

Blood disease–causing and –suppressing transcriptional enhancers: general principles and *GATA2* mechanisms

Emery H. Bresnick¹⁻⁴ and Kirby D. Johnson¹⁻⁴

¹UW-Madison Blood Research Program, ²Department of Cell and Regenerative Biology, ³Wisconsin Institutes for Medical Research, and ⁴UW Carbone Cancer Center, School of Medicine and Public Health, University of Wisconsin–Madison, Madison, WI

Intensive scrutiny of human genomes has unveiled considerable genetic variation in coding and noncoding regions. In cancers, including those of the hematopoietic system, genomic instability amplifies the complexity and functional consequences of variation. Although elucidating how variation impacts the protein-coding sequence is highly tractable, deciphering the functional consequences of variation in noncoding regions (genome reading), including potential transcriptional-regulatory sequences, remains challenging. A crux of this problem is the sheer abundance of gene-regulatory sequence motifs (*cis* elements) mediating protein-DNA interactions that are intermixed in the genome with thousands of look-alike sequences lacking the capacity to mediate functional interactions with proteins *in vivo*. Furthermore, transcriptional enhancers harbor clustered *cis* elements, and how altering a single *cis* element within a cluster impacts enhancer function is unpredictable. Strategies to discover functional enhancers have been innovated, and human genetics can provide vital clues to achieve this goal. Germline or acquired mutations in functionally critical (essential) enhancers, for example at the *GATA2* locus encoding a master regulator of hematopoiesis, have been linked to human pathologies. Given the human interindividual genetic variation and complex genetic landscapes of hematologic malignancies, enhancer corruption, creation, and expropriation by new genes may not be exceedingly rare mechanisms underlying disease predisposition and etiology. Paradigms arising from dissecting essential enhancer mechanisms can guide genome-reading strategies to advance fundamental knowledge and precision medicine applications. In this review, we provide our perspective of general principles governing the function of blood disease–linked enhancers and *GATA2*-centric mechanisms.

Introduction

Representing a human genetics revolution, next-generation DNA sequencing is routinely used in clinical settings to obtain patient-specific insights into disease etiology, progression, and drug sensitivity. Typically, DNA sequences of exons from a limited candidate gene cohort (panel) are analyzed. Alternatively, whole-exome sequencing generates sequences from a much larger gene cohort. Standardized algorithms are deployed to distinguish between innocuous genetic variation and variation that informs clinical medicine. Simultaneously assessing the structural integrity of many protein-coding genes has been transformative. From the perspective of transitioning outside of known territory, however, a major limitation is that these analyses are blind to sequences beyond exons at enhancers, promoters, and chromatin insulators. Genetic variation in noncoding sequences is commonly deemed “variants of undetermined significance.” Because *cis*-element genetic variation can yield phenotypic

consequences as profound as null mutations within a gene, panels and exome sequencing yield incomplete sketches with intrinsic limitations for advancing genome science and patient care.

The shortcomings of gene panels and whole-exome sequencing can be surmounted by whole-genome sequencing, albeit the cost can be prohibitive in clinical contexts. Irrespective of economics, whole-genome sequencing fails to detect or discards sequences from genomic regions with physical properties that create obstacles to sequencing analytical pipelines, for example, repetitive sequences that do not map uniquely to discrete targets. Repetitive sequences, such as retrotransposons, can confer regulatory functions.¹ In pathologies characterized by a low mutant allele burden, mutation detection necessitates a high sequencing depth. Irrespective of obstacles to documenting variation, it remains challenging to definitively ascertain the significance of noncoding region variation. From an acute clinical perspective, the less than optimal genome-reading logistics may yield data that are not deemed beneficial and/or generate more questions than answers. Whole-genome-sequencing data can influence perceptions regarding a patient's health and/or propensity to develop disease, even though the data may not yield high-fidelity predictions nor inform interventions. Although we have only begun to scratch the surface of deciphering noncoding sequence variation, as genome-reading acumen improves, clinically annotated patient-sequence banks will constitute an invaluable resource to advance fundamental and clinical/translational research.

An attractive approach for deciphering genome function involves amalgamating genomic data documenting histone posttranslational modifications, cytosine guanine dinucleotide DNA methylation and hydroxymethylation, chromatin accessibility, transcription factor and coregulator occupancy with evolutionary conservation, and DNA sequence to generate topographic maps genome-wide.² This limited-dimensional analysis can be enhanced via strategies that incorporate 3-dimensional chromosome conformation, for example, HiC³ and Capture C,^{4,5} to reveal the spatial relationship between a putative regulatory sequence and neighboring genes that may not be evident in 2-dimensional space.⁶ Genetic variation may not necessarily impact the nearest-neighbor genes in 2 dimensions. To pinpoint bona fide regulatory elements, combinations of these parameters can yield instructive predictions.⁷⁻¹¹ Inferences regarding potential functionality require direct testing, which is enabled by gene-editing technologies, with zinc-finger nucleases^{12,13} or transcription activator-like effector nucleases (TALENs)¹⁴ and now predominantly clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9)¹⁵⁻¹⁷ or Cas9-like permutations¹⁸ to excise sequences from a genome. Alternatively, one can use designer fusion proteins containing a DNA-binding domain recognizing the sequence of interest fused to a module that activates or represses genes at the docking site.¹⁹⁻²¹ Although rigorous functional analyses at endogenous loci are increasingly feasible, they remain challenging in low-abundant cell populations and contexts that cannot be recapitulated faithfully with cultured cells. Given the swift pace of this technology development, current limitations will likely be surmounted in the near-term, and overcoming difficulties will further transform genome science, clinical genetics, and precision medicine.

Establishing and maintaining cell-type-specific transcriptomes

As genomes generate dynamically regulated transcriptomes, the transcriptome, which can be straightforward to measure, serves as an invaluable proxy of genome functional status. Cell-type-specific transcription factors act in concert with large ensembles of broadly expressed transcription factors and coregulators at enhancers to establish and maintain transcriptomes. Enhancers were discovered as DNA sequences that confer position-independent and orientation-independent expression of genes on plasmids transfected into cells.²² As technologies evolved from plasmids to transgenes integrated at ectopic chromosomal sites and ultimately endogenous loci, it became clear that enhancers can reside within introns or quite far (eg, 100 kb) or close to a gene. Enhancer function in plasmids often does not correlate, quantitatively or qualitatively, with endogenous locus function. Enhancers can activate reporter genes in plasmids hundreds of fold, for example, the β -globin hypersensitive site II enhancer,²³ while contributing incrementally (eg, ~20%) to endogenous locus activity in vivo.^{24,25} Enhancers consist of clustered *cis* elements mediating transcription factor binding, and their sizes range from that approximating nucleosomal DNA (200 bp)^{26,27} to several or many kilobases (superenhancer²⁸). Regardless of whether an enhancer is remote or proximal to a gene and small or large, enhancer-bound transcription factors recruit chromatin-modifying and -remodeling coregulators²⁹ and RNA polymerase II.^{30,31} Through enzymatic functions, in which coregulators posttranslationally modify histones, and interactions with transcription factors, enhancers stimulate higher-order chromatin transitions (looping)^{32,33}; whether sustained or transient enhancer-promoter interactions are essential is still debated.^{34,35} Regardless of sustained vs transient looping, enhancers relocalize loci within the 3-dimensional topography of the nucleus and its functionally distinct subdomains.³⁶⁻³⁸

Knowledge on sequence and chromatin attributes of enhancers has led to strategies to predict enhancers genome-wide. One or more of the following have enhancer-predictive utility, at least for activity in transfection assays and upon transgene integration at ectopic loci: elevated histone H4 monomethylated at lysine 4, histone H3 acetylated at K27, p300 occupancy, chromatin accessibility, higher-order chromatin conformation, enhancer-derived RNAs, and evolutionary conservation.³⁹⁻⁴² Whether these parameters can be used broadly to predict activity at endogenous loci is unclear. Much more importantly is discriminating between essential enhancers exerting critical functions vs modulatory or redundant enhancers not vital for cellular and organismal functions. However, even a modest degree of enhancer activity may control a crucial biological process.

Examples have emerged in which a sequence fulfills the criteria to qualify as an enhancer, having activity in plasmids and/or in transgenes, yet its excision from the genome yields little to no impact on expression of the associated gene.^{24,25,43} In this case, the enhancer might target a gene residing in proximity in 3- but not 2-dimensional space, reinforcing the need to consider chromosomal conformation. Significant caveats with reconstructing 3-dimensionality include the reliance on formaldehyde to crosslink macromolecules in cells, which yields false-positives and incompletely traps conformations. Conformational maps may only inform the specific cellular contexts in which they are generated,

given the genome remodeling intrinsic to cellular processes such as differentiation. The 3-dimensional configuration of loci in erythroleukemia cells has limited utility to inform genome function in hematopoietic progenitor cells. Ascertaining functional implications of an enhancer deletion requires global transcriptional profiling to establish whether the deletion impacts genes proximal to the enhancer and genes predicted to reside in the neighborhood, based on chromatin conformation. Of course, if an enhancer controls expression of a transcription factor, for example, GATA2, many genes will be dysregulated indirectly. Enhancer-dependent cellular phenotypes can also be highly informative. If an enhancer deletion does not influence neighboring genes or genes more broadly and does not elicit cellular phenotypes, either the enhancer is not essential to regulate the neighboring genes, it regulates genes not critical for cellular physiology, or the physiological processes regulated are not recapitulated in the system or are unknown. A negative result may also reflect redundant activity masked by other *cis* elements, nonredundant activity in particular cell types, or contexts distinct from those analyzed or misassignment of the sequence, which is not a bona fide enhancer.

Although the number of enhancers analyzed at endogenous loci *in vivo* is increasing, reports of those demonstrated unequivocally to exert essential activities (eg, required for development or others vital processes) are limited. It is unclear whether most enhancers function *in vivo* as all-or-none switches to convert repressed into active genes, or whether a spectrum of activities exists ranging from switch-like behavior to modulatory adjustments in gene expression, the latter being more difficult to analyze and interpret.

Discovering enhancers essential for hematopoiesis: GATA2 paradigm

Because enhancer activities can be highly context-dependent, it is instructive to consider how they establish/maintain unique protein expression patterns, for example, in stem and progenitor cells. It is not our intent to describe all enhancers studied in the hematopoietic system, but rather to focus on principles illustrated by essential enhancers that control transcription and important biological processes.

GATA2 is required for hematopoietic stem cell (HSC) generation and function,⁴⁴⁻⁴⁷ myeloerythroid progenitor generation and function,^{48,49} the function of committed erythroid precursors,⁴⁹ and even endothelial cell function.⁵⁰⁻⁵⁷ As disrupted expression or mutational alteration of GATA2 are pathogenic, ensuring the fidelity of GATA2 expression and function in these distinct cellular contexts is crucial. GATA2 expression in hemogenic endothelial cells in the aorta gonad mesonephros region (AGM) of the embryo induces HSC emergence.^{54,58} GATA2 expression in HSCs confers long-term repopulating activity, and its expression in myeloerythroid progenitors and erythroid precursors confers differentiation potential.^{45,48,49,59} GATA2 stimulates cellular proliferation and promotes survival,^{44,47,60} although the underlying mechanisms, and whether unifying mechanisms operate in the distinct contexts, are unclear. As GATA2 functions in widely variable regulatory milieus, mechanisms governing its expression are likely to be context-dependent. Alternatively, a comparable

cohort of regulatory factors in distinct contexts might generate a common mechanism.

Gata2 nucleoprotein structure was initially elucidated in a mouse erythroid precursor cell line lacking GATA1 (G1E)^{61,62} that expresses endogenous GATA2. This work led to the discovery of conserved *Gata2* enhancers with essential activities to control embryogenesis, as well as developmental and regenerative hematopoiesis.^{48,51,63-67} The +9.5 and -77 (9.5 and 77 kb downstream and upstream of the transcription start site) enhancers are essential for embryogenesis and hematopoiesis *in vivo* (Figure 1), and their disruption leads to human pathologies including leukemia.^{48,51,68} As GATA1 is required for erythroid and megakaryocytic differentiation,⁶⁹⁻⁷² GATA2-expressing G1E cells propagate in an immature, erythroid precursor state. Activation of a conditional GATA1 allele (ER-GATA1) represses *Gata2* transcription and induces erythroid maturation.^{63,73} GATA1 replaces GATA2 at 5 *Gata2* sites.⁷⁴⁻⁷⁶ These GATA switches correlate with repression, suggesting that GATA2 positively autoregulates *Gata2*, and GATA1 represses *Gata2*, in part, by disrupting positive autoregulation. Reciprocal GATA2 and GATA1 expression patterns occur in diverse mouse and human erythroid systems.⁷⁷⁻⁷⁹

In vitro studies suggested that 1 or more of the *Gata2* GATA switch sites might establish *Gata2* transcription *in vivo* and/or GATA1-instigated *Gata2* repression during erythroid maturation. These possibilities were tested using mice lacking the individual sites. Individual deletions of sites -1.8, -2.8, and -3.9 kb relative to the *Gata2* promoter had little to no consequences for *Gata2* expression, hematopoiesis, and stress erythropoiesis, yet these sites exhibited enhancer attributes.^{43,66,67} As the -1.8-kb deletion resulted in *Gata2* upregulation in erythroblasts where *Gata2* is normally repressed, this site was required for maintenance, but not initiation, of *Gata2* repression.⁶⁶ The -2.8-kb deletion modestly reduced *Gata2* expression in progenitors, suggesting its enhancer function is modulatory, rather than a critical switch.⁶⁷ The -3.9 deletion had little to no impact on *Gata2* expression and hematopoiesis.⁴³

Unlike the -1.8, -2.8, and -3.9 deletions that removed conserved GATA motifs and neighboring sequences representing potential *cis* elements, a 46-bp deletion of the intronic +9.5 site was lethal at approximately embryonic day 13.5 (E13.5).⁵¹ This contrasts with approximately E10.5 lethality of the *Gata2* coding region knockout.⁴⁴ The +9.5 deletion abrogated HSC emergence in the AGM and strongly reduced fetal liver hematopoietic stem and progenitor cells (HSPCs).^{51,58} As the mutant embryos retained abundant primitive erythroid cells, and E9.5 yolk sac generated primitive erythroid colonies with no obvious defects *ex vivo*, the +9.5 deletion selectively impairs definitive hematopoiesis.⁵¹ Although +9.5 intronic localization differs from the -1.8, -2.8, and -3.9 sites, the sites shared GATA factor occupancy, variable degrees of enhancer activity in transfection assays, and enhancer-predictive chromatin attributes. The +9.5 constitutes the sole report of an enhancer essential for triggering stem cell generation.

An evolutionarily conserved GATA factor-occupied sequence (-77) resides downstream of *Rpn1*, the nearest neighbor to the *Gata2* 5' end. As GATA1 represses *Gata2* transcription and does not regulate *Rpn1* expression,⁶³ we hypothesized that -77 is a distal enhancer that controls *Gata2*, collectively with +9.5 or independently in specific contexts.^{65,74} Like the +9.5, a 257-bp deletion of the -77 is embryonic lethal, but -77^{-/-} embryos live longer (lethality after approximately E15.5) than +9.5^{-/-} embryos.⁴⁸ The -77 mutant

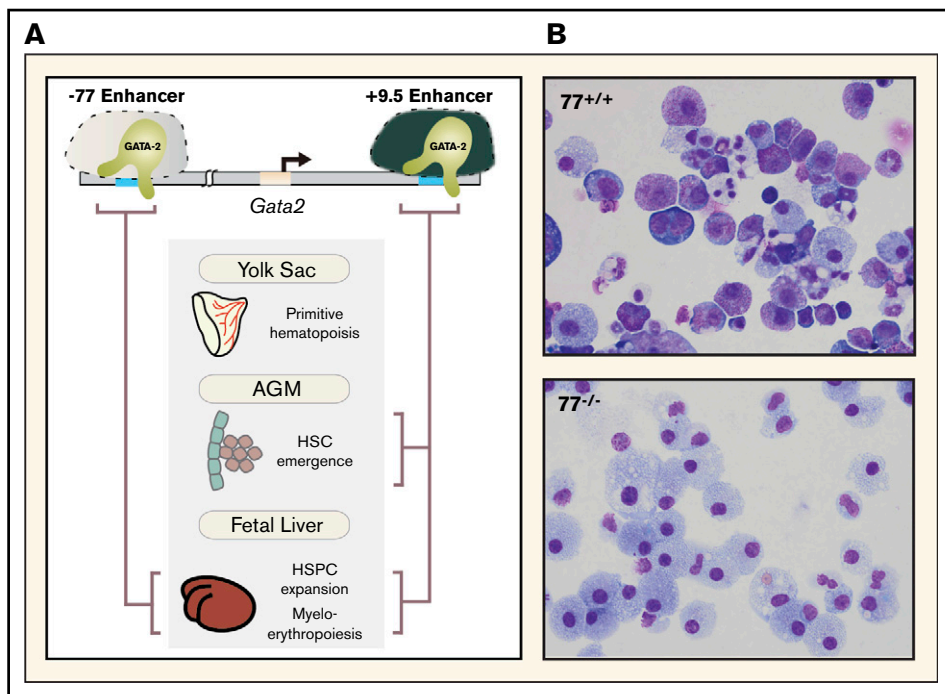


Figure 1. Essential enhancers governing *Gata2* expression and hematopoiesis. (A) The -77 and $+9.5$ enhancers reside 77 and 9.5 kb upstream and downstream, respectively, of the *Gata2* transcriptional start site. These evolutionarily conserved enhancers control mouse and human *Gata2* transcription in specific biological contexts and vital steps in hematopoiesis (depicted with brackets). (B) The photomicrographs depict Giemsa-stained myeloerythroid progeny resulting from ex vivo differentiation of $-77^{+/+}$ or $-77^{-/-}$ fetal liver progenitor cells. AGM, aorta gonad mesonephros region of the embryo proper.

fetal liver myeloerythroid progenitors retain the capacity to undergo monocytic differentiation, while erythroid and granulocytic differentiation is nearly abrogated.^{48,49} Despite the essential $+9.5$ requirement for HSC emergence in the AGM, and essential $+9.5$ and -77 functions, HSC emergence and HSC activity during embryogenesis are unaffected in $-77^{-/-}$ embryos.⁴⁸ Consistent with this activity, -77 confers *Gata2* expression in progenitors, but not in multipotential precursors. Thus, 2 conserved *Gata2* enhancers control distinct processes: $+9.5$ -regulated HSC emergence and -77 -regulated progenitor fate.

It is unknown whether sequences extending beyond the core contribute to core activity. As superenhancers are large enhancers operating under what appear to be similar principles to enhancers (transcription factor occupancy, coregulator recruitment, chromatin looping, etc), this designation does not uniquely inform mechanisms.

Because the $+9.5$, but not the -77 , triggers HSC emergence,^{58,80} it was unclear whether the 2 enhancers ever function collectively. This problem was addressed by analyzing genetic interactions between heterozygous enhancer alleles at distinct anatomical sites and developmental stages. A compound heterozygous mutation eliminating 1 copy of each enhancer on distinct alleles ($-77^{+/-}; +9.5^{+/-}$) is lethal at approximately E14, with no impact on yolk-sac hematopoiesis and HSC emergence and function.⁴⁹ Resembling $-77^{-/-}$ fetal liver myeloid progenitors, $77^{+/-}; +9.5^{+/-}$ progenitors generated predominantly macrophages ex vivo.⁴⁹ The $77^{+/-}; +9.5^{+/-}$ fetal liver lacks megakaryocyte erythrocyte progenitors.⁴⁹ Thus, both enhancers must reside on 1 allele to confer *Gata2* expression to support progenitor function and megakaryocyte erythrocyte progenitor generation. The -77 and $+9.5$ do not interact genetically to control HSC emergence,⁴⁹ illustrating an additional distinction between enhancer requirements for HSC generation vs progenitor generation/function. Interrogating genetic interactions between multiple enhancers

at a locus has broad applicability to elucidate mechanisms operating in distinct contexts in vivo.

In aggregate, these analyses and others demonstrated that essential activities and their quantitative importance cannot be inferred from enhancer attributes conventionally used to predict “enhancers.” Multiple essential enhancers at the same locus can contribute qualitatively unique regulatory modes, and their integrated actions confer expression of GATA2, ensuring its capacity to govern HSPC transitions.

Leveraging essential enhancer attributes to discover enhancer cohorts genome-wide

Because general enhancer attributes do not yield high-fidelity predictions of essential enhancer activities at endogenous loci, can unique attributes of essential enhancers, for example, $+9.5$, be used to identify comparable enhancers? The $+9.5$ core conforms to an E-box (CATCTG) 8-bp spacer GATA motif (AGATAA) composite element^{55,56,81} (Figure 2). This configuration can confer GATA1- or GATA2-dependent enhancer activity in transfection assays,⁸²⁻⁸⁴ and multiple transcription factors (eg, Tal1 and Fli1) and coregulators (eg, Lmo2 and Ldb1) can co-occupy these motifs with the GATA factor.^{82,85,86} As the human genomes contain ~ 8900 CATCTG 6- to 14-bp spacer AGATAA elements,^{87,88} it is instructive to consider what parameters render this sequence, in the context of $+9.5$, an essential enhancer. Is the composite element sufficient for factor binding and activity when situated in accessible chromatin? Do neighboring *cis* elements endow, amplify, or attenuate composite element activity? Does the location relative to gene features (eg, promoter, intron, exon, or distal) dictate activity? Does conservation discriminate functional from nonfunctional elements?

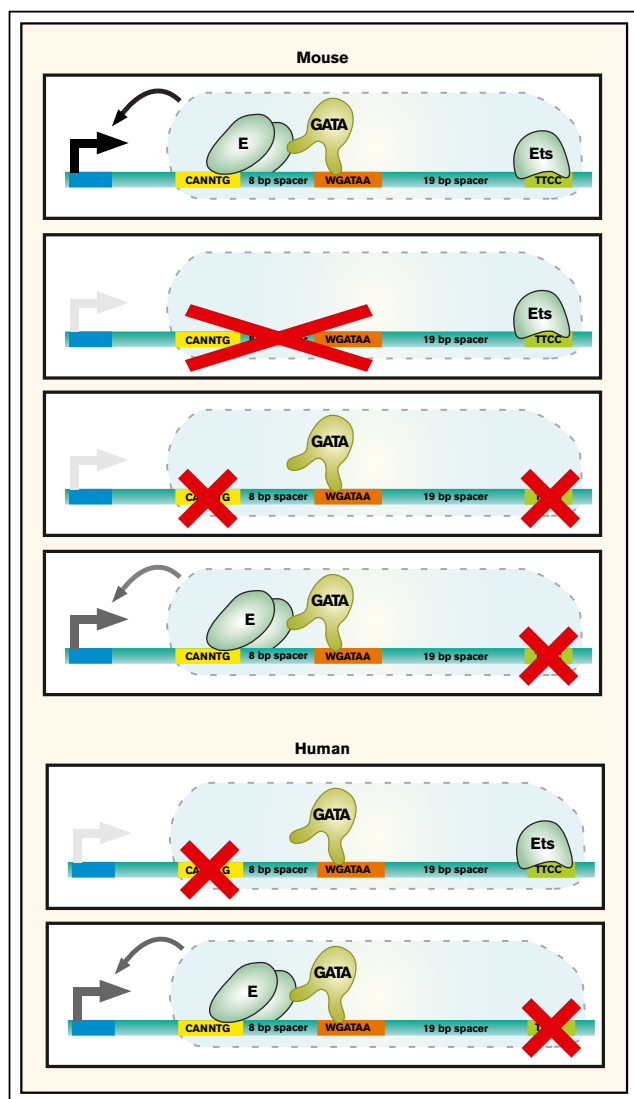


Figure 2. Mouse and human +9.5 enhancer alleles. The wild-type mouse +9.5 configuration is depicted, with 3 mutant mouse alleles below. Ets-only and GATA-only mutants fail to support developmental hematopoiesis and are both embryonic lethal. The Ets motif point mutant that models the human single-nucleotide mutation exhibits largely normal developmental hematopoiesis yet is incompetent to support regeneration of the hematopoietic system post-myeloablation. Human alleles identified in patients with GATA2-deficiency syndrome are depicted. Black arrow extending from +9.5 to promoter, transcriptional activation; gray arrow, partial activation; absence of an arrow, severely attenuated enhancer activity. E, E-box-binding protein.

Hewitt et al⁸⁸ identified all “+9.5-like” composite elements in mouse and human genomes and devised a multifactorial prioritization scheme to parse these potential *cis* elements using parameters characteristic of +9.5, including intronic localization, conservation, GATA2 occupancy, and chromatin attributes. Chromatin immunoprecipitation (ChIP) sequencing data sets (76 histone modification and 38 chromatin occupancy) were used to rank 797 +9.5-like elements, based on their +9.5-like molecular signature. The advent of genetic-editing technologies has transformed our ability to discriminate between potentially

important vs essential enhancers. High- and low-ranked +9.5-like sequences were analyzed using TALENs to excise several elements from their endogenous loci, which identified functional GATA2-activated enhancers; the data set almost certainly harbors many more that remain to be validated.⁸⁸ Deletion of an intronic composite element at the poorly studied *Samd14* gene strongly reduces *Samd14* expression in GATA2-expressing G1E cells, mouse bone marrow, and spleen.^{88,89} Mice lacking this enhancer revealed a *Samd14* function to control erythrocyte regeneration.⁸⁹ Phenylhydrazine- or phlebotomy-induced hemolytic anemia activated the enhancer, increasing *Samd14* expression, stem cell factor-dependent c-Kit signaling, and erythrocyte regeneration. This response confers survival in anemia, thus linking an enhancer mechanism to a vital regenerative process.⁹⁰

Although the strategy described in the prior paragraph is broadly applicable to identify essential enhancers, the exact combination of parameters that enables universal predictions is unknown. The parameters may be context-dependent. For example, in an embryonic stem cell, in which chromatin differs greatly from a differentiated cell, the attributes with enhancer-predictive utility might not extrapolate to all systems. Similarly, examples exist in which factor occupancy of chromatin has a propensity to occur at distal sites in 1 context and promoters in another.⁹¹ Thus, genomic location might constitute a parameter with context-dependent predictive utility.

Recently innovated high-throughput technologies offer new tools, when combined with rigorous locus-specific functional analyses, to identify essential enhancers. HiChIP involves trapping higher-order chromatin interactions in a cell, followed by ChIP to define factor occupancy at chromatin segments engaged in long-range interactions.⁹² By mapping H3K27ac, this approach yields insights into the proximity of potential enhancers to potential target genes.⁹³ Strategies have deployed guide RNAs to direct recruitment of a Kruppel-associated box repressor domain fusion to catalytically inactive Cas9 to chromatin (CRISPR interference).^{19,20,94-96} Gasperini et al used this strategy to analyze functional consequences at 5920 candidate enhancers, profiling >250 000 single-cell transcriptomes, which revealed 470 “enhancer-gene pairs.”⁹⁶ However, gene repression or activation by an artificial fusion protein may occur through diverse mechanisms, including those independent of disrupting endogenous enhancer activity. Even if altered expression reflects disrupted enhancer activity, whether the enhancer is essential, modulatory, or redundant in physiological and pathological contexts requires genome deletion analysis, ideally *in vivo* to interrogate developmental and context-dependent activities. Combining these approaches with single-nucleotide polymorphisms and genetic variation in pathological contexts can uncover enhancers linked to human phenotypes and disease.

Enhancer mechanisms that suppress nonmalignant and malignant blood diseases

Steven Holland at the National Institute of Allergy and Infectious Diseases/National Institutes of Health (NIH), one of the discoverers of germline *GATA2*-coding mutations in patients with immunodeficiency, myelodysplastic syndrome, and acute myeloid leukemia (AML; GATA2-deficiency syndrome),⁹⁷⁻¹⁰² identified a patient with telltale signs of GATA2-related disease, yet lacking *GATA2*-coding

mutations. Sequencing revealed a heterozygous germline 28-bp deletion that disrupts the +9.5 E-box and upstream sequences⁵¹ (Figure 2). Four additional patients harbored a single-nucleotide C-T transition in an Ets motif 23 bp downstream of the 3' end of WGATAA.⁶⁸ As *GATA2* messenger RNA is lower in patient mononuclear cells, and the Ets mutation impairs +9.5 activity in a transfection assay, a haploinsufficiency mechanism of pathogenesis was proposed.⁶⁸

Several hundred adult and pediatric patients with *GATA2* germline mutations have been described, with the single-nucleotide Ets motif mutation being the most common.¹⁰³ Despite multiple conserved +9.5 sequences,^{51,80} patient mutations have been restricted to the Ets motif and the E-box/upstream sequence and have not been detected in another Ets motif upstream of the E-box, GATA motif, or other sequences. Panels and exome sequencing would not detect +9.5 mutations, and whole-genome sequencing is deployed only in limited clinical contexts. Given that +9.5 is a vital determinant of *GATA2* regulation and its disruption creates a disease predisposition,^{51,68,80,103} medical centers (eg, NIH Clinical Center and University of Chicago) screen for +9.5 genetic variation.

As with *GATA2*-coding mutant patients, not all +9.5-mutant patients develop disease, and there is major variability in the disease onset age.^{101,104} These findings suggest a model in which *GATA2* mutations create a disease predisposition insufficient for pathogenesis, which is strongly supported by modeling the Ets mutation in mice.⁸⁰ Ets motif-mutant embryos develop normally, and the adult hematopoietic system in the steady state is normal, including a nearly indistinguishable multipotent hematopoietic precursor (Lin⁻Sca1⁺Kit⁺ [LSK] cell population) transcriptome vs wild-type cells. However, the mutants are hypersensitive to 5-fluorouracil, which ablates progenitors, forcing HSCs to regenerate the hematopoietic system. The mutation corrupts the regenerating LSK cell transcriptome, enhances 5-fluorouracil-induced bone marrow failure, impairs hematopoietic regeneration, and increases lethality. Structure/function analyses revealed that the +9.5 GATA motif is insufficient to support developmental hematopoiesis without the E-box and Ets motifs⁸⁰ (Figure 2).

As the Ets motif mutation sensitizes the hematopoietic system to a secondary insult,⁸⁰ it is instructive to consider the spectrum of insults that impact the mutant human hematopoietic system and whether a predisposition mutation increases the probability of secondary mutations. In principle, a range of genetic and environmental aberrations may trigger the pathogenic consequences of the “silent” *GATA2* mutation. Although these triggers are not established, patients with germline *GATA2* mutations can acquire somatic mutations¹⁰⁵⁻¹⁰⁹ constituting potential triggers, which was reviewed recently.¹⁰³ The triggering mechanism(s) might reduce expression of the heterozygous wild-type allele below a critical threshold, alter function of *GATA2*-regulated genetic network components or impact processes operating in parallel with *GATA2* mechanisms that govern HSPC generation/function.

GATA2 establishes and maintains complex genetic networks.^{48,49,51,58,75,80,87} Network functionality relies on intranetwork circuit integrity and circuit integration. Genetic and environmental aberrations can disrupt network integrity in a nearly infinite number of ways, and pathogenesis may not emerge from a predominant molecular aberration. This model extends the

haploinsufficiency concept to loss-of-function (enhancer mutation) and gain-of-function (*GATA2* overexpression or ectopic signaling that increases *GATA2* activity) scenarios, both corrupting networks that regulate stem/progenitor cells. This new vision of *GATA2*-linked pathogenesis is supported by findings that *GATA2*-coding disease mutations are not strictly loss of function.^{110,111} In a genetic rescue assay in primary progenitor cells, mutants can retain activity or exert activity greater than *GATA2* at select target genes.¹¹⁰ Although the mutants are defective in rescuing erythroid differentiation in progenitors with reduced *GATA2* expression, they can retain or have exaggerated granulopoiesis-inducing activity.¹¹⁰

Analogous to *GATA2*, expression of *Spi1* encoding the myeloid transcription factor PU.1, an Ets family member, must be tightly controlled to ensure normal hematopoiesis.¹¹²⁻¹¹⁷ Although *GATA2* is not a determinant of PU.1 expression in progenitors,^{48,49} it may function with PU.1, positively or negatively, in certain contexts.¹¹³ PU.1 levels are regulated by an enhancer 14 kb upstream of the *Spi1* promoter (upstream regulatory element [URE]).^{114,116,117} Unlike +9.5 and -77, the URE is not essential for survival during development and in adult mice.¹¹⁴ Homozygous deletion of the URE causes a large, but incomplete, decrease in PU.1 expression in bone marrow LSK cells and B220⁺ B cells. *Spi1*^{-/-} mice develop hematopoietic defects, including B-cell lymphoproliferative syndrome, altered early thymocyte development, T-cell lymphoma, and AML. As PU.1 expression is higher in mutant vs wild-type DN1 T cells, this enhancer appears to have context-dependent repressor activity. Unlike +9.5 and -77, in which deletion phenotypes were ~100% penetrant, URE deletion phenotypes vary considerably (eg, 6% to 64%). It was proposed that Wnt signaling targets the URE to induce *Spi1* transcription, thereby generating lymphocyte progenitors, whereas differentiation-associated declines in Wnt signaling downregulate PU.1, facilitating T-cell specification.¹¹⁴ Heterozygous URE-mutant mice, in which PU.1 decreases by ~35%, develop a preleukemia state, and combining this mutation in a *Msh2*^{-/-} background yields AML.¹¹⁷ *Msh2* encodes a DNA mismatch repair component. In humans, a URE single-nucleotide polymorphism impacts protein binding and reporter gene activity and is associated with an approximately twofold lower level of endogenous *Spi1* expression. As disrupting DNA repair machinery triggers leukemogenesis, it will be instructive to assess whether this mechanism can be extrapolated to other predisposition mutations.

Because enhancers consist of clustered *cis* elements, and individual elements can be vital for activity, there is ample opportunity for mutational disruption or generation of transcription factor-binding sites within enhancers. Somatic mutations can create transcription factor-binding sites¹¹⁸ that activate or repress a neighboring gene. If such a change occurs at a chromatin site permissive for factor binding, this may dysregulate genes via multiple mechanisms. The ectopically bound transcription factor might induce assembly of a complex that activates a repressed gene, upregulates an expressed gene, alters expression dynamics, or attenuates transcription by displacing endogenous factors from adjacent or overlapping sites, diverting factors away from prescribed locations or creating inhospitable chromatin. Heterozygous somatic indels in T-cell acute lymphocytic leukemia (T-ALL) cell lines and patient samples generate Myb transcription factor-binding sites upstream of the T-ALL oncogene *TAL1*.¹¹⁸ Ectopic Myb occupancy correlates with occupancy by other transcription factors and acquisition of

superenhancer attributes. A heterozygous single-nucleotide change ~4 kb upstream of the oncogenic *LMO1* locus generates a Myb transcription factor-binding site at a region lacking a known enhancer, which induces *LMO1* overexpression in T-ALL.¹¹⁹ Somatic intronic indels at the T-ALL oncogene *LMO2* elevate *LMO2* expression.¹²⁰ For somatic mutations in heterogeneous cell populations, it cannot be assumed that potentially deleterious alterations are critical, as sequence motifs in chromatin are often inaccessible to binding proteins. Discriminating between chromosomal aberrations with cancer-driving activity vs those merely reflecting genomic instability, and therefore surmounting a major impediment to cancer genome reading, necessitates detailed functional analyses.

Usurping enhancer function as a blood cancer-causing mechanism

As a paradigm-establishing discovery, a translocation links *MYC*-coding sequences to an immunoglobulin H (IgH) 3' enhancer, elevating *MYC* expression as a mechanism instigating Burkitt lymphoma.¹²¹⁻¹²⁴ In multiple myeloma, a t(4:14) translocation involving the *IgH* locus upregulates *FGFR3* and *MMSET* expression via acquisition of *IgH* 3' and intronic enhancers, respectively.^{125,126} *MMSET*, which encodes a histone methyltransferase,¹²⁷ overexpression, but not *FGFR3* overexpression, is implicated in myelomagenesis.^{128,129}

Given cancer cell genomic instability, presumably, oncogenic mechanisms involving a chromosomal rearrangement that leads to enhancer expropriation by a gene, resulting in transcriptional induction and a growth and/or survival advantage, are not rare. The chromatin landscape is crucial for deciphering these scenarios, as insulators¹³⁰ and other elements may negate the actions of surreptitiously introduced enhancers, rendering them inactive or diverting their activities to other genes, while protecting the nearest neighbors.

A scenario analogous to *MYC* has emerged with the human *GATA2* enhancer counterpart (h-77) to the -77 enhancer. Although +9.5 mutations can cause *GATA2*-deficiency syndrome,^{51,68} h-77 point mutations have not been reported. In poor-prognosis 3q21;q26 AML, an inversion repositions ~18 kb of sequence, harboring h-77, ~4 Mb upstream of *GATA2* next to *MECOM* encoding the leukemogenic protein *EVI1*. Studies with human cells and mice indicate that h-77-induced *EVI1* expression, concomitant with *GATA2* loss, constitutes the leukemogenic mechanism.¹³¹⁻¹³³ TALEN-mediated excision of the repositioned h-77 in MUTZ-3 AML cells liberates a maturation blockade, resulting in differentiation into monocyte/macrophage-like cells.¹³² Because deleting -77 strongly reduces *Gata2* expression,⁴⁸ this has important implications for *EVI1* upregulation in AML. Removing h-77 from *GATA2* would decrease *GATA2* levels, raising the question of which factors drive h-77 activity to increase *EVI1* transcription. Other than occupancy by *GATA2* and factors that colocalize with *GATA2* (eg, *LDB1*),⁴⁸ mechanisms underlying -77 (and h-77) activity are unresolved.

Corrupting, creating, and expropriating enhancers: general principles

Genetic and epigenetic aberrations corrupt, create, and expropriate enhancers (Figure 3) to cause, promote, or suppress blood pathologies. The examples described herein highlight the impact of mutations and chromosomal aberrations on enhancer-dependent

oncogenic mechanisms involving HSPCs. Epigenetic mechanisms are also critical, although the complexity of the consequences differs greatly from enhancer corruption, which often dysregulates a limited cadre of nearest-neighbor genes. Altered levels or activity of a chromatin regulator, such as a histone methyltransferase, can elicit broad-sweeping epigenome remodeling over a wide swath of the genome. Ascribing the contribution of individual enhancers and genes to cellular phenotypes is extremely complicated. In the case of *GATA2* enhancer corruption, because *GATA2* regulates a large target gene ensemble, this aberration may derail many cellular processes, secondarily to the primary impact on *GATA2* expression.

Considering that multiprotein complexes drive enhancer function, and posttranslational modifications are prevalent in the proteome, altered expression or activity of enzymes mediating these modifications constitutes another mode of dysregulating enhancer function. Signaling mechanisms regulate proteins occupying enhancers and their partners tethered via protein-protein interactions. An instructive example of how oncogenic signaling corrupts a signal-dependent enhancer mechanism emerged from the extensively studied c-Myc oncoprotein. Notch1 signaling activates a long-range *MYC* enhancer that promotes thymocyte development. Dysregulated Notch1 signaling in T-ALL alters enhancer activity as an oncogenic mechanism.¹³⁴ At first glance, it would seem that disrupting signal-dependent transcriptional mechanisms affects broad target gene ensembles. However, such mechanisms can exert context-dependent influences on transcription factor function at restricted target gene cohorts. Oncogenic Ras signaling induces *GATA2* multisite phosphorylation, which increases its activity at only select target genes.^{60,135} Although many questions exist regarding the mechanistic basis of context-dependent cellular signaling, this is likely related to differential coregulator requirements for transcription factor function at distinct target genes.^{136,137} Chromatin access, complex assembly, and coregulator recruitment and utilization all represent steps in which signaling mechanisms differentially influence different loci in distinct subnuclear environments.

A central question relates to why certain enhancers are essential for transcriptional activation, whereas others are modulatory or seem to lack activity. This is crucial when considering how mutations in specific motifs within an enhancer affect function. A single motif within a cluster of motifs may contribute qualitatively or quantitatively to enhancer function or be redundant with other motifs within the cluster. A single-nucleotide change in a motif might abrogate or attenuate factor binding, enhance binding, or impact binding dynamics and therefore complex assembly. Mutations can generate factor-binding sites permissive for factor occupancy in chromatin and enhancer generation. While considering a cancer cell genome rife with mutations and rearrangements, one can envision that these aberrations will create enhancers, while corrupting others, and expropriate enhancers to genes that are not normally enhancer-dependent. If these events occur in a relatively homogenous cell population, for example, a predominant clone in clonal hematopoiesis, it is feasible to deploy current technologies to map prospective enhancers and gene activity and piece together this muddled landscape. However, considering tumor cell heterogeneity and the diversity of genome scrambling in different cells of a tumor, this is a much more daunting problem to contemplate.

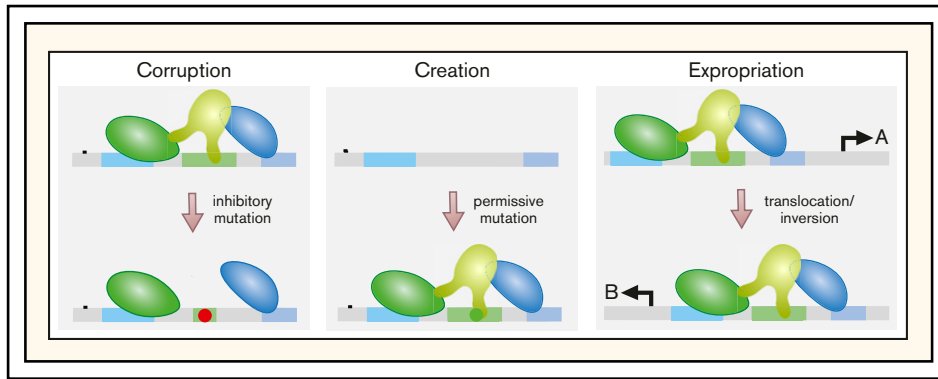


Figure 3. Corrupting, creating, and expropriating enhancers through mutation or chromosomal aberrations. By destroying *cis*-element integrity and impairing factor binding, mutations corrupt enhancer function. Corruption may involve a complete loss of activity (if the factor affected is essential), partial loss of activity (if the remaining enhancer components can confer some degree of transcriptional activation), or acquisition of ectopic activity distinct from that of the wild-type enhancer. Mutations may generate *cis* elements that permit transcription factor binding, following by binding of additional transcription factors and coregulators to generate an enhancer at a site normally lacking an enhancer. Chromosomal inversions or translocations can expropriate a gene's enhancer, transferring it to another gene, thus creating an ectopic gene-regulatory mechanism. A, gene normally controlled by enhancer; B, gene that expropriates the enhancer.

Much more work is required to determine the impact of genetic variation on enhancer corruption, creation, and expropriation *in vivo* and devise strategies to mitigate deleterious actions of rogue enhancers operational in contexts in which they should not exist. Mechanistic advances will elevate genome-reading perspicuity and invariably accelerate clinical genetic and precision medicine opportunities.

Acknowledgments

This work was supported by research funding from the National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases (DK50107 and DK68634).

References

1. Elbarbary RA, Lucas BA, Maquat LE. Retrotransposons as regulators of gene expression. *Science*. 2016;351(6274):aac7247.
2. Davis CA, Hitz BC, Sloan CA, et al. The Encyclopedia of DNA Elements (ENCODE): data portal update. *Nucleic Acids Res*. 2018;46(D1):D794-D801.
3. Lieberman-Aiden E, van Berkum NL, Williams L, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*. 2009;326(5950):289-293.
4. Davies JO, Oudelaar AM, Higgs DR, Hughes JR. How best to identify chromosomal interactions: a comparison of approaches. *Nat Methods*. 2017;14(2):125-134.
5. Oudelaar AM, Davies JOJ, Hanssen LLP, et al. Single-allele chromatin interactions identify regulatory hubs in dynamic compartmentalized domains. *Nat Genet*. 2018;50(12):1744-1751.
6. Dekker J, Belmont AS, Guttman M, et al; 4D Nucleome Network. The 4D nucleome project [published correction appears in *Nature*. 2017;552(7684):278]. *Nature*. 2017;549(7671):219-226.
7. Dickel DE, Ypsilanti AR, Pla R, et al. Ultraconserved enhancers are required for normal development. *Cell*. 2018;172(3):491-499.e415.
8. Osterwalder M, Barozzi I, Tissières V, et al. Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature*. 2018;554(7691):239-243.
9. Dogan N, Wu W, Morrissey CS, et al. Occupancy by key transcription factors is a more accurate predictor of enhancer activity than histone modifications or chromatin accessibility. *Epigenetics Chromatin*. 2015;8:16.
10. Nord AS, Blow MJ, Attanasio C, et al. Rapid and pervasive changes in genome-wide enhancer usage during mammalian development. *Cell*. 2013;155(7):1521-1531.
11. Wu H, Nord AS, Akiyama JA, et al. Tissue-specific RNA expression marks distant-acting developmental enhancers. *PLoS Genet*. 2014;10(9):e1004610.
12. Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci USA*. 1996;93(3):1156-1160.
13. Carroll D. Genome engineering with targetable nucleases. *Annu Rev Biochem*. 2014;83:409-439.
14. Sung YH, Baek IJ, Kim DH, et al. Knockout mice created by TALEN-mediated gene targeting. *Nat Biotechnol*. 2013;31(1):23-24.

15. Knott GJ, Doudna JA. CRISPR-Cas guides the future of genetic engineering. *Science*. 2018;361(6405):866-869.
16. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell*. 2014;157(6):1262-1278.
17. Mali P, Esvelt KM, Church GM. Cas9 as a versatile tool for engineering biology. *Nat Methods*. 2013;10(10):957-963.
18. Liu JJ, Orlova N, Oakes BL, et al. CasX enzymes comprise a distinct family of RNA-guided genome editors [published correction appears in *Nature*. 2019;568(7752):E8-E10]. *Nature*. 2019;566(7743):218-223.
19. Gilbert LA, Larson MH, Morsut L, et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*. 2013;154(2):442-451.
20. Qi LS, Larson MH, Gilbert LA, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*. 2013;152(5):1173-1183.
21. Maeder ML, Linder SJ, Cascio VM, Fu Y, Ho QH, Joung JK. CRISPR RNA-guided activation of endogenous human genes. *Nat Methods*. 2013;10(10):977-979.
22. Banerji R, Rusconi S, Schaffner W. Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell*. 1981;27(2 Pt 1):299-308.
23. Talbot D, Grosveld F. The 5'HS2 of the globin locus control region enhances transcription through the interaction of a multimeric complex binding at two functionally distinct NF-E2 binding sites. *EMBO J*. 1991;10(6):1391-1398.
24. Bender MA, Roach JN, Halow J, et al. Targeted deletion of 5'HS1 and 5'HS4 of the beta-globin locus control region reveals additive activity of the DNaseI hypersensitive sites. *Blood*. 2001;98(7):2022-2027.
25. Fiering S, Epner E, Robinson K, et al. 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*. 1995;9(18):2203-2213.
26. Bulger M, Groudine M. Functional and mechanistic diversity of distal transcription enhancers. *Cell*. 2011;144(3):327-339.
27. Plank JL, Dean A. Enhancer function: mechanistic and genome-wide insights come together. *Mol Cell*. 2014;55(1):5-14.
28. Whyte WA, Orlando DA, Hnisz D, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*. 2013;153(2):307-319.
29. Brownell JE, Zhou J, Ranalli T, et al. Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell*. 1996;84(6):843-851.
30. Johnson KD, Christensen HM, Zhao B, Bresnick EH. Distinct mechanisms control RNA polymerase II recruitment to a tissue-specific locus control region and a downstream promoter. *Mol Cell*. 2001;8(2):465-471.
31. Gurumurthy A, Shen Y, Gunn EM, Bungert J. Phase separation and transcription regulation: are super-enhancers and locus control regions primary sites of transcription complex assembly? *BioEssays*. 2019;41(1):e1800164.
32. Vakoc CR, Letting DL, Gheldof N, et al. Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. *Mol Cell*. 2005;17(3):453-462.
33. Schleif R. DNA looping. *Annu Rev Biochem*. 1992;61:199-223.
34. Chen H, Levo M, Barinov L, Fujioka M, Jaynes JB, Gregor T. Dynamic interplay between enhancer-promoter topology and gene activity. *Nat Genet*. 2018;50(9):1296-1303.
35. Gu B, Swigut T, Spencley A, et al. Transcription-coupled changes in nuclear mobility of mammalian cis-regulatory elements. *Science*. 2018;359(6379):1050-1055.
36. Lee H-Y, Johnson KD, Fujiwara T, Boyer ME, Kim S-I, Bresnick EH. Controlling hematopoiesis through sumoylation-dependent regulation of a GATA factor. *Mol Cell*. 2009;36(6):984-995.
37. Kosak ST, Scalzo D, Alworth SV, et al. Coordinate gene regulation during hematopoiesis is related to genomic organization. *PLoS Biol*. 2007;5(11):e309.
38. Lee HY, Johnson KD, Boyer ME, Bresnick EH. Relocalizing genetic loci into specific subnuclear neighborhoods. *J Biol Chem*. 2011;286(21):18834-18844.
39. Hoffman MM, Ernst J, Wilder SP, et al. Integrative annotation of chromatin elements from ENCODE data. *Nucleic Acids Res*. 2013;41(2):827-841.
40. Arnold CD, Gerlach D, Stelzer C, Boryn LM, Rath M, Stark A. Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science*. 2013;339(6123):1074-1077.
41. Andersson R, Gebhard C, Miguel-Escalada I, et al. An atlas of active enhancers across human cell types and tissues. *Nature*. 2014;507(7493):455-461.
42. Li Y, Shi W, Wasserman WW. Genome-wide prediction of cis-regulatory regions using supervised deep learning methods. *BMC Bioinformatics*. 2018;19(1):202.
43. Sanalkumar R, Johnson KD, Gao X, et al. Mechanism governing a stem cell-generating cis-regulatory element. *Proc Natl Acad Sci USA*. 2014;111(12):E1091-E1100.
44. Tsai FY, Keller G, Kuo FC, et al. An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature*. 1994;371(6494):221-226.
45. Tsai F-Y, Orkin SH. 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*. 1997;89(10):3636-3643.
46. Ling KW, Ottersbach K, van Hamburg JP, et al. GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells. *J Exp Med*. 2004;200(7):871-882.
47. Rodrigues NP, Janzen V, Forkert R, et al. Haploinsufficiency of GATA-2 perturbs adult hematopoietic stem-cell homeostasis. *Blood*. 2005;106(2):477-484.

48. Johnson KD, Kong G, Gao X, et al. Cis-regulatory mechanisms governing stem and progenitor cell transitions. *Sci Adv.* 2015;1(8):e1500503.
49. Mehta C, Johnson KD, Gao X, et al. Integrating enhancer mechanisms to establish a hierarchical blood development program. *Cell Reports.* 2017; 20(12):2966-2979.
50. Linnemann AK, O'Geen H, Keles S, Farnham PJ, Bresnick EH. Genetic framework for GATA factor function in vascular biology. *Proc Natl Acad Sci USA.* 2011;108(33):13641-13646.
51. Johnson KD, Hsu AP, Ryu MJ, et al. Cis-element mutated in GATA2-dependent immunodeficiency governs hematopoiesis and vascular integrity. *J Clin Invest.* 2012;122(10):3692-3704.
52. Kang H, Mesquita WT, Jung HS, Moskvina OV, Thomson JA, Slukvin II. GATA2 is dispensable for specification of hemogenic endothelium but promotes endothelial-to-hematopoietic transition. *Stem Cell Reports.* 2018;11(1):197-211.
53. Butko E, Distel M, Pouget C, et al. Gata2b is a restricted early regulator of hemogenic endothelium in the zebrafish embryo. *Development.* 2015;142(6): 1050-1061.
54. de Pater E, Kaimakis P, Vink CS, et al. Gata2 is required for HSC generation and survival. *J Exp Med.* 2013;210(13):2843-2850.
55. Wozniak RJ, Boyer ME, Grass JA, Lee Y, Bresnick EH. Context-dependent GATA factor function: combinatorial requirements for transcriptional control in hematopoietic and endothelial cells. *J Biol Chem.* 2007;282(19):14665-14674.
56. Khandekar M, Brandt W, Zhou Y, et al. A Gata2 intronic enhancer confers its pan-endothelia-specific regulation. *Development.* 2007;134(9):1703-1712.
57. Lim KC, Hosoya T, Brandt W, et al. Conditional Gata2 inactivation results in HSC loss and lymphatic mispatterning. *J Clin Invest.* 2012;122(10):3705-3717.
58. Gao X, Johnson KD, Chang YI, et al. Gata2 cis-element is required for hematopoietic stem cell generation in the mammalian embryo. *J Exp Med.* 2013; 210(13):2833-2842.
59. Rodrigues NP, Boyd AS, Fugazza C, et al. GATA-2 regulates granulocyte-macrophage progenitor cell function. *Blood.* 2008;112(13):4862-4873.
60. Katsumura KR, Ong IM, DeVilbiss AW, Sanalkumar R, Bresnick EH. GATA factor-dependent positive-feedback circuit in acute myeloid leukemia cells. *Cell Reports.* 2016;16(9):2428-2441.
61. Weiss MJ, Yu C, Orkin SH. Erythroid-cell-specific properties of transcription factor GATA-1 revealed by phenotypic rescue of a gene-targeted cell line. *Mol Cell Biol.* 1997;17(3):1642-1651.
62. Gregory T, Yu C, Ma A, Orkin SH, Blobel GA, Weiss MJ. GATA-1 and erythropoietin cooperate to promote erythroid cell survival by regulating bcl-xL expression. *Blood.* 1999;94(1):87-96.
63. Grass JA, Boyer ME, Pal S, Wu J, Weiss MJ, Bresnick EH. GATA-1-dependent transcriptional repression of GATA-2 via disruption of positive autoregulation and domain-wide chromatin remodeling. *Proc Natl Acad Sci USA.* 2003;100(15):8811-8816.
64. Martowicz ML, Grass JA, Boyer ME, Guend H, Bresnick EH. Dynamic GATA factor interplay at a multicomponent regulatory region of the GATA-2 locus. *J Biol Chem.* 2005;280(3):1724-1732.
65. Grass JA, Jing H, Kim S-I, et al. Distinct functions of dispersed GATA factor complexes at an endogenous gene locus. *Mol Cell Biol.* 2006;26(19):7056-7067.
66. Snow JW, Trowbridge JJ, Fujiwara T, et al. A single cis element maintains repression of the key developmental regulator Gata2. *PLoS Genet.* 2010;6(9): e1001103.
67. Snow JW, Trowbridge JJ, Johnson KD, et al. Context-dependent function of "GATA switch" sites in vivo. *Blood.* 2011;117(18):4769-4772.
68. Hsu AP, Johnson KD, Falcone EL, et al. GATA2 haploinsufficiency caused by mutations in a conserved intronic element leads to MonoMAC syndrome. *Blood.* 2013;121(19):3830-3837, S3831-3837.
69. Pevny L, Simon MC, Robertson E, et al. Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature.* 1991;349(6306):257-260.
70. Pevny L, Lin CS, D'Agati V, Simon MC, Orkin SH, Costantini F. Development of hematopoietic cells lacking transcription factor GATA-1. *Development.* 1995;121(1):163-172.
71. Fujiwara Y, Browne CP, Cunniff K, Goff SC, Orkin SH. Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. *Proc Natl Acad Sci USA.* 1996;93(22):12355-12358.
72. Shivdasani RA, Fujiwara Y, McDevitt MA, Orkin SH. A lineage-selective knockout establishes the critical role of transcription factor GATA-1 in megakaryocyte growth and platelet development. *EMBO J.* 1997;16(13):3965-3973.
73. Crispino JD, Lodish MB, MacKay JP, Orkin SH. Use of altered specificity mutants to probe a specific protein-protein interaction in differentiation: the GATA-1:FOG complex. *Mol Cell.* 1999;3(2):219-228.
74. Bresnick EH, Lee HY, Fujiwara T, Johnson KD, Keles S. GATA switches as developmental drivers. *J Biol Chem.* 2010;285(41):31087-31093.
75. Katsumura KR, Bresnick EH; GATA Factor Mechanisms Group. The GATA factor revolution in hematology. *Blood.* 2017;129(15):2092-2102.
76. Bresnick EH, Katsumura KR, Lee HY, Johnson KD, Perkins AS. Master regulatory GATA transcription factors: mechanistic principles and emerging links to hematologic malignancies. *Nucleic Acids Res.* 2012;40(13):5819-5831.
77. Lugus JJ, Chung YS, Mills JC, et al. GATA2 functions at multiple steps in hemangioblast development and differentiation. *Development.* 2007;134(2): 393-405.
78. Huang J, Liu X, Li D, et al. Dynamic control of enhancer repertoires drives lineage and stage-specific transcription during hematopoiesis. *Dev Cell.* 2016; 36(1):9-23.

79. Gao Z, Huang Z, Olivey HE, Gurbuxani S, Crispino JD, Svensson EC. FOG-1-mediated recruitment of NuRD is required for cell lineage re-enforcement during haematopoiesis. *EMBO J*. 2010;29(2):457-468.
80. Soukup AA, Zheng Y, Mehta C, et al. Single-nucleotide human disease mutation inactivates a blood-regenerative GATA2 enhancer. *J Clin Invest*. 2019; 129(3):1180-1192.
81. Wozniak RJ, Keles S, Lugus JJ, et al. Molecular hallmarks of endogenous chromatin complexes containing master regulators of hematopoiesis. *Mol Cell Biol*. 2008;28(21):6681-6694.
82. Wadman IA, Osada H, Grütz GG, et al. 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*. 1997;16(11):3145-3157.
83. Lécuyer E, Herblot S, Saint-Denis M, et al. The SCL complex regulates c-kit expression in hematopoietic cells through functional interaction with Sp1. *Blood*. 2002;100(7):2430-2440.
84. Xu Z, Huang S, Chang LS, Agulnick AD, Brandt SJ. 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*. 2003;23(21):7585-7599.
85. Wilson NK, Foster SD, Wang X, et al. Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell*. 2010;7(4):532-544.
86. Beck D, Thoms JA, Perera D, et al. Genome-wide analysis of transcriptional regulators in human HSPCs reveals a densely interconnected network of coding and noncoding genes. *Blood*. 2013;122(14):e12-e22.
87. Fujiwara T, O'Geen H, Keles S, et al. Discovering hematopoietic mechanisms through genome-wide analysis of GATA factor chromatin occupancy. *Mol Cell*. 2009;36(4):667-681.
88. Hewitt KJ, Kim DH, Devadas P, et al. Hematopoietic signaling mechanism revealed from a stem/progenitor cell cistrome. *Mol Cell*. 2015;59(1):62-74.
89. Hewitt KJ, Katsumura KR, Matson DR, et al. GATA factor-regulated Samd14 enhancer confers red blood cell regeneration and survival in severe anemia. *Dev Cell*. 2017;42(3):213-225.e214.
90. Bresnick EH, Hewitt KJ, Mehta C, Keles S, Paulson RF, Johnson KD. Mechanisms of erythrocyte development and regeneration: implications for regenerative medicine and beyond. *Development*. 2018;145(1):dev151423.
91. Pilon AM, Ajay SS, Kumar SA, et al; NISC Comparative Sequencing Center. Genome-wide ChIP-Seq reveals a dramatic shift in the binding of the transcription factor erythroid Kruppel-like factor during erythrocyte differentiation. *Blood*. 2011;118(17):e139-e148.
92. Mumbach MR, Rubin AJ, Flynn RA, et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat Methods*. 2016;13(11): 919-922.
93. Mumbach MR, Satpathy AT, Boyle EA, et al. Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements. *Nat Genet*. 2017;49(11):1602-1612.
94. Fulco CP, Munschauer M, Anyoha R, et al. Systematic mapping of functional enhancer-promoter connections with CRISPR interference. *Science*. 2016; 354(6313):769-773.
95. Thakore PI, D'Ippolito AM, Song L, et al. Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat Methods*. 2015;12(12):1143-1149.
96. Gasperini M, Hill AJ, McFaline-Figueroa JL, et al. A genome-wide framework for mapping gene regulation via cellular genetic screens [published correction appears in *Cell*. 2019;176(6):1516]. *Cell*. 2019;176(1-2):377-390.e19.
97. Dickinson RE, Griffin H, Bigley V, et al. Exome sequencing identifies GATA-2 mutation as the cause of dendritic cell, monocyte, B and NK lymphoid deficiency. *Blood*. 2011;118(10):2656-2658.
98. Hahn CN, Chong CE, Carmichael CL, et al. Heritable GATA2 mutations associated with familial myelodysplastic syndrome and acute myeloid leukemia. *Nat Genet*. 2011;43(10):1012-1017.
99. Kazenwadel J, Secker GA, Liu YJ, et al. Loss-of-function germline GATA2 mutations in patients with MDS/AML or MonoMAC syndrome and primary lymphedema reveal a key role for GATA2 in the lymphatic vasculature. *Blood*. 2012;119(5):1283-1291.
100. Hsu AP, Sampaio EP, Khan J, et al. Mutations in GATA2 are associated with the autosomal dominant and sporadic monocytopenia and mycobacterial infection (MonoMAC) syndrome. *Blood*. 2011;118(10):2653-2655.
101. McReynolds LJ, Calvo KR, Holland SM. Germline GATA2 mutation and bone marrow failure. *Hematol Oncol Clin North Am*. 2018;32(4):713-728.
102. Bigley V, Cytlak U, Collin M. Human dendritic cell immunodeficiencies. *Semin Cell Dev Biol*. 2019;86:50-61.
103. Churpek JE, Bresnick EH. Transcription factor mutations as a cause of familial myeloid neoplasms. *J Clin Invest*. 2019;129(2):476-488.
104. Spinner MA, Sanchez LA, Hsu AP, et al. GATA2 deficiency: a protean disorder of hematopoiesis, lymphatics, and immunity. *Blood*. 2014;123(6):809-821.
105. Greif PA, Dufour A, Konstandin NP, et al. GATA2 zinc finger 1 mutations associated with biallelic CEBPA mutations define a unique genetic entity of acute myeloid leukemia. *Blood*. 2012;120(2):395-403.
106. Green CL, Tawana K, Hills RK, et al. GATA2 mutations in sporadic and familial acute myeloid leukaemia patients with CEBPA mutations. *Br J Haematol*. 2013;161(5):701-705.
107. Bödör C, Renneville A, Smith M, et al. Germ-line GATA2 p.THR354MET mutation in familial myelodysplastic syndrome with acquired monosomy 7 and ASXL1 mutation demonstrating rapid onset and poor survival. *Haematologica*. 2012;97(6):890-894.
108. West RR, Hsu AP, Holland SM, Cuellar-Rodriguez J, Hickstein DD. Acquired ASXL1 mutations are common in patients with inherited GATA2 mutations and correlate with myeloid transformation. *Haematologica*. 2014;99(2):276-281.

109. Churpek JE, Pyrtel K, Kanchi KL, et al. Genomic analysis of germ line and somatic variants in familial myelodysplasia/acute myeloid leukemia. *Blood*. 2015;126(22):2484-2490.
110. Katsumura KR, Mehta C, Hewitt KJ, et al. Human leukemia mutations corrupt but do not abrogate GATA-2 function. *Proc Natl Acad Sci USA*. 2018;115(43):E10109-E10118.
111. Zhang SJ, Ma LY, Huang QH, et al. Gain-of-function mutation of GATA-2 in acute myeloid transformation of chronic myeloid leukemia. *Proc Natl Acad Sci USA*. 2008;105(6):2076-2081.
112. DeKoter RP, Singh H. Regulation of B lymphocyte and macrophage development by graded expression of PU.1. *Science*. 2000;288(5470):1439-1441.
113. Walsh JC, DeKoter RP, Lee H-J, et al. Cooperative and antagonistic interplay between PU.1 and GATA-2 in the specification of myeloid cell fates. *Immunity*. 2002;17(5):665-676.
114. Rosenbauer F, Owens BM, Yu L, et al. Lymphoid cell growth and transformation are suppressed by a key regulatory element of the gene encoding PU.1. *Nat Genet*. 2006;38(1):27-37.
115. Steidl U, Rosenbauer F, Verhaak RG, et al. Essential role of Jun family transcription factors in PU.1 knockdown-induced leukemic stem cells. *Nat Genet*. 2006;38(11):1269-1277.
116. Steidl U, Steidl C, Ebralidze A, et al. A distal single nucleotide polymorphism alters long-range regulation of the PU.1 gene in acute myeloid leukemia. *J Clin Invest*. 2007;117(9):2611-2620.
117. Will B, Vogler TO, Narayanagari S, et al. Minimal PU.1 reduction induces a preleukemic state and promotes development of acute myeloid leukemia. *Nat Med*. 2015;21(10):1172-1181.
118. Mansour MR, Abraham BJ, Anders L, et al. Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. *Science*. 2014;346(6215):1373-1377.
119. Li Z, Abraham BJ, Berezovskaya A, et al. APOBEC signature mutation generates an oncogenic enhancer that drives LMO1 expression in T-ALL. *Leukemia*. 2017;31(10):2057-2064.
120. Abraham BJ, Hniz D, Weintraub AS, et al. Small genomic insertions form enhancers that misregulate oncogenes [published correction appears in *Nat Commun*. 2017;8:15797]. *Nat Commun*. 2017;8:14385.
121. Gostissa M, Yan CT, Bianco JM, Cogné M, Pinaud E, Alt FW. Long-range oncogenic activation of Igh-c-myc translocations by the Igh 3' regulatory region. *Nature*. 2009;462(7274):803-807.
122. Dalla-Favera R, Bregni M, Erikson J, Patterson D, Gallo RC, Croce CM. Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. *Proc Natl Acad Sci USA*. 1982;79(24):7824-7827.
123. Erikson J, ar-Rushdi A, Drwinga HL, Nowell PC, Croce CM. Transcriptional activation of the translocated c-myc oncogene in Burkitt lymphoma. *Proc Natl Acad Sci USA*. 1983;80(3):820-824.
124. Siebenlist U, Hennighausen L, Battey J, Leder P. Chromatin structure and protein binding in the putative regulatory region of the c-myc gene in Burkitt lymphoma. *Cell*. 1984;37(2):381-391.
125. Chesi M, Nardini E, Lim RS, Smith KD, Kuehl WM, Bergsagel PL. The t(4;14) translocation in myeloma dysregulates both FGFR3 and a novel gene, MMSET, resulting in IgH/MMSET hybrid transcripts. *Blood*. 1998;92(9):3025-3034.
126. Chesi M, Nardini E, Brents LA, et al. Frequent translocation t(4;14)(p16.3;q32.3) in multiple myeloma is associated with increased expression and activating mutations of fibroblast growth factor receptor 3. *Nat Genet*. 1997;16(3):260-264.
127. Kuo AJ, Cheung P, Chen K, et al. NSD2 links dimethylation of histone H3 at lysine 36 to oncogenic programming. *Mol Cell*. 2011;44(4):609-620.
128. Keats JJ, Reiman T, Maxwell CA, et al. In multiple myeloma, t(4;14)(p16;q32) is an adverse prognostic factor irrespective of FGFR3 expression. *Blood*. 2003;101(4):1520-1529.
129. Santra M, Zhan F, Tian E, Barlogie B, Shaughnessy J Jr. A subset of multiple myeloma harboring the t(4;14)(p16;q32) translocation lacks FGFR3 expression but maintains an IGH/MMSET fusion transcript. *Blood*. 2003;101(6):2374-2376.
130. Chung JH, Whiteley M, Felsenfeld G. A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. *Cell*. 1993;74(3):505-514.
131. Yamazaki H, Suzuki M, Otsuki A, et al. A remote GATA2 hematopoietic enhancer drives leukemogenesis in inv(3)(q21;q26) by activating EVI1 expression. *Cancer Cell*. 2014;25(4):415-427.
132. Gröschel S, Sanders MA, Hoogenboezem R, et al. A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia. *Cell*. 2014;157(2):369-381.
133. Katayama S, Suzuki M, Yamaoka A, et al. GATA2 haploinsufficiency accelerates EVI1-driven leukemogenesis. *Blood*. 2017;130(7):908-919.
134. Herranz D, Ambesi-Impombato A, Palomero T, et al. A NOTCH1-driven MYC enhancer promotes T cell development, transformation and acute lymphoblastic leukemia. *Nat Med*. 2014;20(10):1130-1137.
135. Katsumura KR, Yang C, Boyer ME, Li L, Bresnick EH. Molecular basis of crosstalk between oncogenic Ras and the master regulator of hematopoiesis GATA-2. *EMBO Rep*. 2014;15(9):938-947.
136. DeVilbiss AW, Boyer ME, Bresnick EH. Establishing a hematopoietic genetic network through locus-specific integration of chromatin regulators. *Proc Natl Acad Sci USA*. 2013;110(36):E3398-E3407.
137. DeVilbiss AW, Tanimura N, McIver SC, Katsumura KR, Johnson KD, Bresnick EH. Navigating transcriptional coregulator ensembles to establish genetic networks: a GATA factor perspective. *Curr Top Dev Biol*. 2016;118:205-244.