mol Cell* 36, 682–695. [PubMed: 19941827]

- Yue F, Cheng Y, Breschi A, Vierstra J, Wu W, Ryba T, Sandstrom R, Ma Z, Davis C, Pope BD, Shen Y, Pervouchine DD, Djebali S, Thurman RE, Kaul R, Rynes E, Kirilusha A, Marinov GK, Williams BA, Trout D, Amrhein H, Fisher-Aylor K, Antoshechkin I, DeSalvo G, See LH, Fastuca M, Drenkow J, Zaleski C, Dobin A, Prieto P, Lagarde J, Bussotti G, Tanzer A, Denas O, Li K, Bender MA, Zhang M, Byron R, Groudine MT, McCleary D, Pham L, Ye Z, Kuan S, Edsall L, Wu YC, Rasmussen MD, Bansal MS, Kellis M, Keller CA, Morrissey CS, Mishra T, Jain D, Dogan N, Harris RS, Cayting P, Kawli T, Boyle AP, Euskirchen G, Kundaje A, Lin S, Lin Y, Jansen C, Malladi VS, Cline MS, Erickson DT, Kirkup VM, Learned K, Sloan CA, Rosenbloom KR, Lacerda de Sousa B, Beal K, Pignatelli M, Flicek P, Lian J, Kahveci T, Lee D, Kent WJ, Ramalho Santos M, Herrero J, Notredame C, Johnson A, Vong S, Lee K, Bates D, Neri F, Diegel M, Canfield T, Sabo PJ, Wilken MS, Reh TA, Giste E, Shafer A, Kutayavin T, Haugen E, Dunn D, Reynolds AP, Neph S, Humbert R, Hansen RS, De Bruijn M, Selleri L, Rudensky A, Josefowicz S, Samstein R, Eichler EE, Orkin SH, Levasseur D, Papayannopoulou T, Chang KH, Skoultschi A, Gosh S, Distechi C, Treuting P, Wang Y, Weiss MJ, Blobel GA, Cao X, Zhong S, Wang T, Good PJ, Lowdon RF, Adams LB, Zhou XQ, Pazin MJ, Feingold EA, Wold B, Taylor J, Mortazavi A, Weissman SM, Stamatoyannopoulos JA, Snyder MP, Guigo R, Gingeras TR, Gilbert DM, Hardison RC, Beer MA, Ren B, Mouse EC, 2014. A comparative encyclopedia of DNA elements in the mouse genome. *Nature* 515, 355–364. [PubMed: 25409824]
- Zeng X, Sanalkumar R, Bresnick EH, Li H, Chang Q, Keles S, 2013. jMOSAICS: joint analysis of multiple ChIP-seq datasets. *Genome Biol* 14, R38. [PubMed: 23844871]
- Zhou Y, Zhu S, Cai C, Yuan P, Li C, Huang Y, Wei W, 2014. High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. *Nature* 509, 487–491. [PubMed: 24717434]
- Zon LI, Gurish MF, Stevens RL, Mather C, Reynolds DS, Austen KF, Orkin SH, 1991. GATA-binding transcription factors in mast cells regulate the promoter of the mast cell carboxypeptidase A gene. *J. Biol. Chem* 266, 22948–22953. [PubMed: 1744088]
- Zullo JM, Demarco IA, Pique-Regi R, Gaffney DJ, Epstein CB, Spooner CJ, Luperchio TR, Bernstein BE, Pritchard JK, Reddy KL, Singh H, 2012. DNA sequence-dependent compartmentalization and silencing of chromatin at the nuclear lamina. *Cell* 149, 1474–1487. [PubMed: 22726435]

Highlights

GATA factors occupy a small subset of GATA motifs in chromatin, and amalgamation of these motifs with distinct motifs yields unique composite *cis*-elements.

A *Gata2* E-box-GATA composite *cis*-element (+9.5) is required for hematopoietic stem cell generation

+9.5 mutation underlies an immunodeficiency that progresses to myelodysplastic syndrome and acute myeloid leukemia.

“+9.5-like” E-box-GATA composite elements are enhancers at endogenous loci in hematopoietic stem/progenitor cells and erythroid precursor cells.

“+9.5-like” composite elements comprise the *cis*-regulatory circuitry (HSPC cistrome) governing hematopoietic stem and progenitor cell transitions.

Figure 1 -

Composite E-box-GATA elements regulate diverse aspects of hematopoietic stem/progenitor cell development and function. A multimeric complex, consisting of GATA-2, Scl/TAL-1, E2A, Lmo2, and LDB1, and likely additional components, forms at a highly restricted subset of E-box-GATA elements in the genome. The GATA-2-activated E-box-GATA elements are important constituents of the HSPC cistrome, which orchestrates vital steps in the stem and progenitor cell transitions that give rise to the full complement of blood cells. Examples of genes controlled by GATA-2 through these elements include *Gata2*, *Samd14*, *Epb4.2*, *Bcl2l1*, and *Dapp1*.