CRISPR-ChIP delineates a Menin-dependent oncogenic DOT1L complex in MLL-leukaemia

Omer Gilan[1,2,3*], Charles C. Bell[1,2], Laure Talarmain[1,2], Daniel Neville[3], Kathy Knezevic[1], Daniel Ferguson[3], Marion Boudes[4], Yih-Chih Chan[1,2], Chen Davidovich[4], Enid Y.N. Lam[1,2], Mark A. Dawson[1,2,5, 6*]

[1]Cancer Research Division, Peter MacCallum Cancer Centre, Melbourne, Victoria 3000, Australia
[2]Sir Peter MacCallum Department of Oncology, University of Melbourne, Victoria 3052, Australia
[3]Australian Centre for Blood Diseases, Monash University, Melbourne, Victoria, 3004, Australia
[4]Department of Biochemistry and Molecular Biology, Biomedicine Discovery Institute, Monash University, Clayton, Victoria, 3168, Australia
[5]Department of Clinical Haematology, Peter MacCallum Cancer Centre & Royal Melbourne Hospital, Melbourne, Victoria 3000, Australia
[6]Centre for Cancer Research, University of Melbourne, Melbourne, Victoria 3000, Australia

*Corresponding authors.

Professor Mark A. Dawson
Peter MacCallum Cancer Centre,
305 Grattan Street, Melbourne, VIC 3000, Australia
Email: mark.dawson@petermac.org

Dr Omer Gilan
Australian Centre for Blood Diseases, Monash University,
99 Commercial Rd, Melbourne, VIC 3004, Australia
Email: omer.gilan@monash.edu

34 **Summary**

35 The regulation of all chromatin-templated processes involves the selective recruitment of
36 chromatin factors to facilitate DNA repair, replication, and transcription. Chromatin
37 immunoprecipitation (ChIP) is a critical experimental method used to provide spatiotemporal
38 evidence for the coordination of these chromatin-based events including the dynamic
39 regulation of chromatin modifications at cis-regulatory elements. However, obtaining a global
40 appreciation of all the factors that influence a specific chromatin event has remained
41 challenging. Here, as a proof of concept we demonstrate the utility of coupling unbiased
42 functional genomics with ChIP to identify the factors associated with active transcription.
43 Specifically, we use this method to identify the major chromatin factors associated with the
44 catalysis of two evolutionarily conserved histone modifications; H3K4me3 present at the
45 transcriptional start site and H3K79me2 present through the gene body of actively transcribed
46 genes. With CRISPR-ChIP, we identify all the non-redundant COMPASS complex members
47 required for H3K4me3 and demonstrate that RNA polymerase II is dispensable for the
48 maintenance of H3K4me3. As H3K79me2 has a putative oncogenic function in leukaemia cells
49 driven by MLL-translocations, using CRISPR-ChIP we reveal a functional partitioning of
50 H3K79 methylation into two distinct regulatory units. An oncogenic DOT1L complex, where
51 the malignant driver directs the catalytic activity of DOT1L at MLL-Fusion target genes and a
52 separate endogenous DOT1L complex, where catalytic activity is directed by MLLT10. This
53 functional demarcation provides an explanation for the observed synergy with Menin and
54 DOT1L inhibitors and why Menin inhibition surprisingly controls methylation of H3K79 at a
55 critical subset of genes that sustain MLL-fusion leukaemia. Overall, CRISPR-ChIP provides a
56 powerful tool for the unbiased interrogation of the mechanisms underpinning chromatin
57 regulation.

58
59
60
61
62
63
64
65
66
67
68

Introduction

The transcriptional output of every gene is controlled by the cooperative function of chromatin modifiers, transcription factors and transcriptional co-activators. Our ability to unravel this complexity has emerged through a combination of biochemical approaches integrated with genomic assays such as ChIP-seq and RNA-seq[1]. However, the genomic assays are limited by the fact that they provide information for a single perturbation at a time. Similarly, as chromatin regulators often function within multi-subunit complexes, the biochemical approaches in isolation, are unable to delineate the precise complex components required for a chromatin-based event, such as a chromatin modification. The opportunity to simultaneously assess the contribution of thousands of proteins in the regulation of a defined chromatin-based event in an unbiased high-throughput manner may provide unique insights and accelerate our understanding of chromatin regulation.

To address this need we developed a ChIP-based CRISPR screen (CRISPR-ChIP) which enables the concurrent assessment of thousands of nuclear proteins in the coordinated regulation of chromatin. As a proof of principle, we chose to study the regulation of two histone modifications associated with active transcription. H3K4me3 is catalysed by the COMPASS family of methyltransferases and is primarily located over the transcriptional start site (TSS) of actively transcribed genes. In contrast, H3K79me2 is deposited by DOT1L in a non-redundant manner and is predominantly localised throughout the body of the transcribed locus. Notably, H3K79me2 is one of the few histone modifications directly implicated in cancer, specifically MLL-Fusion leukaemias, where several targeted therapies have been specifically developed and clinically translated to inhibit DOT1L and reduce H3K79methylation. Therefore, we sought to focus on this histone modification and study its regulation in MLL-Fusion leukaemia cells.

Using CRISPR-ChIP, which can couple regulation of chromatin events with a concomitant readout of cell viability, we reveal an unexpected discordance between the functional regulators of H3K79methylation and their essentiality in MLL leukaemia cells. Mechanistically, we demonstrate that underpinning this observation is a functional separation of DOT1L into two complexes in MLL leukaemia cells (i) an endogenous complex containing MLLT10 which deposits H3K79me2 at the majority of expressed genes and (ii) a neomorphic malignant DOT1L complex driven by the MLL-fusion oncogene which localises DOT1L to potentiate gene expression at a subset of target genes required for the maintenance of leukaemia. Additionally, our data provide detailed molecular insights into how Menin inhibitors unexpectedly and selectively influence H3K79methylation at a subset of genes and specify the mechanistic rationale for the synergistic efficacy of Menin and DOT1L inhibition in this disease.

Results
Establishing the CRISPR-ChIP method

Functional genomics using CRISPR screens have provided an unprecedented opportunity to uncover new biological insights in an unbiased manner. Thus far, the majority of CRISPR screening approaches have been restricted to positive and negative selection screens using cell

113   viability as a read out[2]. More recently, these have expanded to FACS-based screens, which
114   enable the selective enrichment of cells expressing a fluorescent protein or cell-surface target
115   that is linked to a specific biological process[2]. As chromatin-based events cannot be easily
116   assessed in this manner we sought to develop CRISPR-ChIP, which directly couples unbiased
117   functional genomics with chromatin immunoprecipitation.

118

119   ChIP involves fragmentation of linker-DNA, which separates neighbouring nucleosomes,
120   using sonication or enzymatic digestion[3]. Whilst micrococcal nuclease can provide mono-
121   nucleosome resolution, methods that involve sonication of crosslinked chromatin usually result
122   in di- and tri-nucleosome fragments[3]. Therefore, a key consideration in developing the method
123   was the need to ensure that the integrated guide RNA (sgRNA) sequence remained coupled to
124   the chromatin template that is enriched in a ChIP assay. To address this technical requirement,
125   we chose to link the sgRNA 50-100bp upstream of the regulatory element being assessed in
126   the ChIP assay (**Fig. 1a**).

127

128   To optimise our method, it was important to determine whether the lentivirally integrated
129   regulatory sequences were modified and regulated in a similar manner to the corresponding
130   endogenous loci. Following separate lentiviral infection, we sorted BFP+ cells expressing
131   integrated regulatory sequences representing (i) a strong promoter (*EF1a* promoter), (ii) a weak
132   promoter (*PGK* promoter), (iii) a developmental promoter (*Meis1* promoter) and (iv) the *Myc*
133   enhancer[4]. We then performed ChIP-qPCR, to measure the enrichment of H3K27ac,
134   H3K4me3, H3K9ac, H3K79me2, and Pol II at both the integrated regulatory sequence and the
135   corresponding endogenous locus in human cells. Overall, our results showed that the
136   chromatinised environment surrounding the virally integrated regulatory element undergoes a
137   similar level of histone modification compared to the endogenous locus (**Extended Data Fig.**
138   **1a-d**). Next, we wanted to address if the histone modifications at our integrated reporter were
139   regulated in a similar manner to the endogenous loci. As we were interested in assessing histone
140   modifications associated with active transcription, we chose to use the full (~1kb) EF1a
141   promoter and assess H3K27ac and H3K79me2 as the enzymes that deposit these modifications
142   are amenable to inhibition with potent and selective small molecules[5,6]. This allowed us to
143   assess the changes in the level of these histone modifications around our integrated promoter
144   sequence and the corresponding endogenous locus. We treated cells expressing the integrated
145   EF1a promoter sequence with a DOT1L inhibitor (SGC0946) or a P300/CBP inhibitor (A-485)
146   and repeated our ChIP-qPCR analysis. Our data confirmed that treatment of the cells with
147   SGC0946 and A-485 resulted in a marked decrease in H3K79me2 and H3K27ac respectively,
148   at both the integrated promoter sequence and at endogenous genes marked by the histone
149   modifications (**Extended Data Fig. 1e-f**). Together these results suggested that our integrated
150   promoter sequence was subjected to the same epigenetic regulatory processes active at
151   endogenously expressed genes.

152

153   Having established the suitability of the integrated promoter sequence to assess changes in
154   histone modifications associated with gene expression we next sought to validate the ability to
155   discern the depletion of sgRNAs against nuclear proteins that regulate a histone modification
156   of interest. Cognisant of the fact that many chromatin regulators are also common-essential

157   proteins, we wanted to understand if we were able to capture major regulators of a histone
158   modification that were also required for the viability of the cell. To achieve this, we mixed two
159   DOT1L guides with two non-targeting guides and introduced these into leukaemia cells
160   sustained by MLL-AF9 (**Fig. 1b**). We then harvested the cells at both day 2 and day 10 after
161   sgRNA infection. Since H3K79me2 has no known demethylase and we had previously shown
162   that this modification is slowly turned over[7], we waited until day 10 to perform ChIP for
163   H3K79me2 to allow for marked decrease in H3K79me2 (**Fig. 1b**). To assess the effects on cell
164   viability, we determined the relative guide abundance of the guides at day 10 relative to day 2.
165   As expected, we found that DOT1L guides were depleted in MLL-Fusion leukaemia cells at
166   day 10 (**Fig. 1c and Extended Data Fig. 1g**). To determine the impact on H3K79me2, we
167   performed a ChIP assay at day 10 in viable cells and compared the representation of the
168   sgRNAs from the ChIP sample relative to the input sample at day 10. Here, we see a marked
169   depletion in DOT1L sgRNAs from the H3K79me2 ChIP material, suggesting that in principle,
170   the CRISPR-ChIP method can reliably identify a known epigenetic regulator of a histone
171   modification even when the regulator is required for cell viability (**Fig. 1c and Extended Data**
172   **Fig. 1g**).

173

174   Another important consideration in developing the method to be suitable for a high-throughput
175   screen with a library of thousands of sgRNAs is that the absolute amount of DNA that is
176   immunoprecipitated in a typical ChIP assay can be relatively low (~0.1-10% of the input DNA
177   for the locus of interest). Therefore, unlike conventional CRISPR screening approaches which
178   harvest the entire cellular genomic DNA, the lower yield in this method is a critical parameter
179   that could directly influence the representation of guides in the library. To establish the
180   technical feasibility of the method, we developed a guide RNA library targeting 1144
181   chromatin and transcriptional regulators (each gene was targeted by 6 guides each)
182   (**Supplementary Table 1**). In our workflow (**Fig. 1d**), the lentiviral sgRNA library is infected
183   into cells expressing Cas9; importantly, we aim for a low multiplicity of infection (MOI ~0.3)
184   to ensure that only a single sgRNA is integrated per cell. Therefore, we are only assessing the
185   contribution of genetic depletion of a single chromatin regulator per cell. Expression of BFP
186   from the integrated reporter enables FACS sorting of cells expressing the reporter construct
187   and sgRNA. The viable BFP+ cells are then harvested at the desired timepoint and from the
188   same population of cells, we apportion one sample to provide us with the input library, while
189   the remaining cells are cross-linked, sonicated and taken forward in a standard ChIP assay for
190   the target of interest (**Fig. 1d**). Both input and ChIP samples are then assessed by next-
191   generation sequencing (NGS) for the depletion or enrichment of guides (using MAGECK) in
192   the ChIP sample relative to input.

193

194   To assess the cell number required to reliably retain library representation, we performed
195   CRISPR-ChIP for H3K27ac from Cas9 negative K562 cells. Since no gene editing occurs, we
196   could determine the extent to which library representation is altered in the ChIP sample
197   (relative to the input library) from different starting cell numbers. Our results revealed that
198   library representation following lentiviral infection was very highly correlated between two
199   independent replicates of the input library (**Fig. 1e**). Importantly, we also found that the two
200   independent replicates of H3K27ac ChIP were highly correlated (**Fig. 1e**). Notably, an

201 excellent correlation in the representation of the sgRNA library between the ChIP and input
202 samples was faithfully maintained from 100 million cells (r=0.985) down to 5 million cells
203 (r=0.958) (**Fig. 1e and Extended Data Fig. 2a**). As ChIP efficiency is dependent on the quality
204 of the antibody and abundance of the target of interest, we wanted to use our data to model the
205 point at which technical limitations in the ChIP could result in substantial skewing in library
206 representation. To address this, we performed *in silico* down sampling of the H3K27ac ChIP
207 sample and plotted the guide representation density. This analysis revealed that substantial
208 skewing only begins to occur when the ChIP sample contains less than 50-fold representation
209 of each sgRNA within the library (**Fig. 1f**).
210
211 The change in abundance of histone modifications is not binary and instead spans a continuous
212 spectrum, we therefore wanted to understand whether CRISPR-ChIP has sufficient sensitivity
213 to capture a quantitative change in the levels of the modification. To achieve this, we infected
214 separate Cas9 negative K562 cells with either a human or mouse guide library and treated the
215 cells harbouring the human library with a DOT1L inhibitor (SGC0946) for up to 96hrs. In
216 parallel, we treated the cells containing the mouse library with vehicle (DMSO). The mouse
217 and human K562 cells were then mixed at 1:1 ratio and the relative guide enrichment from the
218 ChIP for H3K79me2 were determined at various timepoints. This approach allowed us to
219 compare the enrichment of the CRISPR library from SGC0946 treated cells where there is a
220 gradual and progressive reduction in H3K79me2 (**Fig. 1g**) compared to vehicle treated cells
221 where the modification stays constant. Our results showed that the CRISPR-ChIP for
222 H3K79me2 in the treated/vehicle (human/mouse) mix accurately reflected the input guide
223 distribution and enrichment (**Fig. 1h**). Furthermore, we clearly observed a quantitative time-
224 dependent reduction in the representation of the human library with our ChIP for H3K79me2
225 following SGC0946 treatment relative to the vehicle treated mouse library (**Fig. 1h**). Notably,
226 these data closely parallel the change in abundance noted by western blot (**Fig. 1i**). Together
227 these data serve to highlight the fact that CRISPR-ChIP can detect a quantitative alteration in
228 the level of a chromatin modification. Finally, we also performed MAGECK analysis on the
229 CRISPR-ChIP data from Cas9 negative cells to determine the expected background scores
230 generated by the method when no editing occurs (**Extended Data Fig. 2b**). This control
231 analysis revealed the low probability of false positives arising from CRISPR-ChIP providing
232 further confidence that our method was robust.
233

CRISPR-ChIP identifies the major non-redundant regulators of H3K4me3.
235 The tri-methylation of lysine 4 on histone H3 (H3K4me3) is one of the best characterised
236 histone modifications, as it is invariably localised at the transcription start site of actively
237 transcribed genes in all eukaryotes and its abundance is directly correlated with transcriptional
238 activity at the locus[8,9]. The yeast model organism *Saccharomyces cerevisiae* has a single
239 methyltransferase, SET1, that deposits this modification. However, in mammals there are six
240 H3K4 methyltransferases (SET1A/B and MLL1-4), which have all been shown to strongly
241 associate with the initiated form of RNA polymerase II (Pol II)[10,11]. Of these, SET1A/B are
242 thought to be the methyltransferases responsible for the majority of H3K4me3[12], whereas,
243 MLL1/2 deposit H3K4me3 at key developmental genes[13]. The methyltransferases in isolation
244 are not efficient enzymes at depositing H3K4me3, instead their catalytic activity is almost

entirely dependent on their incorporation into a multi-subunit protein complex called COMPASS (containing either SET1A/B) or COMPASS-like (containing one of MLL1-4). The various subunits of the COMPASS and COMPASS-like complexes are essential for the stability and functional integrity of the methyltransferases.

To uncover the non-redundant functional regulators of H3K4me3, the prototypical histone modification associated with active transcription, we performed CRISPR-ChIP for H3K4me3 in K562 cells (**Fig. 2a and Supplementary Table 2**). The top hits in our screen were the four proteins commonly referred to as WRAD (**Fig. 2b**), including WD repeat-containing protein 5 (WDR5), Retinoblastoma Binding Protein 5 (RBBP5), Absent-Small-Homeiotic-2-Like protein (ASH2L) and Dumpy-30 (DPY30)[14]. The four WRAD components are fundamental core proteins that comprise all COMPASS and COMPASS-like complexes and are known to be required for the functional integrity of COMPASS and COMPASS-like complexes, which deposit and maintain of H3K4me3[13,14]. In line with the results of the CRISPR-ChIP screen, we subsequently validated that CRISPR-Cas9 mediated deletion of ASH2L or RBBP5 resulted in a global loss of H3K4me3 by western blot and ChIP-seq (**Fig. 2c-e and Extended Data Fig. 2c**). Consistent with the fact that there is significant redundancy between the catalytic orthologs SETD1A/B, MLL1/2 and MLL3/4, none of these H3K4 methyltransferase enzymes were identified in our screen. However, the screen did identify CXXC1 and WDR82 that are specific subunits of the mammalian SETD1A/B COMPASS complex (**Fig. 2a-b**). These findings are entirely consistent with biochemical and cell biology studies showing that loss of CXXC1 and WDR82 is followed by a very specific global loss of H3K4me3 and not H3K4me1 or H3K4me2[15-17]. A further testament to the validity of this screening approach is the fact that we also identified several members of the mammalian PAF complex which has been carefully characterised through various biochemical approaches in *S. cerevisiae* and shown to interact with Pol II to provide a scaffold for the COMPASS complex and consequent deposition of H3K4me3[18].

Loss of RNA Pol II does not alter global levels of H3K4me3

The reciprocal interplay between Pol II and H3K4me3 has been the subject of intense debate[19-21]. It has been proposed that rather than being instructive for transcription, H3K4me3 may be a consequence of it, and instead, plays a key role in influencing transcriptional consistency[22] and regulating co-transcriptional events such as splicing and termination of transcription[23, 24]. In this regard, a striking observation from the screen was the clear absence of Pol II subunits and the general transcriptional machinery (**Fig. 2a**). To assess whether these are false negative hits, we independently verified the finding that loss of Pol II did not affect global levels of H3K4me3 by depleting the major subunit of Pol II, Rpb1, using two independent sgRNAs. Western blot analysis revealed that depletion of Rpb1 did not impact the global levels of H3K4me3 (**Fig. 2f**). As a distinct control, we included H3K27ac, another dynamic and rapidly turned over histone modification associated with gene activity. Here again, we found that loss of Rpb1 did not affect global levels of H3K27ac (**Fig. 2f**). To confirm these results with an orthogonal approach, we performed ChIP-seq for both Pol II, H3K4me3 and H3K27ac after depletion of Rpb1. These data again confirm that H3K4me3 and Pol II levels at TSS are highly correlated in the control cells (**Extended Data Fig. 2d**). Furthermore, as expected, depletion

289   of Pol II by CRISPR-mediated knockout of Rpb1 resulted in a profound reduction of Pol II at
290   chromatin (**Fig. 2g-h**). However, again we found that the dramatic loss of chromatin bound Pol
291   II had a negligible effect on H3K4me3 genome-wide (**Fig. 2i**).
292
293   It has been suggested by studies in yeast that H3K4me3 levels can persist for several hours
294   after transcription ceases[25]. Therefore, to ascertain whether the retention in H3K4me3 levels
295   after Pol II deletion is also seen with other histone modifications associated with active
296   transcription, we also performed ChIP-seq for H3K27ac following Pol II depletion. Consistent
297   with our observations for H3K4me3, Pol II was also not required for the genome-wide
298   maintenance of H3K27ac (**Fig. 2j**). Taken together, these results suggest that although
299   H3K4me3 is strongly associated with active transcription, a global decrease in chromatin
300   bound Pol II has little influence on H3K4me3 levels.
301

302   **MLLT10 and global H3K79me2 are dispensable for survival of MLL-Fusion leukaemia**
303   **cells**
304   We were encouraged by the fact that in a single experiment, CRISPR-ChIP had identified most
305   of the known regulators of H3K4me3. Therefore, we next wanted to apply the CRISPR-ChIP
306   approach to study the regulation of a histone modification closely associated with the
307   maintenance of malignant gene expression in a unique cancer setting, leukaemias driven by
308   MLL-fusion oncogenes[26]. H3K79 lies in the globular domain of histone 3 and unlike
309   H3K4me3, which is deposited by multiple enzymatic subunits, H3K79methylation is known
310   to be globally regulated by a single enzyme, which in mammals is DOT1L[27]. H3K79me2 is
311   distributed throughout the coding region of actively transcribed genes, where it is thought to
312   play a role in facilitating transcription. Previous proteomic and biochemical studies have shown
313   that the endogenous DOT1L complex is comprised of DOT1L, MLLT10, MLLT1/ENL and
314   MLLT6/AF17[7,28] (**Extended Data Fig. 3a**). Moreover, whilst MLLT1/ENL is also found in
315   the P-TEFb/super elongation complex (SEC), the other components of the DOT1L complex
316   (DOT1L, MLLT10 and MLLT6) are mutually exclusive with P-TEFb/SEC[28-30]. On the other
317   hand, in leukaemia cells containing an MLL-Fusion oncogene, DOT1L has been shown to be
318   recruited by MLL-Fusion partners to increase H3K79methylation at critical MLL-Fusion target
319   genes associated with leukaemia development[31-34]. Whilst DOT1L activity in MLL-Fusion
320   leukaemia has been intensively studied[35], the relative contribution of the MLL-Fusion
321   oncogene to the regulation of H3K79me2 and whether there are other major regulators of this
322   histone modification remains unclear.
323

324   To address these questions, we performed a CRISPR-ChIP screen for H3K79me2 in MLL-AF9
325   leukaemia cells. Our results clearly identified DOT1L and MLLT10 but not the other DOT1L
326   complex members as key non-redundant regulators of H3K79me2 (**Fig. 3a and**
327   **Supplementary Table 3**). Early biochemical studies in yeast established the importance of
328   H2Bub for the methylation of H3K79 by DOT1L[36-38] and more recently the mammalian ring
329   finger protein 20 (RNF20), the major H2B ubiquitin ligase, was shown to be required for the
330   maintenance of H3K79me2 and the DOT1L mediated malignant gene expression program in
331   MLL-AF9 leukaemia cells[39]. In line with these findings, CRISPR-ChIP also identified RNF20
332   as a major non-redundant regulator of H3K79me2 (**Fig. 3a**). Whilst the PAF complex is known

333   to be required for efficient deposition of H3K79me2, our CRISPR-CHIP data specifically
334   identified RTF1, which in contrast to its role in yeast, is a labile PAF complex subunit in
335   mammalian cells where it functions to markedly potentiate transcriptional elongation[40,41].
336   Interestingly, recent structural studies have suggested that the stimulation of Pol II elongation
337   by RTF1 depends on its interaction with other components of the Pol II elongation complex
338   particularly DSIF[41]. Notably, our CRISPR-ChIP screen identified both components of DSIF
339   (SUPT4H1 and SUPT5), Iws1 (Supt6h), Tceb3 (Elongin A) and Ints3 (**Fig. 3a**), which are all
340   known to function as potentiators of Pol II elongation[41-43]. Although, H3K79me2 has been
341   loosely associated with transcriptional elongation, functional experimental evidence linking
342   H3K79me2, and Pol II elongation has been scarce. Rather remarkably, our functional genomics
343   approach identified the key subunits of the (i) PAF-complex[44], (ii) DSIF complex[45], (iii)
344   Elongin complex[46] and (iv) Integrator complex[47], which have all been functionally validated to
345   affect pause release and Pol II elongation rate[48,49]. These results strongly support the notion
346   that the maintenance of H3K79me2 by the DOT1L complex is dependent on efficient Pol II
347   elongation.

348

349   As our CRISPR-ChIP approach was designed to enable us to uncouple the regulation of
350   H3K79me2 from the requirement for malignant cell viability, we next wanted to assess which
351   of the major regulators of H3K79me2 were also required for the survival of the MLL-AF9
352   leukaemia cells. As expected, many of the components required for Pol II elongation are
353   common essential proteins and similarly, RNF20 and DOT1L which are established
354   dependencies in MLL-AF9 leukaemia, also read out in our viability screen (**Fig. 3b and**
355   **Supplementary Table 4**). Surprisingly, while MLLT10 is known to recruit DOT1L to
356   chromatin via its PZP domain[50] and clearly reads out as being required for H3K79me2 levels,
357   it did not appear to be required for leukaemia cell survival (**Fig. 3b and Extended Data Fig.**
358   **3b**). To investigate these unexpected findings further we initially used two independent
359   sgRNAs against each of the components of the DOT1L complex and assessed the effects on
360   both global H3K79me2 levels and cell survival in MLL-AF9 leukaemia cells. These data
361   confirmed that loss of DOT1L results in a global depletion of H3K79me2 and also dramatically
362   reduces the survival of MLL-AF9 leukaemia cells[6,51] (**Fig. 3c-d**). Moreover, we also validated
363   previous results showing that ENL is a dependency in MLL-AF9 leukaemia[52,53] but does not
364   alter global H3K79methylation levels (**Fig. 3c-d**). Notably however, we confirmed that loss of
365   MLLT10 has a profound effect on global H3K79me2 levels, but this is not accompanied by a
366   survival disadvantage for the leukaemia cells (**Fig. 3c-d**). These findings are recapitulated in
367   the Dep-Map database where all the MLL-leukaemia cell lines show a dependency on ENL for
368   survival but only the cell lines, MOLM13 and MOLM14, both derived from the same patient,
369   and SHI-1 (MLL-AF6), which does not aberrantly recruit DOT1L, show a dependency on
370   MLLT10 (**Fig. 3e**). To explore the functional consequences of depletion of the four members
371   of the DOT1L complex, we performed RNA-seq seq analysis following CRISPR/Cas9 deletion
372   of each complex member (**Extended Data Fig. 3a-c**). When we analyse the top 100 genes
373   downregulated by DOT1L loss, we also see a concomitant modest decrease in expression of
374   these genes after depletion of ENL (**Fig. 3f**). Strikingly however, the genes most affected by
375   DOT1L loss in MLL-AF9 leukaemia cells do not change their expression when either MLLT10
376   or MLLT6 are depleted (**Fig. 3f and Extended Data Fig. 3c-e**). Together, these surprising

377   findings raise the prospect that although DOT1L is critical, global H3K79methylation levels
378   maybe dispensable in MLL leukaemia and highlight our incomplete understanding of the
379   regulation of H3K79methylation in this disease.
380

381   **Selective regulation of transcription by MLL-AF9**
382   In order to address the discrepancy between H3K79me2 regulation and leukaemia cell survival
383   by MLLT10 and DOT1L, we first sought to determine the influence of the MLLAF9 oncogene
384   on both H3K79methylation and transcription of its direct target genes.
385

386   Current approaches for assessing the transcriptional influence of chromatin proteins largely
387   rely on coupling methods to map chromatin occupancy (such as ChIP-seq) with changes in
388   transcriptional output (using various RNA-Seq methods) following a perturbation of the target
389   of interest [1]. However, in the absence of small molecule inhibitors, the slow kinetics of genetic
390   depletion experiments (e.g. shRNA/CRISPR) hampers the accurate assessment and separation
391   of primary and secondary transcriptional changes. These limitations can be overcome by
392   genetically engineered approaches that hijack the endogenous protein degradation machinery
393   for rapid degradation of the target of interest by small molecules[54-56]. We adopted this system
394   by generating a murine retroviral model of MLL-AF9 leukaemia engineered to contain a
395   3xFLAG tag for efficient mapping by ChIP-seq and a minimal auxin inducible degron (mAID)
396   for rapid degradation using the plant hormone, auxin (IAA)[55]. To ensure that the MLL-AF9-
397   F3-mAID construct faithfully recapitulated the disease biology induced by MLL-AF9, we
398   transplanted our *in vitro* transformed haematopoietic stem and progenitor cells (HSPC) into
399   C57/BL6 mice (**Fig. 4a**). We generated both MLL-AF9-F3-mAID and MLL-AF9-F3-mAID +
400   mutant NRAS transformed cells and showed that both genotypes engrafted and gave rise to
401   fulminant leukaemia with the expected disease latency (**Extended Data Fig. 4a**). To determine
402   the genome-wide occupancy of MLL-AF9, we performed FLAG ChIP-seq in biological
403   duplicates from MLL-AF9-F3-mAID along with ChIP-seq for Menin and H3K79me2 as well
404   as a FLAG ChIP-seq from non-FLAG tagged MLL-AF9 transformed cells as a negative control
405   (**Fig. 4b-c and Extended Data Fig. 4b-c**). The MLL-AF9-F3-mAID ChIP-seq data shows a
406   high level of enrichment at previously reported MLL-AF9 target genes (e.g., Meis1 and Hoxa9)
407   (**Fig. 4d and Extended Data Fig. 4b-c**). Notably, genes bound by MLL-AF9 were also co-
408   occupied by Menin and H3K79me2, with all three factors displaying highly correlated
409   enrichment (**Fig. 4c-d**).
410

411   To determine the genes directly regulated by MLL-AF9, we treated MLL-AF9-F3-mAID cells
412   expressing the plant Tir1 E3 ligase gene with indole-3-acetic acid (IAA) to facilitate IAA-
413   dependent degradation of MLLAF9 (**Fig. 1e**). Time course analysis of MLL-AF9-F3-mAID +
414   Tir1 cells treated with IAA showed near complete degradation within 2-4 hours of treatment
415   (**Fig. 4f**). MLL-AF9 degradation resulted in a marked proliferation arrest and induction of
416   differentiation, consistent with the role of MLL-AF9 in the maintenance of the leukaemic
417   state[57] (**Extended Data Fig. 4d-e**). To then identify the direct targets of MLL-AF9, we treated
418   cells for 2hrs with IAA and performed FLAG ChIP-seq for MLLAF9-F3-mAID and nascent
419   RNA-seq (4su-seq). FLAG ChIP-seq showed a dramatic global loss of MLL-AF9 from
420   chromatin; however, nascent RNA-seq uncovered a core set of genes (Padj < 0.05) representing

421 only ~2% (260/10221) of the genes bound by MLL-AF9 that were immediately downregulated
422 following MLL-AF9 degradation (**Fig. 4g-h and Extended Data Fig. 4f-h**). To complement
423 the nascent RNA-seq data we also performed ChIP-seq for RNA polymerase II phosphorylated
424 on serine 2 (RNA-PolIIS2ph). While we observed a global loss of MLL-AF9 from chromatin
425 following 2hrs of degradation, global levels of RNA-PolIIS2ph were largely unchanged
426 (**Extended Data Fig. 4i**). However, in agreement with the nascent RNA-seq data, we observed
427 a selective loss of RNA-PolIIS2ph only at the MLL-AF9 bound genes that change their
428 expression 2 hours after degradation of the oncoprotein (**Fig. 4i-j and Extended Data Fig. 4j**).
429 Taken together, these data illustrate that the chromatin localisation of the oncoprotein is
430 insufficient in isolation to assign a regulatory influence on transcription and emphasises the
431 fact that although a large number of active genes are bound by MLL-AF9, only a small subset
432 of these are fully reliant on the MLL-FP for their expression.

433

434 **MLL-AF9 selectively influences H3K79methylation at distinct target genes**
435 Recently published data revealed that inhibition of either Menin or DOT1L with small
436 molecule inhibitors effectively evicted MLL-AF9 from chromatin[56] (**Fig. 5a**). However, the
437 converse, which is the impact of global MLLAF9 eviction on H3K79methylation was not
438 comprehensively assessed. Indeed, we confirmed the previous reports that both Menin
439 (VTP50469) and DOT1L (SGC0946) inhibition caused the global eviction of MLL-AF9 from
440 chromatin after 48 and 72hrs of treatment, respectively (**Fig. 5b, d and Extended Data Fig.
441 5a**). Nonetheless, a major conundrum when assessing the chromatin localisation of
442 transcriptional activators is that it is often hard to distinguish whether displacement of the
443 regulator from chromatin is simply a consequence of a decrease in transcription. Despite,
444 numerous studies assessing the transcriptional effects of MLL fusion proteins such as MLL-
445 AF9, none have specifically addressed or answered this important issue. To resolve this issue,
446 we decided to rapidly degrade the BET bromodomain proteins as targeting these proteins has
447 previously been shown to be effective in MLL leukaemia cells[58,59] and rapid degradation leads
448 to a profound and global loss of transcription[60,61]. Remarkably, we found that even after 8 hours
449 of treatment with the IBET-VHL degrader, MLL-AF9 remained bound to chromatin (**Fig. 5c,
450 e and Extended Data Fig. 5b**). Collectively, these data emphasise two important previously
451 unappreciated facts (i) MLL-FP are bound at chromatin in a manner that is insensitive to the
452 transcriptional status of the locus and (ii) that both H3K79me2 and Menin are specifically
453 important for the chromatin occupancy of MLL-FP.

454

455 Whilst loss of H3K79me2 clearly resulted in a displacement of MLL-AF9, we and others have
456 previously demonstrated that despite rapid enzymatic inhibition of DOT1L, H3K79me2 levels
457 require several days to decrease[7,62]. Therefore, it remained unclear what role H3K79me2 plays
458 in the immediate early transcriptional regulation of MLL-AF9 target genes. Similarly, although
459 Menin inhibition has recently been shown to unexpectedly result in a decrease in H3K79me2
460 levels, the molecular mechanism underpinning this observation is unknown[63]. To address these
461 questions, we first sought to determine the transcriptional changes that are induced following
462 treatment with SGC0946 and VTP50469 for 72hrs and 24hrs, respectively, on the core MLL-
463 AF9 targets we identified. We chose these time points as this is when we and others observed
464 efficient displacement of MLLAF9 from chromatin with these specific inhibitors of DOT1L

465 and Menin[56] (**Fig. 5b, d**). We found that most of the genes repressed after MLL-AF9
466 degradation were also downregulated by both inhibitors, albeit at later time-points (**Fig. 5f and**
467 **Extended Data Fig. 5c**).
468
469 Although the same genes were regulated by MLL-AF9 degradation, Menin and DOT1L
470 inhibition, our data raised the prospect that the time frame in which these genes were altered
471 with each of these interventions was substantially different. Given the slow turnover of
472 H3K79me2 we considered the possibility that the levels of H3K79me2 alone are insufficient
473 to maintain transcription of MLL-FP targets. To explore this further, we first wanted to
474 determine the kinetics of Menin inhibition in relation to DOT1L inhibition at the same genes.
475 We performed primary transcript-qPCR (nascent qPCR) after a time course with VTP50469
476 treatment for up to 24hrs and found that Menin inhibition resulted in rapid downregulation of
477 key targets (*Meis1, Jmjd1c, Epha7*) as early as 2 hrs, which closely paralleled our findings with
478 MLL-AF9 degradation (**Fig. 5g**). To understand how this is reflected at chromatin, we
479 performed ChIP-seq for H3K79me2 and MLLAF9-FLAG after early and late treatment with
480 VTP50469 and SGC0946. Consistent with our previous studies using SGC0946, H3K79me2
481 was only reduced after 48hrs of treatment with no global changes detected after 6hrs (**Fig. 5h-**
482 **i**). Importantly, ChIP-seq analysis shows eviction of MLL-AF9 after 6 and 48hrs of VTP50469
483 treatment with no global changes in H3K79me2 (**Fig. 5h-i and Extended Data Fig. 5d**). These
484 results suggest that transcription of key MLL-FP target genes can be markedly suppressed
485 without any change in H3K79me2.
486
487 Interestingly, although global H3K79me2 levels are not changed with Menin inhibition at 6
488 hours or 48hrs (**Fig. 5h**), we observed a potent suppression of H3K79me2 after 48 hrs of
489 VTP50469 treatment at key MLL-target genes (**Fig. 5i**), raising the prospect that the Menin-
490 MLL-Fusion complex recruits DOT1L to chromatin to deposit H3K79me2 at a small subset of
491 genes. Notably, when we looked at the genes where VTP50469 treatment results in
492 transcriptional repression, we find that MLL-AF9 displacement is accompanied by a profound
493 loss of H3K79me2 (**Fig. 5j and Extended Data Fig. 5e-f**). However, even though VTP50469
494 results in a global displacement of MLL-AF9 (**Fig. 5b, d**) not all MLL-AF9 bound genes show
495 a loss of H3K79me2 after displacement of the oncoprotein with Menin inhibition (**Fig. 5k and**
496 **Extended Data Fig. 5e-f**). Taken together, these results suggested a multifactorial regulation
497 of H3K79methylation in MLL-FP leukemia which cannot be solely explained by the simple
498 model of aberrant recruitment of DOT1L by MLL-AF9.
499
500 **MLLT10 regulates H3K79methylation at regions not occupied by MLL-FP**
501 The selective effects of Menin inhibition and MLLAF9 eviction on H3K79methylation
502 alongside the clear requirement for Menin and MLL-AF9 in the maintenance of leukaemia
503 cells survival, sharply contrasted with the negligible effects on leukaemia cell survival after
504 loss of MLLT10, despite there being a profound global loss of H3K79me2 levels.
505
506 To explore the mechanistic basis for this observation, we performed spike-in normalised ChIP-
507 seq for H3K79me2 after depletion of DOT1L or MLLT10 with two independent sgRNAs.
508 These data provided orthogonal confirmation that both MLLT10 and DOT1L regulate

509 H3K79me2 levels globally (**Extended Data Fig. 6a**). Notably, when we look specifically at
510 our previously annotated MLL-AF9 target genes (**Fig. 4h**), we find that H3K79me2 levels are
511 higher at these genes compared with Non-MLL-AF9 target genes (**Fig. 6a-b**). More surprising
512 is that the fact that whilst H3K79me2 levels are markedly reduced after DOT1L loss, there are
513 no changes in H3K79me2 levels at the major MLL-AF9 target genes including MEIS1 and
514 HOXA9 following MLLT10 loss (**Fig. 6a-c and Extended Data Fig. 6b-c**). Together these
515 findings raised the intriguing prospect that in MLL-AF9 leukaemia, DOT1L is functionally
516 partitioned into two separate complexes; one that contains the malignant MLL-AF9 oncogene
517 and one that is the native endogenous DOT1L complex containing MLLT10.

519 The possibility of two distinct DOT1L complexes within MLL-AF9 leukaemia cells was
520 further supported by the fact that disruption of MLLT10 in the leukaemia cells appeared to
521 increase H3K79me2 levels at known MLL-AF9 target genes (**Fig. 6c**). To explore this further
522 we looked globally to identify genes where H3K79me2 levels were unchanged or increased
523 following MLLT10 depletion. These data clearly identified ~200 genes where H3K79me2
524 levels were elevated rather than depleted following MLLT10 loss (**Fig. 6d**). Remarkably, when
525 we assess the binding of MLL-AF9 across the genome we find that the genes with the greatest
526 chromatin occupancy of MLL-AF9 are the genes where H3K79me2 levels are increased
527 following MLLT10 loss (**Fig. 6e**). In contrast, H3K79me2 levels decrease after MLLT10
528 depletion at the genes with the lowest occupancy of MLL-AF9 (**Fig. 6b, e**). Together these data
529 suggest that loss of MLLT10 results in the disassembly of the native DOT1L complex
530 containing MLLT10, which in turn results in a relative excess of DOT1L that is able to
531 associate with MLL-AF9 and further accentuate H3K79me2 levels at the major target genes of
532 the oncogenic complex. In further support of this contention, is the striking observation that
533 H3K79me2 was selectively increased directly where MLLAF9 is bound but was reduced in the
534 gene body of MLLAF9 target genes following loss of MLLT10 (**Fig. 6f-g and Extended Data
535 Fig. 6d**).

537 To further strengthen these findings, we sought to demonstrate the existence of biochemically
538 distinct complexes using size exclusion chromatography. Whilst our use of an epitope tag to
539 MLL-AF9 enabled our detailed studies of chromatin binding in mouse cells, there are
540 unfortunately no reliable antibody reagents to detect mouse MLLT10, DOT1L, ENL and
541 MLLT6. However, as many of the paradigms established for MLL leukaemias in mouse cells
542 are conserved in human cells and moreover shared by several of the common MLL-fusions we
543 chose to biochemically characterise cellular extracts from the most commonly used human
544 MLL-leukaemia cell line MV4;11 containing an MLL-AF4 fusion. When analysed by western
545 blot these data showed that although MLLT10 and DOT1L co-elute (fraction 3), DOT1L but
546 not MLLT10 co-elutes with MLL1 (**Extended Data Fig. 6f**). Taken together with the ChIP-
547 seq analyses these findings are consistent with the existence of two distinct DOT1L complexes
548 in MLL-FP leukaemia: (i) an oncogenic DOT1L complex containing the MLL-FP and (ii) a
549 native DOT1L complex containing MLLT10.

551 Our earlier data suggested that whilst Menin inhibition has no influence on the regulation of
552 H3K79me2 by the native endogenous DOT1L complex which includes MLLT10, it does

553   dramatically affect DOT1L activity as it displaces the oncogenic DOT1L complex associated
554   with MLL-AF9 from chromatin (**Fig. 5b, j**). Consistent with the fact that there is a reciprocal
555   re-distribution of DOT1L between the oncogenic and endogenous complexes, we find that
556   displacement of chromatin bound MLL-AF9 after Menin inhibition results in an increase in
557   H3K79me2 at genes dependent on the endogenous MLLT10/DOT1L complex (**Fig. 6h-i**).
558   These results further support the fact that DOT1L exists in both a native MLLT10-dependent
559   complex and an MLL-AF9 oncogenic complex, which together cooperate to regulate
560   H3K79methylation at critical MLL-FP target genes. These data also provide a molecular
561   explanation for the differential regulation of H3K79me2 by Menin inhibition at MLL-AF9
562   target genes (**Fig. 5i-k and Extended Data Fig. 5e-f**).
563
564   **Native and oncogenic DOT1L complexes cooperate to regulate MLL-FP target genes.**
565   To demonstrate the potential clinical implications of the partitioning of DOT1L into two
566   functional complexes, we assessed how abolition of the native complex (through MLLT10
567   loss) impacts the transcriptional response of MLLAF9 cells to VTP50469. We treated control
568   or MLLT10 depleted MLLAF9 cells with a low dose of Menin inhibitor (VTP50469 20nM) or
569   DMSO for 48hrs and performed RNA-seq analysis. When we looked at target genes directly
570   bound and regulated by MLLAF9 (including *Meis1*, *Six1*, and *Epha7*), we found that these
571   genes were substantially more downregulated by Menin-inhibition in the MLLT10 depleted
572   cells compared to control cells (**Fig. 7a**). Indeed, the majority of the Menin/MLLAF9 target
573   genes showed a more pronounced downregulation in MLLT10 depleted cells after VTP50469
574   treatment (**Fig. 7b and Extended Data Fig. 7a**). We next explored whether these
575   transcriptional responses would translate to a functional outcome by performing proliferation
576   assays in the MLLT10 depleted vs control cells treated with VTP50469. These data show that
577   whilst cells deficient in MLLT10 have no growth disadvantage in the absence of drug
578   (**Extended Data Fig. 7b**), the resulting absence of a native DOT1L complex, sensitises the
579   cells to Menin inhibitors (**Fig. 7c and Extended Data Fig. 7b**). To provide a mechanistic
580   explanation for this, we looked at H3K79methylation by ChIP-seq and observed that in the
581   absence of the native DOT1L complex (deletion of MLLT10), H3K79me2 levels were
582   markedly more sensitive to treatment with VTP50469 compared to control cells at MLL-FP
583   target genes (**Fig. 7d and Extended Data Fig. 7c**). These results indicate that specific
584   disruption of the native DOT1L complex causes H3K79me2 at MLL-FP targets to be largely
585   maintained by the oncogenic DOT1L complex, which substantially enhances the
586   transcriptional response and sensitivity to Menin inhibition.
587
588   Since there are currently no small molecules available against MLLT10, we instead used
589   DOT1L inhibitors to determine the potential synergy of targeting the Menin-dependent
590   oncogenic DOT1L complex with Menin inhibition alongside inhibitors of the native DOT1L
591   complex. We demonstrate that combination of DOT1L and Menin inhibitors showed strong
592   synergistic efficacy (**Fig. 7e and Extended Data Fig. 7d-e**). Consistent with the greater
593   functional inhibition of MLL-AF9 cells we also noted that the levels of chromatin bound RNA-
594   Pol II at MLL-AF9 target genes were more dramatically reduced in the combination treated
595   cells (**Fig. 7f**). Taken together, these data highlight that both the oncogenic and native DOT1L
596   complexes cooperate in the maintenance of MLL-FP target gene expression and provide the

597 molecular rationale for potential future combination therapy strategies using DOT1L and
598 Menin inhibitors in MLL-FP leukaemia (**Fig. 8**).
599
600 **Discussion**
601
602 Despite being one of the most intensely studied subset of oncoproteins, the molecular
603 mechanisms by which MLL-FP and its cooperating oncogenic cofactors regulate their target
604 genes remain incompletely understood. Our study sought to address this key issue by focussing
605 on the molecular regulation of a histone modification, H3K79me2, which has been a central
606 focus of the oncogenic process mediated by MLL fusion proteins. Identifying the major
607 regulatory proteins required for histone modifications originally involved a candidate approach
608 which was markedly facilitated by yeast genetics coupled to labour-intensive biochemistry[64].
609 However, these biochemical strategies alone are unable to assign a functional role to complex
610 members or identify independent functional regulators. As the cornerstone method for
611 understanding chromatin biology has been ChIP[65], we sought to couple ChIP to high-
612 throughput CRISPR screens to enable a strategy for unbiased discovery of chromatin templated
613 processes. Whilst our CRISPR-ChIP approach can be adapted to study any defined chromatin
614 event, as a proof-of-concept, here we have used it to identify all of the major non-redundant
615 regulators of two highly conserved histone modifications, H3K4me3 and H3K79me2. Using
616 this novel method, we showed that each member of the WRAD complex is essential for
617 efficient deposition of H3K4me3 by the COMPASS and COMPASS-like complexes. CRISPR-
618 ChIP further allowed us to reveal the specific COMPASS complex (SET1A/B) that is likely
619 functioning at the EF1a locus and that Pol II is dispensable for the maintenance of H3K4me3
620 genome-wide in human cells. We subsequently used CRISPR-ChIP to identify MLLT10 and
621 DOT1L as key regulators of H3K79me2 genome-wide. Notably, our unique approach also
622 uncovered an unexpected discrepancy between global H3K79me2 levels and leukaemia cell
623 survival. The prevailing paradigm had suggested that H3K79me2 levels and leukemia cell
624 viability were inexorably linked; however, our unbiased screens had illustrated that global
625 levels of H3K79me2 can be profoundly diminished with loss of MLLT10 with a negligible
626 influence on leukaemia cell survival. We reconcile these observations by illustrating that
627 DOT1L exists in two functionally distinct complexes in MLL-FP leukaemia. The native
628 complex includes MLLT10 which facilitates H3K79 methylation at the majority of
629 transcriptionally active genes but has little influence on the maintenance of the malignant state.
630 In contrast, DOT1L also exists in a neomorphic oncogenic complex with the MLL-FP, which
631 is critical for the maintenance of the malignant transcription program and viability of the
632 leukaemia cells.
633
634 Beyond establishing the molecular regulation of H3K79me2 levels in MLL-leukaemia our data
635 also sought to address another contentious and unresolved issue in the field - is H3K79me2
636 necessary for transcription? Using a combination of approaches to accurately map chromatin
637 occupancy of the MLL-AF9 oncogene and the subsequent rapid and selective degradation of
638 the oncogene coupled to nascent transcriptomics we illustrate several key principles. (i)
639 Although MLL-AF9 is bound to several thousand genes, only ~2% of these genes have a
640 discernible immediate and early decrease in transcriptional output following loss of the

641 oncogene. (ii) The process of transcription per se does not influence the chromatin occupancy
642 of MLL-FP, instead, the chromatin occupancy of this oncogenic fusion is selectively dependent
643 on its interaction with Menin. In this regard, a key insight that our data provides relates to the
644 different kinetics of transcriptional regulation between catalytic inhibition of DOT1L and
645 disruption of Menin-MLL1 interaction. While MLLAF9 degradation and Menin inhibition
646 result in rapid transcriptional repression, DOT1L inhibition takes days to alter transcription.
647 The implications of these findings are important as they illustrate that although H3K79me2 is
648 associated with transcriptional activity, it is not sufficient to maintain transcription and
649 H3K79me2 is not the major regulator of MLL-FP target gene expression. Instead, our data
650 suggests that H3K79me2 is likely required to maintain a chromatin state that is permissive for
651 MLL-AF9 recruitment and binding.
652
653 Finally, our data highlighting the existence of native and oncogenic DOT1L complex provides
654 the molecular explanation for the unexplained observation that Menin inhibition unexpectedly
655 results in a delayed decrease in H3K79me2 at MLL-FP target genes. It also explains why
656 inhibition of Menin in MLL-FP leukaemia results in an increase in gene expression and
657 H3K79me2 levels at genes not bound by Menin or MLL-FP. Understanding how perturbation
658 of one chromatin bound complex results in the dynamic flux of key members between sub-
659 stoichiometric complexes leading to differential gene expression is an important emerging
660 theme in chromatin regulation. Moreover, our findings of functionally distinct DOT1L
661 complexes specifies the molecular rationale for observations of synergy between DOT1L and
662 Menin inhibitors[66]. Taken together, they highlight why targeting both the endogenous and
663 oncogenic DOT1L complexes with combination therapies may increase the therapeutic
664 efficacy and clinical outcomes with these novel therapies against MLL-FP driven leukaemia.
665
666

676   **Author contributions**
677   O.G and M.A.D conceived, designed and supervised the research and wrote the
678   manuscript. O.G, C.C.B, D.N, D.F, K.K, M.B and C.D conducted experiments and/or analysed
679   data. O.G, C.C.B and K.K performed the CRISPR-ChIP screens. E.Y.N.L and L.T led the
680   analysis of the genomic data and CRISPR-ChIP screens, with contribution from Y-C.C.
681

682   **Data and code availability**
683   Sequencing data has been deposited into the sequence read archive, hosted by the National
684   Center for Biotechnology Information. The accession number for the sequencing data reported
685   in this paper is NCBI sequence read archive: GSE192562.
686
687

Materials and Methods

Cell lines

K562, HEK293T, MOLM-13, MV4;11 and drosophila S2 cells were obtained from ATCC. MLL-AF9 leukaemic blasts were generated by magnetic bead selection (Miltenyi Biotec) of c-KIT$^+$ cells from whole mouse bone marrow and subsequent retroviral transduction with an MSCV-FLAGx3-MLL-AF9 construct. K562, MV4;11 and MOLM-13 cells were cultured in RPMI-1640 supplemented with 2mM Glutamax, 100 IU/mL Penicillin, 100 μg/mL Streptomycin and 10% HI-FCS. HEK293T cells were cultured in DMEM supplemented with 2mM Glutamax, 100 IU/mL Penicillin, 100 μg/mL Streptomycin and 10% HI-FCS. MLL-AF9 cells were cultured in RPMI-1640 supplemented with 2 mM Glutamax, 100 IU/mL Penicillin, 100 μg/mL Streptomycin, 10% HI-FCS, IL-3 10 ng/mL, IL-6 10 ng/mL and SCF 50 ng/mL (Peprotech) for the first two weeks and then IL-3 10ng/mL for ongoing culturing. Drosophila S2 cells were cultured in Schneider's Drosophila media (Life Technologies) supplemented with 100 IU/mL Penicillin, 100 μg/mL Streptomycin and 10% HI-FCS. All cell lines were cultured in 5% CO_2 at 37°C with the exception of Drosophila S2, which was cultured at room temperature. All cell lines were subjected to regular mycoplasma testing and underwent short tandem repeat (STR) profiling.

Chemicals

SGC0946 was obtained from structural genome consortium, VTP50469 and A-485 were purchased from Selleck chemicals. For each inhibitor doses used, and durations of treatment are indicated in figure legends. For ChIP-qPCR, A-485 was used at a dose of 3uM for 24hrs. For ChIP-seq experiments, VTP50469 was used at a dose of 500nM and cells typically harvested at 48hrs unless otherwise indicated. SGC0946 was used at a dose of 5uM for 72hrs. DMSO was used as vehicle control for all drug treatment expeirments.

CRISPR sgRNA library

The targeted sgRNA CRISPR library is a custom built ~7239 human and ~7395 mouse sgRNAs Targeting ~1144 chormatin regulators (6sgRNAs per gene), as well as non-targeting and safe targeting control sgRNAs (guide sequences are provided in Table S1). Library sgRNAs were expressed in the CRISPR-ChIP plasmid, which is a modified pKLV-U6gRNA(BbsI)-$_{EF1a}$puro$_{2A}$BFP lentiviral sgRNA expression vector, which encodes puromycin and BFP selection markers obtained from Addgene #50946, a gift from Kosuke Yusa [67] as described. Murine *Meis1* promoter comprised of ~1.4kb upstream of TSS Murine *Myc* enhancer was obtained from H3K27ac ChIP-seq and ATAC-seq data [68] and cloned upstream of a minimal CMV promoter. 500bp of human PGK promoter and 1kb of human EF1a promoter including 5'UTR intron. The cis-elements were cloned into the PKLV guide plasmid ~40bp downstream of the guide sequence.

CRISPR-ChIP screens

K562 and MLLAF9 cells were transduced with a lentiviral vector encoding Cas9-IRES mCherry. mCherry positive cells were sorted as single cells to obtain a clonal population of high cas9 expressing cells.

732

733   Cas9 cells (>100 x 10^6 per cell line) were infected with the pooled lentiviral sgRNA library at
734   a multiplicity of infection of 0.3. The percentage of cells infected was determined by flow
735   cytometry-based evaluation of BFP$^+$ (sgRNA-expressing) cells 48 hours following
736   transduction of the targeted epigenetic library. Infected cells were selected with 2 µg/mL
737   puromycin commencing 48 hours after transduction. Viable cells were collected and sorted at
738   day 6 (Pol II, H3K4me3) and day 9 (H3K79me2) for ChIP as well as library control (input)
739   samples. For the ChIP samples, ~120x10^6 cells were cross linked with 1% formaldehyde
740   (Sigma-aldrich). For the library control samples ~5-10x10^6 cells were harvested for genomic
741   DNA. Lysed cells for ChIP were sonicated to obtain average fragment size of 1-2kb to ensure
742   gRNA remains linked to promoter sequence. ChIP was performed as previously described, with
743   some modifications, ~8-10x10^6 cells' worth of chromatin was incubated with 2-4ug of antibody
744   in separate tubes, between 12-15 ChIP tubes per CRISPR-ChIP experiment was used. 50ul of
745   Protein A or G beads was added to each tube after O/N incubation followed by washing and
746   elution. Elution was performed twice and de-crosslinked O/N. ~12 ChIP tubes were purified
747   using Qiagen PCR purification columns. 3 ChIP tubes were purified across one PCR cleanup
748   column and column elution was performed at least 4 times in 50ul volume with heated
749   nuclease-free water to maximise the yield of ChIP DNA.

750

751   One-step PCR was performed to amplify the guide sequence from ChIP DNA or genomic
752   DNA, containing adaptors for Illumina sequencing. Samples were sequenced with single-end
753   75 bp reads on an Illumina NextSeq500. The sequence reads were trimmed to remove the
754   constant portion of the sgRNA sequences with fastx clipper
755   (http://hannonlab.cshl.edu/fastx_toolkit/) or cutadapt v1.14[69], then mapped to the reference
756   sgRNA library with bowtie2[70]. After filtering to remove multi-aligning reads, the read counts
757   were computed for each sgRNA. MAGECK was used to rank the genes for which targeting
758   sgRNAs were significantly enriched in the ChIP samples compared to the control samples
759   taken from the same pool of growing mutagenized cells. MAGECK analysis of CRISPR-ChIP
760   screens is in the Supplementary Tables 2-3, and CRISPR-dropout analysis is in Supplementary
761   Table 4.

762

763 **CRISPR/Cas9-mediated gene disruption**
764   Single guide RNA (sgRNA) oligonucleotides (Integrated DNA Technologies) were cloned into
765   lentiviral expression vectors pKLV-U6gRNA(BbsI)-$_{PGK}$puro$_{2A}$BFP (Addgene #50946, a gift
766   from Kosuke Yusa)[67] as described. Oligonucleotide sequences are listed in Table 2, with
767   sources as indicated[71,72]. For CRISPR/Cas9-mediated gene disruption, cells were first
768   transduced with the Cas9 expression vector pHRSIN-P$_{SFFV}$-Cas9-P$_{PGK}$-Blasticidin[73] or
769   FUCas9Cherry (a gift from Marco Herold, Addgene #70182)[74], and selected with blasticidin
770   or sorted for mCherry expression respectively. To generate polyclonal populations with
771   targeted gene disruption, cells were subsequently transduced with pKLV-gRNA-
772     $_{PGK}$puro$_{2A}$BFP encoding either gene-specific sgRNAs or with a control sgRNA targeting a
773   'safe' genomic location with no annotated function[71]. Cells infected with pKLV-gRNA-
774     $_{PGK}$puro$_{2A}$BFP were selected with 2 µg/mL puromycin for 72 hours, commencing 48 hours

775 after transduction. Efficient functional CRISPR/Cas9-mediated gene disruption of target genes
776 was confirmed by immunoblot or sanger sequencing using TIDE analysis.
777

Virus production and transduction
779 Lentivirus was produced by triple transfection of HEK-293ET cells with a lentiviral transfer
780 vector and the packaging plasmids psPAX2 and pMD.G at a 0.5:0.35:0.15 ratio. Retrovirus
781 was produced by triple transfection of HEK-293ET cells with a retroviral transfer vector,
782 structural pMD1-gag-pol plasmid and pMD.G envelope plasmid at a 0.75:0.22:0.03 ratio. All
783 transfections were performed using polyethylenimine (PEI). Viral supernatants were collected
784 48 hours following transfection, filtered through a 0.45 μm filter and added to target cells.
785

Flow cytometry
787 Cells were washed in PBS/2% FCS, samples were resuspended in PBS with 2% FCS and
788 filtered. Data were acquired on a BD LSRFortessa or BD FACSymphony and analysed in
789 FlowJo or sorted using a BD FACSFusion sorter.
790

Immunoblotting
792 Cells were lysed in 1% SDS in 100 mM Tris-HCl pH 8.0 with Roche cOmplete EDTA-free
793 protease inhibitor at room temperature. DNA was fragmented by adding 1:100 Benzonase
794 (Sigma). Lysates were heated to 70°C in SDS sample buffer with 50 mM DTT for 10 minutes,
795 separated by SDS-PAGE and transferred to PVDF membrane (Millipore). Membranes were
796 blocked in 5% milk in TBS-T and incubated with indicated antibodies O/N at 4 degrees. Blots
797 were imaged with ECL prime using a gel doc instrument.
798

Size exclusion chromatography
800 MV411 cell extract from 50×10^6 cells in 500 μL was clarified by centrifugation at 21,000 g
801 at 4°C for 10 min. 500 μL sample was then injected into a 500 μL loop and was loaded onto a
802 Superose 6 Increase 10/300 GL column (Cytiva) equilibrated with Buffer A (50 mM HEPES
803 pH 7.5, 250 mM NaCl, 10% glycerol, 0.5% Triton X-100, 1 mM DTT, 1 mM EDTA, 1mM
804 AEBSF, 20 μM leupeptin, 1 μM pepstatin, 1 mM PMSF). Fractions of 300 μL were collected
805 at a flow rate of 0.5 mL/min. Fractions were precipitated by the addition of 4 volumes of cold
806 acetone and stored overnight at -80°C. Proteins were pelleted by centrifugation at 21,000 g at
807 4°C for 20 min. Pellets were washed with cold acetone, air dried, resuspended in Laemmli
808 buffer (10% glycerol, 2% SDS, 60 mM Tris-HCl pH 6.8, 0.02% bromophenol blue, 2.5% 2-
809 mercaptoethanol) and were then subjected to SDS-PAGE and immunoblotting.

Antibodies
811 The following antibodies were used for western blot and/or ChIP analyses: anti-HSP60
812 (sc13966, Santa Cruz Biotechnology), anti-H3K27ac (ab4729, Abcam), anti-H3K4me3
813 (ab8580, Abcam), anti-H3K79me2 (ab3594, Abcam), Anti-H3 total (ab1791, Abcam), anti-
814 SET1A (clone-D3VPS, 61702S, Cell Signalling), anti-MLL1 (D2M7U, 14689S, Cell
815 Signalling), anti-MLL2/KMT2B (D6X2E, 63735S, Cell Signalling), anti-Pol II (CTD4H8, 05-
816 623, Millipore), anti-FLAG M2 antibody (F3165, Sigma). Anti-ENL (ab245520, Abcam), anti-

817 MLLT6 (PA5-30089, ThermoFisher), anti-MLLT10 (PA5-50907, ThermoFisher), anti-
818 DOT1L (ab72454, Abcam) antibodies.

819 **qRT-PCR**
820 mRNA was prepared with a Qiagen RNeasy kit and cDNA synthesis was performed with a
821 SuperScript VILO kit (Life Technologies), per the manufacturers' instructions. Quantitative
822 PCR analysis was performed on an Applied Biosystems StepOnePlus System or Roche
823 LightCycler 480 Real-Time PCR System with SYBR Green reagents. All samples were
824 assayed in triplicate. Relative expression levels were determined with the ΔCt method and
825 normalized to *GAPDH*. RT-qPCR primers are listed in Table 1.

826 **Table 1: ChIP-qPCR and Nascent-qRT-PCR primers**

GRNAEF1a (Primer 1) FW	tgaaaaagtggcaccgagtc
GRNAEF1a (Primer 1) RV	Aacttctcggggactgtgg
FW EF1aPURO vector (primer 2) CHIP	tggaatttgcccttttttgag
RV EF1aPURO vector (primer 2) CHIP	ggcttgtactcggtcatggt
FW Neg ctrl 1 ChIP mouse	TGCCATGTGAGTGCAAGAAC
RV Neg ctrl 1 ChIP mouse	ATGGTGGGAGGAAGAAAGGG
Mouse nascent JMJD1c FW	GTCATCCGAGCCGTGTCC
Mouse nascent JMJD1c RV	ccaaacgcaggatggtttt
Mouse nascent Meis1 FW	GAAGTTGAAGTAGGAAGGGAGCC
Mouse nascent Meis1 RV	ggctttaggatcctcattcaa
Mouse nascent EPHA7 FW	CTAAAGAGCTGCGACCCAAC
Mouse nascent EPHA7 RV	tttgctttcctctggcctta
Mouse nascent GAPDH FW	CCCAGCTTAGGTTCATCAGG
Mouse nascent GAPDH RV	ctcctggcttctgtctttgg
MEIS1 mouse ChIP FW	TCACCACGTTGACAACCTCG
MEIS1 mouse ChIP RV	GCTTTCTGCCACTCCAGCTG
EF1a human ChIP FW	Cctgtctcgctgctttcgat
EF1a human ChIP RV	Tggcccgcatttacaagact
PGK promoter human FW	tcgcacacattccacatcca
PGK promoter human RV	tccatttgtcacgtcctgca
PGK-Puro FW	cttcaaaagcgcacgtctgc
PGK-Puro RV	ggcttgtactcggtcatggt
Meis1-puro FW	TTTCTGCCACTCCAGCTGTC
Meis1-puro RV	ggcttgtactcggtcatggt
Myc-mCMV-puro FW	Tggcttgctcagagacacag

827
828
829
830
831
832
833

Myc-mCMV-puro RV	ggcttgtactcggtcatggt
FW mouse MYC enhancer	Aggggcttgcagagagtaga
RV mouse MYC enhancer	Ggagttgcccttggagatcc

834 **Table 2: sgRNA oligos**

mNT G2 FW	CACCGCTCGTATCTTTTCCCACGGC	835
mNT G2 RV	AAACGCCGTGGGAAAAGATACGAGC	836
DOT1L G1 FW	CACCGCTGCAAACATCACTACGGAG	837
DOT1L G1 RV	AAACCTCCGTAGTGATGTTTGCAGC	838
DOT1L G2 FW	CACCGGTCAATGGTGTGCAGATAGT	839
DOT1L G2 RV	AAACACTATCTGCACACCATTGACC	840
MLLT10 G1 FW	CACCGGTTTATTGCGACGGGCA	841
MLLT10 G1 RV	AAACTGCCCGTCGCAATAAACC	842
MLLT10 G2 FW	CACCGAGCTGTGTCCCCATAAAGA	843
MLLT10 G2 RV	AAACTCTTTATGGGGACACAGCTC	844
MLLT6 G1 FW	CACCGTTGGCGAACTGCACCTC	845
MLLT6 G1 RV	AAACGAGGTGCAGTTCGCCAAC	846
MLLT6 G2 FW	CACCGCTTGTCGGCATCCGTGG	847
MLLT6 G2 RV	AAACCCACGGATGCCGACAAGC	848
MLLT1/ENL G1 FW	CACCGTCCAGGTGAAGTTAGAGCT	849
MLLT1/ENL G1 RV	AAACAGCTCTAACTTCACCTGGAC	850
MLLT1/ENL G2 FW	CACCGCTGGATGTCACACTGCTCA	851
MLLT1/ENL G2 RV	AAACTGAGCAGTGTGACATCCAGC	852
RBBP5 G1 FW	CACCGTGAAGGGAATCGAAACCTC	853
RBBP5 G1 RV	AAACGAGGTTTCGATTCCCTTCAC	854
RBBP5 G2 FW	CACCGTGACACTATGTTATCAG	855
RBBP5 G2 RV	AAACCTGATAACATAGTGTCAC	856
ASH2L G1 FW	CACCGCAGGACCAATGTTACTA	857
ASH2L G1 RV	AAACTAGTAACATTGGTCCTGC	858
ASH2L G2 FW	CACCGCAACGTCTGCCATCACAGT	859
ASH2L G2 RV	AAACACTGTGATGGCAGACGTTGC	860

861

Cell proliferation and dose–response assays

For proliferation assays, cells were seeded at a consistent density prior to treatment in triplicate and treated with DMSO, VTP50469 or SGC0946 over the indicated time period. Drug was refreshed at least every three days. Cell number was calculated using the BD FACSverse (BD Biosciences), CytoFLEX (Beckman Coulter) or Muse cell counter.

sgRNA competition assays

MLL-AF9 Cas9 cells were transduced with lentivirus expressing a gene specific sgRNA. The percentage of BFP-positive cells was measured between day 2 and day 12 after infection and normalised to the percentage of BFP at day 2.

ChIP-sequencing

Chromatin immunoprecipitation was performed as described previously[50]. Briefly, for each ChIP, 20 million cells were crosslinked for 15 mins with 1% formaldehyde. Crosslinked material was sonicated to ~200–1000 bp using the Covaris Ultrasonicator e220. Sonicated material was incubated overnight with each antibody, then incubated for 3 h with Protein A magnetic beads. Beads were washed with low and high salt wash buffers, LiCl buffer and TE, before being eluted and decrosslinked overnight. DNA was purified using Qiagen Minelute columns. All ChIP antibodies were used at ~10ug per IP and are listed under the antibodies section. Sequencing libraries were prepared from eluted DNA using Rubicon ThruPLEX DNA-seq kit. Libraries were size selected between 200–500 bps and sequenced on the NextSeq500 using the 75 bp single-end chemistry. Chromatin from S2 cells was spiked-in to ChIP-seq experiments involving treatment with SGC0946 or knockout of DOT1L complex components. The S2 chromatin was added to chromatin from leukaemia cells prior to sonication at approximately 5-10% of the cell numbers used for the sample.

ChIP-seq analysis

Reads were aligned to the human genome (GRCh37.73) with BWA-mem[51]. Duplicate reads and reads mapping to blacklist regions or the mitochondria were removed. Peak calling was performed with MACS2[52] with default parameters. Genome-browser images of ChIP–seq data was generated by converting the bam files from BWA to TDF files with igvtools and viewing in IGV[53]. ChIP–seq coverage across selected genomic regions was calculated with BEDtools[54]. Average profile plots were generated using the NGSplot software [75].

RNA-sequencing

RNA was extracted using the Qiagen RNeasy kit. RNA concentration was quantified using a Qubit Fluorometer (Thermo Fisher Scientific). Libraries were prepared using QuantSeq 3′ mRNA-seq Library Prep kit (Lexogen). Libraries were sequenced on the NextSeq500 using 75 bp single end chemistry.

RNA-sequencing analysis

Bcl2fastq (Illumina) was used to perform sample demultiplexing and to convert BCL files generated from the sequencing instrument into Fastq files. Reads were aligned to the human genome (G1k V37) using HiSAT2[76] and reads were assigned to genes using htseq-count[77].

906 Differential expression was calculated using DESeq2[78]. Genes with a false discovery rate
907 corrected for multiple testing using the method of Benjamini and Hochberg below 0.05 and a
908 log fold change greater than 1 were considered significantly differentially expressed.
909 Scatterplots depicting log fold change differences between conditions were generated in R
910 using the ggplot2 package and custom R code.
911

912 **Quantification and statistical analysis**
913 Statistical analysis was carried out using GraphPad Prism 9. Details of statistical analysis
914 performed are provided in the figure legends. Data were reported as mean ± SD, SEM or
915 independent replicates shown as individual data points, as indicated in the figure legends.
916 Significance was defined as $p < 0.05$.
917

918

919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944

References:

1. Park, P.J. ChIP-seq: advantages and challenges of a maturing technology. *Nature reviews. Genetics* **10**, 669-680 (2009).

2. Przybyla, L. & Gilbert, L.A. A new era in functional genomics screens. *Nature Reviews Genetics* (2021).

3. Nelson, J.D., Denisenko, O. & Bomsztyk, K. Protocol for the fast chromatin immunoprecipitation (ChIP) method. *Nature Protocols* **1**, 179-185 (2006).

4. Qin, J.Y. et al. Systematic comparison of constitutive promoters and the doxycycline-inducible promoter. *PLoS One* **5**, e10611 (2010).

5. Lasko, L.M. et al. Discovery of a selective catalytic p300/CBP inhibitor that targets lineage-specific tumours. *Nature* **550**, 128-132 (2017).

6. Daigle, S.R. et al. Selective killing of mixed lineage leukemia cells by a potent small-molecule DOT1L inhibitor. *Cancer Cell* **20**, 53-65 (2011).

7. Gilan, O. et al. Functional interdependence of BRD4 and DOT1L in MLL leukemia. *Nature Structural & Molecular Biology* **23**, 673-681 (2016).

8. Shilatifard, A. The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annual review of biochemistry* **81**, 65-95 (2012).

9. Santos-Rosa, H. et al. Active genes are tri-methylated at K4 of histone H3. *Nature* **419**, 407-11 (2002).

10. Soares, L.M. et al. Determinants of Histone H3K4 Methylation Patterns. *Molecular Cell* **68**, 773-785.e6 (2017).

11. Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R. & Young, R.A. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130**, 77-88 (2007).

12. Sze, C.C. et al. Coordinated regulation of cellular identity-associated H3K4me3 breadth by the COMPASS family. *Science advances* **6**, eaaz4764-eaaz4764 (2020).

13. Piunti, A. & Shilatifard, A. Epigenetic balance of gene expression by Polycomb and COMPASS families. *Science* **352**, aad9780 (2016).

14. Ernst, P. & Vakoc, C.R. WRAD: enabler of the SET1-family of H3K4 methyltransferases. *Briefings in Functional Genomics* **11**, 217-226 (2012).

15. Wu, M. et al. Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS. *Mol Cell Biol* **28**, 7337-44 (2008).

16. Lee, J.H. & Skalnik, D.G. Wdr82 is a C-terminal domain-binding protein that recruits the Setd1A Histone H3-Lys4 methyltransferase complex to transcription start sites of transcribed human genes. *Mol Cell Biol* **28**, 609-18 (2008).

17. Schneider, J. et al. Molecular regulation of histone H3 trimethylation by COMPASS and the regulation of gene expression. *Mol Cell* **19**, 849-56 (2005).

18. Krogan, N.J. et al. The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. *Mol Cell* **11**, 721-9 (2003).

19. Howe, F.S., Fischl, H., Murray, S.C. & Mellor, J. Is H3K4me3 instructive for transcription activation? *Bioessays* **39**, 1-12 (2017).

20. Pérez-Lluch, S. et al. Absence of canonical marks of active chromatin in developmentally regulated genes. *Nature genetics* **47**, 1158-1167 (2015).

21. Henikoff, S. & Shilatifard, A. Histone modification: cause or cog? *Trends Genet* **27**, 389-96 (2011).

22. Benayoun, B.A. et al. H3K4me3 breadth is linked to cell identity and transcriptional consistency. *Cell* **158**, 673-88 (2014).

23. Bieberstein, N.I., Carrillo Oesterreich, F., Straube, K. & Neugebauer, K.M. First exon length controls active chromatin signatures and transcription. *Cell Rep* **2**, 62-8 (2012).

24. Terzi, N., Churchman, L.S., Vasiljeva, L., Weissman, J. & Buratowski, S. H3K4 trimethylation by Set1 promotes efficient termination by the Nrd1-Nab3-Sen1 pathway. *Mol Cell Biol* **31**, 3569-83 (2011).

25. Ng, H.H., Robert, F., Young, R.A. & Struhl, K. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell* **11**, 709-19 (2003).

26. Nguyen, A.T., Taranova, O., He, J. & Zhang, Y. DOT1L, the H3K79 methyltransferase, is required for MLL-AF9-mediated leukemogenesis. *Blood* **117**, 6912-22 (2011).

27. Nguyen, A.T. & Zhang, Y. The diverse functions of Dot1 and H3K79 methylation. *Genes & Development* **25**, 1345-1358 (2011).

28. Mohan, M. et al. Linking H3K79 trimethylation to Wnt signaling through a novel Dot1-containing complex (DotCom). *Genes Dev* **24**, 574-89 (2010).

29. Gilan, O. et al. Functional interdependence of BRD4 and DOT1L in MLL leukemia. *Nat Struct Mol Biol* **23**, 673-81 (2016).

30. Lin, C. et al. AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. *Mol Cell* **37**, 429-37 (2010).

31. Daigle, S.R. et al. Selective killing of mixed lineage leukemia cells by a potent small-molecule DOT1L inhibitor. *Cancer cell* **20**, 53-65 (2011).

32. Okada, Y. et al. hDOT1L links histone methylation to leukemogenesis. *Cell* **121**, 167-78 (2005).

33. Zeisig, D.T. et al. The eleven-nineteen-leukemia protein ENL connects nuclear MLL fusion partners with chromatin. *Oncogene* **24**, 5525-32 (2005).

34. Mueller, D. et al. A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification. *Blood* **110**, 4445-54 (2007).

35. Wang, X., Chen, C.W. & Armstrong, S.A. The role of DOT1L in the maintenance of leukemia gene expression. *Curr Opin Genet Dev* **36**, 68-72 (2016).

36. Briggs, S.D. et al. Gene silencing: trans-histone regulatory pathway in chromatin. *Nature* **418**, 498 (2002).

37. Ng, H.H., Xu, R.M., Zhang, Y. & Struhl, K. Ubiquitination of histone H2B by Rad6 is required for efficient Dot1-mediated methylation of histone H3 lysine 79. *J Biol Chem* **277**, 34655-7 (2002).

38. Wood, A. et al. Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter. *Mol Cell* **11**, 267-74 (2003).

39. Wang, E. et al. Histone H2B ubiquitin ligase RNF20 is required for MLL-rearranged leukemia. *Proc Natl Acad Sci U S A* **110**, 3901-6 (2013).

40. Cao, Q.F. et al. Characterization of the Human Transcription Elongation Factor Rtf1: Evidence for Nonoverlapping Functions of Rtf1 and the Paf1 Complex. *Mol Cell Biol* **35**, 3459-70 (2015).

41. Vos, S.M., Farnung, L., Linden, A., Urlaub, H. & Cramer, P. Structure of complete Pol II–DSIF–PAF–SPT6 transcription complex reveals RTF1 allosteric activation. *Nature Structural & Molecular Biology* **27**, 668-677 (2020).

42. Diebold, M.L. et al. The structure of an Iws1/Spt6 complex reveals an interaction domain conserved in TFIIS, Elongin A and Med26. *EMBO J* **29**, 3979-91 (2010).

43. Stadelmayer, B. et al. Integrator complex regulates NELF-mediated RNA polymerase II pause/release and processivity at coding genes. *Nat Commun* **5**, 5531 (2014).

44. Zhu, B. et al. The human PAF complex coordinates transcription with events downstream of RNA synthesis. *Genes Dev* **19**, 1668-73 (2005).

1044   45.   Vos, S.M., Farnung, L., Urlaub, H. & Cramer, P. Structure of paused transcription
1045   complex Pol II–DSIF–NELF. *Nature* **560**, 601-606 (2018).
1046   46.   Kamura, T. et al. The Elongin BC complex interacts with the conserved SOCS-box
1047   motif present in members of the SOCS, ras, WD-40 repeat, and ankyrin repeat families.
1048   *Genes & development* **12**, 3872-3881 (1998).
1049   47.   Pfleiderer, M.M. & Galej, W.P. Structure of the catalytic core of the Integrator
1050   complex. *Mol Cell* **81**, 1246-1259.e8 (2021).
1051   48.   Elrod, N.D. et al. The Integrator Complex Attenuates Promoter-Proximal Transcription
1052   at Protein-Coding Genes. *Mol Cell* **76**, 738-752.e7 (2019).
1053   49.   Chen, F.X. et al. PAF1 regulation of promoter-proximal pause release via enhancer
1054   activation. *Science (New York, N.Y.)* **357**, 1294-1298 (2017).
1055   50.   Chen, S. et al. The PZP Domain of AF10 Senses Unmodified H3K27 to Regulate
1056   DOT1L-Mediated Methylation of H3K79. *Mol Cell* **60**, 319-27 (2015).
1057   51.   Bernt, K.M. et al. MLL-rearranged leukemia is dependent on aberrant H3K79
1058   methylation by DOT1L. *Cancer Cell* **20**, 66-78 (2011).
1059   52.   Wan, L. et al. ENL links histone acetylation to oncogenic gene expression in acute
1060   myeloid leukaemia. *Nature* **543**, 265-269 (2017).
1061   53.   Erb, M.A. et al. Transcription control by the ENL YEATS domain in acute leukaemia.
1062   *Nature* **543**, 270-274 (2017).
1063   54.   Nabet, B. et al. The dTAG system for immediate and target-specific protein
1064   degradation. *Nat Chem Biol* **14**, 431-441 (2018).
1065   55.   Natsume, T., Kiyomitsu, T., Saga, Y. & Kanemaki, M.T. Rapid Protein Depletion in
1066   Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors. *Cell
1067   Reports* **15**, 210-218 (2016).
1068   56.   Olsen, S.N. et al. MLL::AF9 degradation induces rapid changes in transcriptional
1069   elongation and subsequent loss of an active chromatin landscape. *Mol Cell* **82**, 1140-
1070   1155.e11 (2022).
1071   57.   Zuber, J. et al. An integrated approach to dissecting oncogene addiction implicates a
1072   Myb-coordinated self-renewal program as essential for leukemia maintenance. *Genes
1073   Dev* **25**, 1628-40 (2011).
1074   58.   Dawson, M.A. et al. Inhibition of BET recruitment to chromatin as an effective
1075   treatment for MLL-fusion leukaemia. *Nature* **478**, 529-33 (2011).
1076   59.   Zuber, J. et al. RNAi screen identifies Brd4 as a therapeutic target in acute myeloid
1077   leukaemia. *Nature* **478**, 524-8 (2011).
1078   60.   Winter, G.E. et al. BET Bromodomain Proteins Function as Master Transcription
1079   Elongation Factors Independent of CDK9 Recruitment. *Mol Cell* **67**, 5-18.e19 (2017).
1080   61.   Gilan, O. et al. Selective targeting of BD1 and BD2 of the BET proteins in cancer and
1081   immunoinflammation. *Science* **368**, 387-394 (2020).
1082   62.   Kuntimaddi, A. et al. Degree of recruitment of DOT1L to MLL-AF9 defines level of
1083   H3K79 Di- and tri-methylation on target genes and transformation potential.
1084   *CellReports* **11**, 808-820 (2015).
1085   63.   Krivtsov, A.V. et al. A Menin-MLL Inhibitor Induces Specific Chromatin Changes and
1086   Eradicates Disease in Models of MLL-Rearranged Leukemia. *Cancer Cell* **36**, 660-
1087   673.e11 (2019).
1088   64.   Allis, C.D., Jenuwein, T. & Reinberg, D. *Epigenetics*, x, 502 p. (Cold Spring Harbor
1089   Laboratory Press, Cold Spring Harbor, N.Y., 2007).
1090   65.   Furey, T.S. ChIP-seq and beyond: new and improved methodologies to detect and
1091   characterize protein-DNA interactions. *Nat Rev Genet* **13**, 840-52 (2012).
1092   66.   Dafflon, C. et al. Complementary activities of DOT1L and Menin inhibitors in MLL-
1093   rearranged leukemia. *Leukemia* **31**, 1269-1277 (2017).

67. Koike-Yusa, H., Li, Y., Tan, E.P., Velasco-Herrera, M.D.C. & Yusa, K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat. Biotechnol.* **32**, 267-273 (2014).

68. Bell, C.C. et al. Targeting enhancer switching overcomes non-genetic drug resistance in acute myeloid leukaemia. *Nature Communications* **10**, 2723-15 (2019).

69. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal* **17**, 10-12 (2011).

70. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357-359 (2012).

71. Morgens, D.W. et al. Genome-scale measurement of off-target activity using Cas9 toxicity in high-throughput screens. *Nat. Commun* **8**, 15178 (2017).

72. Wang, T., Wei, J.J., Sabatini, D.M. & Lander, E.S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80-4 (2014).

73. Burr, M.L. et al. CMTM6 maintains the expression of PD-L1 and regulates anti-tumour immunity. *Nature* **549**, 101-105 (2017).

74. Aubrey, B.J. et al. An Inducible Lentiviral Guide RNA Platform Enables the Identification of Tumor-Essential Genes and Tumor-Promoting Mutations In Vivo. *Cell Rep.* **10**, 1422-1432 (2015).

75. Shen, L., Shao, N., Liu, X. & Nestler, E. ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics* **15**, 284 (2014).

76. Kim, D., Paggi, J.M., Park, C., Bennett, C. & Salzberg, S.L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907-915 (2019).

77. Anders, S., Pyl, P.T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169 (2014).

78. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

79. Li, W. et al. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biology* **15**, 554 (2014).

1126   **Figure legends**

1127

1128   **Figure 1: Establishing the CRISPR-ChIP method.**

1129   **A)** Schematic of vector utilised for the screen highlighting the proximity of the gRNA sequence
1130   to the cis-regulatory element of interest driving expression of a fluorescent reporter. **B)**
1131   Schematic of CRISPR-ChIP proof-of-concept experiment with a mix of two independent
1132   DOT1L and two non-silencing guides sampled at day 2 and day 10 followed by ChIP-PCR
1133   NGS at day 10 for H3K79me2. **C)** H3K79me2 CRISPR-ChIP performed in murine MLL-AF9
1134   cells at day 10 post-infection with a pool of two independent DOT1L and two independent
1135   control (non-silencing) guides. Fold-change depletion of DOT1L guides normalised to control
1136   guides in the library control sample (input) at day 2 and day 10 (Top panel). Fold-change
1137   depletion of DOT1L sgRNAs in the H3K79me2 ChIP at day 10 vs day 10 library control (input)
1138   (Bottom panel). Guide representation was assessed by next-generation sequencing (NGS). **D)**
1139   Schematic of the CRISPR-ChIP workflow. Cells are infected with a lentiviral CRISPR library
1140   (MOI <0.3) and cultured for a desired time period, cells were then collected as a library control
1141   or crosslinked for ChIP with antibody of interest. sgRNAs are amplified by PCR from ChIP
1142   DNA or library control (input) DNA and sequenced by NGS. Differential enrichment of
1143   sgRNAs between ChIP and library control is determined using MAGECK [79]. **E)** The CRISPR-
1144   ChIP chromatin library was transduced into Cas9 negative K562 cells. H3K27ac ChIP was
1145   performed, and guide representation was assessed by NGS in two sampled replicates of library
1146   control and two independent ChIP samples. Data shown are correlation plots of library control
1147   Rep1 vs library control Rep2 (R=0.998), H3K27ac (50x10^6) Rep1 vs H3K27ac (50x10^6)
1148   Rep2 (R=0.993), and Library control Rep1 vs H3K27ac Rep1 (50x10^6) (R=0.985). **F)** Guide
1149   library was down sampled *in silico* at different fold representations, 10^7 = ~1000-fold
1150   representation down to 10000 guides = ~1fold representation. Guide distribution was plotted
1151   for each sampling event, width of the distribution represents the degree of skewing, with a
1152   wider distribution being more skewed. **G)** Schematic of H3K79me2 CRISPR-ChIP of
1153   SGC0946 timecourse treatment in K562 cells infected with a human guide library and mixed
1154   with DMSO treated K562 cells infected with mouse guide library ~ 1:1 ratio. **H)** Guide
1155   enrichment assessed by NGS and the percentage of human guides remaining in library control,
1156   H3K79me2 in the DMSO or SGC0946 treated samples at the indicated time points is shown.
1157   **I)** Western blot analysis of H3K79me2 levels after treatment with SGC0946 at the indicated
1158   timepoints, antibody against tubulin was used as a loading control.

1159

1160

1161   **Figure 2: Characterising the regulation of H3K4me3 using CRISPR-ChIP.**

1162   **A)** CRISPR-ChIP bubble plot showing the genes required for the maintenance of H3K4me3 at
1163   the EF1a promoter in K562 cells. P-values calculated using MAGECK algorithm. **B)**
1164   Schematic of the COMPASS and COMPASS-like complexes highlighting subunits identified
1165   in the CRISPR-ChIP screen, WRAD complex in green and SET1A/B complex specific
1166   subunits in orange. **C)** Western blot analysis of K562 Cas9 cells infected with control guide
1167   (Safe) or two independent guides targeting ASH2L or RBBP5 using antibodies against
1168   H3K4me3, H3K27ac, Total H3 and HSP60 (loading control). **D)** Profile plot of H3K4me3
1169   ChIP-seq in K562 Cas9 cells transduced with control (NT), ASH2L sgRNA or RBBP5 sgRNA.

1170   **E)** ChIP-seq heatmap of H3K4me3 in K562 Cas9 cells infected with control guides or guides
1171   targeting RBBP5, ASH2L. Showing all genes and ranked based on H3K4me3 levels in control
1172   cells. **F)** Western blot analysis of K562 Cas9 cells infected with control guide (Safe) or two
1173   independent guides targeting Rpb1 (largest subunit of RNA Pol II) using antibodies against
1174   Rpb1, H3K4me3, H3K27ac and HSP60 (loading control). **G)** Profile plot of Pol II ChIP-seq in
1175   K562 Cas9 cells transduced with control (NT), Rpb1 sgRNA (G1) or Rpb1 sgRNA (G2). **H)**
1176   ChIP-seq heatmap of RNA-Pol II in K562 Cas9 cells infected with control guides or two
1177   independent guides targeting Rpb1. Showing all genes and ranked based on RNA-Pol II levels
1178   in control cells. **I)** ChIP-seq heatmap of H3K4me3 in K562 Cas9 cells infected with control
1179   guides or two independent guides targeting Rpb1. Showing all genes and ranked based on
1180   H3K4me3 levels in control cells. **J)** ChIP-seq heatmap of H3K27ac in K562 Cas9 cells infected
1181   with control guides or two independent guides targeting Rpb1. Showing all genes and ranked
1182   based on H3K4me3 levels in control cells.

1183

1184   **Figure 3: MLLT10 and global H3K79me2 are dispensable for the survival of MLL-**
1185   **Fusion leukaemia cells**
1186   **A)** Bubble plot showing the genes required for the maintenance of H3K79me2 at the EF1a
1187   promoter in the CRISPR-ChIP vector in murine MLL-AF9 cells. P-values calculated using
1188   MAGECK algorithm. **B)** Scatter plot of gene dependencies in murine MLL-AF9 cells as
1189   assessed from day 14 of the CRISPR-ChIP screen, highlighted members of the DOT1L
1190   complex (DOTCOM) contained in the library, DOT1L, MLLT10 and MLLT6. **C)** Immunoblot
1191   analysis of H3K79me2, Total H3 and HSP60 in MLL-AF9 Cas9 cells transduced with the
1192   indicated sgRNAs. **D)** sgRNA negative selection competition assays in MLL-AF9 Cas9 cells
1193   transduced with either a non-silencing guide (Control) or guides targeting DOT1L, MLLT10,
1194   MLLT6 or ENL. Percentage of sgRNA positive cells remaining over time. Data represents n=3
1195   experiments. Mean +/- sd. **E)** Analysis of DepMap cell line dependencies comparing MLLT10
1196   vs DOT1L gene effect and MLLT1(ENL) vs DOT1L gene effect across wild-type MLL and
1197   MLL-rearranged acute leukaemia cell lines. Cell lines containing MLL rearrangements are
1198   highlighted in orange. **F)** Correlation plot of RNA-seq LFC between the indicated knockouts
1199   (MLLT10 and ENL) for the top 100 downregulated genes vs DOT1L KO cells.

1200

1201   **Figure 4: Selective regulation of transcription by MLL-AF9.**
1202   **A)** Schematic of retroviral transformation of cKit+ HSPC with MLL-AF9 containing a c-
1203   terminal 3xFLAG and mAID tag. **B)** Venn diagram of MLL-AF9 FLAG ChIP-seq peaks from
1204   two replicates. **C)** Heatmap of MLL-AF9 FLAG, Menin, and H3K79me2 ChIP-seq in
1205   MLLAF9-F3-mAID cells. **D)** Genome browser snapshot of MLL-AF9, Menin and H3K79me2
1206   at the *Meis1* locus. **E)** Schematic of the auxin-inducible degron system. **F)** Western blot
1207   analysis of MLL-AF9 degradation following an auxin (IAA) treatment timecourse using
1208   antibodies against FLAG and HSP60 (loading control). **G)** Heatmap of MLL-AF9 ChIP-seq in
1209   DMSO or IAA (2hrs) from MLLAF9-F3-mAID cells. **H)** Volcano plot of differentially
1210   expressed genes from nascent RNA-seq (4su-seq) in MLLAF9-F3-mAID cells treated with
1211   DMSO or IAA for 2hrs. Downregulated genes in blue (<LFC -0.4) and key direct target genes
1212   highlighted in red. **I)** Average profile plot of RNA-Pol IIS2ph ChIP-seq at highly expressed
1213   non-MLLAF9 and MLL-AF9 target genes in DMSO and IAA treated cells. **J)** Genome browser

1214 snapshot of the JMJD1c locus for MLL-AF9 and Pol IIser2 ChIP-seq and 4su-seq after 2hrs of
1215 DMSO or IAA treatment.
1216

1217 **Figure 5: Selective regulation of H3K79me2 at target genes by MLL-AF9**
1218 **A)** Schematic of MLLAF9 oncogene with some of the associated cofactors along with their
1219 targeted therapies. **B)** Heatmap of MLLAF9 ChIP-seq after treatment with either DMSO,
1220 SGC0946 (5uM) for 72hrs, or VTP50469 (500nM) for 48hrs. **C)** Heatmap of MLLAF9 ChIP-
1221 seq after treatment with either DMSO, IBET-151 (500nM) for 8hrs, or IBET-VHL (500nM)
1222 for 8hrs. **D)** Average profile plot of MLLAF9 ChIP-seq across all genes after treatment with
1223 either DMSO, VTP50469 (500nM) and SGC0946 (5uM). **E)** Genome browser snapshots of the
1224 *Meis1* locus in MLLAF9-F3-mAID cells treated with DMSO, SGC0946, VTP50469 or IBET-
1225 151 and IBET-VHL. **F)** Heatmap of downregulated genes from 3'RNA-seq in MLLAF9-F3-
1226 mAID cells treated with SGC0946 (72hrs), IAA (24hrs), or VTP50469 (24hrs). **G)** Nascent
1227 qPCR (primary transcript) analysis of VTP50469 timecourse treatment (500nM) vs DMSO in
1228 MLLAF9-F3-mAID cells using primers against *Jmjd1c*, *Meis1*, and *Epha7,* data normalised to
1229 housekeeping gene *(Gapdh),* error bars represent mean+/- SD from 3 biological replicates. **H)**
1230 Heatmap of H3K79me2 ChIP-seq analysis in MLLAF9-F3-mAID cells treated with either
1231 DMSO, SGC0946 (6hrs), SGC0946 (48hrs), VTP50469 (6hrs), and VTP50469 (48hrs). **I)**
1232 Genome browser snapshot of the Meis1 locus from the H3K79me2 ChIP-seq in (H). **J)** Average
1233 profile plot of H3K79me2 ChIP-seq across the body of genes downregulated after VTP50469
1234 and bound by MLLAF9 in DMSO, SGC0946, or VTP50469 treatment. **K)** Genome browser
1235 snapshot of exemplar VTP50469 target genes with differential changes in H3K79me2 after
1236 VTP50469, *Meis1* (changed) and *Six1* (unchanged).
1237

1238 **Figure 6: DOT1L functions in distinct native and oncogenic complexes in MLL-FP**
1239 **leukaemia.**
1240 **A)** Average profile plot of H3K79me2 levels at MLL-AF9 target genes in MLL-AF9[FLAG-mAID]
1241 Cas9 cells transduced with either a non-silencing guide (Control) or 2 independent guides
1242 targeting DOT1L (G1 and G2) or MLLT10 (G1 and G2). **B)** Average profile plot of H3K79me2
1243 levels at the top 500 expressed genes <u>not</u> regulated by MLL-AF9 in MLL-AF9[FLAG-mAID] Cas9
1244 cells transduced with either a non-silencing guide (Control) or 2 independent guides targeting
1245 DOT1L (G1 and G2) or MLLT10 (G1 and G2). **C)** Genome browser snapshots of canonical
1246 MLL-AF9 target gene *Meis1* and the *Ef1a* locus used in the CRISPR-ChIP plasmid. Tracks
1247 show levels of H3K79me2 and input in MLL-AF9 Cas9 cells transduced with either a non-
1248 silencing guide (Control) or two independent guides targeting DOT1L or MLLT10. **D)**
1249 Correlation plot of H3K79me2 ChIP-seq peak size in non-silencing control vs MLLT10
1250 knockout (average of two independent guides) showing genes with increased H3K79me2 (red)
1251 or decreased (blue). **E)** Hockey stick plot of genes ranked by [FLAG]MLL-AF9 occupancy; red
1252 dots represent genes that display increased H3K79me2 following MLLT10 KO vs non-
1253 silencing guides (Control) while blue dots represent genes with decreased H3K79me2
1254 following MLLT10 KO. **F)** Profile plot of H3K79me2 distribution across MLLAF9 peaks +/-
1255 2kb in MLLAF9[FLAG-mAID] Cas9 cells transduced with control guides or guides targeting DOT1L
1256 or MLLT10. Red box over the MLL-AF9 peak. **G)** Genome browser snapshot of exemplar
1257 direct and strongly bound MLL-AF9 target gene that shows increased H3K79me2 following

1258 MLLT10 KO in MLL-AF9 cells, *Runx2*. Red box over the MLL-AF9 peak. **H)** Correlation

1259 plot of H3K79me2 normalised count coverage in vehicle treated (DMSO) or Menin inhibitor

1260 (VTP50469) murine MLL-AF9 cells. Genes with a decrease in H3K79me2 following

1261 VTP50469 are highlighted in red and genes with increased H3K79me2 are in blue. **I)** Genome

1262 browser snapshot of an exemplar Menin inhibitor non-responsive gene, *Mxi1*, with no MLL-

1263 AF9 occupancy, showing increased H3K79me2 after VTP50469 treatment and decreased

1264 H3K79me2 in MLLT10 KO cells compared to control cells.

1265

1266 **Figure 7: Disruption of the native DOT1L complex sensitises MLLAF9 leukaemia cells**

1267 **to Menin inhibition.**

1268 **A)** Violin plot of RNA-seq in MLLAF9$^{FLAG-mAID}$ Cas9 cells expressing control guides or

1269 MLLT10 guides (stable knockout of MLLT10) and treated with either DMSO or VTP50469

1270 (20nM). Menin inhibitor target genes that are also bound by MLLAF9 are shown. **B)** Lollipop

1271 plot of RNA-seq logFC of the ratio of Control vs MLLT10 KO after treatment with VTP50469

1272 (20nM) in MLLAF9$^{FLAG-mAID}$ cells. **C)** Proliferation assay in control and MLLT10 KO

1273 MLLAF9$^{FLAG-mAID}$ Cas9 cells treated with either DMSO, 20nM or 40nM VTP50469. Cell

1274 counts shown at day 3 and day 7 and normalised to day 1 counts. n=3 biological replicates. *

1275 p <0.05 ** p<0.01. **D)** Genome browser snapshot of exemplar MLLAF9 targets *Meis1* and

1276 *Runx2* from H3K79me2 ChIP-seq in MLLAF9$^{FLAG-mAID}$ cells with control guides or MLLT10

1277 KO and treated with vehicle (DMSO) or VTP50469 (20nM) for 48hrs. MLLAF9 tracks shown

1278 for reference. **E)** Proliferation assay in MLLAF9$^{FLAG-mAID}$ Cas9 cells treated with either

1279 DMSO, SGC0946 (2uM), VTP50469 (10nM, 20nM or 50nM) and combinations. Cell counts

1280 shown at day 4 and day 6 and normalised to day 1 counts. n=3 biological replicates. * p <0.05.

1281 **F)** Profile plot of RNA-Pol II ChIP-seq enrichment at MLLAF9 target genes after treatment

1282 with either DMSO, SGC0946, VTP50469 or Combo for 48hrs.

1283

1284

1285 **Figure 8:** Schematic to illustrate the distinct function of the native and oncogenic DOT1L

1286 complex.

1287

1288

1289

1290

1291

1292

1293

1294

1295

1296

1297

1298

1299

1300

1301

Supplementary figure legends

Extended Data Figure 1: A) ChIP-qPCR analysis of H3K27ac ChIP performed in parental K562 cells or K562 cells infected with the CRISPR-ChIP plasmid containing an EF1a promoter. Primers for a negative control region (Neg Ctrl), endogenous EF1a or regions spanning the gRNA/EF1a (Primer 1), and EF1a/Puro (Primer 2) within the lentiviral CRISPR-ChIP vector were used, data represents mean +/- SD from two independent biological replicates. **B)** ChIP-qPCR analysis of H3K27ac, H3K4me3, H3K9ac and RNA-Pol II ChIP at the endogenous PGK promoter or integrated PGK promoter-Puro reporter, an intergenic region not enriched in active histone marks was used as a negative control (Neg control). Data represents mean +/- SD from two independent biological replicates. **C)** ChIP-qPCR analysis of H3K27ac, H3K4me3, H3K79me2 and RNA-Pol II ChIP at the endogenous Meis1 promoter or integrated Meis1 promoter (~1.2Kb upstream of TSS)-Puro reporter, an intergenic region not enriched in active histone marks was used as a negative control (Neg control). Data represents mean +/- SD from two independent biological replicates. **D)** ChIP-qPCR analysis of H3K27ac, H3K4me3, and H3K9ac ChIP at the endogenous Myc enhancer or integrated Myc enhancer-mCMV-Puro sequence, an intergenic region not enriched in active histone marks was used as a negative control (Neg control). Data represents mean +/- SD from two independent biological replicates. **E)** ChIP-qPCR analysis of H3K79me2 in K562 cells infected with the CRISPR-ChIP plasmid and treated with DOT1L inhibitor (SGC0946) for 9 days. Primers for Negative control region (Neg Ctrl), Meis1, EF1a, and a region spanning the gRNA/EF1a were used, data represents mean +/- SD from two independent biological replicates. **F)** qPCR analysis of H3K27ac ChIP in K562 cells infected with the CRISPR-ChIP vector and treated with either DMSO or CBP/P300 inhibitor (A-485) for 24hrs. Primers for Negative control region (Neg Ctrl), Meis1, EF1a, and a region spanning the gRNA/EF1a were used, data represents mean +/- SD from two independent biological replicates. **G)** NGS representation of control guides (NT) at day 2, day 10, ChIP input day 10, and H3K79me2 ChIP (Left panel). NGS representation of DOT1L guides at day 2, day 10, ChIP input day 10, and H3K79me2 ChIP (Right panel).

Extended Data Figure 2: A) Correlation plot of library guide counts between H3K27ac ChIP and a library control sample taken from Cas9 negative K562 cells. H3K27ac ChIP was performed from different starting cell number, 5 million, 10 million, 25 million, 50 million and 100 million cells. **B)** Bubble plot of H3K27ac ChIP (50M cells) from Cas9 negative K562 cells analysed using MAGECK. **C)** Correlation plot of RBBP5 KO H3K4me3 LFC vs ASH2L KO H3K4me3 LFC. **D)** Correlation plot of normalised Pol II levels (rpm) vs normalised H3K4me3 levels (rpm).

Extended Data Figure 3: A) Schematic of DOTCOM complex (DOT1L, ENL/MLLT1, MLLT6/AF17, MLLT10). **B)** Guide counts over time from CRISPR dropout screen in MLL-AF9 cells for 2 MLLT10 and 2 DOT1L guides. **C)** PCA analysis of RNA-seq from MLLAF9 Cas9 cells transduced with control (non-targeting), DOT1L G1, two independent ENL guides, two independent MLLT10 guides and two independent MLLT6 guides. **D)** Heatmap of RNA-seq data described above, showing downregulated genes in the various knockouts as indicated.

1346   **E)** Correlation plot of RNA-seq LFC between two independent MLLT6 sgRNAs and DOT1L
1347   KO for the top 100 downregulated genes in the DOT1L KO cells.

1348

1349   **Extended Data Figure 4: A)** Kaplan-meir curve of mice transplanted with MLLAF9$^{FLAG-mAID}$
1350   cells or Nras + MLLAF9$^{FLAG-mAID}$. **B)** Genome browser snapshot of MLLAF9 ChIP-seq
1351   replicates at the meis1 locus. FLAG ChIP from non-FLAG tagged MLLAF9 cells used as a
1352   negative control and input shown. **C)** Genome browser snapshot of MLLAF9 ChIP-seq
1353   replicates at the HOXA cluster. FLAG ChIP from non-FLAG tagged MLLAF9 cells used as a
1354   negative control and input shown. **D)** Proliferation assay of MLLAF9$^{FLAG-mAID}$ Tir1 cells
1355   treated with DMSO or IAA (500uM) over 7 days. **E)** FACS analysis of cell surface Gr1 levels
1356   in unstained, stained, DMSO or IAA treated MLLAF9$^{FLAG-mAID}$ Tir1 cells for 4 days. **F)** Profile
1357   plot of MLLAF9 FLAG ChIP-seq in MLLAF9$^{FLAG-mAID}$ cells treated with DMSO, IAA (2hrs
1358   and 4hrs). **G)** Heatmap of 4su-seq (nascent RNA-seq) in MLLAF9$^{FLAG-mAID}$ Tir1 cells treated
1359   with DMSO or IAA for 2hrs, showing the 260 downregulated genes. **H)** PCA plot of RNA-seq
1360   from MLLAF9$^{FLAG-mAID}$ Tir1 cells treated with DMSO or IAA (2hrs). **I)** Heatmap of RNA Pol
1361   II ser2 ChIP-seq at all genes in MLLAF9$^{FLAG-mAID}$ Tir1 cells treated with DMSO or IAA (2hrs).
1362   **J)** Genome browser snapshots of MLLAF9 targets, *Epha7* and *Baz2b* showing RNA Pol II-
1363   ser2 and MLLAF9 ChIP-seq tracks after treatment with either DMSO or IAA (2hrs).

1364

1365   **Extended Data Figure 5: A)** Heatmap of Menin ChIP-seq in MLLAF9 cells treated with
1366   DMSO, SGC0946 (72hrs) or VTP50469 (48hrs) at all genes. **B)** MLLAF9 FLAG ChIP-seq in
1367   MLLAF9 cells treated with DMSO, IBET151 or IBET-VHL for 8hrs at all genes. **C)** PCA
1368   analysis of RNA-seq data in MLLAF9 cells treated with DMSO, SGC0946 (72hrs), IAA
1369   (24hrs), VTP50469 (24hrs). **D)** Profile plot of MLLAF9 ChIP-seq in MLLAF9 cells treated
1370   with either DMSO, VTP50469 (6hrs), SGC0946 (48hrs) or VTP50469 (48hrs). **E)** Heatmap of
1371   MLLAF9 and H3K79me2 ChIP-seq data in MLLAF9 cells after treatment with VTIP50469
1372   for 48hrs at genes that show decreased H3K79me2 after VTP50469 treatment. **F)** Genome
1373   browser snapshot of ChIP-seq tracks at canonical MLLAF9 target genes, *Jmjd1c* and *Runx2*,
1374   showing MLL-AF9 and H3K79me2 after treatment with VTP50469 or SGC0946 for 48hrs.

1375

1376   **Extended Data Figure 6: A)** Average profile plot of H3K79me2 and input at all genes in
1377   MLL-AF9 cells infected with non-silencing guides (Control), 2 independent DOT1L guides
1378   and 2 independent MLLT10 guides. **B)** Heatmap of H3K79me2 at MLL-AF9 bound genes for
1379   the indicated samples. Genes are ranked from highest to lowest coverage for H3K79me2 in the
1380   control cells. **C)** Genome browser snapshot of a canonical MLLAF9 target, *Hoxa9*, from
1381   H3K79me2 ChIP-seq in MLLAF9 cells transduced with control sgRNA or two independent
1382   guides targeting MLLT10 or DOT1L. **D)** Profile plot of H3K79me2 distribution across
1383   MLLAF9 peaks +/- 10kb in MLLAF9$^{FLAG-mAID}$ Cas9 cells transduced with control guides or
1384   guides targeting DOT1L or MLLT10. **E)** Genome browser snapshots of exemplar direct and
1385   strongly bound MLL-AF9 target genes that show increased H3K79me2 following MLLT10
1386   KO in MLL-AF9 cells, *Jmjd1c and Epha7*. **F)** Immunoblot analysis of gel filtration fractions
1387   (superose 6 column) from nuclear extracts of MV4;11 cells (human MLL-AF4) using
1388   antibodies against DOT1L, MLLT10, MLL1, ENL and MLLT6.

1389

1390 **Extended Data Figure 7: A)** Heatmap of RNA-seq data from MLLAF9 Cas9 cells infected
1391 with control or MLLT10 guides and treated with either DMSO or VTP50469 with 20 or 40nM.
1392 **B)** Proliferation assays using MLL-AF9 Cas9 cells transduced with either a control non-
1393 targeting guide or two independent guides targeting MLLT10 (G1 and G2). Cells were treated
1394 with either DMSO, VTP50469 25nM or 50nM. Plots represent n=3 biological replicates. Bars
1395 represent mean +/- SD **C)** Profile plot of H3K79me2 ChIP-seq in MLLAF9 Cas9 cells infected
1396 with control or MLLT10 guides and treated with DMSO or VTP50469 (20nM) for 48hrs.
1397 Analysis performed across MLLAF9 target genes. **D)** Proliferation assay in MOLM13 cells
1398 treated with either DMSO, SGC0946 (3uM), VTP50469 (100nM) or combination. Mean +/-
1399 SD from n=3 biological replicates. **E)** Proliferation assay in MLLAF9$^{\text{FLAG-mAID}}$ Cas9 cells
1400 treated with either DMSO, SGC0946 (3uM), VTP50469 (50nM) or combination. Cell counts
1401 shown at day 4 and day 6 and normalised to day 1 counts. Mean +/- SD from n=3 biological
1402 replicates. * p <0.05.
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422

Figure 1

A CRISPR-ChIP

B Input

Non-targeting G1/G2
DOT1L G1/G2

Day 2

Day 10

Input

H3K79me2 ChIP

C

Fold Change

Day 2
Day 10

DOT1L G1 G2 G1 G2

ChIP vs Input

Fold Change

Input (Day 10)
H3K79me2 ChIP

DOT1L G1 G2

D CRISPR gRNA library

MOI ~0.3
Puromycin

Cas9+ Cells

Library Control (Input)
Sequence and Map

T_n

Input
Guide A B C D......N
Read count

ChIP - gPCR

T_n

ChIP
Guide A B C D......N
Read count

E

Pearson r = 0.993

Pearson r = 0.998

Pearson r = 0.985

F

Density

Fold representation
10000~1
50000~5
1e+05~10
5e+05~50
1e+06~100
5e+06~500
1e+07~1000

Value

G Chromatin library
mouse   human

DMSO   SGC0946 timecourse

Mix ~50:50

CRISPR-ChIP
H3K79me2

H Library   H3K79me2 CRISPR-ChIP

%human guides remaining

DMSO 24hr 48hr 72hr 96hr
SGC0946 (5µM)

I SGC0946

DMSO 8hrs 24hrs 48hrs 72hrs 96hrs

H3K79me2

Tubulin

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6

Figure 7

A VTP50469 20nM MLL-AF9 Cas9 cells

B VTP50469 20nM MLLAF9 cells

C Day 3 / Day 7

Non-silencing control / MLLT10 KO / Non-silencing control / MLLT10 KO

D H3K79me2 — Meis1, Runx2

E MLLAF9 cells — Day 4, Day 6
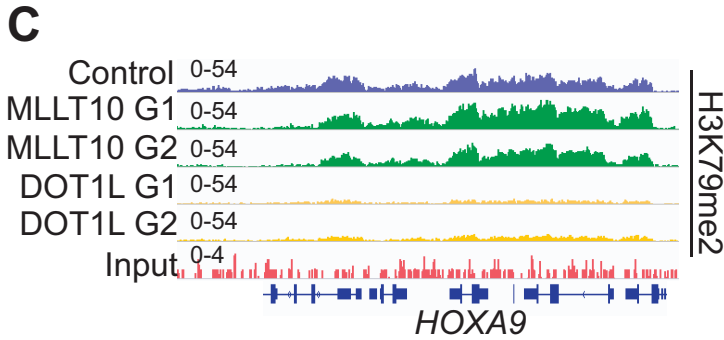
F MLLAF9 target genes
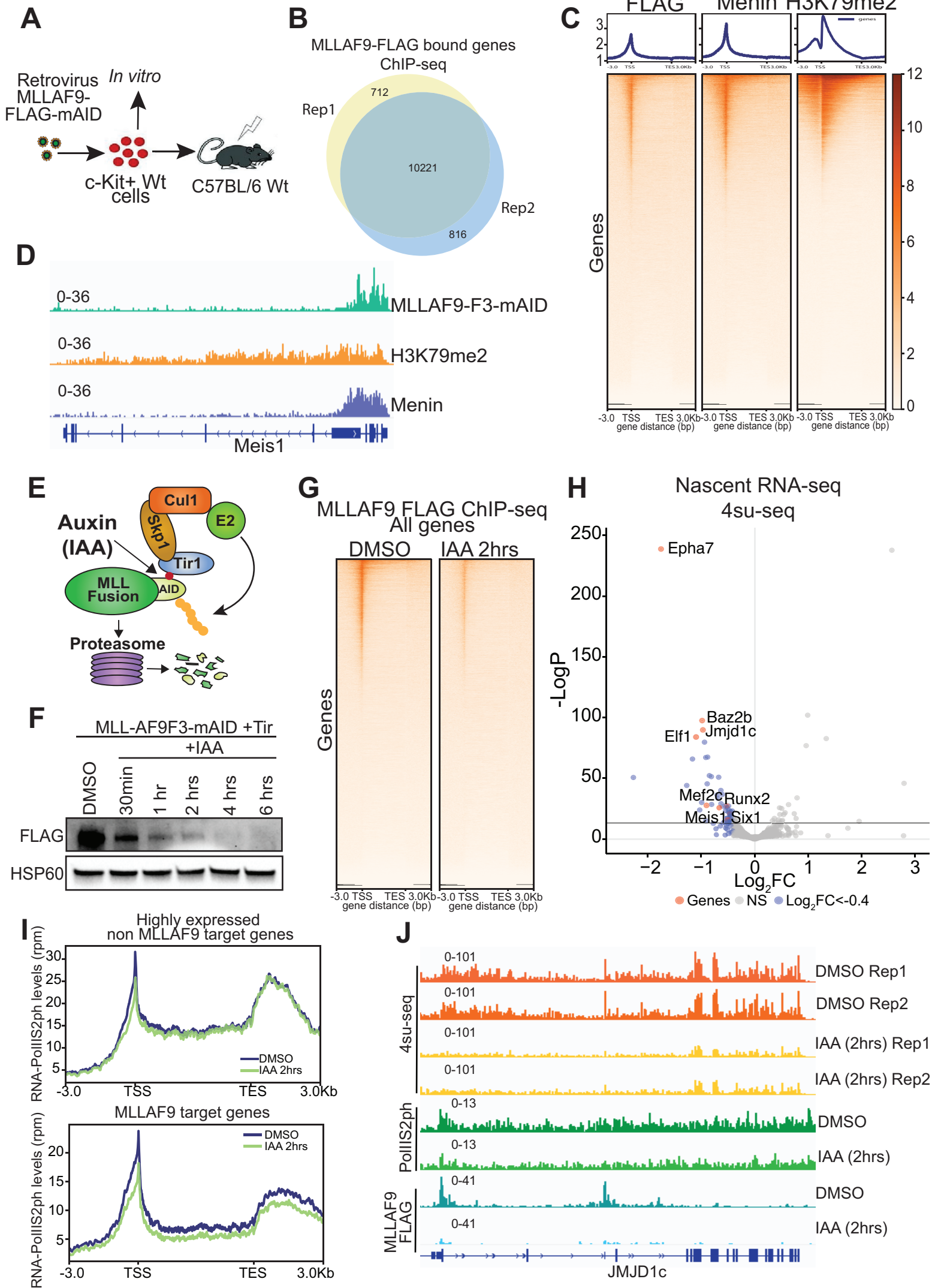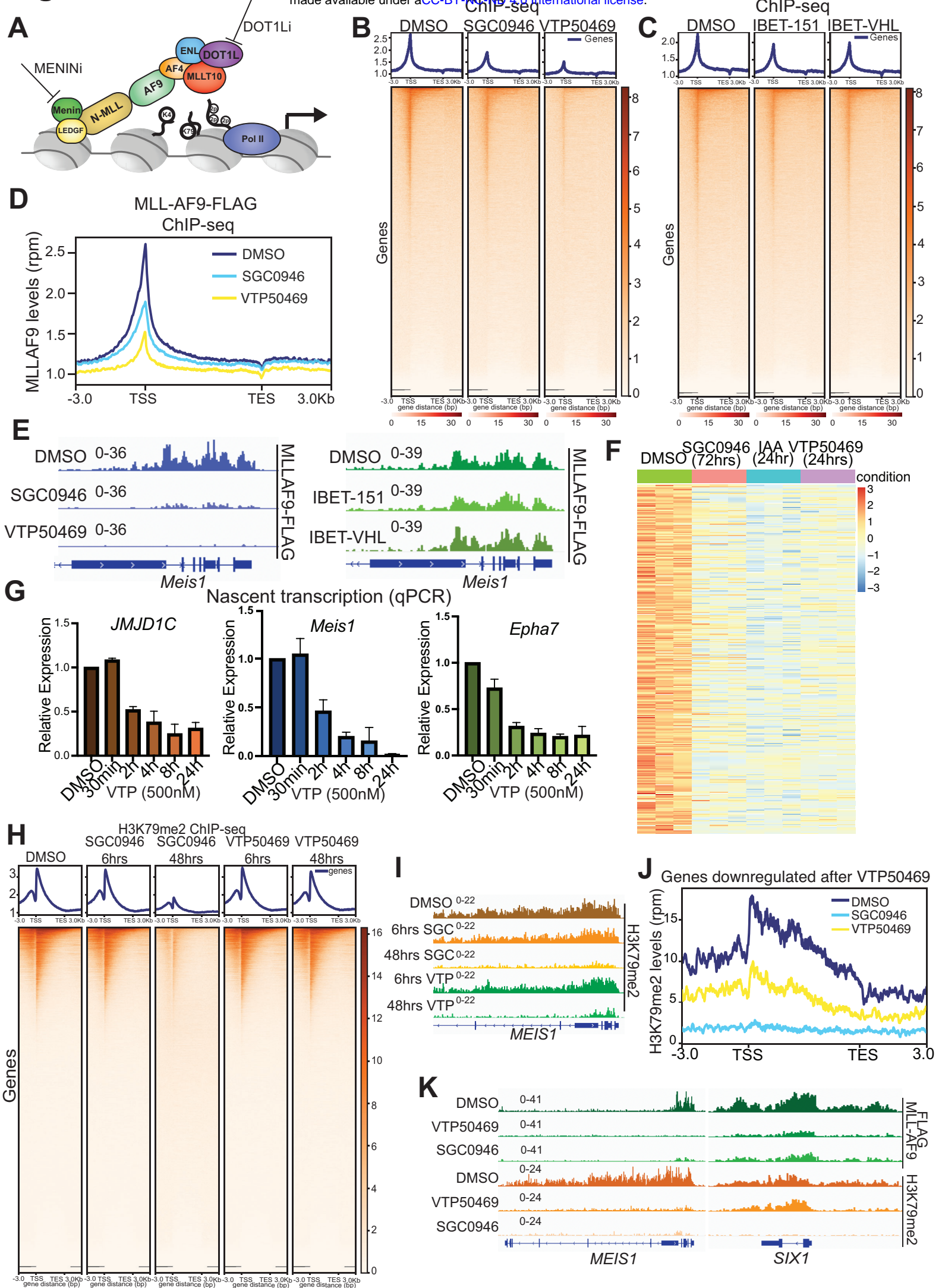
Figure 8
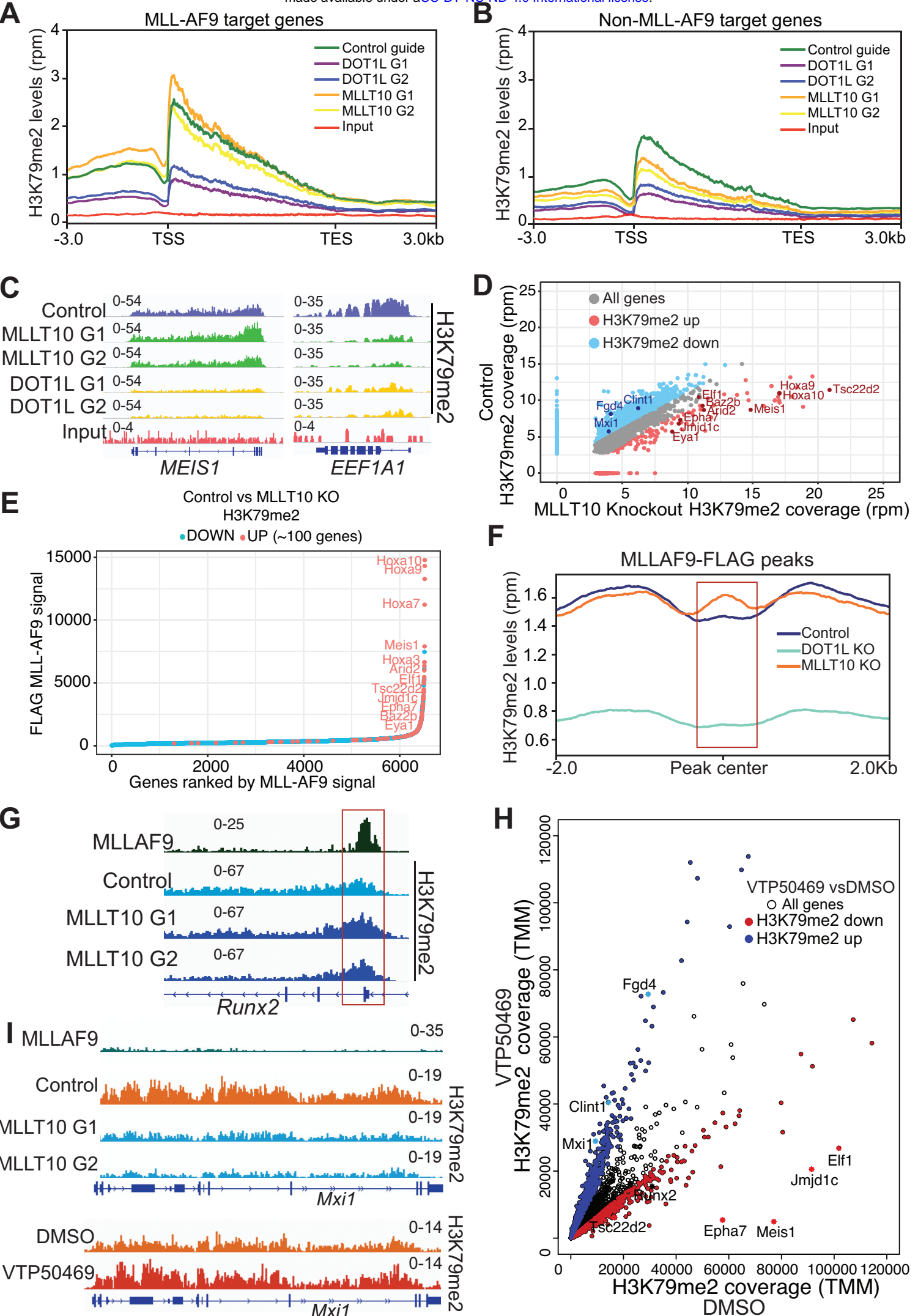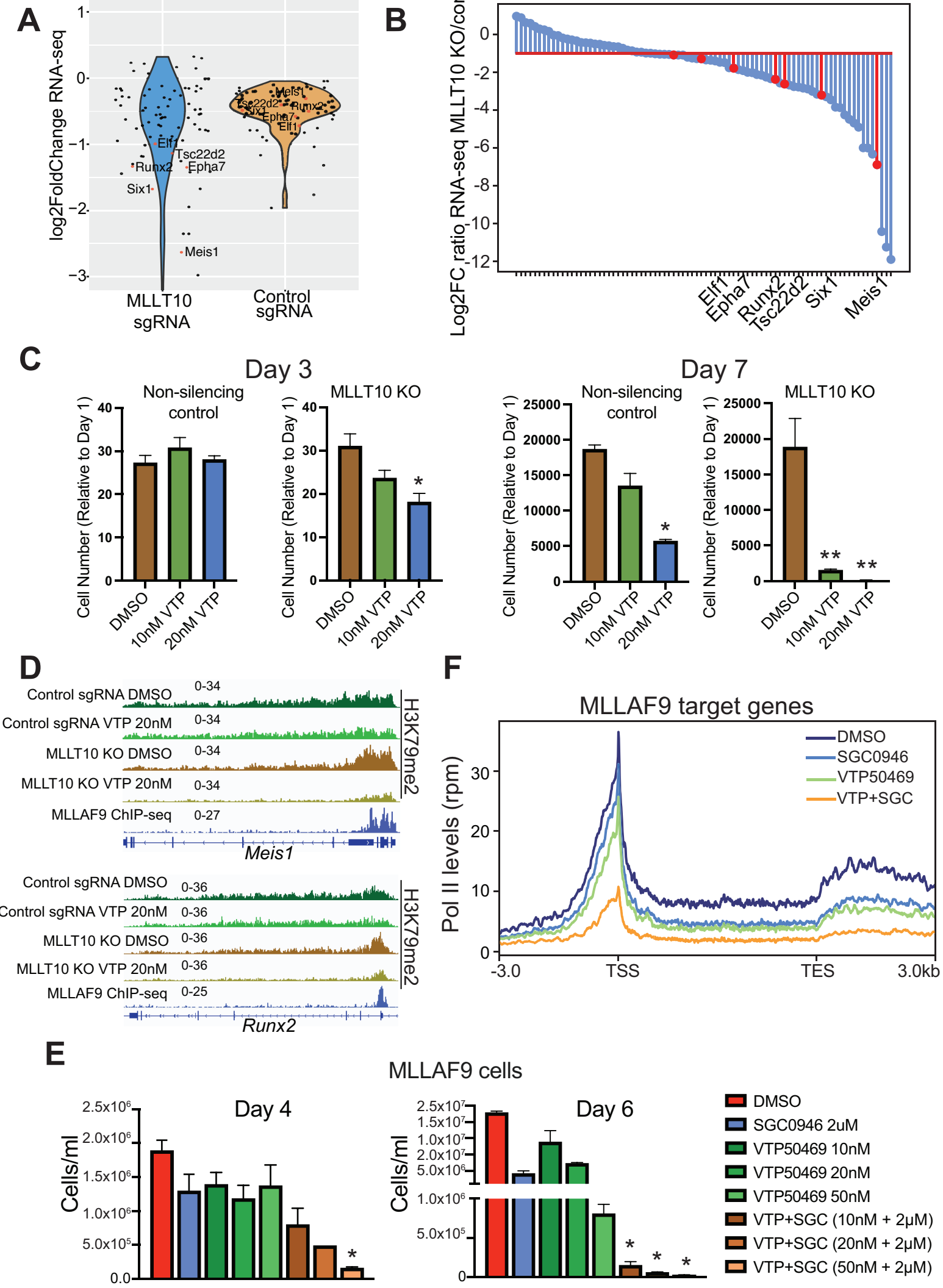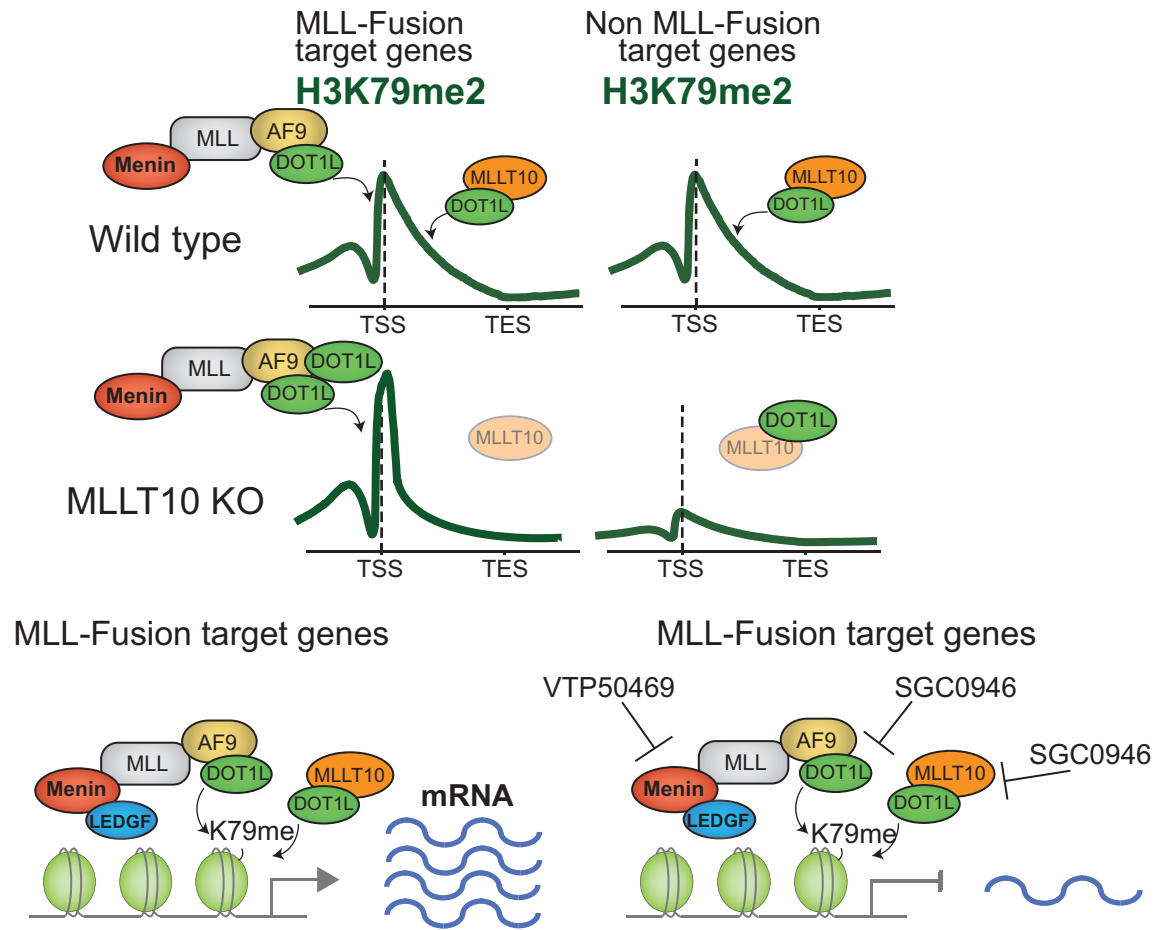
Extended Data Fig 1

Extended Data Fig 2

Extended Data Fig. 3

A

B

C

D

E
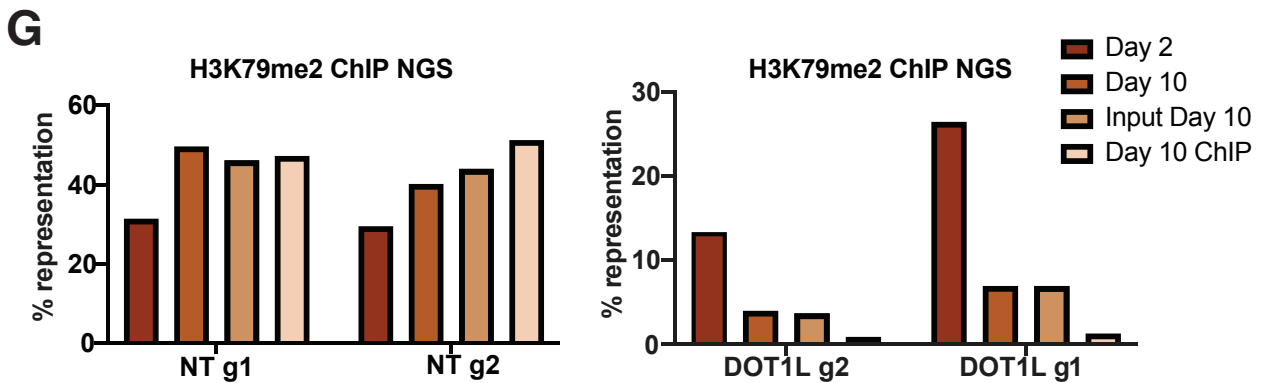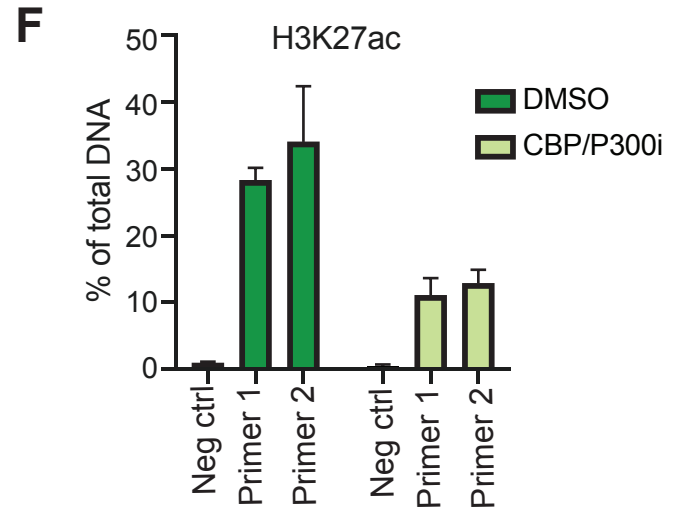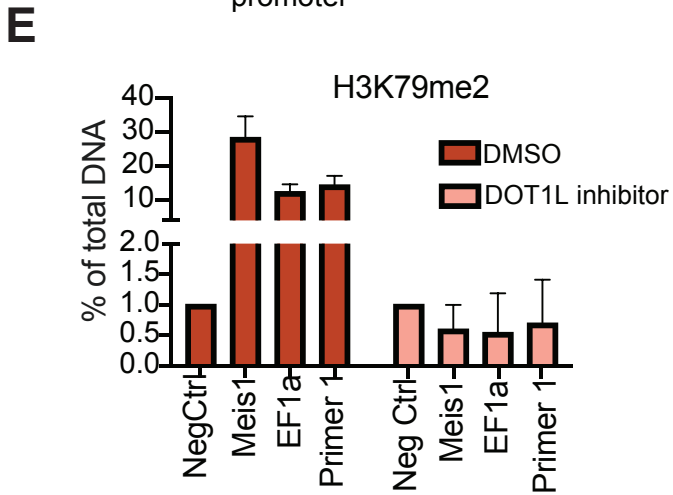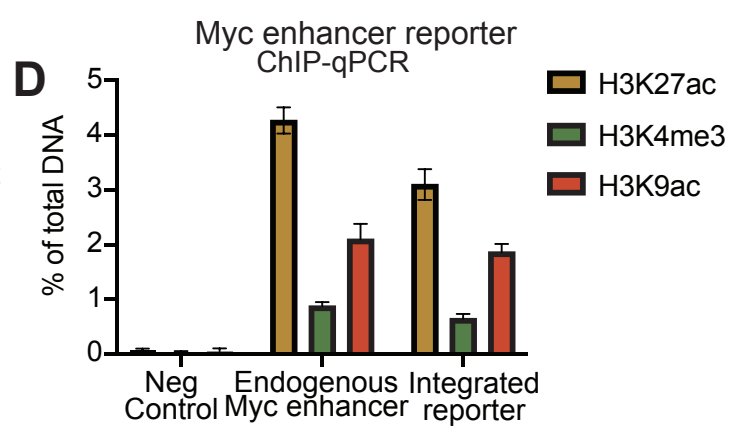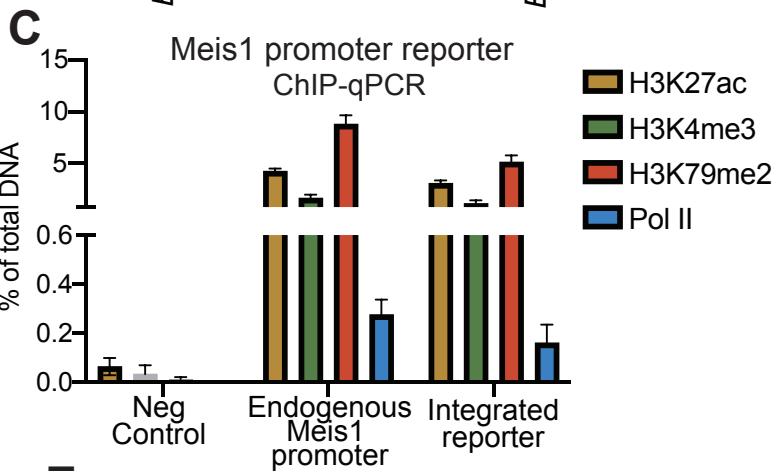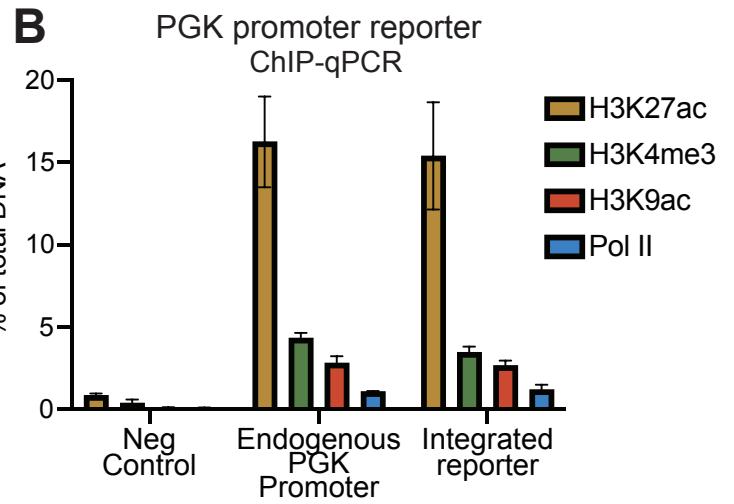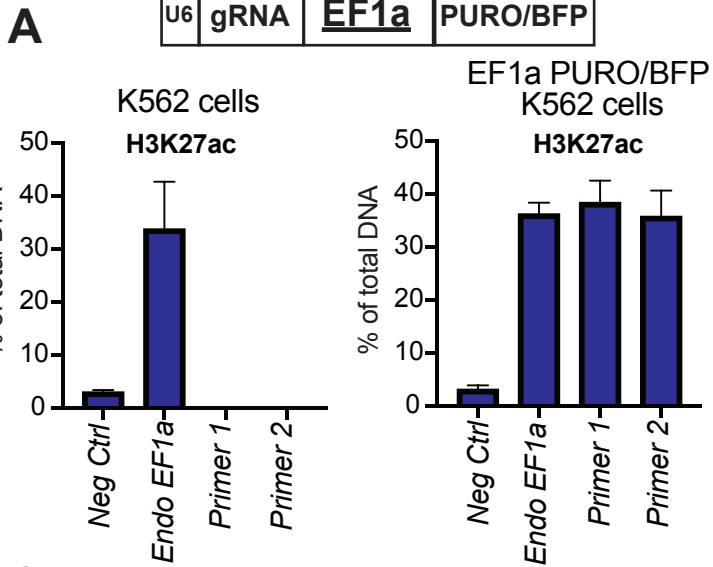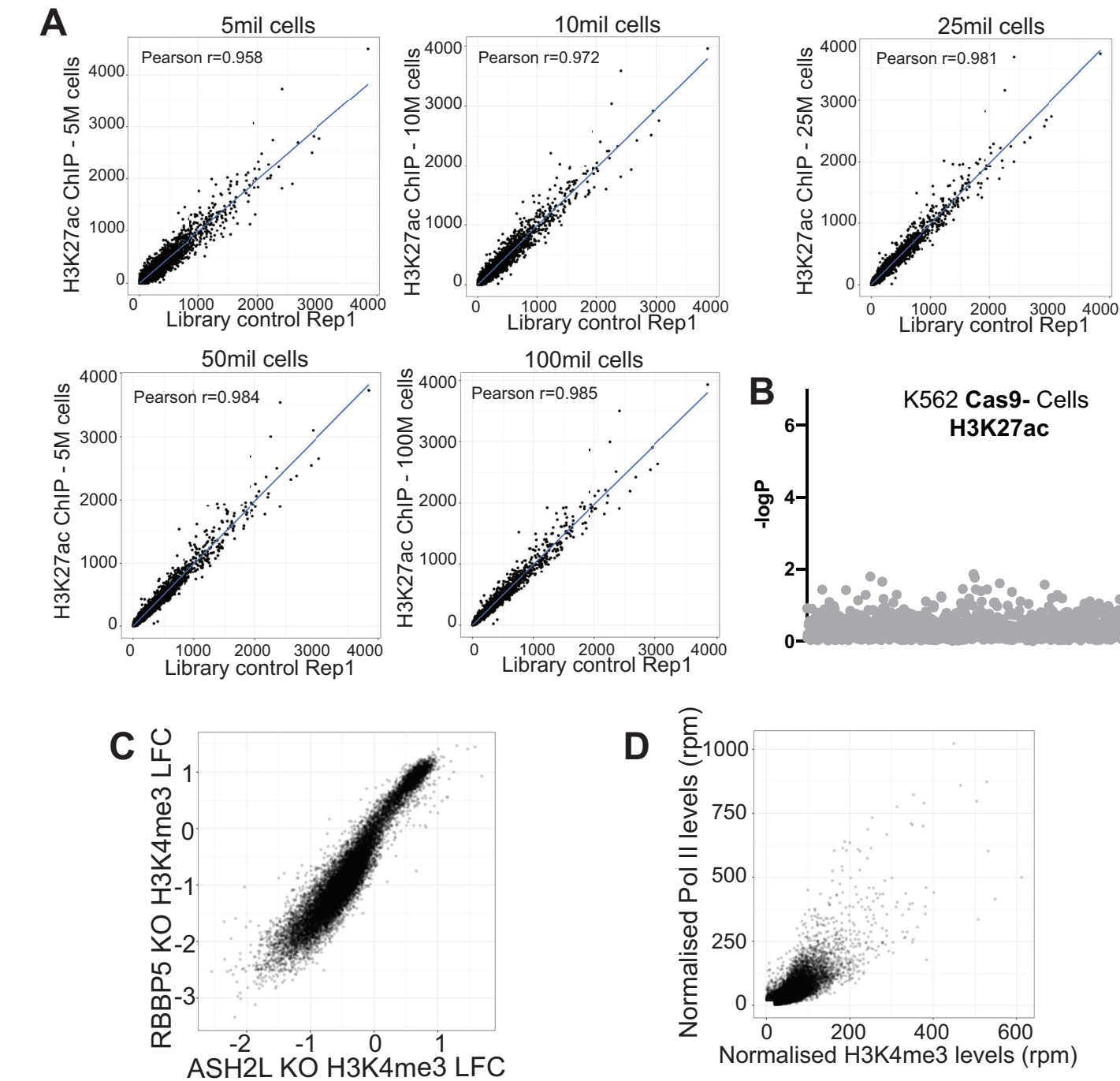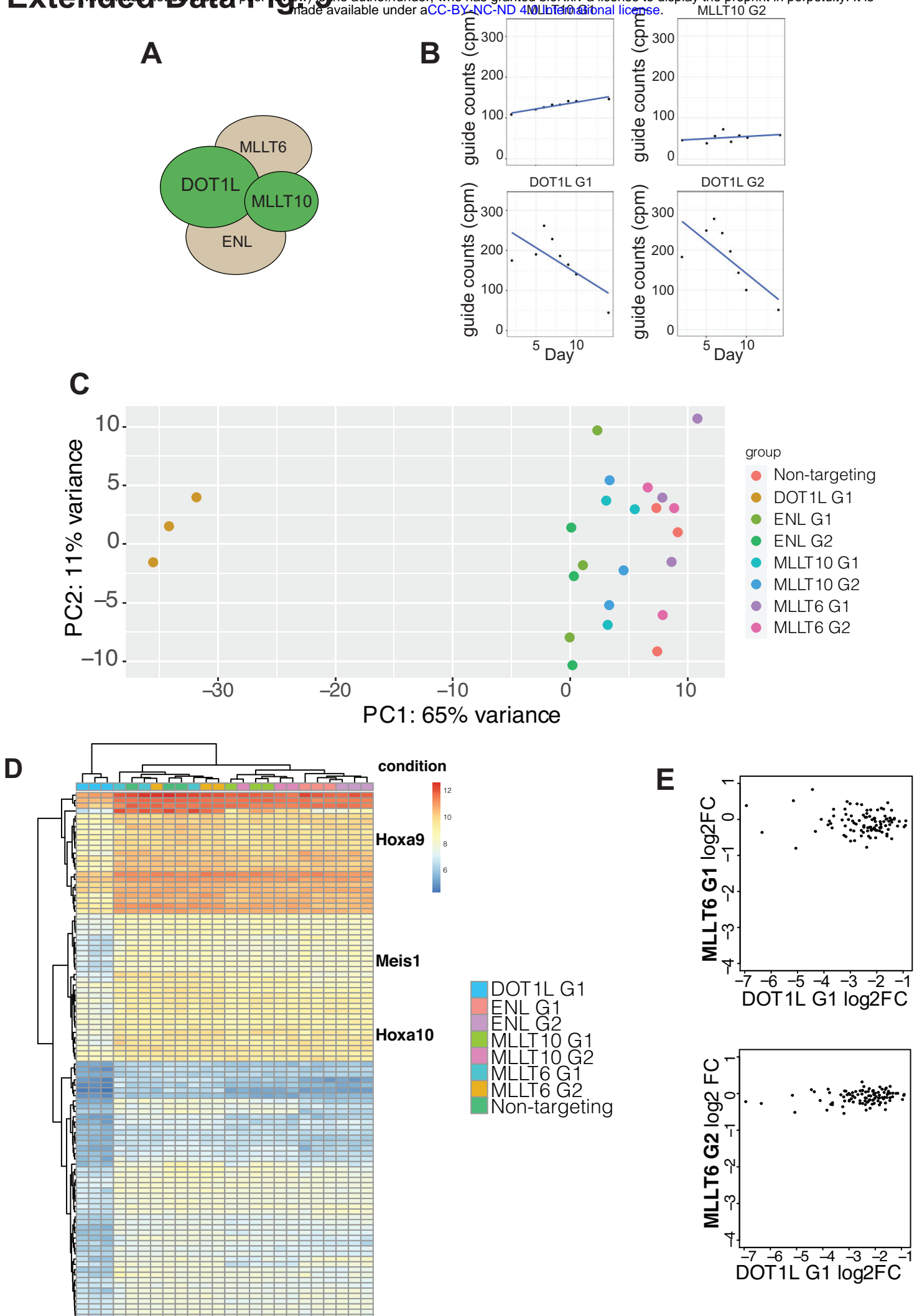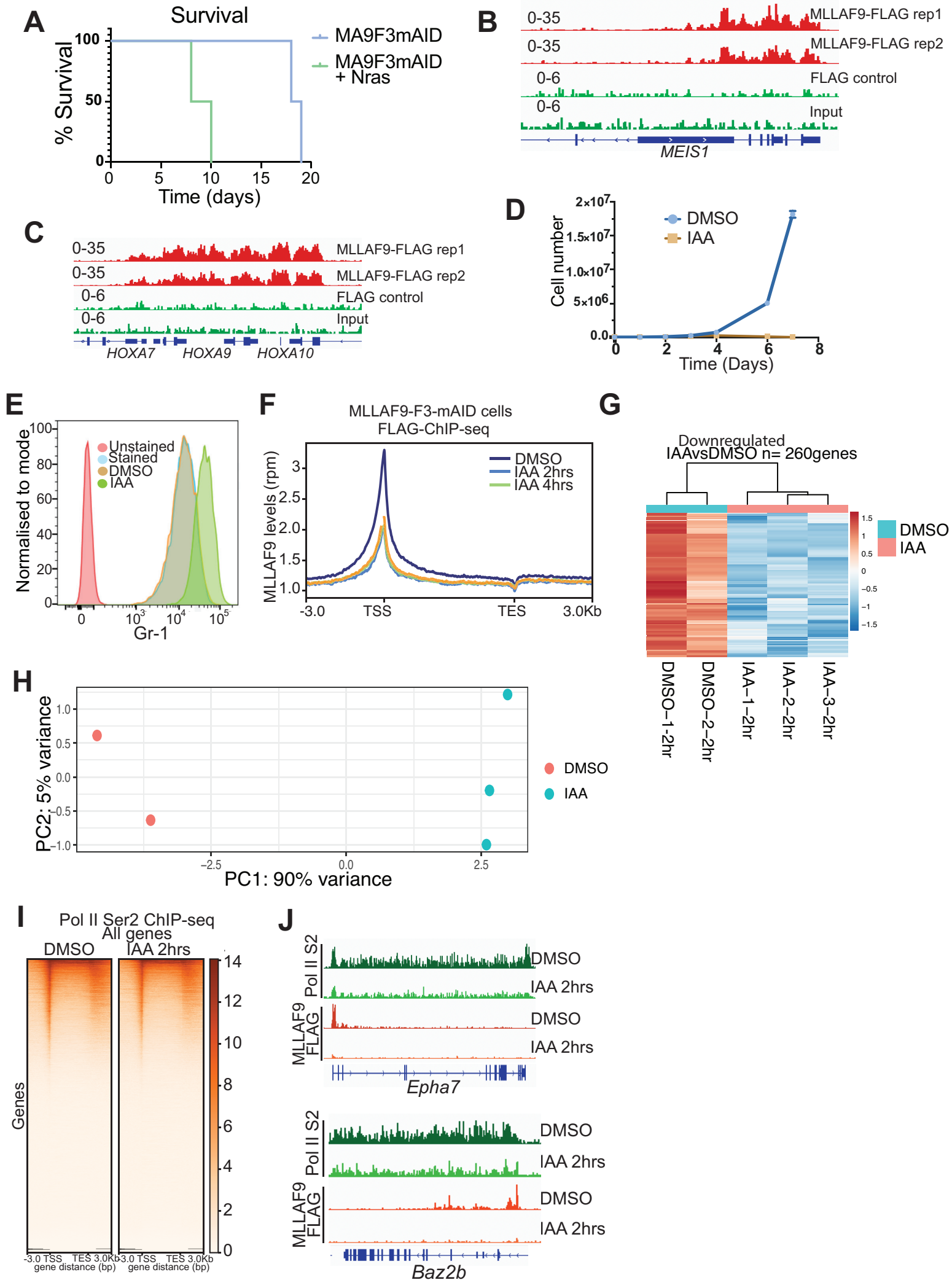
Extended Data Fig 4

Extended Data Fig. 5
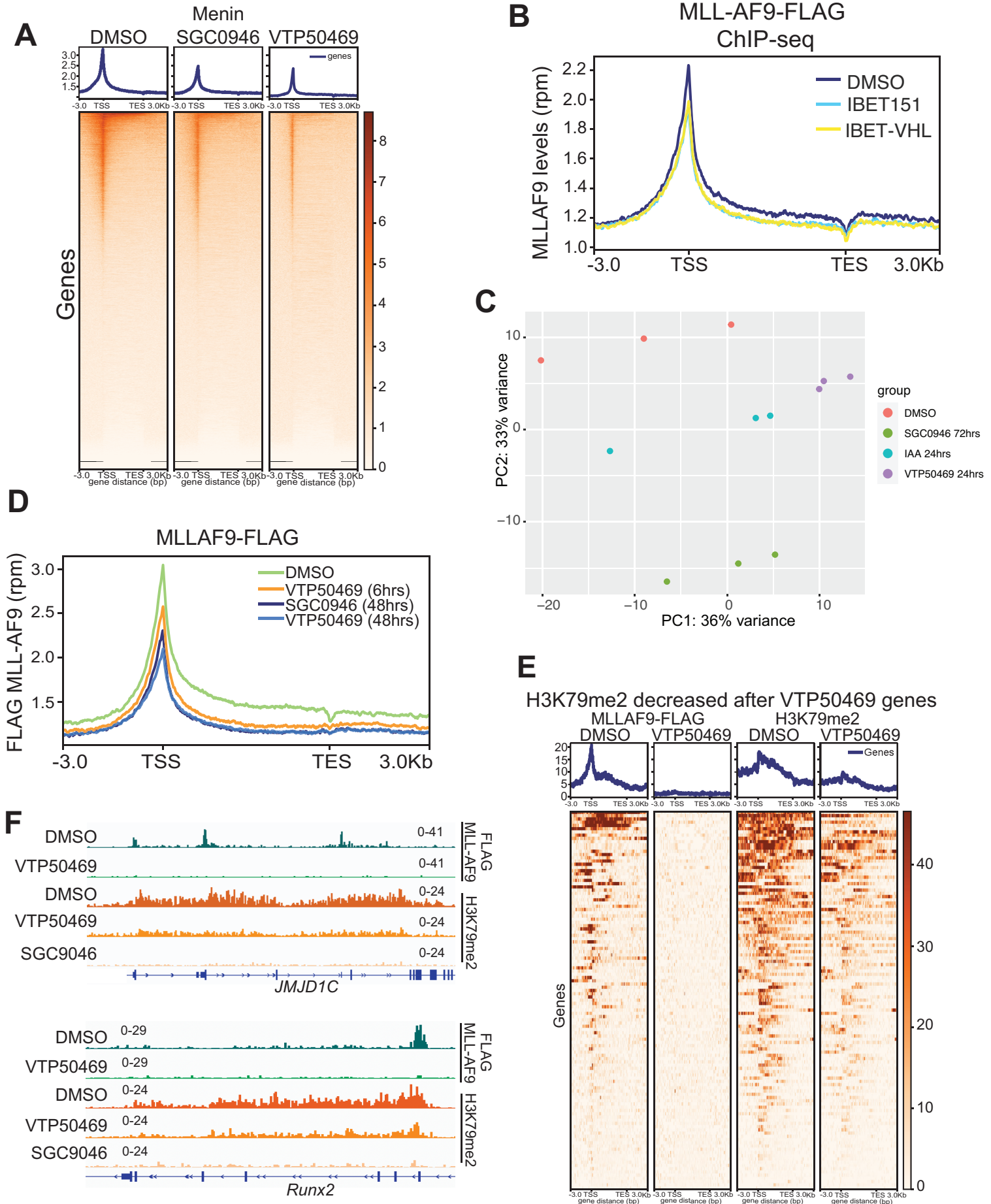
Extended Data Fig. 6

Extended Data Fig. 7