

## Characterizing the molecular impact of *KMT2D* variants on the epigenetic and transcriptional landscapes in Kabuki Syndrome

Youngsook L Jung<sup>1,2</sup>, Christina Hung<sup>1,2</sup>, Jaejoon Choi<sup>1,2</sup>, Eunjung A Lee<sup>1,2</sup> and Olaf Bodamer<sup>1,2,3\*</sup>

<sup>1</sup> Division of Genetics and Genomics, Boston Children's Hospital, Boston, Massachusetts, 02115, USA

<sup>2</sup> Department of Pediatrics, Harvard Medical School, Boston, Massachusetts, 02115, USA

<sup>3</sup> Broad Institute of Harvard University and MIT, Cambridge, Massachusetts, 02142, USA

\*Correspondence: [olaf.bodamer@childrens.harvard.edu](mailto:olaf.bodamer@childrens.harvard.edu) (O.B.)

### **SUMMARY (250 words) 233 words**

Kabuki Syndrome (KS) is a rare, multisystem disorder with a variable clinical phenotype. The majority of KS is caused by dominant loss-of-function mutations in *KMT2D* (lysine methyltransferase 2D). *KMT2D* mediates chromatin accessibility by adding methyl groups to lysine residue 4 of histone 3, which plays a critical role in cell differentiation and homeostasis. The molecular underpinnings of KS remain elusive partly due to a lack of histone modification data from human samples. Consequently, we profiled and characterized alterations in histone modification and gene transcription in peripheral blood mononuclear cells (PBMCs) from 33 patients with *KMT2D* mutations and 36 unaffected healthy controls. Our analysis identified unique enhancer signatures in H3K4me1 and H3K4me2 in KS compared to controls. Reduced enhancer signals were present for promoter-distal sites of immune-related genes for which co-binding of PBMC-specific transcription factors was predicted; thirty-one percent of super-enhancers of normal blood cells overlapped with disrupted enhancers in KS, supporting an association of reduced enhancer activity of immune-related genes with immune deficiency phenotypes. In contrast, increased enhancer signals were observed for

promoter-proximal regions of metabolic genes enriched with *EGR1* and *E2F2* motifs, whose transcriptional levels were significantly increased in KS. Additionally, we identified approximately 100 de novo enhancers in genes, such as in *MYO1F* and *AGAP2*. Together, our results underscore the effect of KMT2D haploinsufficiency on (dys)regulation of enhancer states and gene transcription and provide a framework for the identification of therapeutic targets and biomarkers in preparation for clinical trial readiness.

## KEYWORDS

Kabuki Syndrome, KMT2D, MLL2, MLL4, Histone modification, H3K4me1, H3K4me2, Enhancer, ChIP-seq, RNA-seq

## INTRODUCTION

Kabuki Syndrome (KS) is a rare, heterogeneous, congenital malformation disorder that follows a broad and variable clinical spectrum characterized by a unique facial gestalt, mild to moderate intellectual disability, developmental delays, hearing loss, muscle hypotonia, and structural and/or functional anomalies of cardiac, endocrine, renal and orthopedic systems.(Niikawa et al. 1981; Kuroki et al. 1981; Adam et al. 2019) Abnormal regulation of terminal B-cell differentiation in KS leads to variable immune deficiency with low immunoglobulin (IgG and IgA) levels that may be associated with inadequate antibody response to vaccines and increased susceptibility to recurrent and potentially life-threatening infections.(Y. Li et al. 2011; Micale et al. 2011; Banka et al. 2012; Ming et al. 2005; Lindsley et al. 2016; Stagi et al. 2016; Adam, Hudgins, and Hannibal 2011) KS occurs in approximately 1:32,000 individuals and currently has no approved therapies, despite a significant disease burden.(Adam et al. 2019; Theodore-Oklota et al. 2022) Although it is known that KS is due to dominant loss-of-function mutations in autosomal *KMT2D* (lysine methyltransferase 2D, 80-90%, KS1, MIM: 147920) or X-linked *KDM6A* (lysine demethylase 6A, 5-8%, KS2, MIM: 300867) respectively,(Ng et al. 2010; Lederer et al. 2012) its underlying disease mechanisms are not fully elucidated. Most importantly, there is a limited understanding of the

downstream pathways that are affected in KS and how they relate to the clinical phenotype spectrum.

Transcriptional regulation is a highly specialized and temporally orchestrated, tissue-specific process during early embryogenesis and development, which determines cell fate, cell cycle progression, stem cell function, and ultimately normal embryogenesis.(Fasciani et al. 2020) The complexity and importance of perfectly concerted transcriptional control necessitate a number of regulatory mechanisms, including chromatin and histone modifications. The addition or removal of histone marks, such as acetyl- and methyl-groups, changes the conformation of chromatin and renders genes accessible or inaccessible to transcription factor (TF) complexes.(Fasciani et al. 2020)

For gene activation, pioneer TFs, along with other cell-type-specific TFs, recruit KMT2D, a lysine histone methyl-transferase to prime enhancer regions through the addition of one methyl group (monomethylation) to the fourth lysine of histone 3 (H3K4), thereby promoting open chromatin.(Ng et al. 2010; Froimchuk, Jang, and Ge 2017) KMT2D associates with WRAD (WDR5, RbBP5, ASH2L, DPY30), KDM6A, NCOA6, PA1, and PTIP into one protein complex (ASCOM complex) which is critical for its H3K4 methyltransferase activity and stabilization of KDM6A (also known as UTX).(Froimchuk et al. 2017) Additionally, KMT2D facilitates the binding of the two H3K27 acetyltransferases CREB-binding protein (CBP) and p300 on enhancers, thereby promoting tissue-specific gene expression.

Loss of KMT2D function is therefore expected to alter chromatin states, in particular enhancer states, and to subsequently cause gene dysregulation. In addition to the association with KS, mutations in *KMT2D* have been described in association with various cancers. Considered a tumor suppressor gene, *KMT2D* is one of the most frequently mutated genes in cancer for reasons yet unknown.(J. Zhang et al. 2015; Ortega-Molina et al. 2015) In B-cell lymphoma and leukemia, reduced activity of KMT2D leads to tumor proliferation and disruption of tumor suppressor pathways, contrasting

other cancers, including breast cancer, in which reduced expression of *KMT2D* causes decreased proliferation, suggesting diverging tissue-type-dependent effects.

There have been recent advances in understanding the specific role of *KMT2D* in the KS disease process.(Fahrner et al. 2019; Carosso et al. 2019; Huisman et al. 2021; Bjornsson et al. 2014) Previous studies largely focusing on human blood DNA methylation profiles showed that the DNA methylation profiles differ between KS and unaffected controls.(Aref-Eshghi et al. 2017, 2018; Sobreira et al. 2017) While differentially methylated regions may be useful as future biomarkers, they are of limited use for elucidating regulatory mechanisms in KS pathology as *KMT2D* mutations primarily affect histone modifications and only subsequently exert their effect on DNA methylation.

In this study, we concomitantly profiled genome-wide histone modification along with gene expression in peripheral blood from 33 KS individuals with deleterious *KMT2D* mutations to study the impact on chromatin states and gene regulation. By integrating epigenetic and transcriptomic data, we identified changes in histone modification, gene regulation and regulatory pathways in KS. Importantly, we found unique chromatin changes that may serve as future drug targets or biomarkers, preparing for KS for clinical trial readiness.

## **MATERIAL AND METHODS**

### **Subjects**

Individuals of any age with a molecularly confirmed and clinical diagnosis of KS, due to likely pathogenic or pathogenic variants in *KMT2D*, and age/gender-matched healthy controls were eligible to participate in this study following written informed consent.

Individuals with *KMT2D*-related disorder due to missense variants in exons 38 and 39 of *KMT2D*,(Cuvertino et al. 2020) as well as individuals with KS affected by an additional genetic disorder, were excluded from participation in this study. All subjects were deidentified and family IDs are only known to the principle investigator (OB) and the

immediate research staff. The study was approved by the Institutional Review Board of Boston Children's Hospital (IRB-A00026691).

The molecular diagnosis of KS was confirmed by genetic testing through CLIA-certified clinical laboratories. Parental testing was conducted and showed de novo occurrence of KMT2D variants in the majority of cases. The interpretation of variant classification reflects the classification reported by the diagnostic laboratory and follows the guidelines of The American College of Medical Genetics and Genomics and the Association for Molecular Pathology.(Richards et al. 2015)\_The variant types were annotated and the genomic positions were extracted using the Mutalyzer tool.(Lefter et al. 2021)

### **Sample preparation, library construction and sequencing for ChIP-seq data**

PBMCs were isolated from NaHep samples using the Lymphoprep<sup>TM</sup>/SepMate<sup>TM</sup> system (StemCell Technologies) according to the manufacturer's specifications and cryopreserved in 10% DMSO v/v FBS. For cell fixation, cells were thawed, counted and 6 million cells per sample were fixed in a 1% formaldehyde solution for 15 min at room temperature. The fixation was stopped using 0.125M Glycine for 5 min and cell pellets were washed in 0.5% Igepal in PBS and 1mM PMSF in 0.5% Igepal-PBS. Snap frozen cell pellets were then shipped on dry ice to ActiveMotif for chromatin extraction and Histone Path<sup>TM</sup> ChIP-Seq experiments. Sequences were generated on the illumina NextSeq 500 platform. Takeda Pharmaceutical, Cambridge, MA, provided financial support for the ChIP-Seq experiments. Raw data was returned for further bioinformatics analyses as described in the following sections.

### **Sample preparation, RNA-extraction, library construction and sequencing for RNA-seq data**

Total RNA was extracted from whole blood following the manufacturer's protocol for PAXgene Blood RNA Kit (Qiagen). Following the assessment of RNA concentration and

quality (RIN>7.2) by Agilent 2100 or Fragment Analyzer, the illumina TruSeq Stranded Total RNA with Ribo-Zero Globin library preparation was used, which includes ribosomal RNA and globin mRNA removal. A minimum of 10Gb of 100bp paired-end reads were generated per sample. All library preparations and sequencing was conducted by BGI Americas on the illumina HiSeq 4000 platform. Raw data was returned for further bioinformatics analysis as described in the following sections.

### **Data processing and analysis of ChIP-seq data**

For ChIP-seq data, the reads were mapped to the GRCh38 human reference sequence using BWA (Burrows-Wheeler Aligner) with default parameters.(H. Li and Durbin 2009) Only uniquely mappable reads were used for the downstream analysis. Significantly enriched regions were detected and normalized IP (immunoprecipitation) enrichment profiles were generated using MACS2 (Model-bases analysis of ChIP-seq 2) call peak function with default parameters.(Feng et al. 2012) To assess the data quality, the FRiP (fraction of reads in peaks) scores were calculated and correlations were measured between samples. The samples with low FRiP scores or correlation efficiencies were excluded from the downstream analysis.

For the principal component analysis (PCA), peaks from each individual (both KS and controls) were used to identify the union peak regions. Among the union peak regions, the peaks showing consistent signal intensities across both cohorts (KS vs control) were excluded after calculating a coefficient of variation (standard deviation over mean). The PCA analysis was performed using the read counts in the remaining differing peak sets.

For the detection of differential peaks between groups (KS vs control), Diffbind R package was used for the union peak sets.(Ross-Innes et al. 2012; Gupta et al. 2007) For visualization of the differential peaks in a heatmap, values in each row were normalized. The differential peaks were divided into two groups--a group with increased and a group with decreased signals in KS compared to control. For both groups, the enriched gene ontology terms were identified using the Great tool with default

parameters.(McLean et al. 2010) For motif analysis, the TSS (transcription start site)-proximal and TSS-distal regions were further defined. TSS information was obtained from the UCSC genome browser table for GRCh38. Peaks within a 2 kb distance from TSS were defined as TSS-proximal whereas the peaks > 2 kb from TSS were deemed TSS-distal. DNA sequences were extracted within a 100 bp window from peak summits and known motifs were matched using MEME TOMTOM.(Gupta et al. 2007) We obtained the motif sequences for the search from JASPAR (Khan et al. 2018) and Jolma et al in 2013.(Jolma et al. 2013) The genomic locations of TF binding were predicted using MEME FEMO(Bailey et al. 2009) and used for the pathway enrichment analysis for a subset of TFs. To associate increased H3K4me2 sites in KS with respect to EGR1 binding, we used the publicly available EGR1 ChIP-seq data set from GM12878. To associate changes in histone modifications observed in our KS cohort with publicly available DNA methylation changes in KS, DNA methylation data from individuals with KS and control were downloaded from Sobreira et al (Sobreira et al. 2017) (GSE116300) and reanalyzed after performing lift over from hg18 to hg38.

We identified super-enhancers based on the intensity and breadth of H3K27ac signal in GM12878 following the methods described in Whyte et al (Whyte et al. 2013) and determined the intersection of the top 2000 differential peaks of H3K4me1 and H3K4me2 in KS compared with control and the super-enhancer regions.

### **Analysis of RNA-seq data**

RNA sequence reads were mapped using STAR aligner with default parameters.(Dobin et al. 2013) The transcriptional reads for each annotated gene were estimated using Htseq.(Anders, Pyl, and Huber 2015) Differentially expressed genes between KS and controls were determined using DESeq2 after adjusting for sex, age and batch effect.(Love, Huber, and Anders 2014) To associate genes with peaks of H3K4me1/2/3, the closest gene was assigned for each peak. If the gene has multiple H3K4me1/2/3 peaks, the peak with the lowest p-value was assigned to the gene. The TF list was obtained from AnimalTFDB.(Hu et al. 2019)



To infer transcriptional regulators such as TFs or chromatin regulators underlying differential expressions in KS, we used the method by Qin et al (Qin et al. 2020) which integrates TF binding sites from public ChIP-seq data in different tissue and cell types with differentially expressed genes. Briefly, using ~10,000 available public ChIP-seq data sets, the regulatory potential score of each TF is calculated for differentially expressed genes with the assumption that the effect of the TF binding on gene expression exponentially decays with distance and that the contribution of multiple binding sites is additive. We used differentially expressed genes in KS compared to control with q-value (or FDR, false discovery rate) = 0.1.

## RESULTS

### Summary of cohort, variants and datasets

To generate RNA-seq and ChIP-seq data, we collected peripheral blood samples from 33 KS individuals with mutations in *KMT2D* (females: 23, mean age: 9.4 years, range: 0.9-33.3 years) and from 36 healthy individuals as age or sex-matched controls (females: 22, mean age: 30.8 years, range: 4.9-64.8 years) from either healthy siblings (age-matched), parents (sex-matched) of KS individuals or unrelated age-matched healthy children (see details in **Table S1**).

The variants in *KMT2D* of our cohort are summarized in **Table 1** and **Figure 1A**.

The majority of variants were truncating (60.6%) and located before the SET domain, suggesting a loss of function of *KMT2D*. *KMT2D* has 54 exons. Although the variants were found across the *KMT2D* gene on chromosome 12, the distribution in the genomic positions indicates potential mutation hotspots depending on variant types. Frameshift variants were more prevalent in exon 10, while splice site variants clustered in introns 37-38, nonsense mutations in exon 39 and missense mutations in exon 48 (**Figure 1B**). In particular, exons 39 and 48 were most frequently mutated (36%, 12/33), consistent with the prior studies. (Hannibal et al. 2011; Banka et al. 2012). A total of 8 variants including 6 nonsense mutations (more than half, 6/8) were identified in exon 39, which



encodes long polyglutamine tracts, and 4 variants were found in exon 48. Interestingly, most missense mutations (5/6) were located in F/Y-rich domains (FYRN and FYRC), transcription activation domains involved in TF bindings.(Park et al. 2010)

To characterize the alterations of chromatin and transcriptional states in KS and better understand the underlying regulatory mechanisms in KS, we generated histone modification profiles using ChIP-seq for H3K4me1, H3K4me2 and H3K4me3 which are known to be modified by KMT2D, along with RNA-seq data to study gene transcription. The overview of this study design is shown in **Figure 1C**.

### **Distinctive enhancer signature in KS**

We hypothesized that mutations in *KMT2D* will alter chromatin states and thereby affect the chromatin landscape in KS. H3K4me1 is found in primed enhancer regions that mark regulatory regions for enhancer activation in the later developmental stages or stimuli. These H3K4me1 marks will subsequently induce the enrichment for H3K27ac at the same enhancer regions. The H3K4me1 marks are typically located in promoter-distal sites, whereas H3K4me2, which is also associated with enhancer regions, enriched at both promoter-proximal regions as well as distal regions. In contrast, H3K4me3 is a marker for active promoters.

The principal component analysis for variable genomic regions among individuals revealed distinctive enhancer mark profiles of H3K4me1 (**Figure S1, Methods**) and H3K4me2 (**Figure 2A**), separating KS from control samples with greater distinction for H3K4me2. In contrast, H3K4me3 data could not be grouped according to the presence of KMT2D mutations. Rather, H3K4me3 data showed a tendency to cluster depending on genetic background (**Figure S1**). Similarly, transcriptional data (**Figure S1**) did not separate KS samples from controls. This suggests that *KMT2D* mutations affect enhancer states to a greater degree than promoter states, and that the enhancer signature of KS is different from healthy controls.

### **Regulatory program underlying altered enhancer signals in KS**

We further characterized the alterations of enhancer marks H3K4me1 and H3K4me2 in KS compared to control. H3K4me2, an enhancer mark, was more distinctive between KS and control in PCA analysis. Hierarchical clustering results identified 2149 significantly changed H3K4me2 sites (814 decreased vs 1335 increased sites,  $q = 0.001$ , **Figure 2B, Methods**) confirming grouping according to genotypes, which is in alignment with the PCA result. Because loss of function of *KMT2D* is expected to cause a decrease of H3K4me2, a higher number of increased sites was unexpected but was consistent with a recent report in mouse B-cells.(Luperchio et al. 2021)

Furthermore, detailed examinations of H3K4me2 alterations revealed the association with two different target groups. Decreased sites were located at promoter-distal regions (**Figure 2B**) with enrichment of immune-related GO terms (**Figure 2C**) and were predicted to be bound by blood-cell-specific TFs through motif analysis. These blood-specific TFs, *ERG* (ETS Transcription Factor), *ETS1* (ETS Proto-Oncogene 1), *Fli1* (Friend Leukemia Integration 1) and *RUNX1* (Runt-Related Transcription Factor 1), play a crucial role in blood differentiation and development (**Figure 2B, Table S2**). A representative of significantly decreased H3K4me2 signals is shown in *TLR* (Toll-like receptor 9, top 4) where the decreased peak is located at the downstream intergenic region of the gene (**Figure 2D**). TLR plays an important role in the immune system and is implicated in the pathogenesis of autoimmune diseases.(Baccala et al. 2007) Similarly, we observed decreased H3K4me1 signals in promoter-distal sites of immune-related genes (**Figure S2**). In general, for the significantly changed H3K4me1 regions ( $q = 0.05$ ), the changes of H3K4me1 signals in KS compared to control were correlated well with the changes of H3K4me2 signals ( $r = 0.92$ , changes of 99% peaks regions in the same direction, **Figure S2**).

While decreased H3K4me2 signals could be explained by the loss of function of *KMT2D*, the observed increases in H3K4me2 sites were unexpected. Therefore, we performed more in-depth analyses to explore underlying mechanisms for the increase in H3K4 dimethylation. Contrasting to the GO term associations of decreased H3K4me2 marks, the increased H3K4me2 peaks were enriched for metabolism-related GO terms

and were found near promoter regions with the enrichment of the motifs for TFs associated with proliferation and inflammation such as the *EGR1* and *E2F* family (**Figure 2B**, **Table S2**). The expression changes of the TFs enriched in the increased H3K4me2 sites revealed that the expression levels of *EGR1* and *E2F2* increased by more than two-fold in KS patients compared to controls based on RNA-seq data (**Figure 3A**). In addition, differential expression of genes involved in cell cycles including *E2F3*, *E2F4* and *CCND1* were found to be increased, although the changes did not reach statistical significance.

We next examined the relationship between *EGR1* binding and H3K4me2 changes in KS using publicly available *EGR1* ChIP-seq data profiled in blood cells (**Methods**). The proportion of increased H3K4me2 sites at the promoter-proximal regions was drastically different between *EGR1*-bound sites and *EGR1*-unbound sites (**Figure 3B**). Among the *EGR1*-bound H3K4me2 sites, about 90% were increased in H3K4me2 signals, whereas among the H3K4me2 sites not targeted by *EGR1*, only half (about random expectation) showed increased H3K4me2 in KS, which is statistically significant (Odds ratio > 5 and  $p$ -value <  $10^{-15}$  by Fisher's exact test). Furthermore, the *EGR1* target genes with the increased H3K4me2 showed a significant increase in expression (**Figure S2**) and were enriched in the pathways associated with hypoxia, cell cycles, Notch signaling and transcriptional regulation by the RUNX family (**Figure S2**). Altogether, these results suggest that KMT2D haploinsufficiency results in 1) reduced enhancer activity of immune-related genes and 2) overexpression of *EGR1* and *E2F2* associated with more open chromatin states, leading to the upregulation of target genes.

### **De novo enhancer sites in KS individuals**

Among the significantly increased H3K4me2 sites, we identified 135 de novo H3K4me2 peaks in KS (**Table S3**). One of the examples of de novo H3K4me2 peaks in KS is located in the intragenic region of *MYO1F* (Myosin 1F) (**Figure 2D**, top increased H3K4me2 site in KS compared to control). Several prior studies reported *MYO1F* as one of the most differentially hypomethylated regions in KS samples compared to

controls. (Aref-Eshghi et al. 2017; Butcher et al. 2017; Aref-Eshghi et al. 2018; Sobreira et al. 2017) Comparisons with the previously published DNA methylation profile confirmed that the de novo H3K4me2 peak precisely overlapped with the hypomethylation region in KS (**Figure 3C**). Another example of de novo H3K4me2 sites is the gene *AGAP2* (ArfGAP With GTPase Domain 2, **Figure S3**). Similarly, several previous studies comparing blood DNA methylation profiles between KS and controls reported *AGAP2* as one of the hypomethylated regions in KS. The de novo H3K4me2 peak in *AGAP2* overlapped with the hypomethylated region in KS. In general, de novo H3K4me2 peaks in KS were less methylated in KS compared to controls based on available DNA methylation data (**Figure S3**). These histone and DNA methylation patterns can potentially serve as biomarkers for KS for clinical trial readiness.

### **Association with super-enhancers**

KMT2D deposits H3K4me1 and H3K4me2 marks at genomic targets, followed by p300 interaction mediated through the ASCOM complex, along with co-binding of TFs. This leads to H3K27ac and a fully active enhancer state. (Xu et al. 2018) The loss of function of KMT2D is expected to alter chromatin states in active enhancers as well as super-enhancers, which is supported by a previous study. (Alam et al. 2020)

Super-enhancers are a set of enhancers typically identified by the strong and broad signals of H3K27ac and are known to play a critical role in development, cell identity and disease. (Hnisz et al. 2013; Lovén et al. 2013; Whyte et al. 2013) In our study, we found that approximately 31% of super-enhancers in lymphoblast cells (**Figure 4A**), which are defined by H3K27ac (**Methods**), overlapped with differential H3K4me1 and/or H3K4me2 sites in KS individuals compared to normal controls where the H3K4me1/me2 signal intensities were significantly reduced for the majority of them (**Figure 4B**). Genes whose super-enhancers overlap with differential H3K4me1/2 in KS include several important genes previously associated with mechanistically overlapping clinical syndromes or presentations. (**Figure 4C, Table S4**). *RAP1A* is associated with KS-like phenotypes and plays an important role in the RAS/MAPK signaling pathway. (Bögershausen et al. 2015; Tsai et al. 2018) *CHD7* is a major cause of CHARGE syndrome characterized by an overlapping clinical phenotype with

KS.(Butcher et al. 2017) *PAX5* is a master regulator in B cells and is known to be a co-binding factor of *KMT2D* and implicated in the B cell development defect. (Lindsley et al. 2016) The results indicate that *KMT2D* mutations might be associated with disruption of a substantial fraction of the super-enhancers, implicated in KS pathogenesis.

### **Relationship between transcriptional changes and enhancer signal changes**

To explore the expression changes downstream of changes in enhancer signals in KS, we examined the correlation between H3K4me2 and transcription changes in KS individuals compared to controls (**Methods**). For the differentially expressed genes ( $q < 0.1$ ), the fold changes of H3K4me2 were modest but positively correlated with the fold changes of expression levels (Pearson's  $r = 0.4$ ), with approximately 70% of genes displaying changes of expression and H3K4me2 levels in the same direction as illustrated in **Figure 5A**. The downregulated genes with decreased H3K4me2 signals in KS include genes associated with immunity, linking this H3K4me2 signature to the KS phenotype. For instance, *TNNT1* was reported as one of the 20 most significantly dysregulated genes in *Mll2* knockout mice and plays a role in musculoskeletal development.(Issaeva et al. 2007) *IR23R* and *MXRA8* are implicated in autoimmune disease and immune response, respectively.(Peloquin et al. 2016; R. Zhang et al. 2019) Several genes including *TUSC3* are frequently dysregulated in cancer.(Gu et al. 2016) This result indicates that changes in enhancer marks could explain expression changes for a majority of the target genes. However, additional factors besides H3K4me2 likely also contribute to the transcriptional changes in KS, given that 30% of genes do not show expression changes in the same direction as H3K4me2.

### **Transcriptional regulators underlying expression changes in KS**

In the previous sections, we report that (1) TFs EGR1 and E2F2 experienced substantial transcriptional changes in KS and (2) for a subset of genes, their expression changes cannot be explained by H3K4me2 changes alone. These observations prompted us to search for other TFs, which could globally affect the expression changes of the TF target genes. A total of 65 TFs exhibited noticeable expression changes in KS patients ( $q$ -value  $< 0.1$  or  $\log_2$  fold change  $> 1$ , **Figure 5B, Methods**) including KLF1,

GATA1 and TAL1, master regulators in erythroid cells. A recent study showed the association of KLF1 in autoimmune disease.(Teruya et al. 2018) SOX5 is associated with differentiation in B cells and T helper 17 cells. Its overexpression is implicated in autoimmune disease as well.(Tanaka et al. 2014) YBX TF families are involved in T cell differentiation; NFIX and SMAD1 are associated with hematopoietic stem cell regulation.(Martynoga et al. 2013) These TF expression changes were significant compared to the expectation from the background gene expression changes ( $p < 10^{-5}$  by rank sum test, **Figure S3, Methods**).

To infer transcriptional regulators, such as TFs or chromatin regulators, underlying differential expressions in KS, we computed regulatory potential scores by integrating transcription regulator binding sites from public Chip-seq data and differentially expressed genes in KS. In brief, using about 10,000 publicly available ChIP-seq data, the regulatory potential by each transcriptional regulator was calculated for differentially expressed genes (DEGs, **Methods**). The top 50 TFs underlying DEGs in KS include TAL1, KIF1 and GATA1, master regulators in erythroid cells and SMAD1 for hematopoietic cells (**Figure 5C**). We found KMT2A (lysine-specific methyltransferase 2A; another methyltransferase for H3K4me), CEBPA (CCAAT enhancer binding protein alpha) and KDM1A (lysine-specific histone demethylase 1A) which are associated with enhancer function. However, KMT2D itself was not identified, likely due to sparsity in available ChIP-seq datasets. EGR1 and RUNX1, which are TFs predicted from the differential H3K4me2 sites in **Figure 2B**, were also included in the list. Therefore, the disrupted expression states of these TFs likely regulate downstream differential expression in KS. Although the gene expression changes may not be a direct consequence of *KMT2D* gene mutations, these results support that expression changes are a consequence of expression changes of TFs in KS.

## DISCUSSION



Although *KMT2D* is the major causal gene for KS and its primary role in chromatin states is known, the effect of haploinsufficiency of *KMT2D* on histone modification and subsequent transcriptional (dys)regulation has not been studied in samples from patients with KS. To examine the alterations in the chromatin landscape in KS, we profiled H3K4 methylation states from blood samples of 33 KS patients with pathogenic variants in *KMT2D* along with transcription data and systematically identified altered chromatin and expression states in KS.

Our analysis of H3K4me1/2/3 profiles from *KMT2D* haploinsufficient patients revealed an enhancer landscape (H3K4me1/2) in KS distinctive from unaffected controls. Interestingly, the H3K4me2 profile was more impacted compared to H3K4me1, with better separation between KS and control groups. Although both H3K4me1 and H3K4me2 are associated with enhancers, they target substantially different genomic locations as reported in previous studies.(Pekowska et al. 2010) We also observed a small number of KS-specific alterations in H3K4me3 peaks at promoter sites. However, in our dataset, the majority of H3K4me3 changes are explained by genetic background rather than *KMT2D* genotype. It is notable that the reported roles of *KMT2D* in literature are somewhat inconsistent, as some studies describe *KMT2D* as a methyltransferase specifically acting on a single mark (e.g., H3K4me3 or H3K4me1), while others describe *KMT2D* to act on enhancer marks (H3K4me1/2) or all forms of H3K4me.(Fahrner et al. 2019; Carosso et al. 2019; Huisman et al. 2021; Zhang et al. 2015; Ortega-Molina et al. 2015) Despite the variations and discrepancies in the literature, most studies report greater changes in H3K4me1/2 than H3K4me3. In line with the reported literature, our results show that *KMT2D* haploinsufficiency has a greater impact on H3K4me1/2, thus enhancer states.

More specifically, we found distal enhancer signals of immune-related genes or immune pathways were frequently reduced in KS, suggesting dysregulation of the associated immune-related genes/pathways, thus linking the distal enhancer states with the immunodeficiency phenotype frequently observed in individuals with KS. In addition, we observed promoter-proximal enhancer signals to be increased in genes associated with



proliferation, inflammation and metabolism pathways, which is in line with previous studies in lymphoma, which showed that loss of KMT2D promotes proliferation and tumorigenesis. In contrast, results from the mouse *Kmt2d* model indicated that *Kmt2d* haploinsufficiency in neuronal cells might also affect metabolic pathways but in the opposite direction, leading to reduced proliferation. Taking together our and previous findings, a conclusion is emerging that the effect of KMT2D haploinsufficiency might be tissue- and cell-type specific.

Based on the active enhancer mark of H3K27ac profiled in GM12578, we showed that approximately one third of super-enhancer regions overlapped with altered enhancer regions in KS when compared to control, suggesting a potential impact on the regulation of super-enhancers in KS. Prior studies on the role of KMT2D in super-enhancers during tumorigenesis showed that KMT2D plays a role in maintaining super-enhancers of tumor suppressors and that the loss of KMT2D resulted in dysregulation of glycolytic genes in tumors.[reference] Our observations indicate possible disruptions of super-enhancer regulation in KS to also occur in the non-cancer context, where they are not limited to tumor suppressors or glycolytic genes. Many of the genes associated with disturbed super-enhancers in KS included genes with immune functions or B-cell differentiation. Intriguingly, super-enhancer regions of the genes implicated in KS-like phenotypes, such as *RAP1A* and *CHD7*, overlapped with altered enhancers in KS, implying a possible mechanical link with KS. Further studies are required to determine the extent of the impact on super-enhancers using H3K27ac profiled in KS individuals by haploinsufficiency of KMT2D and the functional studies on the super-enhancers of these genes.

Despite the variations in the mutational landscape of KMT2D, the resulting phenotypes in the enhancer landscape, in particular in H3K4me2, were more consistent among KS. In the current study of 33 individuals, we did not observe a clear correlation between variant types and phenotypes in chromatin or transcriptional landscape in KS. Future studies using a larger number of samples and profiles might allow us to detect the association between genomic variations in KMT2D and chromatin phenotypes.

Our RNA-seq comparison between KS and control revealed that a significant number of TFs, which are expressed in our cohort and experienced transcriptional changes, are involved in differentiation and key regulation of blood cells. A recent study in B-cells of the mouse *Kmt2d* model also indicated the association between *Kmt2d* haploinsufficiency and overall expression changes in TFs.(Luperchio et al. 2021) Although the precise estimation of expression changes of TFs in KS was limited by the heterogeneity of cell populations in our blood samples, our results support the association of transcriptional changes in TFs with dysregulation of blood cell differentiation, including known B-cell differentiation defects reported in KS phenotype(Pilarowski et al. 2020; Lindsley et al. 2016), as well as a broader range of dysregulation of the TF target genes in general.

It is pertinent to note that the data in this study represents the average signals across different cell types, making it challenging to gain insights into cell-type-specific regulation mechanisms. Profiling chromatin accessibility and transcription at a single cell level using emerging sequencing techniques in the future will allow us to better understand the regulation mechanism underlying KS in a cell-type-specific manner and to uncover novel cell types, genes and pathways that contribute to the KS etiology as well as biomarkers.

In summary, we present the first genome-wide chromatin maps of histone modifications from a large number of KS individuals with *KMT2D* mutations. Our analysis suggests unique enhancer signatures in KS with significant changes in gene regulation and regulatory pathways associated with immune, cell cycle, and proliferation as well as TF dysregulation in KS (**Figure 6**). The KS-specific chromatin changes may serve as future drug targets or biomarkers.

## **Acknowledgments**

We are particularly grateful to the patients with KS, their families and the healthy volunteers for their participation. In addition, we thank Takeda Pharmaceuticals for the

financial support for the ChIP-seq data generation. Lastly, we are grateful to Tara Daly, the Roya Kabuki Program Manager at Boston Children's Hospital for her assistance in coordinating the study visits.

## FIGURE LEGENDS

**Figure 1. Summary of *KMT2D* mutations from KS individuals and datasets generated from this study.** **A.** Proportions of *KMT2D* variant types from our KS cohort. Red, frame shift; blue, nonsense; green, missense; orange, splice site mutations. **B.** Variants and their mutation types across the *KMT2D* gene from our KS cohort. The genomic positions of each mutation site are shown along with the *KMT2D* gene where the transcription start direction is indicated with an arrow. The leftmost position indicates the transcription start site (TSS). (2) denotes mutations from two KS individuals. FYRN: F/Y-rich N-terminus domain, FYRC: F/Y-rich C-terminus domain. The same color codes are used as in A. **C.** Overview of this study design.

**Figure 2. Alterations of enhancer state in KS individuals compared to controls.** **A.** PCA analysis of variable H3K4me2 peaks for KS and control. Each dot represents an individual sample. Blue, KS; gray, control. **B.** Differential H3K4me2 peaks between KS and control. Left: row-normalized H3K4me2 signal heatmap of differential signals between KS and control. The hierarchical clustering results are shown on the top of the heatmap plot. Blue, high signal intensity; yellow, low signal intensity. Middle: distribution of distances of the differential peaks from TSSs. Right: top TF motifs predicted in differential H3K4me2 peaks between KS and control. Orange, decreased H3K4me2 peaks in KS compared to control; green, increased peaks in KS. **C.** Top gene ontology (GO) terms enriched for differential H3K4me2 peaks between KS and control. **D.** Representative of differential H3K4me2 sites between KS and control. Left: a representative of decreased H3K4me2 peaks in KS for *TLR9*. Blue, KS; gray, control. A gray box indicates significantly decreased H3K4me2 peaks in KS. An arrow indicates the transcriptional direction. Right: a representative of increased H3K4me2 peaks in KS for *MYO1F*. A gray box indicates significantly increased H3K4me2 peaks in KS.

**Figure 3. Increased H3K4me2 sites in KS compared to control. A.** Expression changes of *EGR1* and *E2F2* in KS whose motifs were most enriched for increased H3K4me2 peaks in KS compared to control. The bar graph shows the expression levels of these genes significantly increased in KS. P-values from a negative binomial model. **B.** Fraction of the increased H3K4me2 peaks in KS for *EGR1* targets vs. non-*EGR1* targets. P-values by Fisher's exact test. **C.** Differential DNA methylation profile of *MYO1F*. From the top, H3K4me2 profiles from KS individual and its control, and differential DNA methylation profile between KS and control, respectively. An arrow indicates the transcriptional direction. A gray box indicates the genomic regions overlapping with significantly increased H3K4me2 peaks and hypomethylated regions in KS compared to control.

**Figure 4. Association between altered enhancer sites in KS and super-enhancer regions. A.** Fraction of super-enhancers overlapping with altered H3K4me1 and/or me2 peaks in KS compared to control. **B.** Fold changes of altered signals of H3K4me1 and H3K4me2 within the super-enhancers. P-value by t-test. **C.** H3K27ac signal intensities for active enhancers. Gray, non-super-enhancers; black, super-enhancers; blue, super-enhancers overlapping with altered H3K4me1/me2 in KS.

**Figure 5. Transcriptional changes in KS. A.** Relationship between H3K4me2 signal changes and expression changes in KS compared to control. Each dot represents each gene with significant expression changes ( $q = 0.1$ ). Blue, changes in H3K4me2 and expression levels in the same direction; gray, changes in H3K4me2 and expression levels in the opposite direction. The blue line shows the regression line with a Pearson correlation efficiency of  $r$  value ( $r = 0.4$ ). **B.** Volcano plot of expression changes in KS compared to control for TFs. Blue, TFs exhibiting substantial transcriptional changes in KS with  $q$ -value  $< 0.1$  or fold changes  $> 2$ ; gray, rest of TFs expressed in blood. Light blue circle, ZNF family; green, erythroid cell master regulators; yellow, hematopoietic cell regulators; orange, cell cycle genes; blue, B, Th and T cell regulation; purple, dysregulated gene in other tissues. **C.** Transcription factors regulating differentially expressed genes in KS compared to control. The y-axis represents the significance of

prediction based on binding from public ChIP-seq data near the differentially expressed genes. The x-axis shows the significance of expression changes in KS for the gene. Blue, top 50 regulators underlying differentially expressed genes in KS; red, top 50 regulators underlying differentially expressed genes in KS and showing greater than 2-fold changes in expressions.

**Figure 6. Molecular processes underlying dysregulation in KS.** Schematic diagram for the underlying transcriptional regulation program in KS. Haploinsufficiency of *KMT2D* directly affects the *KMT2D* targets for enhancers/super-enhancers along with blood cell regulating TFs for immune-related genes. Upregulation of cell cycle genes, indirectly caused by haploinsufficiency of *KMT2D*, is associated with more open chromatin states near target gene promoters, resulting in the upregulation of metabolic genes. Finally, multiple dysregulated TFs affect a vast number of the target genes, which subsequently causes changes in blood cell differentiation.

## TABLES

**Table 1. Features of subjects and mutations in *KMT2D***

\*\* indicates a twin.

| Subject | Sex    | Age at enrollment | <i>KMT2D</i> variant (NM_003482.3)          | Variant type              | Genomic position in hg19 | Parental origin |
|---------|--------|-------------------|---|---------------------------|--------------------------|-----------------|
| S1      | male   | 4-10 years        | c.839+1del, intron 6                        | splice site               | 12:49447259              | n.d.            |
| S2      | female | 0-3 years         | c.12844C>T (p.Arg4282*), exon 39            | nonsense                  | 12:49425644              | de novo         |
| S3      | male   | 0-3 years         | c.5104C>T (p.Arg1702*), exon 21             | nonsense                  | 12:49438067              | maternal        |
| S4      | male   | 11-15 years       | c.13450C>T (p.Arg4484*), exon 39            | nonsense                  | 12:49425038              | n.d.            |
| S5      | female | 16-30 years       | c.6183+3G>T, intron 29                      | splice site               | 12:49435697              | n.d.            |
| S6      | female | 11-15 years       | c.16437del (p.Asn5480Thrfs*7), exon 53      | frameshift                | 12:49415910              | n.d.            |
| S7      | male   | 16-30 years       | c.9703_9704del (p.Lys3235Aspfs*15), exon 34 | frameshift                | 12:49431435              | n.d.            |
| S8      | female | 11-15 years       | c.10740+1G>A, intron 38                     | splice site               | 12:49427849              | n.d.            |
| S9      | female | 0-3 years         | c.2657del (p.Pro886Leufs*44), exon 10       | frameshift                | 12:49444813              | de novo         |
| S10     | female | 16-30 years       | c.7481dup (p.Ala2496Serfs*10), exon 31      | frameshift                | 12:49434073              | de novo         |
| S11     | female | 11-15 years       | c.13689C>T (p.Pro4563=), exon 41            | likely to affect splicing | 12:49424534              | n.d.            |
| S12     | male   | 0-3 years         | c.15641G>T (p.Arg5214Leu), exon 48          | missense                  | 12:49420108              | de novo         |
| S13     | male   | 11-15 years       | c.16052G>A (p.Arg5351Gln), exon 50          | missense                  | 12:49418361              | n.d.            |
| S14     | female | 4-10 years        | c.6086del (p.Pro2029Leufs*18), exon 28      | frameshift                | 12:49435896              | n.d.            |

|              |        |             |  |             |                      |          |
|--------------|--------|-------------|--|-------------|----------------------|----------|
| <b>S15</b>   | male   | 11-15 years | c.12395del (p.His4132Profs*10), exon 39          | frameshift  | 12:49426093          | n.d.     |
| <b>S16</b>   | female | 4-10 years  | c.13450C>T (p.Arg4484*), exon 39                 | nonsense    | 12:49425038          | n.d.     |
| <b>S17</b>   | female | 16-30 years | c.10507+2T>G, intron 37                          | splice site | 12:49428191          | n.d.     |
| <b>S18</b>   | male   | 4-10 years  | c.12592C>T (p.Arg4198*), exon 39                 | nonsense    | 12:49425896          | de novo  |
| <b>S19</b>   | male   | 4-10 years  | c.2533del (p.Arg845Glyfs*85), exon 10            | frameshift  | 12:49444938          | n.d.     |
| <b>S20</b>   | female | 0-3 years   | c.10507+2T>G, intron 37                          | splice site | 12:49428191          | n.d.     |
| <b>S21</b>   | female | 11-15 years | c.11263C>T (p.Gln3755*), exon 39                 | nonsense    | 12:49427225          | n.d.     |
| <b>S22</b>   | female | 16-30 years | c.10813C>T (p.Gln3605*), exon 39                 | nonsense    | 12:49427675          | de novo  |
| <b>S23</b>   | female | 0-3 years   | c.6295C>T (p.Arg2099*), exon 31                  | nonsense    | 12:49435258          | de novo  |
| <b>S24</b>   | female | 31-50 years | c.16469_16472del (p.Lys5490Argfs*18), exon 53    | frameshift  | 12:49415876_49415879 | de novo  |
| <b>S25**</b> | female | 4-10 years  | c.15467A>G (p.Tyr5156Cys), exon 48               | missense    | 12:49420282          | de novo  |
| <b>S26**</b> | female | 4-10 years  | c.15467A>G (p.Tyr5156Cys), exon 48               | missense    | 12:49420282          | de novo  |
| <b>S27</b>   | female | 0-3 years   | c.5645-2A>G, intron 25                           | splice site | 12:49436663          | maternal |
| <b>S28</b>   | female | 16-30 years | c.15536G>A (p.Arg5179His), exon 48               | missense    | 12:49420213          | n.d.     |
| <b>S29</b>   | female | 0-3 years   | c.349C>T (p.Gln117*), exon 3                     | nonsense    | 12:49448362          | de novo  |
| <b>S30</b>   | female | 11-15 years | c.8859_8861delinsCA (p.Lys2953Asnfs*51), exon 34 | frameshift  | 12:49432278_49432280 | n.d.     |
| <b>S31</b>   | male   | 4-10 years  | c.16295G>A (p.Arg5432Gln), exon 51               | missense    | 12:49416416          | de novo  |
| <b>S32</b>   | male   | 4-10 years  | c.2713del (p.Glu905Asnfs*25), exon 10            | frameshift  | 12:49444756          | de novo  |



|            |        |            |   |            |                              |      |
|------------|--------|------------|---|------------|------------------------------|------|
| <b>S33</b> | female | 4-10 years | c.11386_11411del<br>(p.Gln3796Cysfs*207), exon 39 | frameshift | 12:494270<br>80_49427<br>105 | n.d. |
|------------|--------|------------|---|------------|------------------------------|------|

## REFERENCES

- Adam, Margaret P., Siddharth Banka, Hans T. Bjornsson, Olaf Bodamer, Albert E. Chudley, Jaqueline Harris, Hiroshi Kawame, et al. 2019. "Kabuki Syndrome: International Consensus Diagnostic Criteria." *Journal of Medical Genetics* 56 (2): 89–95.
- Adam, Margaret P., Louanne Hudgins, and Mark Hannibal. 2011. "Kabuki Syndrome." In *GeneReviews®*, edited by Margaret P. Adam, Holly H. Ardinger, Roberta A. Pagon, Stephanie E. Wallace, Lora J. H. Bean, Karen W. Gripp, Ghayda M. Mirzaa, and Anne Amemiya. Seattle (WA): University of Washington, Seattle.
- Alam, Hunain, Ming Tang, Mayinuer Maitituoheti, Shilpa S. Dhar, Manish Kumar, Chae Young Han, Chandrashekar R. Ambati, et al. 2020. "KMT2D Deficiency Impairs Super-Enhancers to Confer a Glycolytic Vulnerability in Lung Cancer." *Cancer Cell* 37 (4): 599–617.e7.
- Anders, Simon, Paul Theodor Pyl, and Wolfgang Huber. 2015. "HTSeq--a Python Framework to Work with High-Throughput Sequencing Data." *Bioinformatics* 31 (2): 166–69.
- Aref-Eshghi, Erfan, David I. Rodenhiser, Laila C. Schenkel, Hanxin Lin, Cindy Skinner, Peter Ainsworth, Guillaume Paré, et al. 2018. "Genomic DNA Methylation Signatures Enable Concurrent Diagnosis and Clinical Genetic Variant Classification in Neurodevelopmental Syndromes." *American Journal of Human Genetics* 102 (1): 156–74.
- Aref-Eshghi, Erfan, Laila C. Schenkel, Hanxin Lin, Cindy Skinner, Peter Ainsworth, Guillaume Paré, David Rodenhiser, Charles Schwartz, and Bekim Sadikovic. 2017. "The Defining DNA Methylation Signature of Kabuki Syndrome Enables Functional Assessment of Genetic Variants of Unknown Clinical Significance." *Epigenetics: Official Journal of the DNA Methylation Society* 12 (11): 923–33.
- Baccala, Roberto, Kasper Hoebe, Dwight H. Kono, Bruce Beutler, and Argyrios N. Theofilopoulos. 2007. "TLR-Dependent and TLR-Independent Pathways of Type I Interferon Induction in Systemic Autoimmunity." *Nature Medicine* 13 (5): 543–51.
- Bailey, Timothy L., Mikael Boden, Fabian A. Buske, Martin Frith, Charles E. Grant, Luca Clementi, Jingyuan Ren, Wilfred W. Li, and William S. Noble. 2009. "MEME SUITE: Tools for Motif Discovery and Searching." *Nucleic Acids Research* 37 (Web Server issue): W202–8.
- Banka, Siddharth, Ratna Veeramachaneni, William Reardon, Emma Howard, Sancha Bunstone, Nicola Ragge, Michael J. Parker, et al. 2012. "How Genetically Heterogeneous Is Kabuki Syndrome?: MLL2 Testing in 116 Patients, Review and Analyses of Mutation and Phenotypic Spectrum." *European Journal of Human Genetics: EJHG* 20 (4): 381–88.
- Bjornsson, Hans T., Joel S. Benjamin, Li Zhang, Jacqueline Weissman, Elizabeth E. Gerber, Yi-Chun Chen, Rebecca G. Vaurio, Michelle C. Potter, Kasper D. Hansen, and Harry C. Dietz. 2014. "Histone Deacetylase Inhibition Rescues Structural and Functional Brain Deficits in a Mouse Model of Kabuki Syndrome." *Science Translational Medicine* 6 (256): 256ra135.
- Bögershausen, Nina, I-Chun Tsai, Esther Pohl, Pelin Özlem Simsek Kiper, Filippo Beleggia, E. Ferda Percin, Katharina Keupp, et al. 2015. "RAP1-Mediated MEK/ERK Pathway Defects in Kabuki Syndrome." *Journal of Clinical Investigation*. <https://doi.org/10.1172/jci80102>.
- Butcher, Darci T., Cheryl Cytrynbaum, Andrei L. Turinsky, Michelle T. Siu, Michal Inbar-Feigenberg, Roberto Mendoza-Londono, David Chitayat, et al. 2017. "CHARGE and Kabuki Syndromes: Gene-Specific DNA Methylation Signatures Identify Epigenetic Mechanisms Linking These Clinically Overlapping Conditions." *American Journal of Human Genetics*

100 (5): 773–88.

- Carosso, Giovanni A., Leandros Boukas, Jonathan J. Augustin, Ha Nam Nguyen, Briana L. Winer, Gabrielle H. Cannon, Johanna D. Robertson, et al. 2019. "Precocious Neuronal Differentiation and Disrupted Oxygen Responses in Kabuki Syndrome." *JCI Insight* 4 (20). <https://doi.org/10.1172/jci.insight.129375>.
- Cuvertino, Sara, Verity Hartill, Alice Colyer, Terence Garner, Nisha Nair, Lihadh Al-Gazali, Natalie Canham, et al. 2020. "A Restricted Spectrum of Missense KMT2D Variants Cause a Multiple Malformations Disorder Distinct from Kabuki Syndrome." *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 22 (5): 867–77.
- Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. 2013. "STAR: Ultrafast Universal RNA-Seq Aligner." *Bioinformatics* 29 (1): 15–21.
- Fahrner, Jill A., Wan-Ying Lin, Ryan C. Riddle, Leandros Boukas, Valerie B. DeLeon, Sheetal Chopra, Susan E. Lad, Teresa Romeo Luperchio, Kasper D. Hansen, and Hans T. Bjornsson. 2019. "Precocious Chondrocyte Differentiation Disrupts Skeletal Growth in Kabuki Syndrome Mice." *JCI Insight* 4 (20). <https://doi.org/10.1172/jci.insight.129380>.
- Fasciani, Alessandra, Sarah D'Annunzio, Vittoria Poli, Luca Fagnocchi, Sven Beyes, Daniela Michelatti, Francesco Corazza, et al. 2020. "MLL4-Associated Condensates Counterbalance Polycomb-Mediated Nuclear Mechanical Stress in Kabuki Syndrome." *Nature Genetics* 52 (12): 1397–1411.
- Feng, Jianxing, Tao Liu, Bo Qin, Yong Zhang, and Xiaole Shirley Liu. 2012. "Identifying ChIP-Seq Enrichment Using MACS." *Nature Protocols* 7 (9): 1728–40.
- Froimchuk, Eugene, Younghoon Jang, and Kai Ge. 2017. "Histone H3 Lysine 4 Methyltransferase KMT2D." *Gene* 627 (September): 337–42.
- Gupta, Shobhit, John A. Stamatoyannopoulos, Timothy L. Bailey, and William Stafford Noble. 2007. "Quantifying Similarity between Motifs." *Genome Biology* 8 (2): R24.
- Gu, Ye, Qian Wang, Kang Guo, Weizhao Qin, Wenting Liao, Shuang Wang, Yanqing Ding, and Jie Lin. 2016. "TUSC3 Promotes Colorectal Cancer Progression and Epithelial-Mesenchymal Transition (EMT) through WNT/ $\beta$ -Catenin and MAPK Signalling." *The Journal of Pathology* 239 (1): 60–71.
- Hannibal, Mark C., Kati J. Buckingham, Sarah B. Ng, Jeffrey E. Ming, Anita E. Beck, Margaret J. McMillin, Heidi I. Gildersleeve, et al. 2011. "Spectrum of MLL2 (ALR) Mutations in 110 Cases of Kabuki Syndrome." *American Journal of Medical Genetics. Part A* 155 (7): 1511–16.
- Hnisz, Denes, Brian J. Abraham, Tong Ihn Lee, Ashley Lau, Violaine Saint-André, Alla A. Sigova, Heather A. Hoke, and Richard A. Young. 2013. "Super-Enhancers in the Control of Cell Identity and Disease." *Cell* 155 (4): 934–47.
- Hu, Hui, Ya-Ru Miao, Long-Hao Jia, Qing-Yang Yu, Qiong Zhang, and An-Yuan Guo. 2019. "AnimalTFDB 3.0: A Comprehensive Resource for Annotation and Prediction of Animal Transcription Factors." *Nucleic Acids Research* 47 (D1): D33–38.
- Huisman, Christian, Young A. Kim, Shin Jeon, Bongjin Shin, Jeonghoon Choi, Su Jeong Lim, Sung Min Youn, et al. 2021. "The Histone H3-Lysine 4-Methyltransferase Mll4 Regulates the Development of Growth Hormone-Releasing Hormone-Producing Neurons in the Mouse Hypothalamus." *Nature Communications* 12 (1): 256.
- Issaeva, Irina, Yulia Zonis, Tanya Rozovskaia, Kira Orlovsky, Carlo M. Croce, Tatsuya Nakamura, Alex Mazo, Lea Eisenbach, and Eli Canaani. 2007. "Knockdown of ALR (MLL2) Reveals ALR Target Genes and Leads to Alterations in Cell Adhesion and Growth." *Molecular and Cellular Biology* 27 (5): 1889–1903.
- Jolma, Arttu, Jian Yan, Thomas Whittington, Jarkko Toivonen, Kazuhiro R. Nitta, Pasi Rastas, Ekaterina Morgunova, et al. 2013. "DNA-Binding Specificities of Human Transcription Factors." *Cell* 152 (1-2): 327–39.

- Khan, Aziz, Oriol Fornes, Arnaud Stigliani, Marius Gheorghe, Jaime A. Castro-Mondragon, Robin van der Lee, Adrien Bessy, et al. 2018. "JASPAR 2018: Update of the Open-Access Database of Transcription Factor Binding Profiles and Its Web Framework." *Nucleic Acids Research* 46 (D1): D1284.
- Kuroki, Y., Y. Suzuki, H. Chyo, A. Hata, and I. Matsui. 1981. "A New Malformation Syndrome of Long Palpebral Fissures, Large Ears, Depressed Nasal Tip, and Skeletal Anomalies Associated with Postnatal Dwarfism and Mental Retardation." *The Journal of Pediatrics* 99 (4): 570–73.
- Lederer, Damien, Bernard Grisart, Maria Cristina Digilio, Valérie Benoit, Marianne Crespín, Sophie Claire Ghariani, Isabelle Maystadt, Bruno Dallapiccola, and Christine Verellen-Dumoulin. 2012. "Deletion of KDM6A, a Histone Demethylase Interacting with MLL2, in Three Patients with Kabuki Syndrome." *American Journal of Human Genetics* 90 (1): 119–24.
- Lefter, Mihai, Jonathan K. Vis, Martijn Vermaat, Johan T. den Dunnen, Peter E. M. Taschner, and Jeroen F. J. Laros. 2021. "Next Generation HGVS Nomenclature Checker." *Bioinformatics*, February. <https://doi.org/10.1093/bioinformatics/btab051>.
- Li, Heng, and Richard Durbin. 2009. "Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform." *Bioinformatics* 25 (14): 1754–60.
- Lindsley, Andrew W., Howard M. Saal, Thomas A. Burrow, Robert J. Hopkin, Oleg Shchelochkov, Pooja Khandelwal, Changchun Xie, et al. 2016. "Defects of B-Cell Terminal Differentiation in Patients with Type-1 Kabuki Syndrome." *The Journal of Allergy and Clinical Immunology* 137 (1): 179–87.e10.
- Li, Yun, Nina Bögershausen, Yasemin Alanay, Pelin Ozlem Simsek Kiper, Nadine Plume, Katharina Keupp, Esther Pohl, et al. 2011. "A Mutation Screen in Patients with Kabuki Syndrome." *Human Genetics* 130 (6): 715–24.
- Love, Michael I., Wolfgang Huber, and Simon Anders. 2014. "Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2." *Genome Biology* 15 (12): 550.
- Lovén, Jakob, Heather A. Hoke, Charles Y. Lin, Ashley Lau, David A. Orlando, Christopher R. Vakoc, James E. Bradner, Tong Ihn Lee, and Richard A. Young. 2013. "Selective Inhibition of Tumor Oncogenes by Disruption of Super-Enhancers." *Cell* 153 (2): 320–34.
- Luperchio, Teresa Romeo, Leandros Boukas, Li Zhang, Genay Pilarowski, Jenny Jiang, Allison Kalinousky, Kasper D. Hansen, and Hans T. Bjornsson. 2021. "Leveraging the Mendelian Disorders of the Epigenetic Machinery to Systematically Map Functional Epigenetic Variation." *eLife* 10 (August). <https://doi.org/10.7554/eLife.65884>.
- Martynoga, Ben, Juan L. Mateo, Bo Zhou, Jimena Andersen, Angeliki Achimastou, Noelia Urbán, Debbie van den Berg, et al. 2013. "Epigenomic Enhancer Annotation Reveals a Key Role for NFIX in Neural Stem Cell Quiescence." *Genes & Development* 27 (16): 1769–86.
- McLean, Cory Y., Dave Bristor, Michael Hiller, Shoa L. Clarke, Bruce T. Schaar, Craig B. Lowe, Aaron M. Wenger, and Gill Bejerano. 2010. "GREAT Improves Functional Interpretation of Cis-Regulatory Regions." *Nature Biotechnology* 28 (5): 495–501.
- Micale, Lucia, Bartolomeo Augello, Carmela Fusco, Angelo Selicorni, Maria N. Loviglio, Margherita Cirillo Silengo, Alexandre Reymond, et al. 2011. "Mutation Spectrum of MLL2 in a Cohort of Kabuki Syndrome Patients." *Orphanet Journal of Rare Diseases* 6 (June): 38.
- Ming, Jeffrey E., Karen L. Russell, Donna M. McDonald-McGinn, and Elaine H. Zackai. 2005. "Autoimmune Disorders in Kabuki Syndrome." *American Journal of Medical Genetics. Part A* 132A (3): 260–62.
- Ng, Sarah B., Abigail W. Bigham, Kati J. Buckingham, Mark C. Hannibal, Margaret J. McMillin, Heidi I. Gildersleeve, Anita E. Beck, et al. 2010. "Exome Sequencing Identifies MLL2 Mutations as a Cause of Kabuki Syndrome." *Nature Genetics* 42 (9): 790–93.
- Niikawa, N., N. Matsuura, Y. Fukushima, T. Ohsawa, and T. Kajii. 1981. "Kabuki Make-up Syndrome: A Syndrome of Mental Retardation, Unusual Facies, Large and Protruding Ears,

- and Postnatal Growth Deficiency.” *The Journal of Pediatrics* 99 (4): 565–69.
- Ortega-Molina, Ana, Isaac W. Boss, Andres Canela, Heng Pan, Yanwen Jiang, Chunying Zhao, Man Jiang, et al. 2015. “The Histone Lysine Methyltransferase KMT2D Sustains a Gene Expression Program That Represses B Cell Lymphoma Development.” *Nature Medicine* 21 (10): 1199–1208.
- Park, Sangho, Ute Osmers, Gayathree Raman, Rebecca H. Schwantes, Manuel O. Diaz, and John H. Bushweller. 2010. “The PHD3 Domain of MLL Acts as a CYP33-Regulated Switch between MLL-Mediated Activation and Repression.” *Biochemistry* 49 (31): 6576–86.
- Pekowska, Aleksandra, Touati Benoukraf, Pierre Ferrier, and Salvatore Spicuglia. 2010. “A Unique H3K4me2 Profile Marks Tissue-Specific Gene Regulation.” *Genome Research* 20 (11): 1493–1502.
- Peloquin, Joanna M., Gautam Goel, Eduardo J. Villablanca, and Ramnik J. Xavier. 2016. “Mechanisms of Pediatric Inflammatory Bowel Disease.” *Annual Review of Immunology* 34 (May): 31–64.
- Pilarowski, Genay O., Tareian Cazares, Li Zhang, Joel S. Benjamin, Ke Liu, Sajjeev Jagannathan, Nadeem Mousa, et al. 2020. “Abnormal Peyer Patch Development and B-Cell Gut Homing Drive IgA Deficiency in Kabuki Syndrome.” *The Journal of Allergy and Clinical Immunology* 145 (3): 982–92.
- Qin, Qian, Jingyu Fan, Rongbin Zheng, Changxin Wan, Shenglin Mei, Qiu Wu, Hanfei Sun, et al. 2020. “Lisa: Inferring Transcriptional Regulators through Integrative Modeling of Public Chromatin Accessibility and ChIP-Seq Data.” *Genome Biology* 21 (1): 32.
- Richards, Sue, Nazneen Aziz, Sherri Bale, David Bick, Soma Das, Julie Gastier-Foster, Wayne W. Grody, et al. 2015. “Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology.” *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 17 (5): 405–24.
- Ross-Innes, Caryn S., Rory Stark, Andrew E. Teschendorff, Kelly A. Holmes, H. Raza Ali, Mark J. Dunning, Gordon D. Brown, et al. 2012. “Differential Oestrogen Receptor Binding Is Associated with Clinical Outcome in Breast Cancer.” *Nature* 481 (7381): 389–93.
- Sobreira, Nara, Martha Brucato, Li Zhang, Christine Ladd-Acosta, Chrissie Ongaco, Jane Romm, Kimberly F. Doheny, et al. 2017. “Patients with a Kabuki Syndrome Phenotype Demonstrate DNA Methylation Abnormalities.” *European Journal of Human Genetics: EJHG* 25 (12): 1335–44.
- Stagi, Stefano, Anna Virginia Gulino, Elisabetta Lapi, and Donato Rigante. 2016. “Epigenetic Control of the Immune System: A Lesson from Kabuki Syndrome.” *Immunologic Research* 64 (2): 345–59.
- Tanaka, Shigeru, Akira Suto, Taro Iwamoto, Daisuke Kashiwakuma, Shin-Ichiro Kagami, Kotaro Suzuki, Hiroaki Takatori, et al. 2014. “Sox5 and c-Maf Cooperatively Induce Th17 Cell Differentiation via ROR $\gamma$ t Induction as Downstream Targets of Stat3.” *Journal of Experimental Medicine*. <https://doi.org/10.1084/jem.20130791>.
- Teruya, Shuzo, Tomohisa Okamura, Toshihiko Komai, Mariko Inoue, Yukiko Iwasaki, Shuji Sumitomo, Hirofumi Shoda, Kazuhiko Yamamoto, and Keishi Fujio. 2018. “Egr2-Independent, Klf1-Mediated Induction of PD-L1 in CD4<sup>+</sup> T Cells.” *Scientific Reports* 8 (1): 7021.
- Theodore-Oklot, Christina, Deborah S. Hartman, Deborah L. Hoffman, and Hans T. Björnsson. 2022. “A Qualitative Study to Characterize the Humanistic Burden of Kabuki Syndrome in the United States and Canada.” *Advances in Therapy* 39 (1): 619–31.
- Tsai, I-Chun, Kelly McKnight, Spencer U. McKinstry, Andrew T. Maynard, Perciliz L. Tan, Christelle Golzio, C. Thomas White, et al. 2018. “Small Molecule Inhibition of RAS/MAPK Signaling Ameliorates Developmental Pathologies of Kabuki Syndrome.” *Scientific Reports* 8 (1): 10779.

Whyte, Warren A., David A. Orlando, Denes Hnisz, Brian J. Abraham, Charles Y. Lin, Michael H. Kagey, Peter B. Rahl, Tong Ihn Lee, and Richard A. Young. 2013. "Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes." *Cell* 153 (2): 307–19.

Xu, Jing, Yu Liu, Chong Chen, and Ting Niu. 2018. "KMT2D Is a Haploinsufficient Tumor Suppressor in Acute Leukemia." *Blood* 132 (November): 1511.

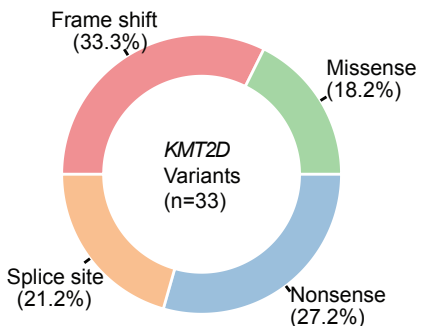
Zhang, Jiyuan, David Dominguez-Sola, Shafinaz Hussein, Ji-Eun Lee, Antony B. Holmes, Mukesh Bansal, Sofija Vlasevska, et al. 2015. "Disruption of KMT2D Perturbs Germinal Center B Cell Development and Promotes Lymphomagenesis." *Nature Medicine* 21 (10): 1190–98.

Zhang, Rong, James T. Earnest, Arthur S. Kim, Emma S. Winkler, Pritesh Desai, Lucas J. Adams, Gaowei Hu, et al. 2019. "Expression of the Mxra8 Receptor Promotes Alphavirus Infection and Pathogenesis in Mice and *Drosophila*." *Cell Reports* 28 (10): 2647–58.e5.

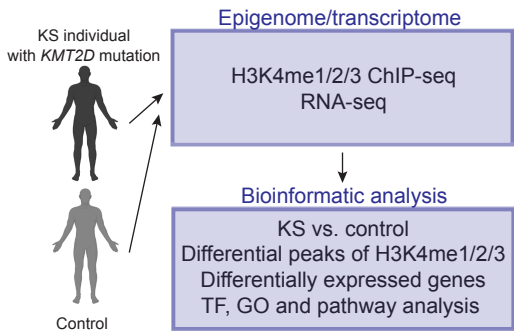


# Figure 1

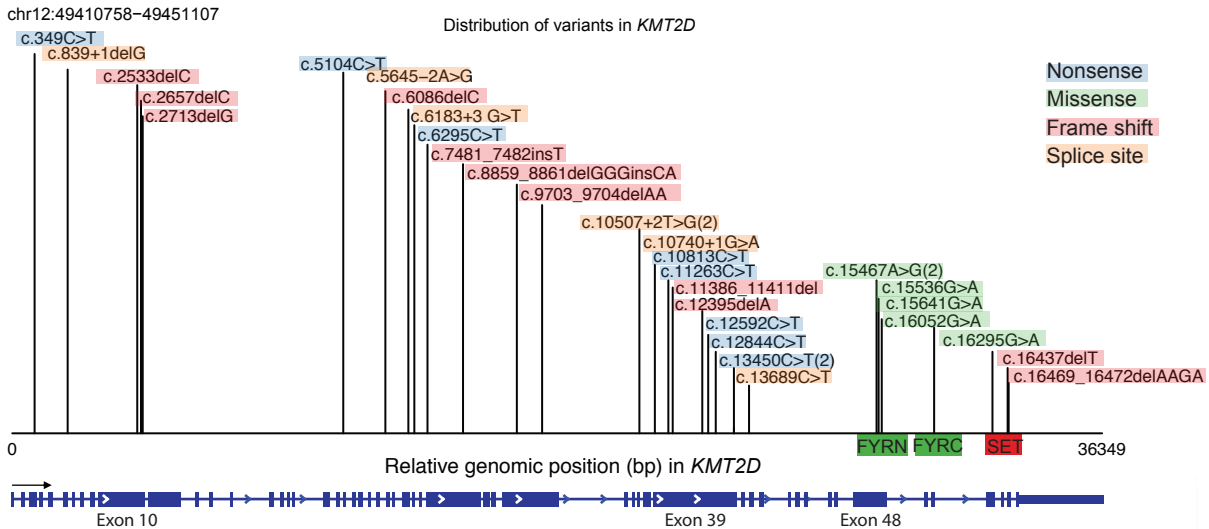
**A**



**C**

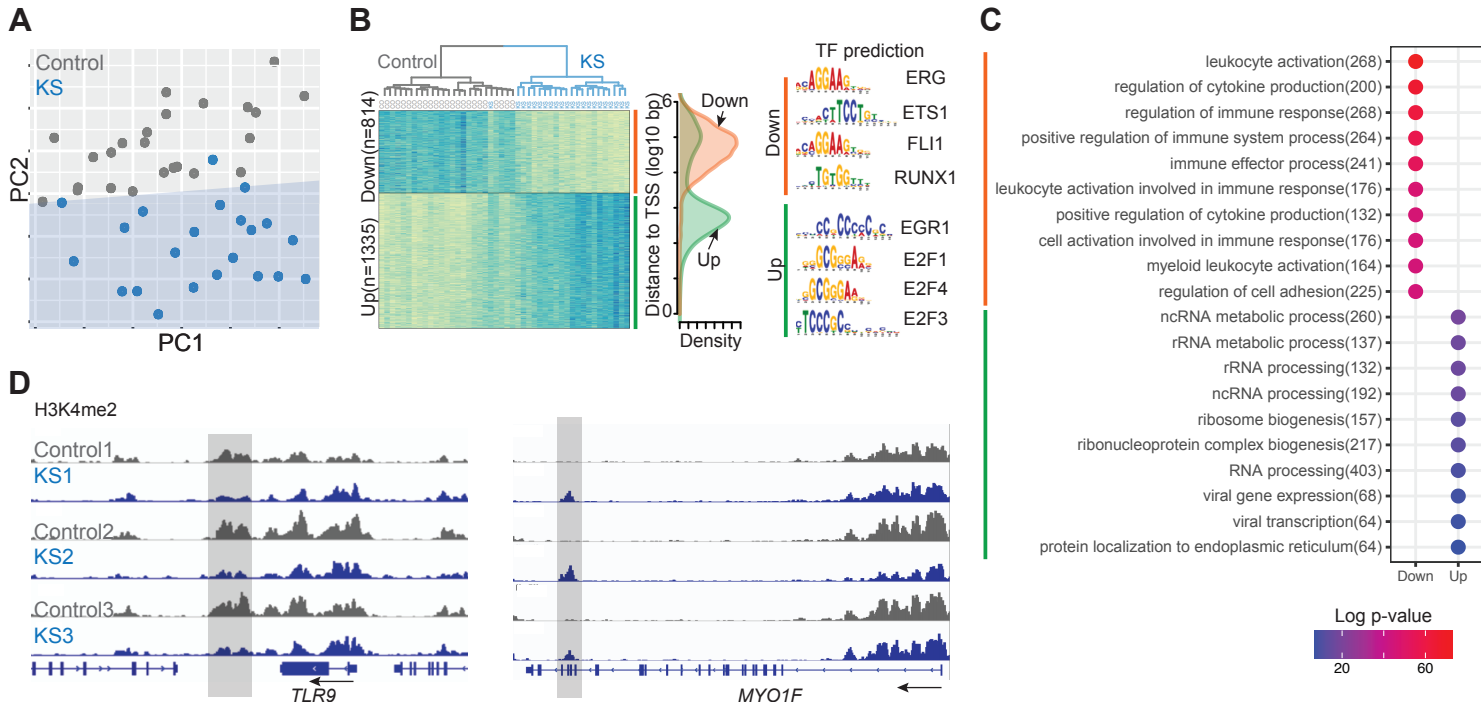


**B**



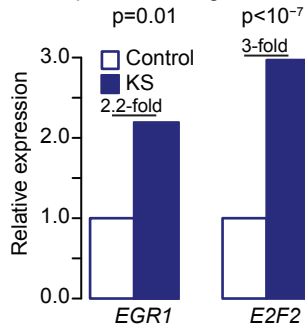


# Figure 2

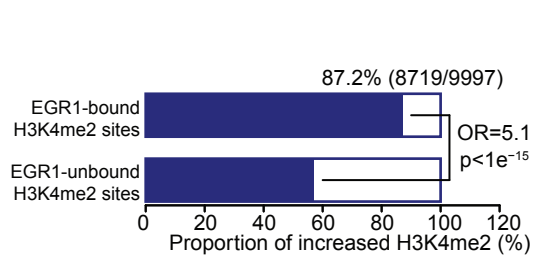


# Figure 3

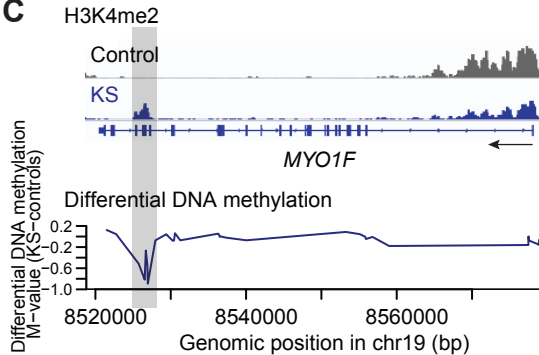
## A Expression changes in KS



## B EGR1-bound vs. EGR1-unbound sites

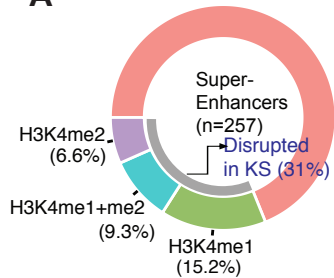


## C

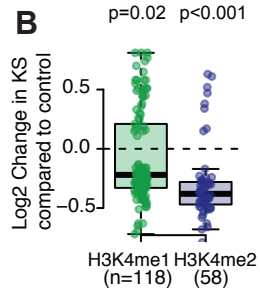


# Figure 4

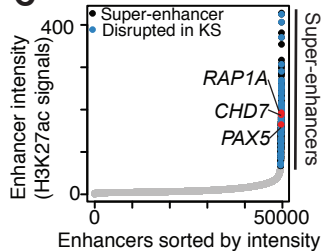
**A**



**B**

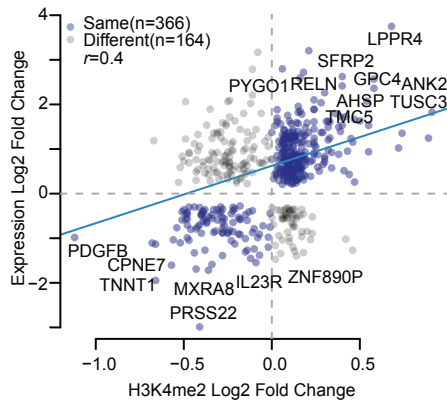


**C**

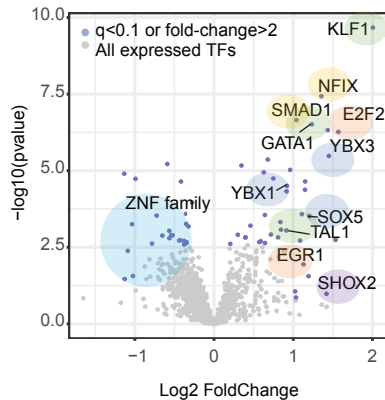


# Figure 5

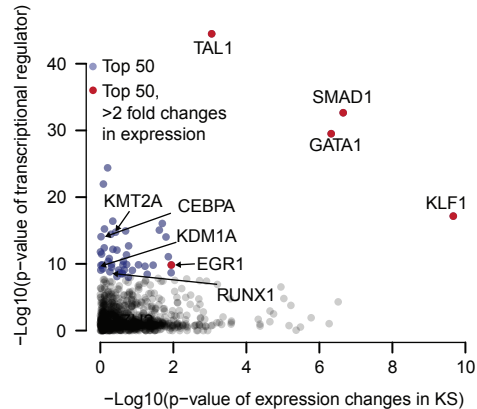
## A



## B



## C



# Figure 6

