H3K27me3-rich genomic regions can function as silencers to repress gene expression via chromatin interactions

Yichao Cai[1]*, Ying Zhang[2]*, Yan Ping Loh[2], Jia Qi Tng[2], Mei Chee Lim[2,3], Zhendong Cao[2,4], Anandhkumar Raju[5], Shang Li[3,6], Lakshmanan Manikandan[5], Vinay Tergaonkar[5], Greg Tucker-Kellogg[1,7†], Melissa Jane Fullwood [2,5,8†]

[1]Department of Biological Sciences, National University of Singapore, 16 Science Drive 4, 117558 Singapore.
[2]Cancer Science Institute of Singapore, National University of Singapore, 14 Medical Drive, 117599 Singapore.
[3]Cancer and Stem Cell Biology Programme, Duke-NUS Medical School, 8 College Road, 169857 Singapore.
[4]Department of Cancer Biology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA19104, USA.
[5]Institute of Molecular and Cell Biology, Agency for Science, Technology and Research (A*STAR), 61 Biopolis Drive, Proteos, 138673 Singapore.
[6]Department of Physiology, Yong Loo Lin School of Medicine, National University of Singapore, 2 Medical Drive,117597 Singapore.
[7]Computational Biology Programme, National University of Singapore, 6 Science Drive 2, 117546 Singapore
[8]School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, 637551 Singapore.

*These authors made equal and critical contributions
[†]Correspondence should be sent to:
(1) Melissa J. Fullwood, Cancer Science Institute Singapore (CSI) and School of Biological Sciences, Nanyang Technological University, Email: mfullwood@ntu.edu.sg; Telephone: (65) 6516 5381; Fax: (65) 6873 9664
(2) Greg Tucker-Kellogg, Department of Biological Sciences, National University of Singapore. Email: dbsgtk@nus.edu.sg, Telephone: (65) 6516 4740

Abstract

Gene repression and silencers are poorly understood. H3K27me3 is a repressive histone modification; we reason that H3K27me3-rich regions (MRRs) of the genome defined from clusters of H3K27me3 peaks may be used to identify silencers that can regulate gene expression via proximity or looping. We found that MRRs are associated with chromatin interactions and tend to interact preferentially with each other. EZH2 inhibition or knockout showed that H3K27me3 was not required for maintenance of chromatin interactions, but genes at or looping to MRRs were upregulated upon loss of H3K27me3. To understand the function of MRRs, we used CRISPR to excise components of MRRs at interaction anchors and functionally characterized the knockouts in cellular assays and xenograft models. MRR removal can lead to upregulation of interacting target genes, altered chromatin interactions, changes in phenotype associated with cell identity, and altered xenograft tumor growth. Our results characterize silencers and their mechanisms of functioning.

Introduction

Gene transcription is controlled largely by transcription factors that bind to enhancers and promoters to regulate genes[1]. Transcription factors can bind to enhancers in the genome, and enhancers distal to genes can loop to gene promoters ("chromatin interactions")[2]. By contrast, mechanisms for gene repression are much less well understood. Silencers, which are regions of the genome that are associated with gene silencing, are less well characterized than enhancers, especially in humans. Distant silencers are thought to loop over to target genes to silence them[3,4], but such examples have not been well explored in human systems.

Heterochromatin and gene silencing are associated with both H3K27me3 and H3K9me3 marks. The H3K9me3 mark is associated with constitutive heterochromatin, which tends to occur at regions with tandem repeats such as centromeres and telomeres, while H3K27me3 is associated with facultative heterochromatin[5]. DNA regions marked by H3K27me3 are known to harbor silencers, but H3K27me3 marks are insufficient to describe silencers[6]. Enhancer of zeste homolog 2 (EZH2), the catalytic component of polycomb repressive complex 2 (PRC2) complex, deposits histone H3K27me3 in the genome. EZH2 is dysregulated or mutated in numerous cancers, such as lymphoma, which shows gain of function mutations in EZH2[7]. EZH2 inhibitors have been developed for clinical purposes[8], however, due to the ubiquity of H3K27me3 marks throughout the human genome, it is unclear how EZH2 inhibitors can be used in a specific manner in the genome.

We reasoned that clusters of H3K27me3 peaks with strong H3K27me3 signal may be a rich source of silencers, on the basis of several lines of evidence. First, super-enhancers, which are regions defined by high levels of H3K27ac, transcription factors and mediators, have been demonstrated to form at and drive key cell identity genes and oncogenes in tumors[9-11]. Similarly, broad H3K4me3 regions have been used to identify tumor suppressor genes[12]. Super-enhancers and broad H3K4me3 regions show chromatin looping to target genes[13,14]. While it is yet unclear whether super-enhancers and broad H3K4me3 regions are distinctly different types of regulatory elements, they have been useful in identifying gene targets important in controlling cell identity as well as therapeutic vulnerabilities of cancers[10,15].

Here, we characterized H3K27me3-rich regions (MRRs) defined from clusters of H3K27me3 peaks in the genome. We found that chromatin interactions are denser within constituent peaks of MRRs than in typical H3K27me3 peaks. We

experimentally tested examples of MRRs for silencer properties and found that they can serve as silencers in the genome and have functional roles in controlling cell identity. We found that MRRs can regulate gene expression via proximity or looping. Moreover, genes at or looping to MRRs were more upregulated upon loss of H3K27me3 through EZH2 inhibition or knockout than genes at typical H3K27me3 regions; chromatin interactions associated with MRRs did not change, however, suggesting that H3K27me3 is not required for the maintenance of chromatin interactions in differentiated cells.

Results

Identification and characterization of H3K27me3-rich regions (MRRs) in the human genome

We identified highly H3K27me3-associated regions from cell lines using H3K27me3 ChIP-seq data[16] in the following manner: we first identified H3K27me3 peaks, then clustered nearby peaks, and ranked the clustered peaks by average H3K27me3 signals levels. The top clusters with the highest H3K27me3 signal were called H3K27me3-rich regions (MRRs) and the rest "typical H3K27me3" regions (Figure 1A, 1B). This is similar to how super-enhancers were defined[9,11].

Many MRR-associated genes in different cell lines are known or predicted tumor suppressor genes[17] (Figure S1A). For example, *NPM1*, the most commonly mutated gene in leukemia[18-21], overlaps with an MRR in the leukemic cell line K562. *FAT1*, which is frequently mutated in CLL and can act as a tumor suppressor through inhibiting Wnt signaling[22,23] also overlaps with an MRR in K562. Gene ontology analysis showed that MRR-related genes are enriched in developmental and differentiation processes, while genes associated with typical H3K27me3 peaks are enriched in cell metabolism and transportation processes (Figure S1B, S1C).

ChIP-seq signals of EZH2 showed high overlap with typical H3K27me3, MRR and constituent peaks of MRRs, consistent with EZH2's role in H3K27me3 mark deposition (Figure 1C; Figure S1D, S1E). Notably, constituent peaks of MRRs had higher H3K27me3 and EZH2 signals than typical H3K27me3 peaks.

We overlapped MRRs with high-resolution *in situ* Hi-C data[24], and found that H3K27me3 peaks within MRRs had a higher density of chromatin interactions than typical H3K27me3 peaks in both K562 and GM12878 (Figure 1D; Figure S1F, S1G). The involvement of chromatin interactions in MRRs was similar to super-enhancers compared with typical enhancers[13], suggesting chromatin interactions might be important within strong histone modification regions.

MRRs were different in different cell lines (Figure 1E), and MRRs unique to individual cell lines were the most common among the lines we studied (Figure S1H,S1I). For example, the cadherin-like coding gene *CPED1* is covered by a broad MRR in GM12878, but overlaps with a super-enhancer in K562. Conversely, the gene for *DENND2D* is associated with an MRR but overlaps with a super-enhancer in GM12878. In addition, MRRs were different in different cell lines, and most MRRs were unique to cell lines (Figure S1J). The specificity and uniqueness of MRRs suggested they might be primed for specific regulation in different cell contexts.

Analysis of cell line expression data showed that genes which transit from MRR-associated to H3K27ac peak-associated in a different cell line were up-regulated, while genes transit from super enhancer-associated to H3K27me3 peak-associated as accompanied were down-regulated (Figure 1F). Further, genes whose

expression was more cell line-specific were associated with more MRRs than those genes with lower expression specificity (Figure S1K).

In summary, we defined MRRs using H3K27me3 ChIP-seq signal, and showed that MRRs might be involved in specific gene repression related to development, differentiation and cancer via chromatin interactions.

H3K27me3-rich regions (MRRs) preferentially associate with MRRs in the human genome via chromatin interactions

We assigned chromatin states at Hi-C interaction anchors using H3K27me3 and H3K27ac peaks: active (A) anchors overlap with H2K27ac peaks, repressive (R) anchors overlap with H3K27me3 peaks, bivalent (B) anchors overlap with both H3K27me3 and H3K27ac peaks, and quiescent (Q) anchors overlap with neither peak (Figure 2A). We further defined the chromatin state pair of an interaction as the chromatin states of its anchors. We calculated the expected number of interactions for each state pair under a homogeneous model, and compared those expectations to the actual number of observations.

Interactions between anchors of the same state (AA, RR, and BB) were more likely to interact with each other, while interactions with highly different chromatin state pairs (e.g., AR, BQ) less likely (Figure 2B, left), regardless of cell line. When grouped into typical H3K27me3 peaks (T) versus high H3K27me3 regions or MRRs (H), the high H3K27me3 regions showed a preference for interactions with other MRRs (Figure 2B, right). In keeping with A/B chromatin compartments of the nucleus, this 'like-like' preference indicated that loci of similar chromatin state were prone to interact with each other. This like-like preference is confirmed in 4C validation experiments on selected loci (Figure S2A-S2E). The tendency of MRRs to interact with MRRs suggested that deeply-repressed genomic regions were closely packed in 3D space.

To further explore the potential regulatory role of MRRs in chromatin interactions, we identified the subset of MRR-anchored interactions where at least one anchor peak overlapped a gene transcription start site, and grouped them according to whether the MRR anchor was proximal or distal to the TSS anchor (Figure 2C, 2D; Figure S2A-S2F, S2G). Both proximal and distal gene looping occur for MRR-anchored interactions, but some MRRs are large enough that both anchors occur in the same MRR. While proximal looping genes are a subset of the genes within MRRs, distal looping genes are only identified by chromatin interactions (Figure 2D, right). Gene ontology analysis shows that MRR-associated genes in the context of chromatin interactions are involved in developmental and differentiation processes (Figure S2H-S2I). Example loci of both a proximal looping gene (*GSE1*) and a distal looping gene (*PPP3CC*) are shown in Figure 2E and 2F, while Figure 2G and 2H show examples of internal looping, all in K562 cells.

Next, we performed Circular Chromosome Conformation Capture (4C) experiments on selected loci to investigate the associated chromatin interactions in a comprehensive and high resolution manner. We and annotated the interactions based on the chromatin state of the anchor distal from the bait in K562 (Figure 2I and Figure S2J-L), and across multiple cell lines (Figure S2M-N). The interaction profiles of 4C baits of different states were largely dominated by interacting regions of the same state as the baits.

We also carried out 4C experiments on the same bait across different cell lines. The interactions and the chromatin state at the bait locus varied in different cell lines, but the interaction profile maintained a preference for the same chromatin state

as the bait (Figure S2M, S2N). As a further test of this concept, the extensive BB long-range interactions (green arcs) connecting *PSMD5* and *TOR1A* in K562 were validated using reciprocal 4C bait design. When the *PSMD5* bait region was A (active) in either GM12878 or HAP1 cells, the BB interactions were largely reduced and other types of interactions started to appear (Figure S2M).

CRISPR excision of an MRR looping to *FGF18* leads to *FGF18* gene upregulation and altered cell adhesion

Next, we asked if MRRs function as silencers to regulate gene expression. We selected 2 MRRs for functional testing based on the H3K27me3 signal, the presence of Hi-C anchors and the number of Hi-C anchors they associated with whether the genes were involved in cell identity (Supplementary Text). Briefly, there are 974 MRRs in K562 (Figure S3A) and of those MRRs, 237 MRRs are associated with genes. Among these, 130 MRRs show proximal looping to genes (MRRs overlap with target gene promoters), 111 MRRs show distal looping to genes (MRR loops over to the promoter of target gene by long-range chromatin interactions) and 51 MRRs show internal looping to genes (Part of the MRR overlaps with the target gene promoter and the other part of the MRR loops over to the promoter of the target gene by long-range chromatin interactions). From this list, we selected MRR1, an internal looping example which showed many loops to *FGF18*, a fibroblast growth factor involved in cell differentiation and cell-to-cell adhesion[25,26] (Figure 3A) and MRR2, an internal looping example which showed many loops to *IGF2*, an imprinted gene known to be associated with genomic silencers[27] and involved in growth, development and cancer[28] (Figure 4A).

We designed the CRISPR deletion site at a 1 kb region in MRR1 (termed “MRR1-A1”) located in the *FBXW11* intronic region that was associated with one of two Hi-C anchors that loop over to *FGF18* (Figure 3A). This region has high H3K27me3 as validated by ChIP-qPCR (Figure S3B). We performed 4C using MRR1-A1 as the viewpoint to detect all the regions that have chromatin interactions with this region in wild-type K562. The 4C-seq results showed that this region indeed had chromatin interactions with *FGF18* and several other genes such as *NPM1* and *UBTD2* (Figure 3A).

After that, we performed CRISPR deletion and generated three knock out (KO) clones (Figure S3C). Through aligning RNA-seq data of one KO clone and 4C-seq data (Figure 3A), we found that *FGF18* and *UBTD2* were both upregulated upon CRISPR deletion. This indicated that these two genes may be silenced by MRR1-A1. We further quantified the gene expression changes through RT-qPCR, which indicated that *FGF18* and *UBTD2* were indeed upregulated significantly after CRISPR deletion in the three KO clones while *FBXW11* gene expression was not affected (Figure 3B). To explore the phenotype of KO cells, we performed gene ontology (GO) analysis which showed that KO cells may undergo cell adhesion and cell differentiation (Figure 3C). We observed that the KO cells show increased adhesion to the surface of the cell culture plate while wild type cells are suspension cells (Figure 3D). Therefore, we performed adhesion assays using fibronectin-coated plates. We found that KO cells show greater adhesion to have more adhesion ability to fibronectin-coated plates (FN) than control cells (Figure 3E). This result suggested that deletion of MRR1-A1 leads to cell identity changes such as cell adhesion and cell differentiation.

To explore if these two target genes were controlled by H3K27me3, we checked the gene expression changes in the DMSO condition and EZH2 inhibition

by GSK343 condition. We found that *FGF18* and *UBTD2* significantly upregulated upon EZH2 inhibition which suggested that these two genes are controlled by MRRs which is consistent with CRISPR results (Figure 3F). To explore if MRR controlling *FGF18* expression is cell type specific, we called MRRs in seven cell lines and results showed that *FGF18* MRR are specific to two of the seven cell lines, K562 and GM12878 (Figure S3D). In summary, this example demonstrated we showed that MRR1-A1 can function as a silencer and can involve looping over to *FGF18* and *UBTD2* to repress gene expression which leads to cell identity changes (Figure 3G).

CRISPR excision of an MRR looping to *IGF2* leads to *IGF2* gene upregulation and altered chromatin interactions

Next, we designed another 1 kb deletion in MRR2 (termed "MRR2-A1") located in an intergenic region 10 kb away from the long non-coding RNA *H19* that was associated with one of three Hi-C anchors that loop over to *IGF2* (Figure 4A) following ChIP-qPCR confirmation of high H3K27me3 signal (Figure S4A).

4C-seq results showed that this region indeed loops to *IGF2* (Figure 4A). RT-qPCR in CRISPR KO cells (Figure S4B) and vector control cells showed that *IGF2* was upregulated in all three KO cells (Figure 4B). *H19* showed upregulation in the RNA-seq data, but this was not consistent across all three KO cells (Figure S4C), indicating that *H19* was not the target gene. *IGF2* expression significantly upregulated upon *EZH2* inhibition which suggested that *IGF2* expression are controlled by MRRs which is consistent with CRISPR results (Figure 4C). Gene ontology of the KO RNA-seq data showed changes in pathways such as cell communication and activation (Figure S4D). The *IGF2* MRR has cell line specificity and is specific to K562 out of seven cell lines examined (Figure S4E).

We asked if CRISPR deletion of a silencer could lead to changes in chromatin interactions. We performed 4C-seq using *IGF2* gene as the viewpoint in one of the KO clones and wild-type cells. We found that several chromatin interactions were lost in KO cells including both repressive interactions and active chromatin interactions (Figure 4D), indicating that CRISPR deletion of genomic regions can lead to alterations of chromatin interactions in the local vicinity. We speculate that MRR2 stabilized many chromatin interactions with other genomic regions (Figure 4A) that collectively formed a large interacting structure with *IGF2*, and deletion of the MRR may lead to destabilization of such chromatin interactions, potentially leading those genomic regions to form chromatin interactions with other regions besides *IGF2*, thus leading to changes of chromatin interactions to *IGF2* (Figure 4E). In summary, this example demonstrated that MRR2-A1 can function as a silencer and looping over to *IGF2* to repress gene expression. MRR2-A1 deletion leads to *IGF2* upregulation and changes of chromatin interactions at *IGF2* gene (Figure 4E).

To explore MRR functions *in vivo*, we injected KO cells into SCID (severe combined immunodeficiency) mouse models. Our results indicated that MRR1-A1 KO cells and MRR2-A1 KO cells both showed tumor growth inhibition (Figure 4F). While *IGF2* and *FGF18* do not have any known tumor suppressor role that has been reported, RNA-seq analysis of MRR1-A1 KO and MRR2-A1 KO cells both showed a number of downstream targets of *IGF2* and *FGF18* that were upregulated and have tumor suppressor roles, such as cell adhesion related gene *SPARC*[29] and estrogen receptor *ESRRB*[30]. Taken together, our results suggested that MRRs can function as tumor suppressors *in vivo*.

MRR-associated gene expression levels are highly susceptible to EZH2 perturbation

In order to investigate the effects of H3K27me3 on MRR-associated chromatin interactions and associated gene expression, we eliminated or reduced H3K27me3 by CRISPR knock-out of EZH2 in HAP1 cells (a near haploid Chronic Myeloid Leukemia derived cell line) and EZH2 inhibitor treatment in K562 cells.

In HAP1 EZH2 KO cells, both EZH2 and H3K27me3 were depleted (Figure 5A). To interrogate the gene expression changes of MRR-related genes, we performed RNA-seq in HAP1 WT and EZH2 KO cells. Our RNA-seq and RT-qPCR results indicated that up-regulation of H3K27me3-associated genes was prominent, while genes associated with H3K27ac peaks (super enhancers or typical enhancers) underwent minimal net change (Figure 5B; Figure S5A) in HAP1 EZH2-KO cells. Notably, MRR-associated genes were the most strongly upregulated as compared with other categories (typical H3K27me3, super-enhancer and typical enhancers) (Figure 5B). Similarly, EZH2 inhibition in K562 also induced H3K27me3 depletion and significant up-regulation of MRR-associated genes as compared with other categories (Figure S5B, S5C). This indicated that MRR-associated genes were highly susceptible to EZH2 inhibition compared to genes associated with typical H3K27me3.

HAP1 EZH2 KO cells showed slower growth rate and more cell adhesion compared with HAP1 WT cells (Figure S5D, S5E). The induced cell adhesion was in concordance with the significantly up-regulated cell adhesion related genes in RNA-seq of HAP1 and K562 cells (Figure S5F-S5L). We then examined the results of EZH2 inhibition on H3K27me3 landscape via ChIP-seq in HAP1 cells. H3K27me3 peaks could be barely found in HAP1 EZH2 KO cells, while the numbers of H3K4me3 peaks were comparable before and after EZH2 KO (Figure 5C).

Based the results of genes up-regulation and H3K27me3 depletion after EZH2 inhibition, we selected candidate MRR-associated genes to examine whether their interactions had been changed by EZH2 inhibition. Surprisingly, although *IGF2*, *FGF18* and *HOXD13* were up-regulated in HAP1 EZH2 KO cells, their promoter-related interaction profiles stayed highly similar to HAP1 WT cells (Figure 5D, 5E). Around these genes, H3K27me3 histone marks were depleted while H3K4me3 stayed comparable with the wild-type cells (Figure 5D; Figure S5M).

Similarly, chromatin interactions of these genes were largely unaffected in K562 after EZH2 inhibition (Figure S5N, S5O, and S5P). In addition, the reciprocal interactions between CRISPR excision sites and promoters of *FGF18* and *IGF2* remain largely unchanged after GSK343 treatment in K562 cells (Figure S5Q). We also designed 4C on other genes that are not associated with MRRs in HAP1 and K562 cells, all of which showed highly similar interaction profiles around the 4C baits at gene promoters or at distal interacting regions (Figure S5R and S5S). Taken together, these results showed that chromatin interactions of MRR-associated genes stayed highly similar despite the depletion of H3K27me3 and up-regulation of the genes.

Discussion

Here, we demonstrated that there are H3K27me3-rich regions of the genome (called "MRRs"), and they are associated with chromatin interactions. MRRs tend to preferentially associate with other MRRs. These highly repressive regions can control gene regulation by acting as silencers looping over to gene promoters. Interestingly, chromatin interactions do not change upon H3K27me3 alteration, and appear to act as "scaffolds" to guide histone modification loading and unloading onto the genome, and contact between H3K27me3 regions with target genes, indicating that H3K27me3 is not required for the maintenance of chromatin interactions in differentiated cells. However, chromatin interactions change upon alterations of DNA regions, such as excision of MRR genomic DNA. Our work addresses a few major questions in chromatin biology.

First, silencers are poorly characterized in the human genome and there are very few known examples that have been definitively shown to silence target genes, such as the human synapsin I gene[31] and the human *BDNF* gene[32]. Our work demonstrates that H3K27me3-rich regions can contain silencers in the human genome, and provides a method for the identification of such silencer regions. Our work also characterizes such potential silencer regions and indicates that silencers may be of functional significance. These silencers are important for our understanding of how genes are repressed and the consequences of such repression on cell biology, and for our understanding of the potential for EZH2 inhibitors in drug development.

A second major question when discussing chromatin interactions and gene activation or repression is the mechanism of functioning. A recent study on polycomb repressive complex 2 reported that the deposition and spreading of H3K27me3 marks are dependent on the contact of PRC2 nucleation sites[33]. Our results suggest that H3K27me3 depletion by EZH2 inhibition will lead to loss of H3K27me3 marks at chromatin interactions, but not necessarily alter interactions themselves. Gene expression changes would still be expected to occur because of the loss of H3K27me3 either in proximity or brought into close proximity by looping.

Our findings support a "stable looping" model which holds that looping events, once formed in cells upon cellular development and differentiation, are stable. These results are in concordance with other studies on transcription activation, that such as TNF-alpha responsive enhancers, which are already in contact with their target genes before the signaling[34].

A third question is whether H3K27me3 is required for maintenance of chromatin interactions. Depletion of *Eed* in embryonic stem cells reduces chromatin interactions between specific regions but does not lead to systemic change of topologically-associated domains[35]. Here, high-resolution characterization by 4C of the results of depletion of EZH2 on several different loci, some with very high levels of H3K27me3, do not show changes in chromatin interactions, suggesting that H3K27me3 is not required for maintenance of looping chromatin interactions. This could be a reflection of the cell types we studied (K562 and HAP1, which are differentiated cells) compared with embryonic stem cells, which may have different requirements for maintenance of chromatin interactions.

In conclusion, maintenance of cellular identity requires that the right genes are expressed and other genes are silenced. The results add an additional dimension to the epigenetic code by involving chromatin interactions as structural scaffolds for epigenetic marks, which could potentially assist in guiding the epigenetic marks to specific regions of the genome. H3K27me3-rich regions demonstrate strong cellular

control of particular genes and can act as silencers. Abrogation of H3K27me3-rich regions demonstrates that these regions are functional in the genome. Just as the concept of “super-enhancers” has been useful in identifying oncogenes and therapeutic vulnerabilities in cancer cells, the concept of H3K27me3-rich regions may be useful in identifying genes of key cellular identity and establishment of cancer potential in the future.

Methods

We performed Hi-C interaction analysis, ChIP-seq, RNA-seq, gene expression analyses, cell culture, RT-qPCR, CRISPR excision, 4C, 3C, xenograft models, western blot, adhesion assays, and growth curves as described in the **Supplementary Methods**. A list of all libraries used and generated is provided in **Supplementary Table 1**. A list of all the primers used is provided in **Supplementary Table 2**.

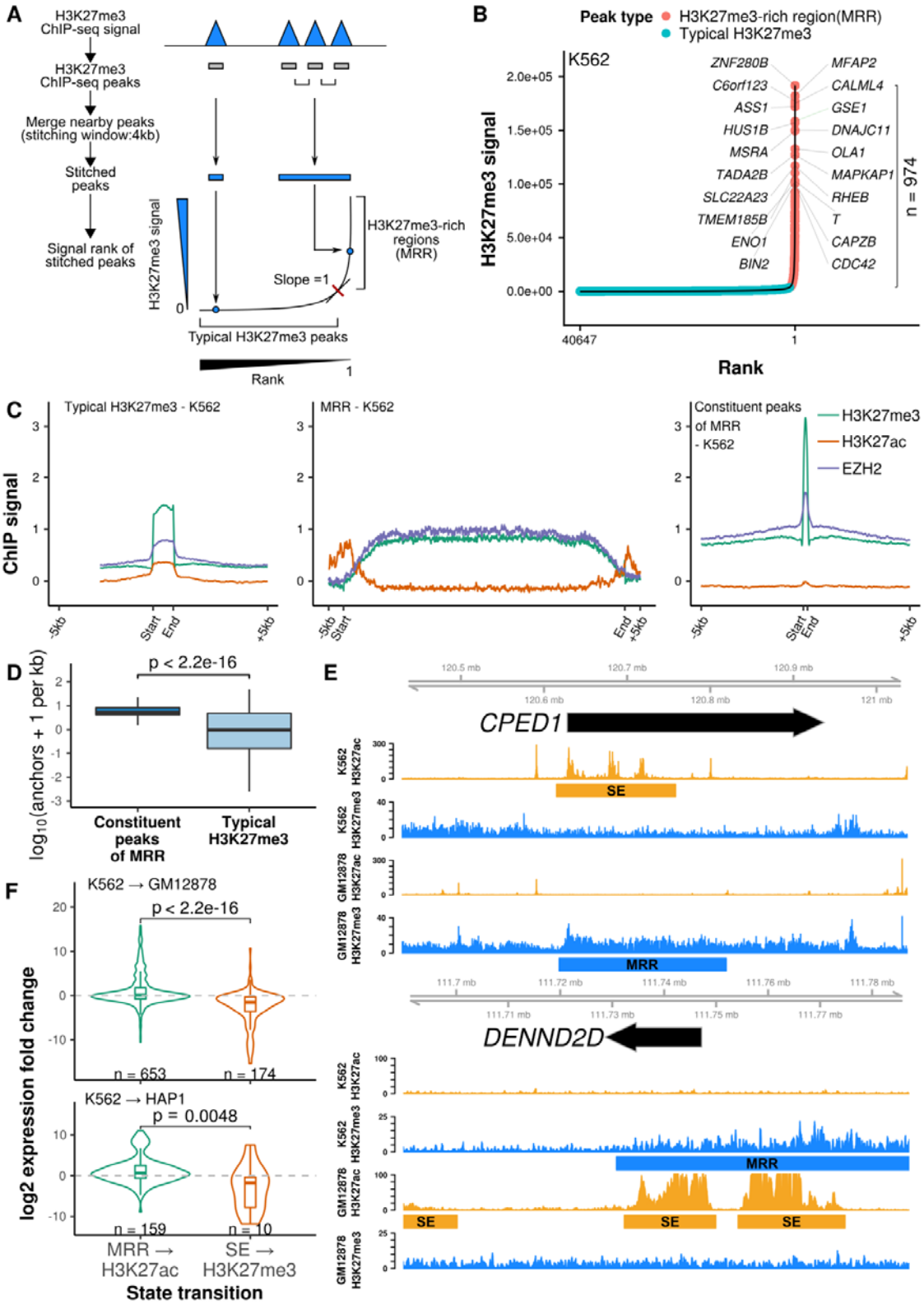

Figure 1. Definition of H3K27me3-rich regions (MRRs) and their characterization.

A. H3K27me3 ChIP-seq peaks within 4kb are stitched together and the stitched peaks ranked according to their H3K27me3 signal. The rank-ordered signal with a slope of 1 is used as cut-off for defining H3K27me3-rich MRRs. **B.** H3K27me3-rich regions (MRRs) and typical H3K27me3 peaks in K562 and their associated genes. A representative overlapping gene from each of the top 10 MRRs are shown. **C.** ChIP-seq signal on typical H3K27me3, MRR and constituent peaks of MRR regions in K562. Peaks are scaled to the same median length of peaks in typical H3K27me3 (1070 bp), MRR (92170 bp) or constituent peaks (203 bp), and the plot expanded by 5kb on both sides of the peak. **D.** Constituent peaks of MRRs have more Hi-C interactions compared to typical H3K27me3. Constituent peaks are peaks that form super peaks. “Super” and “typical” peaks are called by ROSE. Wilcoxon test *p* values are as indicated. **E.** Example of *CPED1* and *DENND2D* and their associated MRR/SE in different cell lines. MRR and SE could be interchangeable in different cell lines. SE, super enhancers; MRR, H3K27me3-rich regions. Expression level of *CPED*1 is 107.826 and 0.029 in K562 and GM12878, respectively; expression level of *DENND2D* is 0.002 and 78.004 (expression in RPKM). **F.** Expression changes associated with state transition between different cells. MRR→H3K27ac, gene associated peaks change from MRR into either super-enhancers or typical enhancers; SE→H3K27me3, gene associated peaks change from super-enhancers into either MRR or typical H3K27me3 peaks. K562→GM12878/K562→HAP1, direction of state transition between two indicated cell line. Expression data is from Epigenetic RoadMap and in-house HAP1 RNA-seq. Wilcoxon test *p* values are as indicated.

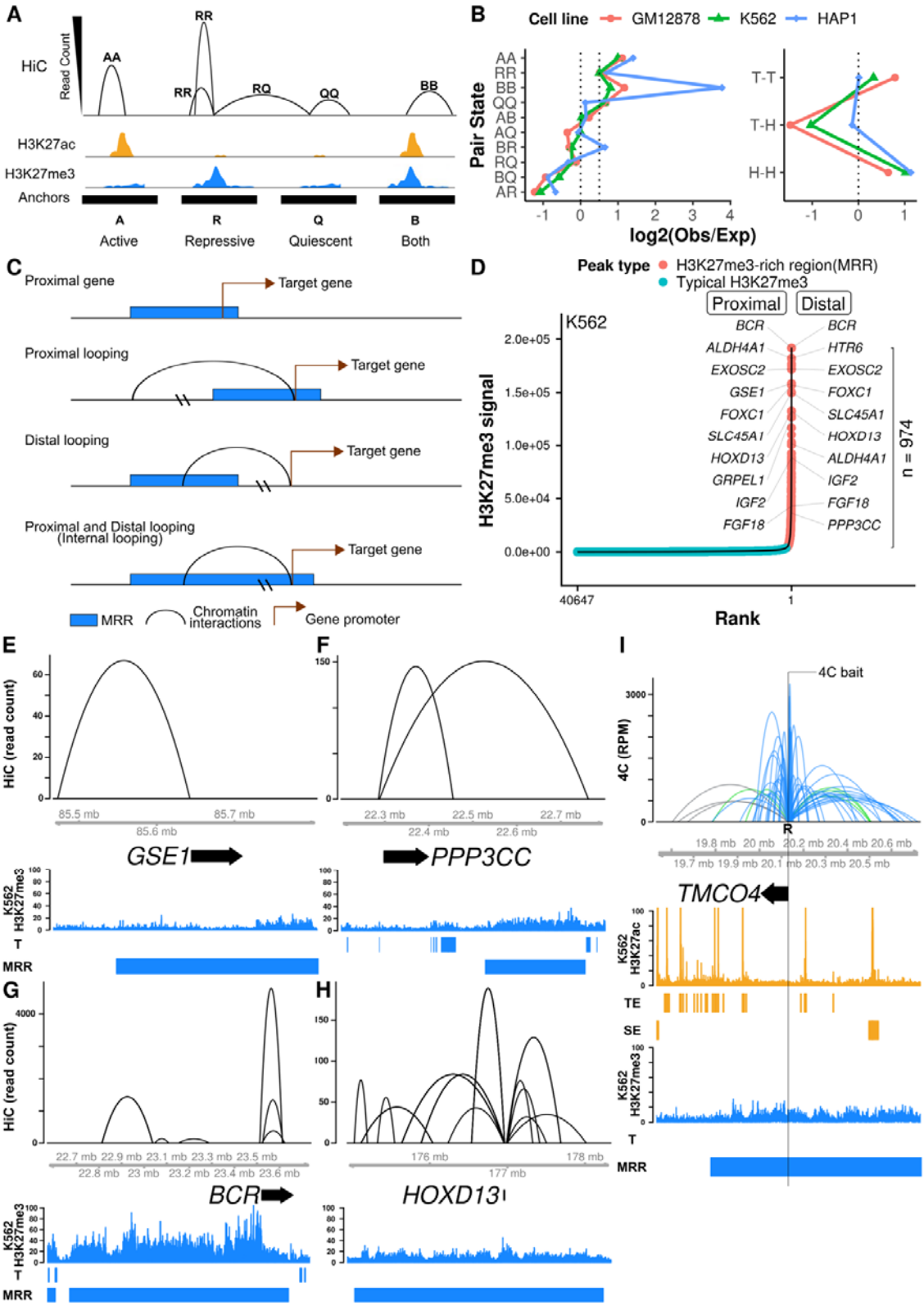

Figure 2. H3K27me3-rich regions (MRRs) preferentially associate with MRRs in the human genome via chromatin interactions.

A. Schematic plot of how different categories of Hi-C interactions are defined. Hi-C anchors are classified by whether they overlap with H3K27me3 or H3K27ac peaks. A (active), overlap with only H3K27ac peaks; R (repressive), overlap with only H3K27me3 peaks; Q (quiescent), overlap with neither H3K27ac nor H3K27me3 peaks; B (both), overlap with both H3K27ac and H3K27me3 peaks. The height of Hi-C interactions (arcs) represents the highest read counts in the interacting regions. **B.** Observed/expected ratio of Hi-C interactions in different categories. Left: categories of chromatin pair states. Right: T (typical H3K27me3) or H (MRR) peaks. The expected interactions are calculated from the marginal distributions of different anchors. **C.** Different categories of MRR associated with genes. **D.** H3K27me3-rich regions (MRRs) and typical H3K27me3 peaks in K562 and their associated genes through chromatin interactions. Peaks overlapping with Hi-C interactions are labeled with associated genes: for peaks labeled “proximal”, the gene TSS and peak occupy the same Hi-C anchor; “distal” peaks are connected to the gene via Hi-C interactions. **E.** Example of MRR (rank: 6) with proximal looping to gene *GSE1*. **F.** Example of MRR (rank: 85) with distal looping to gene *PPP3CC*. **G & H.** Examples of two MRR (rank: 1st and 10th) with internal looping in K562. **I.** Example of *TMCO4* 4C showing extensive internal looping within a MRR in K562. The colors of 4C interactions are based on the distal interacting regions to the 4C bait. Blue: repressive; orange: active; green: both; grey: quiescent. The state of the 4C bait is labeled by text. Each ChIP-seq tracks contains ChIP signal and peaks. TE, typical enhancer; SE, super-enhancer; T, typical H3K27me3; MRR, H3K27me3-rich region.

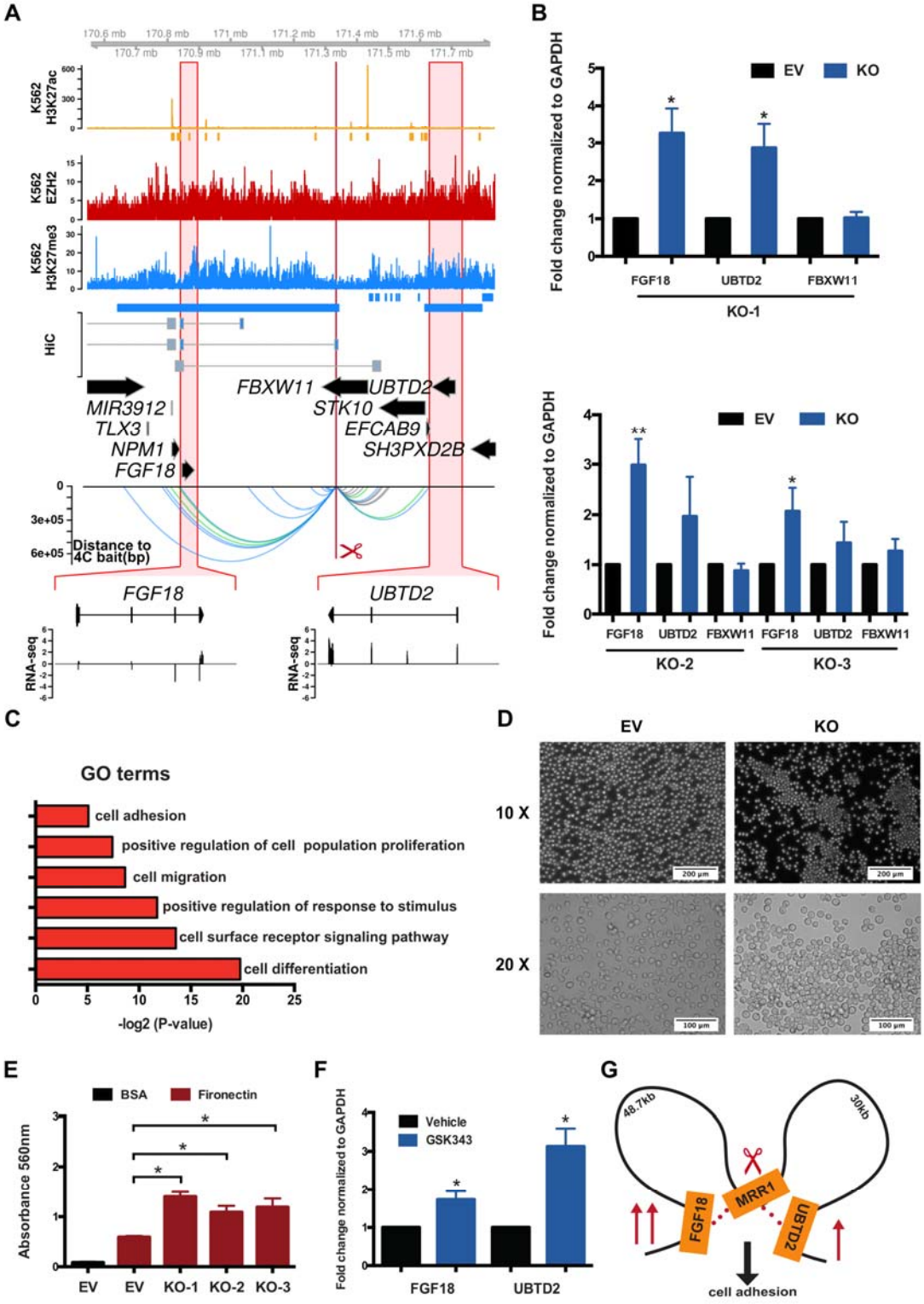

Figure 3. CRISPR characterization of an MRR in K562 cells at *FGF18*.
A. Screenshot showing EZH2 ChIP-seq, H3K27me3 ChIP-seq, H3K27ac ChIP-seq and chromatin interactions as identified from previously published Hi-C data[18], gene information, and 4C performed on the CRISPR-excised region in wild-type cells confirming chromatin interactions to *FGF18*, as well as showing chromatin interactions to *UBTD2*. The regions highlighted in the red boxes are shown in more detail, with RNA-seq was shown as one CRISPR knockout clone over wild-type at *FGF18* and *UBTD2*. The blue bar shows the predicted whole MRR. The red box with the red scissors indicates the region which was excised. **B.** RT-qPCR of *FGF18*, *UBTD2* and *FBXW11* in three different CRISPR-excised clones (KO-1, KO-2, KO-3) as compared with vector control cells ("Empty Vector", "EV"). The data shown is average + standard error. P value less than 0.05 is shown as *. P value less than 0.01 is shown as **. **C.** Gene Ontology (GO) was performed using significant DE genes in the RNA-seq data which was shown as –log2(p value). **D.** Light microscopy photos of empty vector (EV) and CRISPR knockout cells (KO) showing increased cell adhesion and clumping in the KO clones. 10X and 20X magnification were shown. **E.** A fibronectin adhesion assay showed increased adhesion of the three CRISPR knockout cells (KO) as compared with empty vector (EV). BSA was used as a negative control. The data shown is average + standard error. P value less than 0.05 is shown as *. **F.** RT-qPCR of *FGF18* and *UBTD2* expression upon DMSO treatment (vehicle) and GSK343 treatment in K562 cells. The data shown is average + standard error. P value less than 0.05 is shown as *. **G.** Cartoon schematic summary of the results showing that MRR1 interacts with *FGF18* and *UBTD2* in wild-type cells, and CRISPR excision (as indicated by the red scissor) leads to upregulation of *FGF18* and a modest upregulation of *UBTD2*. The cellular consequences of MRR1-A1 excision includes increased cell adhesion.

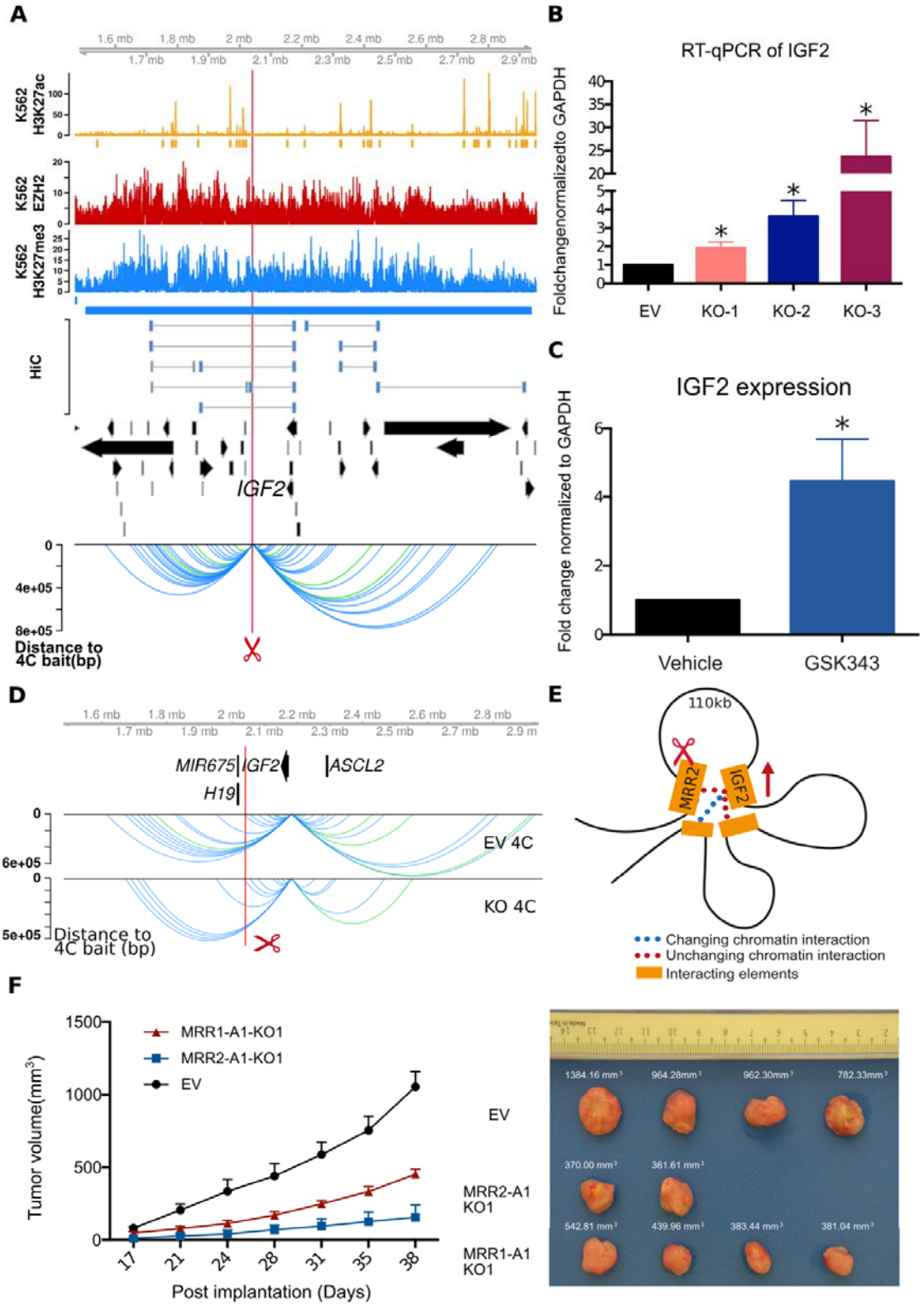

Figure 4. CRISPR characterization of an MRR in K562 cells at *IGF2*.
A. Screenshot showing EZH2 ChIP-seq, H3K27me3 ChIP-seq, H3K27ac ChIP-seq and chromatin interactions as identified from previously published Hi-C data[18], gene information, and 4C performed on the CRISPR-excised region in wild-type cells confirming chromatin interactions to *IGF2*. The blue bar shows the predicted whole MRR. The red box with the red scissor indicates the region which was excised. **B.** RT-qPCR of *IGF2* in three different CRISPR-excised clones (KO-1, KO-2, KO-3) as compared with vector control cells ("Empty Vector", "EV"). The data shown is average + standard error. P value less than 0.05 is shown as *. **C.** RT-qPCR of *IGF2* expression upon DMSO treatment (vehicle) and GSK343 treatment. The data shown is average + standard error. P value less than 0.05 is shown as *. **D.** Chromatin interactions landscape changes at the viewpoint of *IGF2* gene region in knock out cells (KO) compared to control cells. Red box with scissor highlights the CRISPR excised region. **E.** Cartoon schematic summary of the results showed that MRR2 interacts with *IGF2* in control cells, and CRISPR excision (as indicated by the red scissor) leads to upregulation of *IGF2* and chromatin interactions changes at *IGF2* region. Red dots show the unchanging chromatin interactions and blue dots show the changing interactions (either loss or gain) in KO cells. **F.** Tumor growth in SCID (Severe Combined Immunodeficiency) mice injected with MRR1-A1 knock out cells, MRR2-A1 knock cells and empty vector cells (EV), n=5 for each group. The left panel shows the tumor growth curve, and data shown as tumor volume (average + standard error) with different post implantation days. The right panel was tumor pictures of four injected mice in each group at day 38 which labelled with exact tumor volume. For mice injected with MRR2-A1 cells, only two grew tumors.

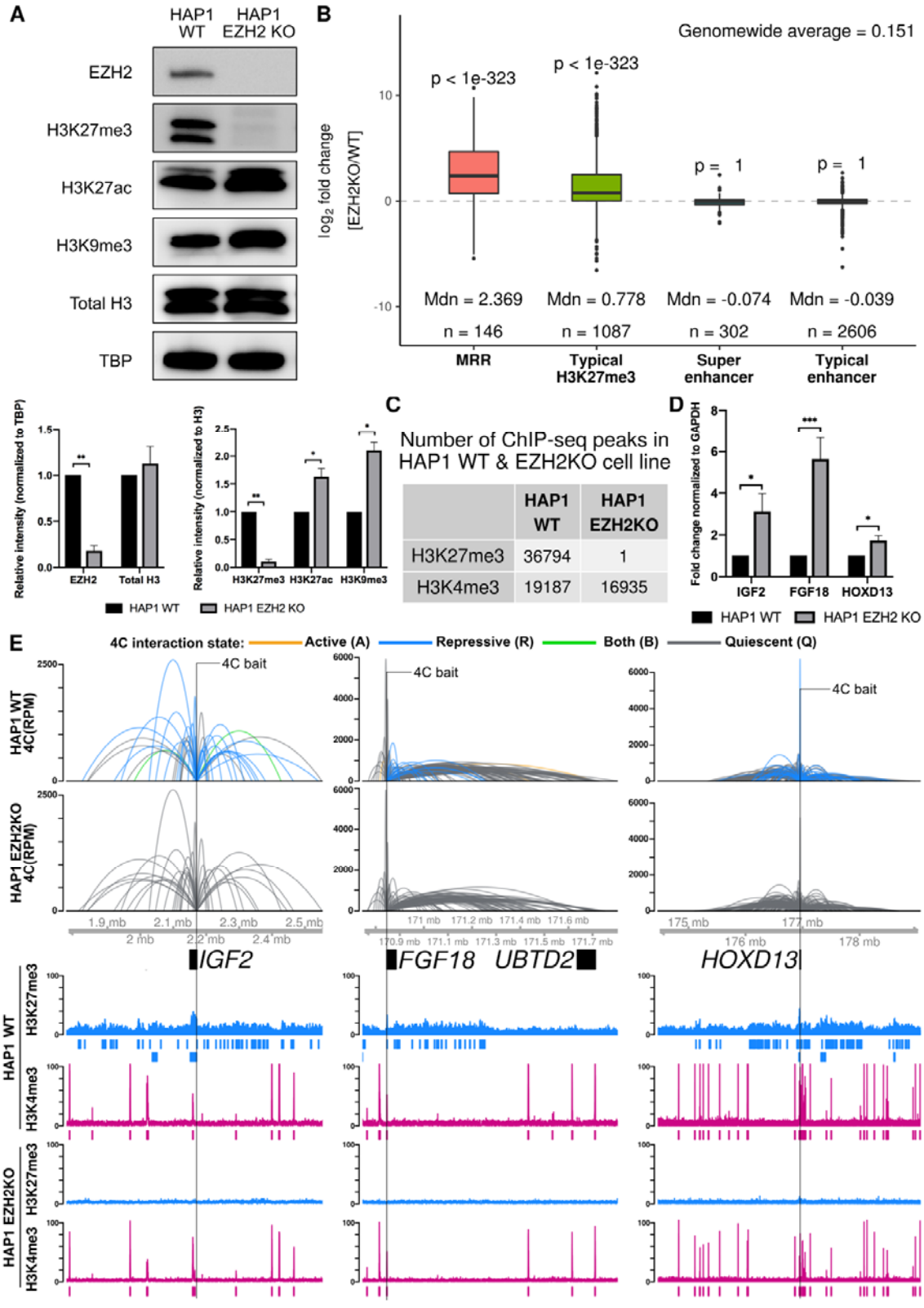

Figure 5. MRR-associated gene expression levels are highly susceptible to EZH2 perturbation.

A. Representative western blot of histone marks in HAP1 wild type (WT) & EZH2 knock-out (EZH2 KO) cells. Bottom panels showed the quantification of the western blots. **B.** Expression changes of genes associated with different types of peaks in HAP1 *EZH2* KO cells. Genes included: 1) Genes with TSS overlapped with different peaks; 2) Genes associated with different peaks through Hi-C interaction. One-tail wald test was used for testing significantly up-regulation. All the P values of genes in each category are aggregated. **C.** Number of ChIP-seq peaks found in HAP1 WT and EZH2 KO cells. **D.** RT-qPCR of *IGF2*, *FGF18* and *HOXD13* in HAP1 WT and EZH2 KO cells. **E.** 4C results of *IGF2*, *FGF18* and *HOXD13* in HAP1 WT and EZH2 KO cells. The colors of 4C interactions are based on the distal interacting regions to the 4C bait. Blue: repressive; orange: active; green: both; grey: quiescent. The height of the 4C is shown in Reads Per Million (RPM). The ChIP signal and peaks of H3K27me3 and H3K4me3 are shown. H3K27me3 ChIP-seq peaks in HAP1 WT contains typical H3K27me3 (top peak track) and MRR (bottom peak track) peaks.

Acknowledgements
We would like to thank all members of the Fullwood Lab and Ah Jung Jeon for helpful comments. This research is supported by the National Research Foundation (NRF) Singapore through an NRF Fellowship awarded to M.J.F (NRF-NRFF2012-054) and NTU start-up funds awarded to M.J.F. This research is supported by the RNA Biology Center at the Cancer Science Institute of Singapore, NUS, as part of funding under the Singapore Ministry of Education Academic Research Fund Tier 3 awarded to Daniel Tenen (MOE2014-T3-1-006). This research is supported by an Singapore MOE Academic Research Research Fund (T1) grant to G.T-K. This research is supported by the National Research Foundation Singapore and the Singapore Ministry of Education under its Research Centres of Excellence initiative.

Author contributions
C.Y.C., Z.Y., L.Y.P., T.J.Q., M.J.F. and G.T-K. conceived of the research. C.Y.C., Z.Y., L.Y.P., T.J.Q., M.J.F. and G.T-K. contributed to the study design. C.Y.C. performed bioinformatics analysis. Z.Y. and L.S. designed CRISPR knock out experiments. Z.Y. performed CRISPR knock out, 4C, RNA-seq, ChIP-qPCR and other functional experiments. L.Y.P. performed *EZH2* inhibitor and HAP1 *EZH2* knockout experiments and 4C experiments. T.J.Q. performed ChIP-seq and ChIP-qPCR experiments. C.Z. and L.M.Q. performed 4C experiments. R.A., L.M. and T.V. designed xenograft experiments. R.A. performed xenograft experiments. C.Y.C., Z.Y., M.J.F. and G.T-K. reviewed the data and wrote the manuscript. All authors reviewed and approved of the manuscript.

Data deposition
The list of libraries used in the study are detailed in Supplementary Table S1. All datasets have been deposited into GEO.

Author information
The authors declare that they have no competing interests.
Correspondence and requests for materials should be addressed to mfullwood@ntu.edu.sg and dbsgtk@nus.edu.sg.

References
1 Bradner, J. E., Hnisz, D. & Young, R. A. Transcriptional Addiction in Cancer. *Cell* **168**, 629-643, doi:10.1016/j.cell.2016.12.013 (2017).
2 Babu, D. & Fullwood, M. J. 3D genome organization in health and disease: emerging opportunities in cancer translational medicine. *Nucleus* **6**, 382-393, doi:10.1080/19491034.2015.1106676 (2015).
3 Kolovos, P., Knoch, T. A., Grosveld, F. G., Cook, P. R. & Papantonis, A. Enhancers and silencers: an integrated and simple model for their function. *Epigenetics Chromatin* **5**, 1, doi:10.1186/1756-8935-5-1 (2012).
4 Mifsud, B. *et al.* Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nat Genet* **47**, 598-606, doi:10.1038/ng.3286 (2015).
5 Trojer, P. & Reinberg, D. Facultative heterochromatin: is there a distinctive molecular signature? *Mol Cell* **28**, 1-13, doi:10.1016/j.molcel.2007.09.011 (2007).

6 Schouten, M. J. & Bruinvels, J. Endogenously formed norharman (beta-carboline) in platelet rich plasma obtained from porphyric rats. *Pharmacol Biochem Behav* **24**, 1219-1223 (1986).
7 Morin, R. D. *et al.* Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nat Genet* **42**, 181-185, doi:10.1038/ng.518 (2010).
8 Gulati, N., Beguelin, W. & Giulino-Roth, L. Enhancer of zeste homolog 2 (EZH2) inhibitors. *Leuk Lymphoma* **59**, 1574-1585, doi:10.1080/10428194.2018.1430795 (2018).
9 Whyte, W. A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307-319, doi:10.1016/j.cell.2013.03.035 (2013).
10 Loven, J. *et al.* Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **153**, 320-334, doi:10.1016/j.cell.2013.03.036 (2013).
11 Hnisz, D. *et al.* Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934-947, doi:10.1016/j.cell.2013.09.053 (2013).
12 Chen, K. *et al.* Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes. *Nat Genet* **47**, 1149-1157, doi:10.1038/ng.3385 (2015).
13 Cao, F. *et al.* Super-Enhancers and Broad H3K4me3 Domains Form Complex Gene Regulatory Circuits Involving Chromatin Interactions. *Sci Rep* **7**, 2186, doi:10.1038/s41598-017-02257-3 (2017).
14 Huang, J. *et al.* Dissecting super-enhancer hierarchy based on chromatin interactions. *Nat Commun* **9**, 943, doi:10.1038/s41467-018-03279-9 (2018).
15 Sharifnia, T. *et al.* Small-molecule targeting of brachyury transcription factor addiction in chordoma. *Nat Med* **25**, 292-300, doi:10.1038/s41591-018-0312-3 (2019).
16 Consortium, E. P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74, doi:10.1038/nature11247 (2012).
17 Davoli, T. *et al.* Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome. *Cell* **155**, 948-962, doi:10.1016/j.cell.2013.10.011 (2013).
18 Kunchala, P., Kuravi, S., Jensen, R., McGuirk, J. & Balusu, R. When the good go bad: Mutant NPM1 in acute myeloid leukemia. *Blood Rev* **32**, 167-183, doi:10.1016/j.blre.2017.11.001 (2018).
19 Ziai, J. M., Siddon, A. J., Education Committee of the Academy of Clinical Laboratory, P. & Scientists. Pathology Consultation on Gene Mutations in Acute Myeloid Leukemia. *Am J Clin Pathol* **144**, 539-554, doi:10.1309/AJCP77ZFPUQGYGWY (2015).
20 Sportoletti, P. *et al.* Npm1 is a haploinsufficient suppressor of myeloid and lymphoid malignancies in the mouse. *Blood* **111**, 3859-3862, doi:10.1182/blood-2007-06-098251 (2008).
21 Hirsch, S. *et al.* Circular RNAs of the nucleophosmin (NPM1) gene in acute myeloid leukemia. *Haematologica* **102**, 2039-2047, doi:10.3324/haematol.2017.172866 (2017).
22 Messina, M. *et al.* Genetic lesions associated with chronic lymphocytic leukemia chemo-refractoriness. *Blood* **123**, 2378-2388, doi:10.1182/blood-2013-10-534271 (2014).
23 de Bock, C. E. *et al.* The Fat1 cadherin is overexpressed and an independent prognostic factor for survival in paired diagnosis-relapse samples of precursor

B-cell acute lymphoblastic leukemia. *Leukemia* **26**, 918-926, doi:10.1038/leu.2011.319 (2012).

24 Rao, S. S. *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665-1680, doi:10.1016/j.cell.2014.11.021 (2014).

25 Shimokawa, T. *et al.* Involvement of the FGF18 gene in colorectal carcinogenesis, as a novel downstream target of the beta-catenin/T-cell factor complex. *Cancer Res* **63**, 6116-6120 (2003).

26 Jeon, E. *et al.* Investigating the role of FGF18 in the cultivation and osteogenic differentiation of mesenchymal stem cells. *PLoS One* **7**, e43982, doi:10.1371/journal.pone.0043982 (2012).

27 Constancia, M. *et al.* Placental-specific IGF-II is a major modulator of placental and fetal growth. *Nature* **417**, 945-948, doi:10.1038/nature00819 (2002).

28 Ravenel, J. D. *et al.* Loss of imprinting of insulin-like growth factor-II (IGF2) gene in distinguishing specific biologic subtypes of Wilms tumor. *J Natl Cancer Inst* **93**, 1698-1703, doi:10.1093/jnci/93.22.1698 (2001).

29 Bhoopathi, P. *et al.* SPARC overexpression inhibits cell proliferation in neuroblastoma and is partly mediated by tumor suppressor protein PTEN and AKT. *PLoS One* **7**, e36093, doi:10.1371/journal.pone.0036093 (2012).

30 Madhu Krishna, B. *et al.* Estrogen receptor alpha dependent regulation of estrogen related receptor beta and its role in cell cycle in breast cancer. *BMC Cancer* **18**, 607, doi:10.1186/s12885-018-4528-x (2018).

31 Li, L., Suzuki, T., Mori, N. & Greengard, P. Identification of a functional silencer element involved in neuron-specific expression of the synapsin I gene. *Proc Natl Acad Sci U S A* **90**, 1460-1464, doi:10.1073/pnas.90.4.1460 (1993).

32 Zuccato, C. *et al.* Widespread disruption of repressor element-1 silencing transcription factor/neuron-restrictive silencer factor occupancy at its target genes in Huntington's disease. *J Neurosci* **27**, 6972-6983, doi:10.1523/JNEUROSCI.4278-06.2007 (2007).

33 Oksuz, O. *et al.* Capturing the Onset of PRC2-Mediated Repressive Domain Formation. *Mol Cell* **70**, 1149-1162 e1145, doi:10.1016/j.molcel.2018.05.023 (2018).

34 Jin, F. *et al.* A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* **503**, 290-294, doi:10.1038/nature12644 (2013).

35 Denholtz, M. *et al.* Long-range chromatin contacts in embryonic stem cells reveal a role for pluripotency factors and polycomb proteins in genome organization. *Cell Stem Cell* **13**, 602-616, doi:10.1016/j.stem.2013.08.013 (2013).