

# Promoter-bound METTL3 maintains myeloid leukaemia by m<sup>6</sup>A-dependent translation control

Isaia Barbieri<sup>1\*</sup>, Konstantinos Tzelepis<sup>2\*</sup>, Luca Pandolfini<sup>1\*</sup>, Junwei Shi<sup>3†</sup>, Gonzalo Millán-Zambrano<sup>1</sup>, Samuel C. Robson<sup>1†</sup>, Demetrios Aspris<sup>2</sup>, Valentina Migliori<sup>1</sup>, Andrew J. Bannister<sup>1</sup>, Namshik Han<sup>1</sup>, Etienne De Braekeleer<sup>2</sup>, Hannes Ponstingl<sup>2</sup>, Alan Hendrick<sup>4</sup>, Christopher R. Vakoc<sup>3</sup>, George S. Vassiliou<sup>2,5,6§</sup> & Tony Kouzarides<sup>1§</sup>

**N<sup>6</sup>-methyladenosine (m<sup>6</sup>A) is an abundant internal RNA modification in both coding<sup>1</sup> and non-coding RNAs<sup>2,3</sup> that is catalysed by the METTL3–METTL14 methyltransferase complex<sup>4</sup>. However, the specific role of these enzymes in cancer is still largely unknown. Here we define a pathway that is specific for METTL3 and is implicated in the maintenance of a leukaemic state. We identify METTL3 as an essential gene for growth of acute myeloid leukaemia cells in two distinct genetic screens. Downregulation of METTL3 results in cell cycle arrest, differentiation of leukaemic cells and failure to establish leukaemia in immunodeficient mice. We show that METTL3, independently of METTL14, associates with chromatin and localizes to the transcriptional start sites of active genes. The vast majority of these genes have the CAATT-box binding protein CEBPZ present at the transcriptional start site<sup>5</sup>, and this is required for recruitment of METTL3 to chromatin. Promoter-bound METTL3 induces m<sup>6</sup>A modification within the coding region of the associated mRNA transcript, and enhances its translation by relieving ribosome stalling. We show that genes regulated by METTL3 in this way are necessary for acute myeloid leukaemia. Together, these data define METTL3 as a regulator of a chromatin-based pathway that is necessary for maintenance of the leukaemic state and identify this enzyme as a potential therapeutic target for acute myeloid leukaemia.**

To identify RNA-modifying enzymes that are necessary for the survival and proliferation of acute myeloid leukaemia (AML) cells, we performed two independent CRISPR screens. First, we performed an *ex vivo* genome-wide CRISPR dropout screen (Screen 1) using Cas9-expressing mouse primary leukaemia cells driven by a KMT2A–MLT3 (also known as *MLL-AF9*) fusion gene and an *Flt3* internal tandem duplication<sup>6</sup> (Fig. 1a). This screen identified 1,550 dropout targets with a false discovery rate (FDR) of 0.25 (Supplementary Table 1), including 75 genes encoding possible RNA-modifying enzymes whose expression is necessary for growth of primary leukaemia cells (see Methods; Supplementary Table 2).

Cas9-induced indel mutations cause stronger negative selection in dropout screens when critical functional domains are targeted<sup>7</sup>. Therefore, we constructed a custom domain-focused guide RNA (gRNA) library including the 75 RNA-modifying enzymes identified above (Screen 2). This identified, with high confidence, 46 potential RNA-modifying enzymes whose catalytic activity is required for leukaemia cell growth (Fig. 1b and Supplementary Table 2). Although the two screens were highly correlated overall, comparison between them highlights targets whose catalytic activity is specifically required for leukaemia cell growth (Extended Data Fig. 1a).

Two *Mettl* gene family members, *Mettl3* and *Mettl16*, scored very highly, whereas *Mettl4* and *Mettl1* showed substantial but lower negative selection. METTL3 and METTL14 form a complex that catalyses RNA adenosine N<sup>6</sup>-methylation<sup>4</sup>. METTL16 is also an m<sup>6</sup>A methyltransferase<sup>8</sup>. This modification is present in mRNAs<sup>1</sup>, pre-microRNAs<sup>2</sup> and long non-coding RNAs<sup>3</sup>, and it affects mRNA stability<sup>9,10</sup> and translation<sup>11</sup>. Notably, the m<sup>6</sup>A demethylase gene *FTO*, which is required for human leukaemia cell growth<sup>12</sup>, was not identified in Screen 1; this may be explained by the heterogeneous genetic background of human AML cell lines.

We validated our results using growth competition assays with individual gRNAs targeting the regions encoding for the catalytic domains of Mettl3 and Mettl16 (as in Screen 2) in mouse AML cells (Extended Data Fig. 1b). Furthermore, negative selection of gRNAs targeting either early exons (as in Screen 1) or the catalytic domain of Mettl3 was validated in different mouse primary leukaemia cell lines (Extended Data Fig. 1c). Finally, disruption of the catalytic domain of Mettl3 strongly suppressed primary mouse AML cell colony formation (Fig. 1c and Extended Data Fig. 1d). By contrast, targeting *Mettl3* in non-transformed NIH3T3 and primary haematopoietic cells had no significant effect (Extended Data Fig. 1e, f). Our findings indicate that these genes are specifically essential for cell survival in AML and not for general cellular viability.

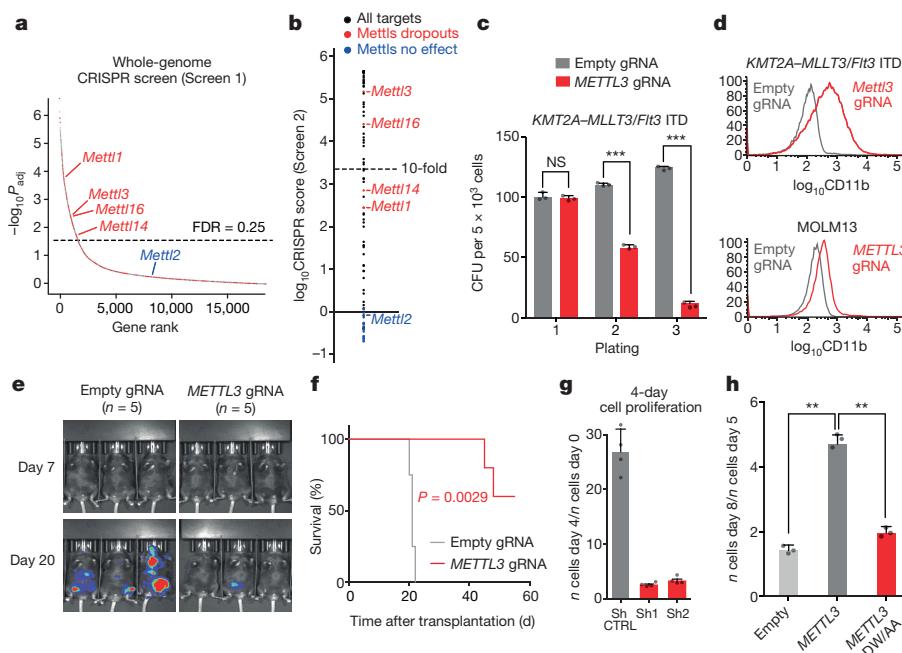
We next targeted *METTL1*, *METTL3*, *METTL14* and *METTL16* in ten human AML cell lines and ten cell lines from heterogeneous cancer types. All four METTLs showed negative selection in all AML cell lines tested (Extended Data Fig. 1g), but displayed varying degrees of negative selection in non-AML tumours (Extended Data Fig. 2a). These differences are not due to variable editing levels across cell lines (Extended Data Fig. 2b).

Disruption of *METTL3* reverses the myeloid differentiation block that is characteristic of AML in both mouse and human AML cells (Fig. 1d and Extended Data Fig. 2c, d). Increased expression of CD11b, a granulocytic differentiation marker<sup>13</sup>, occurred in all METTL3-domain-knockout cells analysed, consistent with METTL3 loss promoting AML cell differentiation. Strikingly, targeting the methyltransferase domain of METTL3 markedly impaired engraftment of human leukaemic cells into immunocompromised mice (Fig. 1e and Extended Data Fig. 2e), with grafted mice surviving significantly longer than controls ( $P=0.0029$ ; Fig. 1f). An independent genetic approach, using human MOLM13 leukaemia cells harbouring inducible *METTL3*-specific short hairpin RNAs (shRNAs), was used to validate our findings. These cells showed near-complete loss of METTL3 mRNA and protein (Extended Data Fig. 3a, b) and markedly reduced proliferation (Fig. 1g) upon

<sup>1</sup>The Gurdon Institute and Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QN, UK. <sup>2</sup>Haematological Cancer Genetics, Wellcome Trust Sanger Institute, Cambridge CB10 1SA, UK. <sup>3</sup>Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, New York 11724, USA. <sup>4</sup>Storm Therapeutics Ltd, Moneta Building (B280), Babraham Research Campus, Cambridge CB22 3AT, UK. <sup>5</sup>Wellcome Trust-MRC Cambridge Stem Cell Institute, University of Cambridge, Cambridge CB2 0XY, UK. <sup>6</sup>Department of Haematology, Cambridge University Hospitals NHS Trust, Cambridge CB2 0QQ, UK. <sup>†</sup>Present addresses: Department of Cancer Biology, Abramson Family Cancer Research Institute, Perelman School of Medicine, University of Pennsylvania, 421 Curie Boulevard, Philadelphia, Pennsylvania 19104, USA (J.S.); School of Pharmacy & Biomedical Science, St Michael's Building, University of Portsmouth, White Swan Road, Portsmouth, UK (S.C.R.).

\*These authors contributed equally to this work.

§These authors jointly supervised this work.



**Figure 1 | METTL3 is essential for AML cells both *in vivo* and *in vitro*.** **a**, Dropout  $P$  values of the genome-wide screen in leukaemic cells driven by *KMT2A-MLLT3* and *Fit3* internal tandem duplication (*KMT2A-MLLT3 Fit3* ITD cells) are displayed. Discontinuous blue line shows the 25% FDR threshold. RNA enzymes are shown as red dots. **b**, CRISPR score for the 75 RNA enzymes (black circles) or the METTL family members (red/blue) as controls in AML-AF9 RN2C cells. **c**, Colony formation assay of *KMT2A-MLLT3 Fit3* ITD cells with Cas9 targeting the catalytic domain of METTL3 showing decreased replating ability compared with control cells. Mean + s.d. of three independent replicates; CFU, colony-forming units. (\*\* $P < 0.001$ ;  $t$ -test). **d**, CD11b expression in METTL3 (catalytic domain-specific)-targeted cells (*KMT2A-MLLT3 Fit3* ITD mouse cells

induction of shRNAs by doxycycline. Similar results were obtained using the human AML cell line THP1 (Extended Data Fig. 3c). Notably, ectopic expression of METTL3 (Extended Data Fig. 3d) fully rescued the proliferation defect, whereas a catalytically inactive mutant failed to do so (Fig. 1h), confirming that loss of growth was due to lack of METTL3 catalytic activity.

RNA sequencing (RNA-seq) of METTL3 knock-down (METTL3-KD) cells showed altered expression of transcripts, both upregulated ( $n = 167$ ) and downregulated ( $n = 180$ ; Extended Data Fig. 3e and Supplementary Table 3). Gene ontology analysis of differentially expressed genes revealed downregulation of cell cycle pathway genes and upregulation of haematopoietic cell differentiation genes, mirroring our findings obtained using CRISPR-Cas9 and flow cytometry analyses (Extended Data Fig. 3f and Supplementary Table 4). These data demonstrate that METTL3 is required for leukaemic cell growth and that its inactivation induces differentiation of *KMT2A-MLLT3*-driven AML cells.

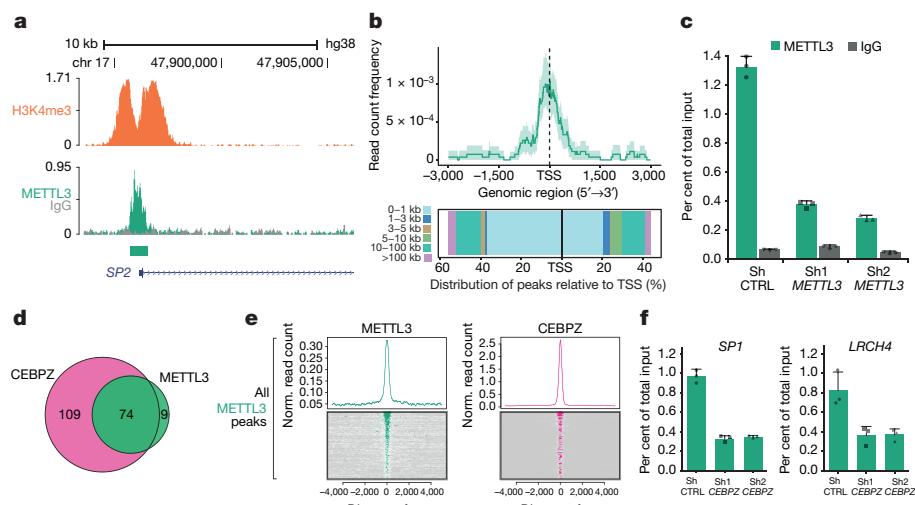
Next we investigated METTL3 expression in different cancer samples from The Cancer Genome Atlas (TCGA)<sup>14</sup>. The highest expression of METTL3 (but not METTL14) was identified in AML cells, diffuse large B-cell lymphoma and prostate adenocarcinoma cells (Extended Data Fig. 4a). Strikingly, these three cancer types correspond to the cell lines that are sensitive to METTL3 targeting (Extended Data Figs 1g, 2a). Furthermore, overexpression of METTL3 increases proliferation in human AML cell lines (Extended Data Fig. 4b). These observations are consistent with METTL3 playing an oncogenic role in AML.

Cytoplasmic signalling enzymes can directly regulate genes by binding chromatin<sup>15</sup>. We therefore tested whether METTL3 binds chromatin. We detected both METTL3 and its partner METTL14 associated with chromatin fractions from MOLM13 cells (Extended Data Fig. 4c).

and MOLM13 human cells) measured by flow cytometry 8 days (mouse) and 6 days (human) after infection. **e**, Bioluminescence imaging of mice transplanted with luciferase-expressing MOLM13 cells transduced with the indicated gRNAs. **f**, Kaplan–Meier plot showing the survival time of mice from Fig. 1e. A log-rank test was performed. ( $n = 5$ ). **g**, Proliferation assay of METTL3-KD or control cells measured between day 4 and day 8 after doxycycline induction. Mean + s.d. of three independent replicates is shown. **h**, Proliferation assay of *Npm1c/Fit3ltd/+*/*Rosa26Cas9/+* mouse leukaemia cells transduced with gRNA targeting the catalytic domain of Mettl3 and plasmids expressing wild-type METTL3 or a catalytically inactive mutant. Mean + s.d. of three independent replicates is shown (\*\* $P < 0.01$ ;  $t$ -test).

To identify genomic loci bound by METTL3 and/or METTL14, we performed chromatin immunoprecipitation (ChIP) using antibodies against METTL3, METTL14 and trimethylation of lysine 4 on histone H3 (H3K4me3). This identified 126 METTL3 and 119 METTL14 genomic peaks (Fig. 2a, Extended Data Fig. 4d and Supplementary Table 5). Both METTL3 and METTL14 localized mainly to the transcriptional start sites (TSSs) of coding genes characterized by bimodal H3K4me3 peaks (for example, Fig. 2a, b and Extended Data Fig. 4e, f). Unexpectedly, the two METTLs did not bind the same TSSs (Extended Data Figs 4g, 5a), suggesting that they have distinct roles on chromatin. Here, we focus on METTL3 and its functions. We validated several METTL3 peaks by ChIP with quantitative PCR (ChIP-qPCR) using two different METTL3-specific antibodies (Extended Data Fig. 5b, c) and confirmed that METTL3-KD led to reduction of the METTL3 ChIP signal at specific promoters (Fig. 2c and Extended Data Fig. 5d).

We interrogated sequences under METTL3 peaks for enriched motifs and identified the CCAAT box as the top hit (Extended Data Fig. 6a). This motif binds the NFYA–NFYB complex<sup>16</sup>, which associates with H3R2me2s (symmetric dimethyl-histone H3 (Arg2)) and the WD repeat protein WDR5 on active promoters<sup>17</sup>. Another significantly enriched METTL3 peak motif is that for the transcription factor KLF9 ( $P = 1 \times 10^{-4}$ ). Using existing ChIP with sequencing (ChIP-seq) datasets<sup>5</sup>, we observed a high correlation between METTL3 binding sites and all of these factors (NFYA–NFYB, H3R2me2s, WDR5 and KLF9; see Extended Data Fig. 6b). METTL14 peaks showed no correlation with the same factors, except for WDR5 (Extended Data Fig. 6b). Co-occupancy by all of these factors on 447 promoters overlaps with METTL3 binding (Extended Data Fig. 6c, d), but their combined predictive power for defining METTL3 bound sites is limited.

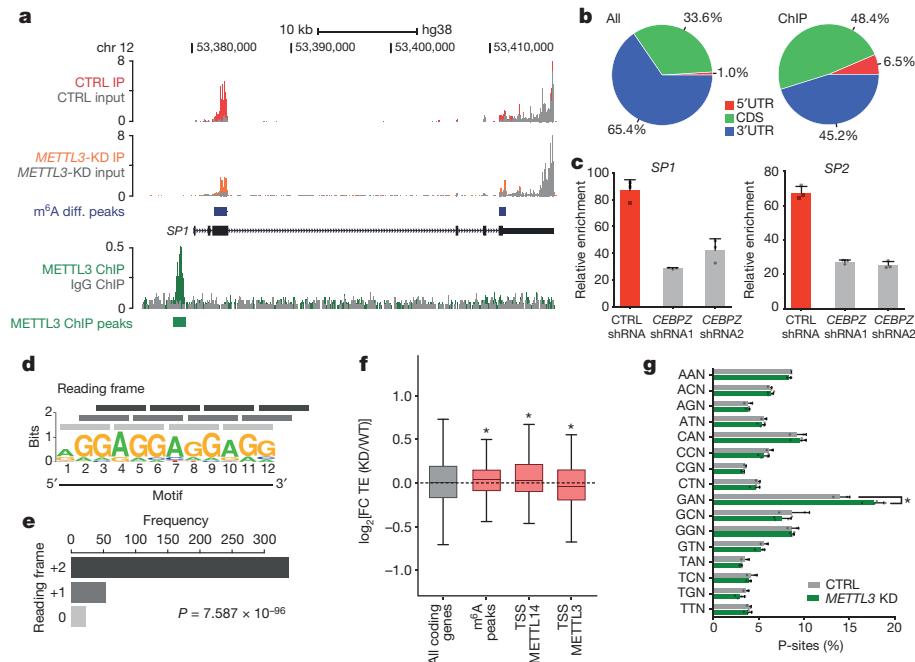


**Figure 2 | METTL3 localizes to specific TSSs on chromatin.** **a**, Genomic visualization of METTL3 and H3K4me3 ChIP-seq dataset at the *SP2* locus. **b**, Distribution of METTL3 ChIP-seq reads centred on TSSs (upper panel) and histogram of ChIP-seq reads distribution relative to TSSs (lower panel). **c**, METTL3 ChIP-seq validation by ChIP-qPCR on the *SP2* TSS in control or *METTL3*-KD MOLM13 cells, showing a specific reduction in METTL3 binding in *METTL3*-KD cells. Mean + s.d. of three technical replicates; experiment performed independently three times.

We then selected the genes that were most tightly co-expressed with METTL3 (top 2.5th percentile) in normal and cancer cell lines<sup>18</sup>. Eleven AML ENCODE ChIP-seq datasets are available for these co-expressed

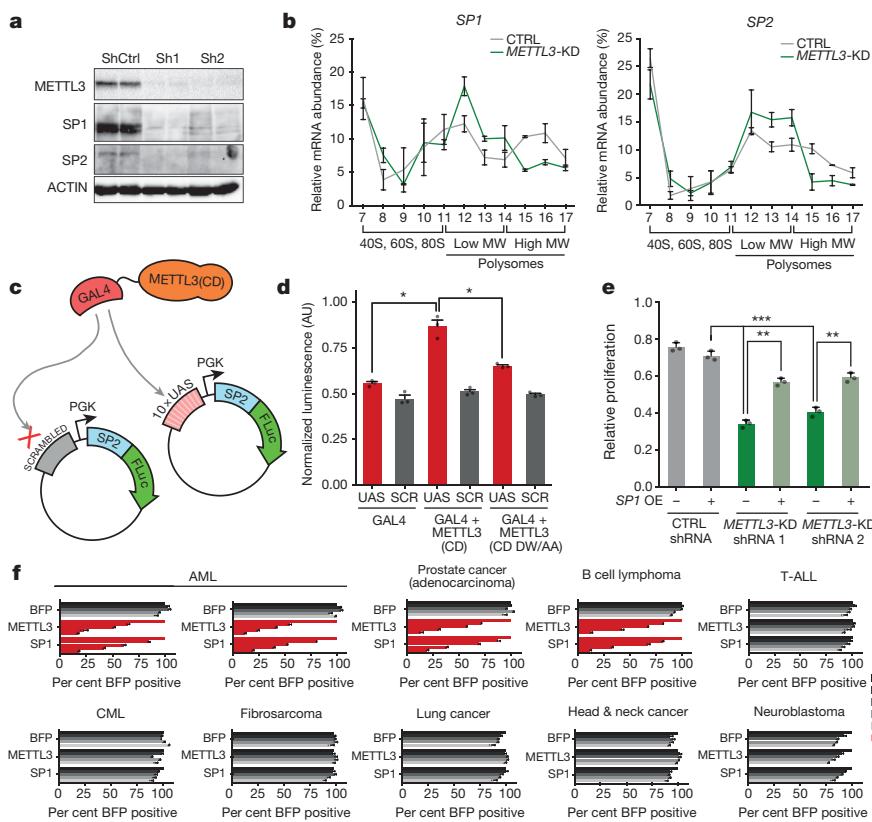
**d**, Venn diagram showing the overlap between CEBPZ and METTL3 ChIP-seq peaks. **e**, Distribution and heatmaps of normalized ChIP-seq reads for CEBPZ centred on METTL3 peaks. **f**, ChIP-qPCR of METTL3 binding on target TSSs in MOLM13 cells, expressing a control shRNA or one of two independent shRNAs against CEBPZ, showing a specific reduction of METTL3 binding in CEBPZ-KD cells. Mean + s.d. of three technical replicates is shown; experiment was performed independently three times.

factors<sup>5</sup>, including WDR5 (Extended Data Fig. 7a). Amongst them, CEBPZ shows high co-expression (Extended Data Fig. 7b) and ChIP co-localization with METTL3 (Fig. 2d, e and Extended Data Fig. 7c, d),



**Figure 3 | Transcripts derived from METTL3-bound promoters harbour m<sup>6</sup>A within their coding sequence (CDS).** **a**, Genomic visualization of the m<sup>6</sup>A immunoprecipitation normalized signal in *METTL3*-KD or control MOLM13 cells on the *SP1* transcript (upper tracks), along with the genomic visualization of the METTL3 ChIP-seq. **b**, Pie charts of the distribution of METTL3-dependent m<sup>6</sup>A peaks within the whole transcriptome or METTL3 chromatin target mRNAs. **c**, m<sup>6</sup>A immunoprecipitation followed by qPCR for m<sup>6</sup>A peaks of *SP1* and *SP2*. The plot shows the m<sup>6</sup>A immunoprecipitation signal over total input in MOLM13 cells expressing a control shRNA or shRNAs targeting CEBPZ. Mean + s.d. of three technical replicates shown; experiment

performed independently twice. **d**, Motif enriched in mRNAs from METTL3-bound TSSs. **e**, Reading frame distribution of the [GAG]<sub>n</sub> motif on transcripts produced at METTL3-bound TSSs. Significance was obtained by multinomial test. **f**, Box plot showing the difference in translational efficiency (TE) between *METTL3*-KD and control (WT) cells. The distributions of log<sub>2</sub>[fold change (FC)] for all coding genes, mRNAs harbouring METTL3-dependent m<sup>6</sup>A and mRNAs originating from METTL3-bound or METTL14-bound TSSs are shown (\*P < 0.05; Wilcoxon test). FC, fold change. **g**, Frequency of P-site occupancy of GAN codons in *METTL3*-KD or control MOLM13 cells (\*P < 0.05; t-test).



**Figure 4 | Negative effect of METTL3 depletion on the translation efficiency of genes necessary for AML growth.** **a**, Western blot showing METTL3, SP1, SP2 and ACTIN protein levels in MOLM13 cells infected with METTL3-specific (Sh1, Sh2) or control (ShCtrl) TET-inducible shRNAs 8 days after doxycycline treatment. Two independent biological replicates are shown. For gel source data see Supplementary Information. **b**, SP1 and SP2 mRNAs in each ribosome fraction were quantified through qPCR and plotted as a percentage of the total. Data are from two independent polysome-profiling experiments. Mean  $\pm$  s.e.m. are shown. **c**, Schematic representation of the engineered reporter system. **d**, Firefly luciferase activity from UAS or scrambled (SCR) sequence carrying plasmid in the presence of GAL4 either alone or fused with METTL3

wild-type (CD) or inactive (CD DW/AA) catalytic domain ( $*P < 0.05$ ; *t*-test). The mean  $\pm$  s.d. of three independent transfections of HT-29 cells is shown. **e**, Proliferation assay of MOLM13 cells infected with *METTL3*-specific or control TET-inducible shRNAs and with an *SP1* expression vector between day 3 and day 6 after doxycycline treatment. Mean  $\pm$  s.d. of three independent replicates is shown. **f**, Competitive co-culture assay showing negative selection of blue fluorescent protein (BFP)<sup>+</sup> human tumour cell lines upon targeting of *METTL3* or *SP1* by CRISPR-Cas9. Results were normalized to day 4 for each gRNA. Mean  $\pm$  s.d. of two independent infections is shown. CML, chronic myeloid leukaemia; T-ALL, T cell acute lymphoblastic leukaemia.

but not with METTL14 (Extended Data Fig. 7e). Notably, CEBPZ is a CCAAT-box binding factor and was amongst the top dropouts in the genomic CRISPR-Cas9 screen (Supplementary Table 1 and Extended Data Fig. 7f). *CEBPZ* knockdown in MOLM13 cells (Extended Data Fig. 7g) impaired cell proliferation (Extended Data Fig. 7h), as shown for *METTL3*.

To test whether CEBPZ is involved in recruiting METTL3 to TSSs, we performed METTL3 ChIP-qPCR experiments after *CEBPZ* knockdown. Figure 2f and Extended Data Figure 7i show reduction of chromatin-bound METTL3 at several promoters in *CEBPZ*-KD cells. Thus, METTL3 and METTL14 are recruited to highly specific chromatin sites, with METTL3 localizing to TSSs with a well-defined signature of transcription factors and histone modifications. This suggests that METTL3 regulates the expression of mRNAs derived from its chromatin target genes.

Transcriptional profiling of METTL3-depleted cells showed that the levels of mRNAs transcribed from METTL3-bound genes were not affected (Extended Data Fig. 7j and Supplementary Table 3; odds ratio (OR) = 0.75; *P* = 1). Given that METTL3 catalytic activity is essential for AML cell growth (Fig. 1h), we investigated post-transcriptional events that could be regulated by METTL3-mediated RNA methylation.

To identify RNAs methylated by METTL3 in AML cells, we performed RNA-immunoprecipitation linked to high throughput sequencing using an m<sup>6</sup>A-specific antibody (m<sup>6</sup>A-IP). We used RNA isolated

from MOLM13 cells (control) and from METTL3-depleted cells. This process identified 4,085 METTL3-dependent m<sup>6</sup>A peaks on poly-A<sup>+</sup>-enriched RNA (Supplementary Table 6, Extended Data Fig. 8a, b). As expected, we observed a widespread reduction in m<sup>6</sup>A upon *METTL3*-KD (Extended Data Fig. 8c).

We then considered the distribution of m<sup>6</sup>A in transcripts derived specifically from METTL3-bound genes, 72.4% of which contained m<sup>6</sup>A peaks, compared to 38.4% of all transcripts (OR = 4.2; *P* =  $5 \times 10^{-4}$ ). The levels of m<sup>6</sup>A on mRNA encoding the transcription factor SP1 are shown in Fig. 3a. The majority of METTL3-dependent m<sup>6</sup>A occurs within the coding regions of transcripts of METTL3-bound genes, in contrast to its distribution in the general transcriptome, where it is enriched within mRNA 3' untranslated regions (UTRs) (Fig. 3b and Extended Data Fig. 8d).

To investigate whether promoter-bound METTL3 is required for m<sup>6</sup>A modification of associated transcripts, we measured m<sup>6</sup>A modification in these mRNAs following *CEBPZ*-KD, as this leads to loss of METTL3 at TSSs (Fig. 2f). *CEBPZ*-KD reduced m<sup>6</sup>A in relevant mRNAs (Fig. 3c and Extended Data Fig. 8e). By contrast, no changes were observed in control mRNAs such as *GAPDH*. As expected, *CEBPZ*-KD did not affect mRNA levels of METTL3 chromatin targets or levels of *METTL3* itself (Extended Data Fig. 8f).

Analysis of the whole coding sequences of these transcripts revealed enrichment of a [GAG]<sub>n</sub> motif (Fig. 3d). This motif, while common

throughout coding transcripts, is significantly overrepresented amongst transcripts derived from METTL3-bound genes ( $P=0.0098$ ; Extended Data Fig. 8g). Notably, the motif's reading frame (+2) is preserved throughout transcripts derived from METTL3-bound genes (Fig. 3e). Although the relevance of this motif remains unclear, its presence suggests that these GAG-rich transcripts require m<sup>6</sup>A modification for translational efficiency. We therefore performed ribosome footprinting (RFP) of control and *METTL3*-KD MOLM13 cells, to evaluate their translational output (Extended Data Fig. 9a–c). Although mRNAs marked with m<sup>6</sup>A generally tended to show increased translational efficiency following *METTL3*-KD<sup>19</sup>, transcripts derived from genes harbouring METTL3 on their promoter were translated less efficiently (Fig. 3f and Supplementary Table 7). We therefore mapped ribosomal pausing sites on mRNAs produced from METTL3-bound genes. P-site codon occupancy in these transcripts indicated that four codons—GAG, GAT, GAC, GAA (GAN codons)—are more occupied by ribosomes in *METTL3*-KD cells than in control cells (Fig. 3g and Extended Data Fig. 9d). The same was not observed as a general feature throughout the transcriptome (Extended Data Fig. 9e, f) and is not due to general overrepresentation of GAN codons throughout transcripts (Extended Data Fig. 9g).

To better characterize how METTL3 regulates translation, we focused on two genes encoding the transcription factors SP1 and SP2, which have promoters occupied by METTL3. The levels of SP1 and SP2 proteins are reduced upon METTL3 depletion (Fig. 4a), but their mRNA levels are unaffected (Extended Data Fig. 10a). Indeed, a similar reduction in SP1 protein levels is observed upon *CEBPZ*-KD (Extended Data Fig. 10b). We next investigated how *METTL3*-KD affects the association of SP1 and SP2 mRNA with polysomes (Extended Data Fig. 10c). Upon METTL3 depletion, there was a specific shift of SP1 and SP2 transcripts towards lower molecular weight polysomes (Fig. 4b), indicating less efficient translation, whereas there were no differences for control mRNAs (Extended Data Fig. 10d).

The above results demonstrate that METTL3 affects translation of the mRNAs whose promoters it occupies. We next aimed to prove that recruitment of METTL3 to promoters is sufficient for this effect. For this, we used a reporter system (Fig. 4c) consisting of a plasmid harbouring GAL4 binding sites upstream of a constitutive promoter that expresses an SP2 (m<sup>6</sup>A peak)–luciferase mRNA fusion. We expressed the METTL3 wild-type catalytic domain, or an inactive point mutant<sup>20</sup>, in-frame with the GAL4 DNA-binding domain. Consistently, only recruitment of wild-type METTL3–GAL4 enhanced luciferase activity (Fig. 4d and Extended Data Fig. 10e), without affecting luciferase mRNA levels (Extended Data Fig. 10f). These data corroborate the theory that promoter-bound METTL3 augments translation of its target genes.

Consistent with depletion of SP1 and SP2 protein in *METTL3*-KD cells, genes directly bound and regulated by these factors<sup>5</sup> were also generally downregulated in METTL3-depleted MOLM13 cells (Extended Data Fig. 10g). Amongst these is the *c-MYC* oncogene, whose promoter is bound by both SP1 and SP2 (Extended Data Fig. 10h and Supplementary Table 3). We tested whether loss of SP1 could explain the proliferation defect caused by METTL3 depletion. Figure 4e shows that ectopic overexpression of SP1 in *METTL3*-KD cells (Extended Data Fig. 10i) rescues cell growth. Strikingly, CRISPR targeting of SP1 is lethal in cell lines that are sensitive to METTL3 inactivation, but not in other cell lines (Fig. 4f). These observations suggest that the sensitivity of leukaemic cells to loss of METTL3 is dictated by the requirement of SP1 for cell growth.

Here we define a set of RNA-modifying enzymes that are necessary for AML leukaemia and identify a leukaemic pathway for the METTL3 RNA methyltransferase. In this pathway, METTL3 is stably recruited by CEBPZ to promoters of a specific set of active genes, resulting in m<sup>6</sup>A methylation of the respective mRNAs and increased translation. One important target is *SP1*, an oncogene in several cancers<sup>21</sup>, which regulates *c-MYC* expression<sup>22</sup>. Consistent with these findings, it has been reported that METTL3 can methylate its targets co-transcriptionally<sup>23</sup>.

METTL3 affects mRNA translation in numerous ways, including by promoting RNA loading onto ribosomes<sup>24</sup> and recruiting specific m<sup>6</sup>A reader proteins (for example, YTHFD1)<sup>25</sup>. The findings presented here provide substantial insight into the mechanisms through which METTL3 promotes post-transcriptional regulation of gene expression. Our data identify a new paradigm for selecting RNAs to be modified, namely the stable recruitment of the RNA-modifying enzyme to specific genomic loci. They also demonstrate a mechanism by which m<sup>6</sup>A can affect translation, namely relief of ribosome stalling at GAN codons of specific transcripts.

The pathway described here is critical for AML leukaemia, as three of its components are required for AML cell growth: (i) the m<sup>6</sup>A RNA methyltransferases METTL3; (ii) the transcription factor CEBPZ, which targets this enzyme to promoters; and (iii) SP1, whose translation is dependent upon the m<sup>6</sup>A modification by METTL3. Together, these observations define METTL3 enzymatic activity as a new candidate target for the treatment of AML.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 7 February; accepted 25 October 2017.

Published online 27 November 2017.

- Dominissini, D. et al. Topology of the human and mouse m<sup>6</sup>A RNA methylomes revealed by m<sup>6</sup>A-seq. *Nature* **485**, 201–206 (2012).
- Alarcón, C. R., Lee, H., Goodarzi, H., Halberg, N. & Tavazoie, S. F. N<sup>6</sup>-methyladenosine marks primary microRNAs for processing. *Nature* **519**, 482–485 (2015).
- Patil, D. P. et al. m<sup>6</sup>A RNA methylation promotes XIST-mediated transcriptional repression. *Nature* **537**, 369–373 (2016).
- Liu, J. et al. A METTL3–METTL14 complex mediates mammalian nuclear RNA N<sup>6</sup>-adenosine methylation. *Nat. Chem. Biol.* **10**, 93–95 (2014).
- Dunham, I. et al.; ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
- Tzelepis, K. et al. A CRISPR dropout screen identifies genetic vulnerabilities and therapeutic targets in acute myeloid leukemia. *Cell Rep.* **17**, 1193–1205 (2016).
- Shi, J. et al. Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat. Biotechnol.* **33**, 661–667 (2015).
- Pendleton, K. E. et al. The U6 snRNA m<sup>6</sup>A methyltransferase METTL16 regulates SAM synthetase intron retention. *Cell* **169**, 824–835 (2017).
- Batista, P. J. et al. m<sup>6</sup>A RNA modification controls cell fate transition in mammalian embryonic stem cells. *Cell Stem Cell* **15**, 707–719 (2014).
- Wang, Y. et al. N<sup>6</sup>-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nat. Cell Biol.* **16**, 191–198 (2014).
- Meyer, K. D. et al. 5' UTR m<sup>6</sup>A promotes cap-independent translation. *Cell* **163**, 999–1010 (2015).
- Li, Z. et al. FTO Plays an oncogenic role in acute myeloid leukemia as a N<sup>6</sup>-methyladenosine RNA demethylase. *Cancer Cell* **31**, 127–141 (2017).
- Ripperger, T. et al. The heteromeric transcription factor GABP activates the ITGAM/CD11b promoter and induces myeloid differentiation. *Biochim. Biophys. Acta* **1849**, 1145–1154 (2015).
- Weinstein, J. N. et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat. Genet.* **45**, 1113–1120 (2013).
- Dawson, M. A. et al. JAK2 phosphorylates histone H3Y41 and excludes HP1α from chromatin. *Nature* **461**, 819–822 (2009).
- Ronchi, A. E., Bottardi, S., Mazzucchelli, C., Ottolenghi, S. & Santoro, C. Differential binding of the NFE3 and CP1/NFY transcription factors to the human gamma- and epsilon-globin CCAAT boxes. *J. Biol. Chem.* **270**, 21934–21941 (1995).
- Migliori, V. et al. Symmetric dimethylation of H3R2 is a newly identified histone mark that supports euchromatin maintenance. *Nat. Struct. Mol. Biol.* **19**, 136–144 (2012).
- Uhlen, M. et al. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).
- Slobodin, B. et al. Transcription impacts the efficiency of mRNA translation via co-transcriptional N<sup>6</sup>-adenosine methylation. *Cell* **169**, 326–337 (2017).
- Fustin, J. M. et al. RNA-methylation-dependent RNA processing controls the speed of the circadian clock. *Cell* **155**, 793–806 (2013).
- O'Connor, L., Gilmour, J. & Bonifer, C. The role of the ubiquitously expressed transcription factor Sp1 in tissue-specific transcriptional regulation and in disease. *Yale J. Biol. Med.* **89**, 513–525 (2016).
- Geltinger, C., Hörtogel, K. & Polack, A. TATA box and Sp1 sites mediate the activation of c-myc promoter P1 by immunoglobulin kappa enhancers. *Gene Expr.* **6**, 113–127 (1996).
- Knuckles, P. et al. RNA fate determination through cotranscriptional adenosine methylation and microprocessor binding. *Nat. Struct. Mol. Biol.* **24**, 561–569 (2017).

24. Lin, S., Choe, J., Du, P., Triboulet, R. & Gregory, R. I. The m<sup>6</sup>A methyltransferase METTL3 promotes translation in human cancer cells. *Mol. Cell* **62**, 335–345 (2016).
25. Wang, X. et al. N<sup>6</sup>-methyladenosine modulates messenger RNA translation efficiency. *Cell* **161**, 1388–1399 (2015).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank K. H. Che for help generating the RNA enzyme list. The Kouzarides laboratory is supported by grants from Cancer Research UK (grant reference RG17001) and ERC (project number 268569), in addition to benefiting from core support from the Wellcome Trust (Core Grant reference 092096) and Cancer Research UK (grant reference C6946/A14492). I.B. is funded by a Kay Kendall Leukaemia Fund project grant (grant reference RG88664). G.M.-Z. is funded by an EMBO fellowship (ALTF907-2014). G.S.V. was funded by a Wellcome Trust Senior Fellowship in Clinical Science (WT095663MA) and Cancer Research UK Senior Cancer Research Fellowship (C22324/A23015). The Vassiliou laboratory is supported by grants from the Kay Kendall Leukemia Fund and Bloodwise, as well as core funding from the Sanger Institute (WT098051). C.R.V. and J.S. are funded by a translational research grant from Northwell Health.

**Author Contributions** I.B., K.T and J.S. designed, performed and validated the CRISPR screens. K.T. and E.D.B. performed the phenotypic analysis

of human mouse targeted cells. H.P. performed bioinformatic analysis of genome-wide CRISPR screens. I.B. and L.P. generated the conditional KD cells, and performed and validated the RNA-seq, ChIP-seq, RNA-IP and riboprofiling experiments. L.P., S.C.R. and N.H. performed bioinformatic analyses of datasets. N.H. generated the expression profiles from the TCGA dataset. G.M.-Z. performed and analysed the polysome fractionation experiments. I.B., L.P., K.T. and D.A. performed the rescue experiments and the luciferase assays. V.M., A.J.B. and A.H. took part in the validation of ChIP-seq and RNA-IP experiments. I.B., K.T. and L.P. designed experiments and interpreted results. C.R.V., G.S.V. and T.K. devised and supervised the project. A.J.B., G.S.V. and T.K. wrote the manuscript with contributions from all authors.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to T.K. ([t.kouzarides@gurdon.cam.ac.uk](mailto:t.kouzarides@gurdon.cam.ac.uk)) or G.S.V. ([gsv20@sanger.ac.uk](mailto:gsv20@sanger.ac.uk)).

**Reviewer Information** *Nature* thanks K. Adelman, R. Agami, R. Levine and the other anonymous reviewer(s) for their contribution to the peer review of this work.

## METHODS

**Cell culture.** MOLM13, THP-1, MV4-11, NOMO-1, HL-60, EOL-1, KG-1, RN2c, HEL, JURKAT, LOUCY and K562 cells were cultured in RPMI1640 (Invitrogen) supplemented with 10% FBS and 1% penicillin/streptomycin/glutamine. OCI-AML2 and OCI-AML3 cells were cultured in Alpha-MEM supplemented with 20% FBS and 1% penicillin/streptomycin/glutamine. 293T, mouse immortalized fibroblasts NIH-3T3, SU-DHL-4, HT1080, SBC-3, DETROIT-562, FADU, SH-SY5Y and HT-29 cells were cultured in DMEM (Invitrogen), supplemented with 10% FBS and 1% penicillin/streptomycin/glutamine. 293T, FADU and HT-29 cells were transfected using Lipofectamine 2000 reagent (Invitrogen) according to the manufacturer's instructions. All human cancer cell lines were obtained from the Sanger Institute Cancer Cell Collection and tested to be negative for mycoplasma contamination. Human cell lines used are not listed in the cross-contaminated or misidentified cell lines database curated by the International Cell Line Authentication Committee (ICLAC).

**Isolation of haematopoietic progenitors.** *Flt3<sup>ITD/+</sup>* mice<sup>26</sup> were kindly provided by G. Gilliland and crossed with *Rosa26<sup>Cas9/+</sup>* mice<sup>6</sup>. Freshly isolated bone marrow from 6- to 10-week-old female *Rosa26<sup>Cas9/+</sup>*, *Flt3<sup>ITD/+</sup>*; *Rosa26<sup>Cas9/+</sup>* or moribund *Npm1<sup>fl/fl-cA/+</sup>*; *Flt3<sup>ITD/+</sup>* mice was used. Bone marrow cells were exposed to erythrocyte lysis (BD PharmLyse, BD Bioscience), followed by magnetic bead selection of Lin<sup>-</sup> cells using the Lineage Cell Depletion Kit (Miltenyi Biotec) according to the manufacturer's instructions. Lin<sup>-</sup> cells were cultured in X-VIVO 20 (Lonza) supplemented with 5% BIT serum substitute (Stem Cell Technologies), 10 ng ml<sup>-1</sup> IL3 (Peprotech), 10 ng ml<sup>-1</sup> IL6 (Peprotech) and 50 ng ml<sup>-1</sup> of SCF (Peprotech). Retrovirus constructs pMSCV-MLL-AF9-IRES-YFP and pMSCV-MLL-ENL-IRES-Neo were used with package plasmid psi-Eco to produce retrovirus. 293T cells (Life Technologies) were cultured and prepared for transduction in 10-cm plates as described above. For virus production, 5 µg of the above plasmids and 5 µg psi-Eco packaging vector were transfected drop-wise into the 293T cells using 47.5 µl TransIT LT1 (Mirus) and 600 µl Opti-MEM (Invitrogen). Transduction of primary mouse cells was performed in 6-well plates as mentioned above. After transduction, cells were sorted for YFP (for *KMT2A-MLLT3*) or selected with neomycin (for *KMT2A-MLLT1*).

**Generation of genome-wide mutant libraries and screening.** *Ex-vivo* CRISPR screens were performed using the previously reported WT Sanger genome-wide CRISPR library<sup>6</sup>, which uses five gRNAs for each of the 18,424 targeted genes.  $3.0 \times 10^7$  cells were infected with a pre-determined volume of the genome-wide gRNA lentiviral supernatant that gave rise to 30% transduction efficiency measured by BFP expression. Two independent infections were conducted for the mouse primary AML cells. Two days after transduction, the cells were selected with puromycin at 1.5 µg ml<sup>-1</sup>.

Genomic DNA extraction and Illumina sequencing of gRNAs were conducted as described previously<sup>6</sup>. For the primary mouse AML screen, 19-bp single-end sequencing was performed with the custom sequencing primer 5'-TCTTCGATCTTGTGAAAGGACGAAACACCG-3'. The number of reads for each guide was counted with an in-house script. Enrichment and depletion of guides and genes were analysed using the MAGeCK statistical package<sup>27</sup> by comparing read counts from each cell line with counts from matching plasmids as the initial population.

The 296 RNA enzyme list, which has been interrogated for dropouts, was compiled as follows: 171 SAM-binding proteins without known histone methyltransferase activity; 65 ATP-dependent helicases with reported RNA interaction in at least 3 out of 4 CLIP data sets<sup>28–31</sup>; 12 pseudouridylases; 48 annotated RNA enzymes not included in the previous lists (Supplementary Table 2).

**Lentiviral vector production and infection.** For virus production, 293T cells were transfected with the appropriate lentiviral vector (PLKO.1 for shRNA, LRG and LentiCRISPR-v2 for CRISPR-Cas9) together with the packaging plasmids PAX2 and VSVg at a 1:1.5:0.5 ratio. Supernatant was harvested 48 and 72 h after transfection.

$1 \times 10^6$  cells and viral supernatant were mixed in 2 ml culture medium supplemented with 8 µg ml<sup>-1</sup> (human) or 4 µg ml<sup>-1</sup> (mouse) polybrene (Millipore), followed by spinfection (60 min, 900g, 32 °C) and further incubated overnight at 37 °C. The medium was refreshed on the following day and the transduced cells were cultured further.

**Pooled domain-focused CRISPR gRNA construction.** For the RNA enzyme domain-focused CRISPR screen, five gRNAs were designed to target the catalytic domain of each protein based on the NCBI database annotation. gRNAs were synthesized in a pooled format on an array platform (Twist Bioscience) and PCR-cloned into a lentiviral gRNA expression vector (LRG, Addgene: 65656). To check the representation and identity of individual gRNAs within the lentiviral pool, a deep-sequence analysis was performed. This confirmed that 100% of the designed gRNAs were cloned in the LRG vector and the abundance of >95% of the gRNA constructs was within fivefold of the mean.

**Pooled gRNA screening, MiSeq library construction and data analysis.** Virus containing the domain-focused gRNA library against RNA modifying enzymes was generated as described above. A serial dilution of this virus in correlation with the GFP<sup>+</sup> cell population was used to estimate the viral titre multiplicity of infection (MOI). In the initial infected cell population, the total number of RN2c cells corresponded to an approximately one thousand-fold representation of each gRNA. To ensure that a single gRNA was transduced per cell, the viral volume for infection corresponded to an MOI of approximately 0.4–0.5. At 2 days post-infection, a portion of the RN2c cells was harvested and used as a reference time point. At the end point of the negative selection experiment, a portion of the RN2c cells was harvested again and saved for gRNA abundance quantification through deep sequencing analysis.

The gRNA cassette MiSeq deep sequencing library was constructed using a similar method as previously described<sup>7</sup>. Genomic DNA was extracted using QiAamp DNA mini kit (Qiagen 51304), following the manufacturer's protocol. In order to maintain approximately one thousand-fold gRNA library representation, 25 parallel PCR reactions were performed to amplify the gRNA cassette using the 2× High Fidelity Phusion Master Mix (Thermo Scientific F-548). PCR products were subjected to Illumina MiSeq library construction and sequencing. First, the PCR product was end repaired with T4 DNA polymerase (New England BioLabs, NEB), DNA polymerase I (NEB), and T4 polynucleotide kinase (NEB). Then, an A overhang DNA fragment was ligated with diversity-increased barcoded Illumina adaptors followed by seven pre-capture PCR cycles. The barcoded libraries were pooled at an equal molar ratio and subjected to massively parallel sequencing through a MiSeq instrument (Illumina) using paired-end 150-bp sequencing (MiSeq Reagent Kit v2; Illumina MS-102-2002).

The sequence data were de-barcoded and trimmed to contain only the gRNA sequence, and subsequently mapped to the reference gRNA library without allowing any mismatches. The read counts were calculated for each individual gRNA. To calculate the relative abundance of each gRNA, the read counts of each gRNA were divided by the total read counts of the pooled gRNA library.

**Validation of the catalytic domain-specific screen.** RN2C cells or NIH3t3 mouse fibroblast cells were infected with LRG lentiviral vectors expressing GFP and a single gRNA targeting the catalytic domain of the indicated RNA enzymes and controls. The percentage of GFP<sup>+</sup> cells was measured at day 2 after infection as a baseline by flow cytometry. The percentage of GFP<sup>+</sup> cells was measured again at day 10 and day 12 for RN2C or NIH3t3 cells, respectively. Three gRNAs from Screen 2 were used for each target and one gRNA for the *rosa26* locus as a negative control. The gRNA sequences are listed in the Supplementary Information. The data were analysed using Flowjo software.

**gRNA competition.** gRNA competition assays were performed using single and dual gRNA vectors as described previously<sup>2</sup>. For the validation of individual target genes, one gRNA was derived from the CRISPR library used in the screens and another gRNA was designed using WTSI Genome Editing website (<http://www.sanger.ac.uk/htgt/wge/>). Viral supernatants were collected 48 h after transfection. All transfections and viral collections were performed in 24-well plates and transduction was performed as mentioned above. For gRNA-BFP competition assays, flow cytometry analysis was performed on 96-well plates using a LSRFortessa instrument (BD). Gating was performed on live cells using forward and side scatter, before measuring of BFP<sup>+</sup> cells. The gRNA sequences are listed in the Supplementary Information.

Efficiency of genome editing in the pool of sgRNA-targeted cells was evaluated by tracking of indels by decomposition (TIDE)<sup>32</sup>.

**Replating assays.** For re-plating assays, 5,000 lineage negative cells and primary murine AML cells expressing Cas9 were plated in three wells of a 6-well-plate of M3434 methylcellulose (Stem Cell Technologies) after infection with lentiviral vectors expressing gRNAs targeting the catalytic domain of METTL3 or empty vectors. The colonies were counted 7 days later and a further 5,000 cells re-seeded and re-counted after a week until no colonies were observed (for the wild type) or until the third replating (for primary mouse AMLs).

**Flow cytometry analyses of AML cells.** Cas9-expressing cells were transduced with gRNA vectors targeting the catalytic domain of METTL3 or empty vectors and stained with anti-mouse CD11b PE/Cy5 (Biolegend, 101210) and anti-human CD11b PE (eBiosciences, 9012-0118). Data were analysed using LSRFortessa (BD) and FlowJo.

**Mouse whole-body bioluminescent imaging.** For *in vivo* experiments, MOLM13 cells expressing Cas9 were transduced with a firefly luciferase-expressing plasmid (System Biosciences). After propagation, the cells were transduced with a lentivirus (either empty or expressing *Mettl3* gRNA; day 0) and selected with puromycin from day 2 to day 5. At day 5 post-transduction, the cells were suspended in fresh medium without puromycin. At day 6,  $1 \times 10^5$  cells were transplanted into female 6-week old *Rag2<sup>-/-</sup> Il2rg<sup>-/-</sup>* mice by tail-vein injection. At day 14 post-transplant,

the tumour burdens of the animals were detected using IVIS Lumina II (Caliper) with Living Image version 4.3.1 software (PerkinElmer). In brief, 100 µl of 30 mg ml<sup>-1</sup> d-luciferin (BioVision) was injected into the animals intraperitoneally. Ten minutes after injection, the mice were maintained in general anaesthesia by isoflurane and put into the IVIS chamber for imaging. The detected tumour burdens were measured and quantified by the same software. Mice were culled when the tumour burden was 10<sup>8</sup> photons per second or higher. The welfare of diseased mice was assessed blindly by qualified animal technicians from the Sanger mouse facility. All animal studies were carried out in accordance with the Animals (Scientific Procedures) Act 1986, UK and approved by the Ethics Committee at the Sanger Institute. Randomization and blinding were not applied. No statistical methods were used to predetermine sample size.

**Generation of conditional knock-down cells.** MOLM13 or THP1 cells (5 × 10<sup>5</sup>) were infected as described above using PLKO-TETOn-Puro lentiviral vectors expressing shRNAs against the coding sequence of human METTL3, CEBPZ or a scrambled control. Twenty-four hours after spinfection, the cells were replated in fresh medium containing 1 µg ml<sup>-1</sup> of puromycin and kept in selection medium for 7 days. shRNA was induced by treatment with 200 ng ml<sup>-1</sup> doxycycline for the indicated times. The shRNA sequences are listed in the Supplementary Information.

**Proliferation assays.** MOLM13 and THP1 control or METTL3-KD cells (5 × 10<sup>4</sup>; 4 days after doxycycline induction) were seeded in 2 ml complete RPMI medium and counted 4 days after plating using the Countess II cell counter. For SP1 rescue experiments, the number of doxycycline-treated cells was normalized to their untreated counterparts.

**Rescue experiments.** cDNA was obtained by reverse transcription of MOLM13 cell RNA with Supercrt III (ThermoFisher Scientific), then the SP1 full-length coding sequence was amplified by PCR and cloned into pHIV-ZsGreen plasmid (Addgene 18121) by Gibson assembly (Gibson Assembly Cloning Kit, NEB), using the HpaI site.

MOLM13 cells were transduced with the lentiviral constructs, then GFP<sup>+</sup> cells were isolated by flow cytometry sorting after 5 days and employed in proliferation assays as described in the previous section.

For rescue experiments with either active or catalytically mutant METTL3, pcDNA3/Flag-METTL3 plasmid (Addgene 53739) was mutagenized at the positions D394A (GAC/GCC) and W397A (TGG/GCG) as described<sup>33</sup> with QuikChange II Site-Directed Mutagenesis Kit (Stratagene).

*Npm1c*<sup>Fli3td/+</sup>/*Rosa26*<sup>Cas9/+</sup> cells were transduced with an empty or domain-specific *Mettl3* gRNA lentiviral vector and selected with puromycin until day 5 post-transduction. Then cells were electroporated in Buffer R (Invitrogen) with the plasmids encoding wild-type METTL3, catalytically inactive METTL3 or an empty vector as a control. In each replicate, 150,000 cells were electroporated with 500 ng of each plasmid and plated in triplicate at 50,000 cells per well in a 6-well plate. Electroporation was performed using the Neon Transfection System (Thermo Fisher Scientific) as previously described<sup>34</sup>. Three days after electroporation (day 8), live cells were measured per well per condition.

**METTL3 overexpression and proliferation experiments.** AML cells were electroporated in Buffer R (Invitrogen) with plasmids encoding wild-type METTL3 or an empty vector as a control. In each replicate, 200,000 cells were electroporated with 500 ng of each plasmid and plated in triplicate at 50,000 cells per well in a 6-well plate. Two days after electroporation (day 3), live cells were measured per well per condition.

**Chromatin immunoprecipitation.** Chromatin immunoprecipitation was performed as previously described<sup>35</sup>. First MOLM13 cells were cross-linked with 0.5% formaldehyde for 10 min. 20 × 10<sup>6</sup> cells were used for each immunoprecipitation with 3 µg of specific antibodies or IgG. Immunoprecipitated DNA was purified with ChIP DNA Clean & Concentrator Columns (Zymo) and either amplified for massive parallel sequencing or analysed on an ABI 7900 real-time PCR machine, Fast SybrGreen PCR mastermix according to the manufacturer's instructions. Primer sequences are listed in the Supplementary Information.

**ChIP sequencing.** For the METTL3, METTL14, H3K4me3 and IgG ChIP-seq experiments, 5, 5, 2 and 3 independent biological replicates were used, respectively. Single-end 50-bp libraries were prepared using the Bioo Scientific NEXTflex ChIP-Seq kit according to the manufacturer's recommendations. Reads were sequenced using HiSeq 1500 and multiplexed reads were split on the basis of their barcodes using Illumina Basespace. Reads were trimmed to remove the TruSeq adaptor using trim\_galore ([http://www.bioinformatics.babraham.ac.uk/projects/trim\\_galore](http://www.bioinformatics.babraham.ac.uk/projects/trim_galore)) with parameters '-q 0 -a AGATCGGAAGAGCACACGTCTGAACCTCCAGTCAC-phred33-fastqc'. Read quality was assessed using FastQC. Ribosomal contamination was removed by first mapping the reads to hg38 rDNA sequences using bwa<sup>36</sup> (default parameters), and then removing all reads that mapped. Trimmed non-ribosomal reads were mapped to the hg38 genome using bwa with parameters '-n 3 -k 2 -R 300 -t 4'. Multiple reads mapping to a single genomic locus were treated as PCR duplicates, and were removed using samtools rmdup<sup>37</sup>. Mapped reads

were filtered to remove reads mapping to more than one unique genomic locus (multihits) by keeping only reads with flag XT:A:U in the output bam file from bwa. Reads were further filtered to remove those with mapping quality less than 20 using samtools. Genome coverage bedgraph files were generated using genomeCoverageBed from the bedtools<sup>38</sup> suite of tools. Coverage files were converted to bigwig format using bedGraphToBigWig. Peaks were called against input sample using MACS2 using default parameters<sup>39</sup>. Peaks from all replicates were merged to give a master list of potential binding loci per condition, and read count (normalized by overall read depth of the library) for each replicate was calculated using the GenomicRanges package in R<sup>40</sup>. Peaks were treated as potential binding loci if all replicates showed normalized score greater than 1 and did not overlap a peak called in the IgG. Genomic annotation of ChIP peaks and reads and gene set operations were performed using the R packages ChIPseeker<sup>41</sup> and VennDiagram<sup>42</sup>, respectively. HOMER tool suite<sup>43</sup> was used for DNA motif discovery coupled with hypergeometric enrichment calculations (or binomial) to determine motif enrichment. **m<sup>6</sup>A RNA immunoprecipitation and sequencing.** Total RNA was isolated from MOLM13 control or METTL3-KD cells (two independent biological replicates for each shRNA) eight days after doxycycline administration using the RNAeasy midi kit (Qiagen). Successively polyA<sup>+</sup> RNA was purified from 300 µg total RNA using the NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs). 500 ng of polyA<sup>+</sup> purified RNA was used for each immunoprecipitation reaction. m<sup>6</sup>A RNA immunoprecipitation was performed using the Magna MeRIP m<sup>6</sup>A kit (Millipore) according to the manufacturer's instructions.

Single-end 50-bp stranded libraries were prepared using the Bio Scientific NEXTflex Rapid Directional RNA kit according to the manufacturer's recommendations. Reads were sequenced using HiSeq 4000 and multiplexed reads were split on the basis of their barcodes using Illumina Basespace. Reads were trimmed to remove the TruSeq adaptor using trim\_galore with parameters '-q 0 -a AGATCGGAAGAGCACACGTCTGAACCTCCAGTCAC-phred33-fastqc'.

Read quality was assessed using FastQC. Ribosomal contamination was removed by first mapping the reads to hg38 rDNA sequences using bwa (default parameters), and removing all reads that mapped. Trimmed non-ribosomal reads were mapped to the hg38 genome using Tophat2 with parameters '-no-coverage-search-max-multihits 300-report-secondary-alignments-read-mismatches 2'. Gene annotations from Ensembl v86 were used to direct transcript alignment. Mapped reads were filtered to remove those mapping to more than one unique genomic locus (multihits) by keeping only reads with flag NH:i:1 in the output bam file from Tophat. Reads were further filtered to remove the ones with mapping quality less than 20 using samtools. Genome coverage bedgraph files were generated for forward and reverse strand reads using genomeCoverageBed from the bedtools suite of tools. Coverage files were converted to bigwig format using bedGraphToBigWig. Transcript assembly was performed using cufflinks2 and a single transcript database was generated using cuffmerge<sup>44</sup>. Statistical analysis of differentially methylated peaks in control and METTL3-KD cells was performed using the R package MeTDiff<sup>45</sup>. Metagene plots were generated by RNAModR package (<https://github.com/mevers/RNAModR>), and m<sup>6</sup>A motifs were analysed according to published protocols<sup>46</sup>. To evaluate the statistical significance of [GAG]<sub>n</sub> motif enrichment, individual motif occurrences were searched throughout the human transcriptome with FIMO program (MEME suite<sup>47</sup>).

**Ribosome profiling.** Ribosome profiling was carried out using the Illumina TruSeq Ribo Profiler (Mammalian) kit. 5 × 10<sup>7</sup> MOLM13 control (two pooled independent biological replicates) or METTL3-KD cells (two pooled independent biological replicates for each shRNA) were treated 5 or 8 days after doxycycline induction with 0.1 mg ml<sup>-1</sup> of cycloheximide for 1 min and the RPF fraction of mRNA was isolated following the manufacturer's instruction. Ribosomal RNA was removed using the Illumina RiboZero Kit.

Single-end 50-bp stranded libraries were prepared following the manufacturer's recommendations. Reads were sequenced using HiSeq 1500 and multiplexed reads were split on the basis of their barcodes using Illumina Basespace. Reads were trimmed to remove the TruSeq adaptor using trim\_galore with parameters '-q 0 -a AGATCGGAAGAGCACACGTCTGAACCTCCAGTCAC-phred33-fastqc'. Read quality was assessed using FastQC. Ribosomal contamination was removed by first mapping the reads to hg38 rDNA sequences using bwa (default parameters), and removing all reads that mapped. Trimmed non-ribosomal reads were mapped to the hg38 genome using Tophat2 with parameters '-no-coverage-search-max-multihits 300-report-secondary-alignments-read-mismatches 2'. Gene annotations from Ensembl v86 were used to direct transcript alignment. Mapped reads were filtered to remove those mapping to more than one unique genomic locus (multihits) by keeping only reads with flag NH:i:1 in the output bam file from Tophat. Reads were further filtered to remove the ones with mapping quality less than 20 using samtools. Genome coverage bedgraph files were generated for forward and reverse strand reads using genomeCoverageBed from the bedtools suite of tools. Coverage files were converted to bigwig format using bedGraphToBigWig.

Transcript assembly was performed using cufflinks and a single transcript database was generated using cuffmerge. Statistical analysis of differentially translated genes in control and *METTL3*-KD cells was performed using the R package xtail<sup>48</sup>. In order to estimate the offset value relative to the read start required to localize the position of P-sites (in our case ~12 nt, as already described in the literature<sup>49</sup>), we employed the function psite of the plastid Python library<sup>49</sup>.

**Polysome fractionation.**  $5 \times 10^7$  MOLM13 control (two pooled independent biological replicates) or *METTL3*-KD cells (two pooled independent biological replicates) were treated 8 days after doxycycline induction with 0.1 mg ml<sup>-1</sup> cycloheximide for 5 min at 37 °C, then they were lysed and polysomes were fractionated on a sucrose gradient as described<sup>50</sup>. Relative RNA abundance in each fraction was then quantified by RT-qPCR.

**RNA sequencing.** Input samples from the m<sup>6</sup>A RNA immunoprecipitation and riboprofiling experiments were combined to give a set of transcriptional profiling samples to treat as an RNA-seq experiment (control  $n = 4$ , *METTL3*-KD  $n = 6$ ). Gene counts were calculated at the transcript level for the combined transcript database from cuffmerge using summarizeOverlaps from the GenomicAlignments<sup>40</sup> package in R using a generalized linear model fit that takes into account possible batch effects of the two distinct experiments. Differential gene expression analysis was conducted using DESeq2<sup>51</sup>. Differentially expressed genes were identified as those with RPKM greater than 1 for wild-type or knockdown showing differential expression greater than twofold (up or down) with a Benjamini–Hochberg corrected *P* value less than 0.05 unless stated otherwise.

Gene set enrichment analysis was performed using the R package GAGE<sup>52</sup>.

**Construction of the engineered translation reporter system.** A DNA sequence carrying ten GAL4 recognition motifs (or a scrambled version of the same sequence) was *in vitro* synthesized and cloned into a pMirGlo plasmid (Promega) upstream of the constitutive PGK promoter by cutting with BglII and ligating with T4 DNA ligase (NEB). The sequence of SP2 known to harbour the m<sup>6</sup>A peak and the [GAN]<sub>n</sub> motif was amplified by PCR and subcloned in frame at the N terminus of firefly luciferase by Gibson assembly (Gibson Assembly Cloning Kit, NEB), using the ApaI site.

To generate the GAL4–*METTL3*(CD) construct, the CREB coding sequence in plasmid pcDNA1-GAL4-CREB-S133A (Addgene 46770) was swapped with the *METTL3* catalytic domain obtained by plasmid pcDNA3/Flag–*METTL3* (Addgene 53739) using restriction enzymes XbaI and EcoRI. Wild-type *METTL3* was mutagenized at the positions D394A (GAC/GCC) and W397A (TGG/GCG) as described<sup>33</sup> with the QuikChange II Site-Directed Mutagenesis Kit (Stratagene). Cloning primer sequences are listed in the Supplementary Information.

**Luciferase assay.** 293T cells were co-transfected with the GAL4–*METTL3* construct and UAS-SP2-Luc or SCR-SP2-Luc vectors expressing the Renilla luciferase as transfection control. Twenty-four hours after transfection, the firefly and Renilla luciferase activities were measured using the Dual-Luciferase Reporter Assay System (Promega) on CLARIOstar microplate reader (BMG Labtech).

Alternatively, transfected cells were lysed and total RNA extracted using the Qiagen RNAeasy Mini kit according to the manufacturer's instructions.

**Western blotting.** Western blotting was performed as previously described<sup>53</sup>. For the isolation of nuclei from the cytoplasmic fraction, cells were incubated in hypotonic buffer (20 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl<sub>2</sub>) and the cytoplasmic membranes disrupted by adding 0.5% NP40. Pelleted nuclei were successively lysed in IPH buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM EDTA, 0.5% NP-40) for 10 min on ice. Nucleoplasmic and chromatin fractions were separated by 15 min centrifugation at 4 °C. The chromatin fraction was resuspended and sonicated in Laemmli buffer before loading on SDS-PAGE.

**RT-qPCR.** Total RNA from MOLM13, THP1 or 293T cells was purified using the RNAeasy mini kit according to the manufacturer's instructions. One microgram of purified total RNA was retrotranscribed using the high-capacity cDNA reverse transcription kit (Applied Biosystems). The levels of specific RNAs were measured using an ABI 7900 real-time PCR machine and Fast SybrGreen PCR mastermix according to the manufacturer's instructions. Firefly luciferase levels were normalized on Renilla luciferase levels while *METTL3* mRNA levels were normalized on *ACTB* mRNA. To quantify polysome fractionation and m<sup>6</sup>A-RIP experiments, we used probes from Universal ProbeLibrary (Roche) with TaqMan Fast Advanced Master Mix (ThermoFisher). Primer sequences are listed in the Supplementary Information.

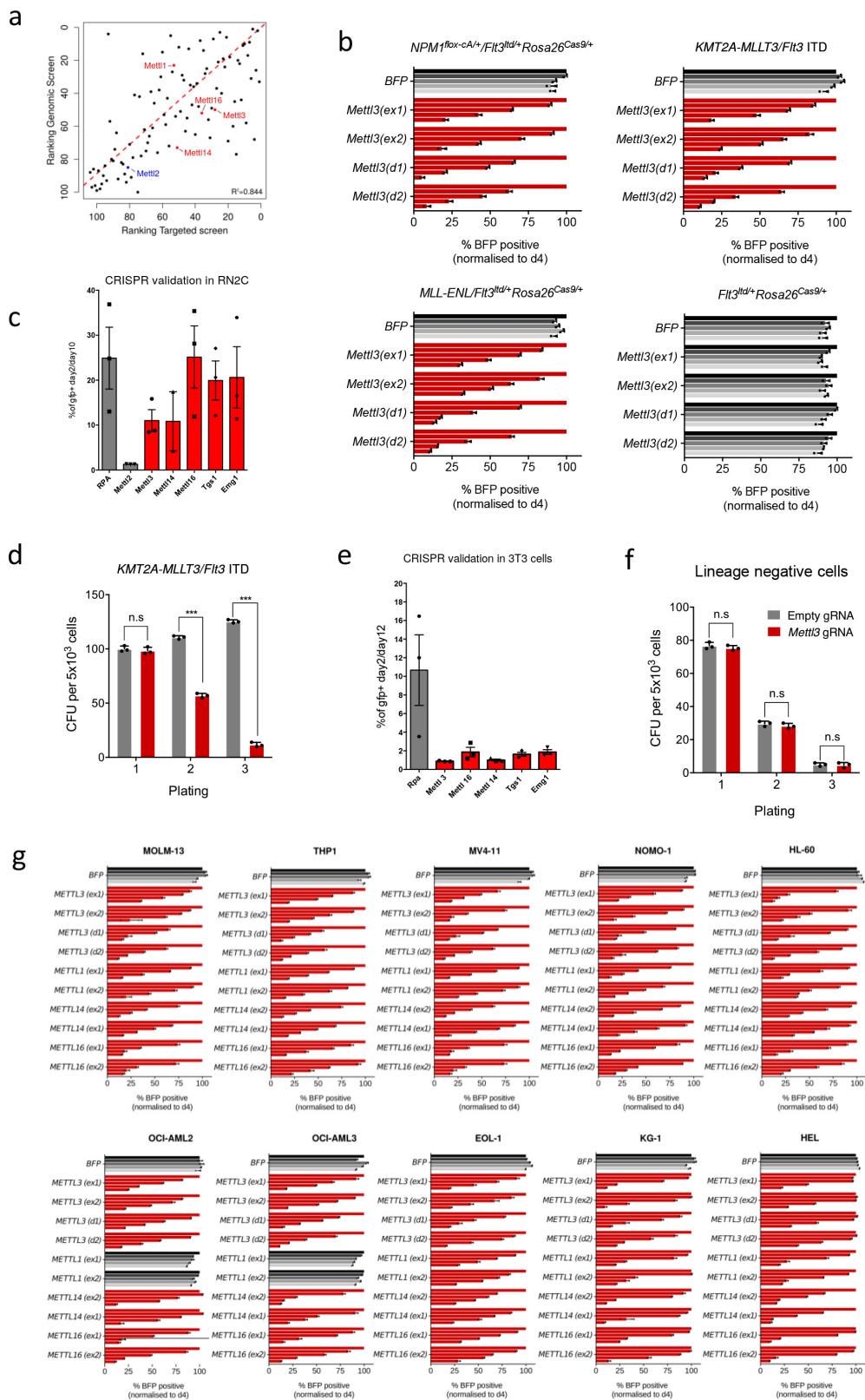
**Antibodies.** For the ChIP experiments the following antibodies were used: anti-METTL3 from Bethyl Laboratories (A301-568A), anti-METTL3 from Bethyl Laboratories (A301-567A), rabbit polyclonal anti-METTL14 from Abcam (ab98166), anti-H3k4me3 from Abcam (ab8580) and IgG Isotype Control (ab171870).

Western blot experiments were performed using the following antibodies: anti-METTL3 from Bethyl Laboratories (A301-568A), anti-METTL14 from Abcam (ab98166), anti-histone H3 from Active Motif (39763), anti-WDR5 from Abcam (ab178410), anti-CEBPZ from Abcam (ab176579), anti-SP1 from Abcam (ab13370), anti-SP2 from Abcam (ab137238) and anti-ACTIN from Abcam (ab8227).

**Statistical analysis.** All general statistical analyses were performed using either a two-tailed Student's *t*-test or a Wilcoxon test (when distributions were assessed not to be normal and homoscedastic) at a confidence interval of 95%, unless otherwise specified. No statistical methods were used to predetermine sample size.

**Data availability.** Raw sequencing data have been deposited in the Gene Expression Omnibus database with accession number GSE94613. All other data are available from the corresponding author upon reasonable request.

26. Lee, B. H. et al. FLT3 mutations confer enhanced proliferation and survival properties to multipotent progenitors in a murine model of chronic myelomonocytic leukemia. *Cancer Cell* **12**, 367–380 (2007).
27. Li, W. et al. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* **15**, 554 (2014).
28. Castello, A. et al. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* **149**, 1393–1406 (2012).
29. Kwon, S. C. et al. The RNA-binding protein repertoire of embryonic stem cells. *Nat. Struct. Mol. Biol.* **20**, 1122–1130 (2013).
30. Gerstberger, S., Hafner, M. & Tuschi, T. A census of human RNA-binding proteins. *Nat. Rev. Genet.* **15**, 829–845 (2014).
31. Baltz, A. G. et al. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol. Cell* **46**, 674–690 (2012).
32. Brinkman, E. K., Chen, T., Amendola, M. & van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* **42**, e168 (2014).
33. Fustin, J.-M. et al. RNA-methylation-dependent RNA processing controls the speed of the circadian clock. *Cell* **155**, 793–806 (2013).
34. Gundry, M. C. et al. Highly efficient genome editing of murine and human hematopoietic progenitor cells by CRISPR/Cas9. *Cell Reports* **17**, 1453–1461 (2016).
35. Dawson, M. A. et al. Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *Nature* **478**, 529–533 (2011).
36. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
37. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
38. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
39. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
40. Lawrence, M. et al. Software for computing and annotating genomic ranges. *PLOS Comput. Biol.* **9**, e1003118 (2013).
41. Yu, G., Wang, L.-G. & He, Q.-Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015).
42. Chen, H. & Boutros, P. C. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* **12**, 35 (2011).
43. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
44. Trapnell, C. et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).
45. Cui, X. et al. MeTDiff: a novel differential RNA methylation analysis for MeRIP-seq data. *IEEE/ACM Trans. Comput. Biol. Bioinform.* <https://doi.org/10.1101/TCBB.2015.240335> (2015).
46. Dominissini, D., Moshitch-Moshkovitz, S., Salmon-Divon, M., Amariglio, N. & Rechavi, G. Transcriptome-wide mapping of N<sup>6</sup>-methyladenosine by m<sup>6</sup>A-seq based on immunocapturing and massively parallel sequencing. *Nat. Protocols* **8**, 176–189 (2013).
47. Bailey, T. L. et al. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* **37**, W202–W208 (2009).
48. Xiao, Z., Zou, Q., Liu, Y. & Yang, X. Genome-wide assessment of differential translations with ribosome profiling data. *Nat. Commun.* **7**, 11194 (2016).
49. Dunn, J. G. & Weissman, J. S. Plastid: nucleotide-resolution analysis of next-generation sequencing and genomics data. *BMC Genomics* **17**, 958 (2016).
50. Panda, A. C., Martindale, J. L. & Gorospe, M. Polysome fractionation to analyze mRNA distribution profiles. *Bio Protoc.* **7**, e2126 (2017).
51. 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).
52. Luo, W., Friedman, M. S., Shadden, K., Hankenson, K. D. & Woolf, P. J. GAGE: generally applicable gene set enrichment for pathway analysis. *BMC Bioinformatics* **10**, 161 (2009).
53. Wyspiarska, B. S. et al. BET protein inhibition shows efficacy against JAK2V617F-driven neoplasms. *Leukemia* **28**, 88–97 (2014).

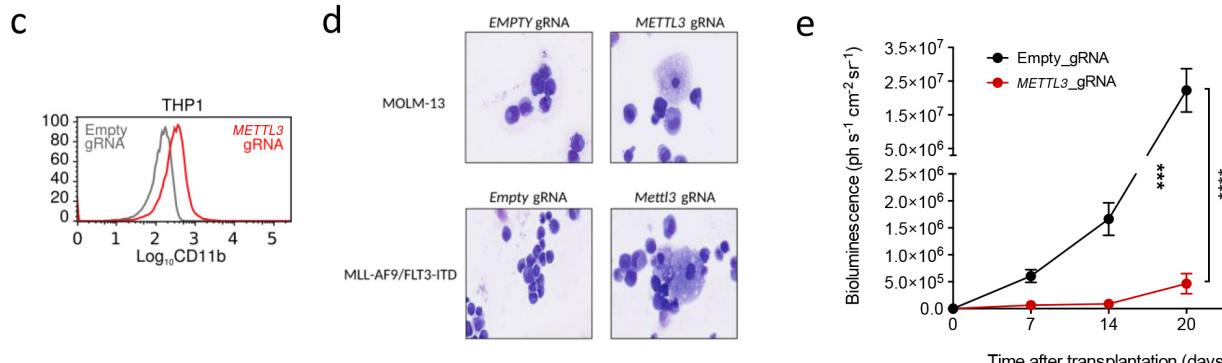
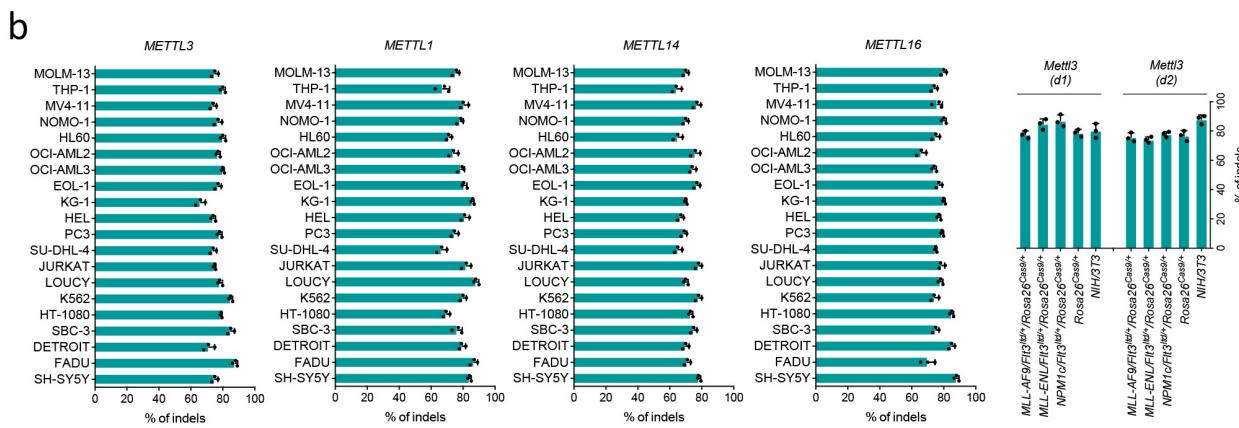
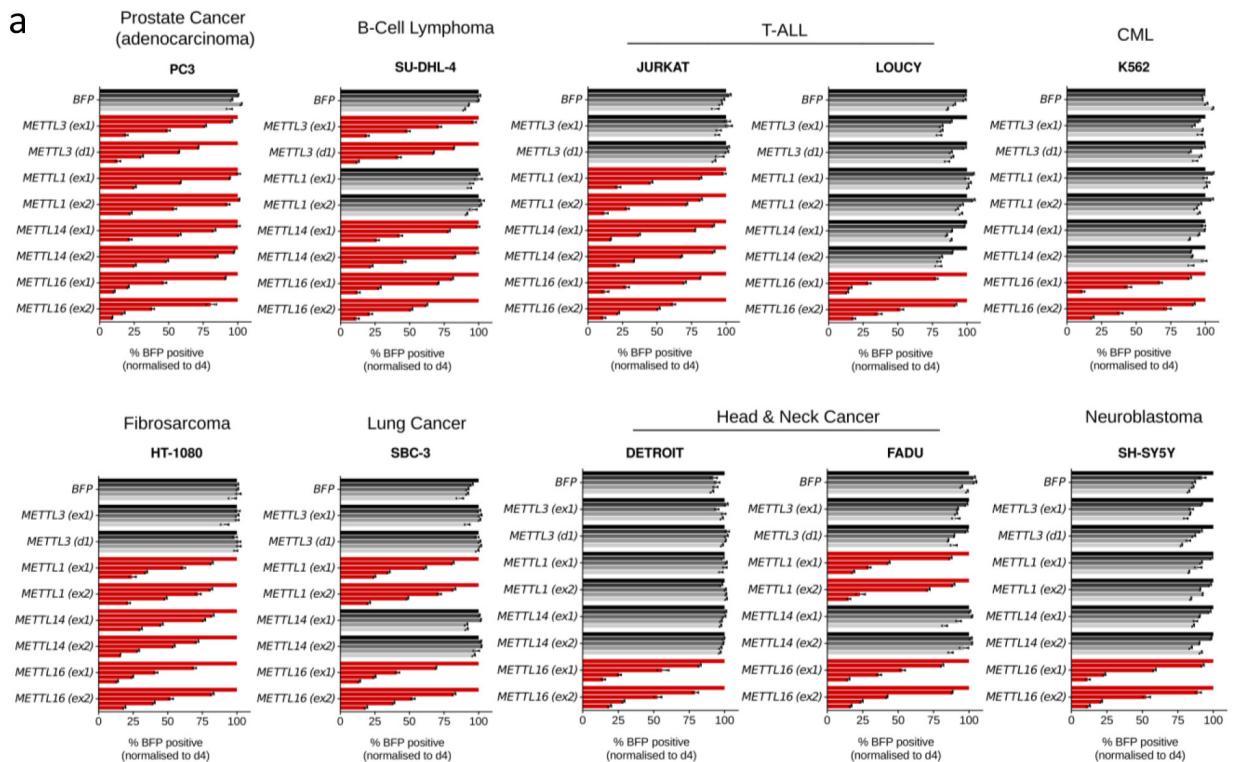


Extended Data Figure 1 | See next page for caption.

**Extended Data Figure 1 | Validation of CRISPR screens.** Related to Fig. 1. **a**, Correlation between gene rankings from the two independent CRISPR–Cas9 screens. Goodness of fit is calculated as Pearson Correlation Coefficient. **b**, Average ratio of the percentage of GFP-positive RN2C cells between day 2 and day 10 after infection with lentiviral vectors expressing GFP and individual gRNAs against the indicated targets. The mean + s.e.m. depletion of three different gRNAs against the catalytic domain of the targets is shown. gRNA targeting the *Rosa26* locus was a negative control. Rpa (replication protein A) is a positive control. **c**, Competitive co-culture assay showing negative selection of BFP<sup>+</sup> MOLM13 or *KMT2A–MLLT3* primary mouse cells upon targeting of METTL3 by CRISPR–Cas9. Cells were transduced with lentiviruses expressing four different gRNAs targeting the 5' exons or the catalytic domain of METTL3 and the BFP-positive fraction was compared with the non-transduced population. Results were normalized to those at day 4 for each gRNA. The mean ± s.d. of two independent infections is shown. **d**, Colony formation assay of *KMT2A–MLLT3 Flt3* ITD Cas9-expressing cells targeting *Mettl3* (catalytic domain-specific) or control,

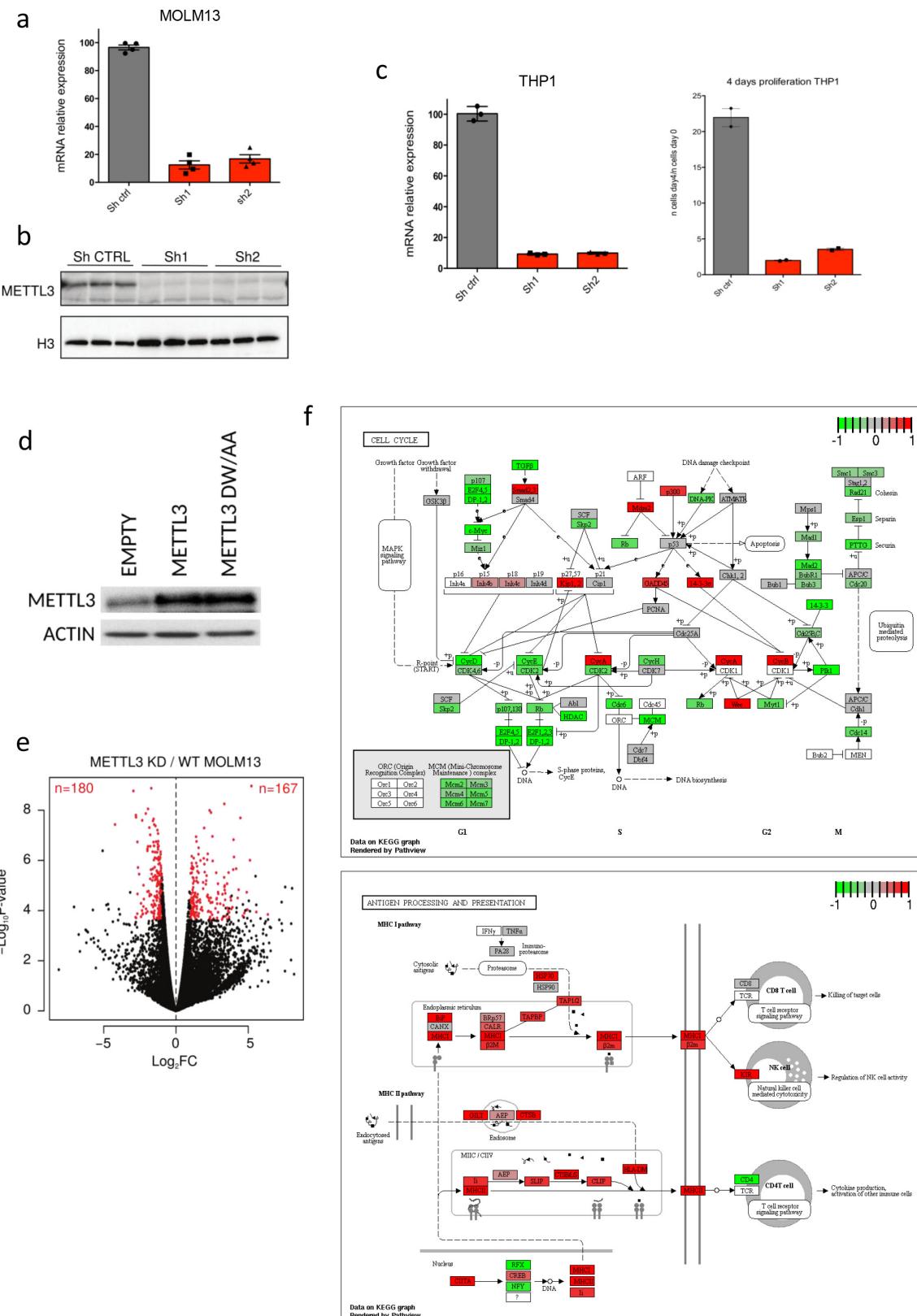
showing decreased replating ability. CFU: colony-forming units; \*\*\* $P < 0.001$ ; *t*-test. The mean + s.d. of three independent experiments is shown. **e**, Average ratio of the percentage of GFP-positive NIH-3T3 mouse fibroblasts between day 2 and day 12 after infection with lentiviral vectors expressing GFP and individual gRNAs against the indicated targets. The mean ± s.e.m. depletion of three different gRNAs against the catalytic domain of the targets is shown. Rpa is a positive control. **f**, Colony formation assay of lineage-negative haematopoietic Cas9-expressing cells targeting *Mettl3* (catalytic domain-specific; right) or control.

\*\*\* $P < 0.001$ ; *t*-test. The mean + s.d. of three independent experiments is shown. **g**, Competitive co-culture assay showing negative selection of BFP<sup>+</sup> AML cell lines upon targeting of *METTL3*, *METTL1*, *METTL14* and *METTL16* by CRISPR–Cas9 using two independent gRNAs for each target. Cells were transduced with lentiviruses expressing BFP and four different gRNAs targeting the 5' exons or the catalytic domain of each target and the BFP-positive fraction was compared with the non-transduced population. Results were normalized to those at day 4 for each gRNA. The mean + s.d. of two independent infections is shown.



**Extended Data Figure 2 | Effects of targeting METTL factors in human cancer cell lines.** Related to Fig. 1. **a**, Competitive co-culture assay showing negative selection of BFP<sup>+</sup> human cancer cell lines upon targeting of METTL3, METTL1, METTL14 and METTL16 by CRISPR–Cas9 using two independent gRNAs for each target. The experiment was performed as described above. **b**, Efficiency of genome editing for gRNAs targeting METTL3, METTL1, METTL14 and METTL16 was measured across the indicated 20 human cell lines through TIDE analysis. Efficiency of targeting was also measured in mouse primary cell lines for gRNAs

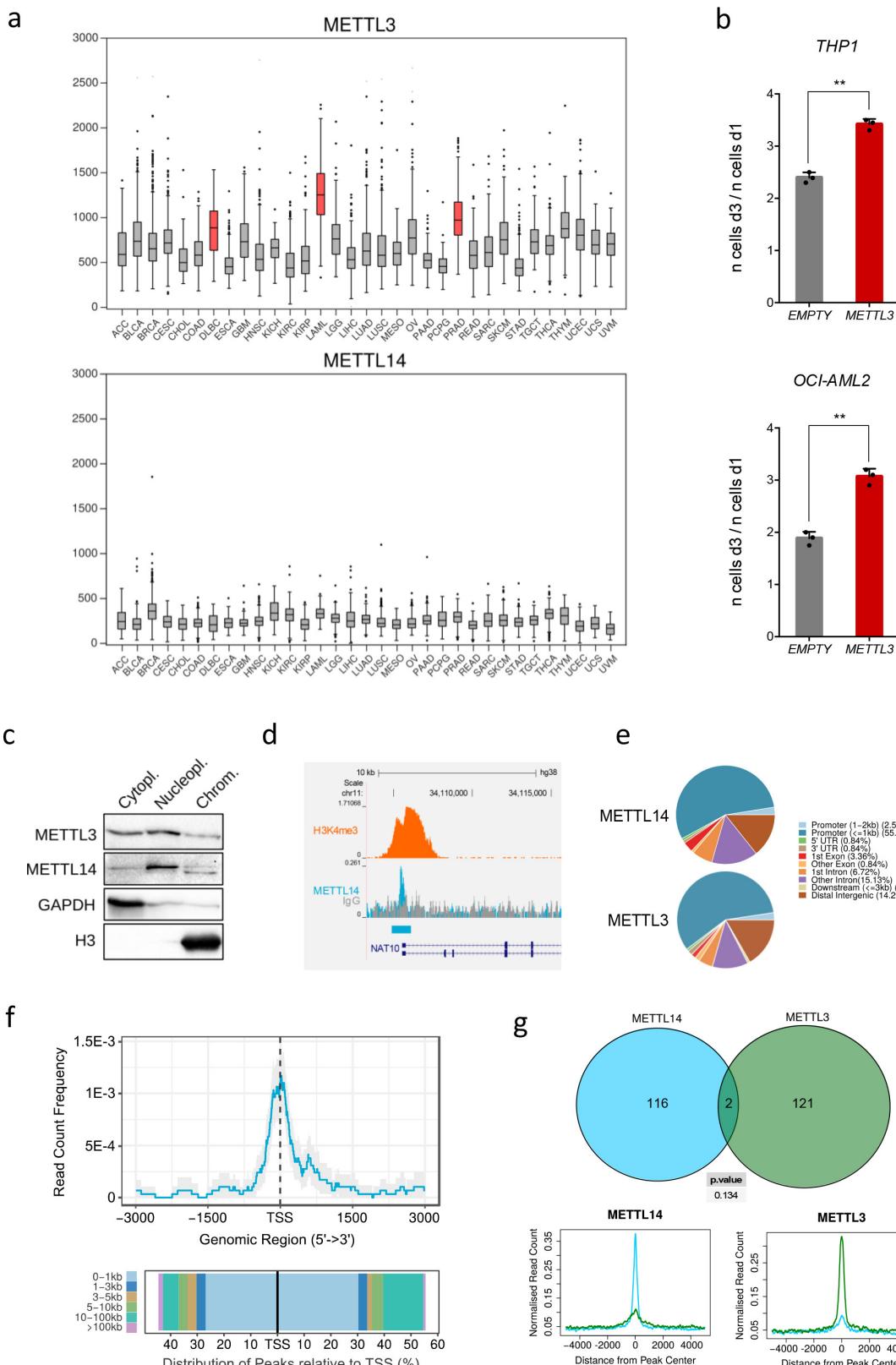
targeting *Mettl3*. **c**, CD11b expression in METTL3 (catalytic domain-specific) targeted cells (THP1 human cell line) was measured by flow cytometry 6 days after infection. **d**, Haematoxylin and eosin staining of human and mouse AML cell lines infected with a control gRNA or gRNAs targeting the catalytic domain of METTL3. **e**, Time course quantification of luminescence from mice transplanted with luciferase-labelled MOLM13 cells targeting METTL3 using gRNAs specific for the catalytic domain or control (\*\*P < 0.001).



Extended Data Figure 3 | See next page for caption.

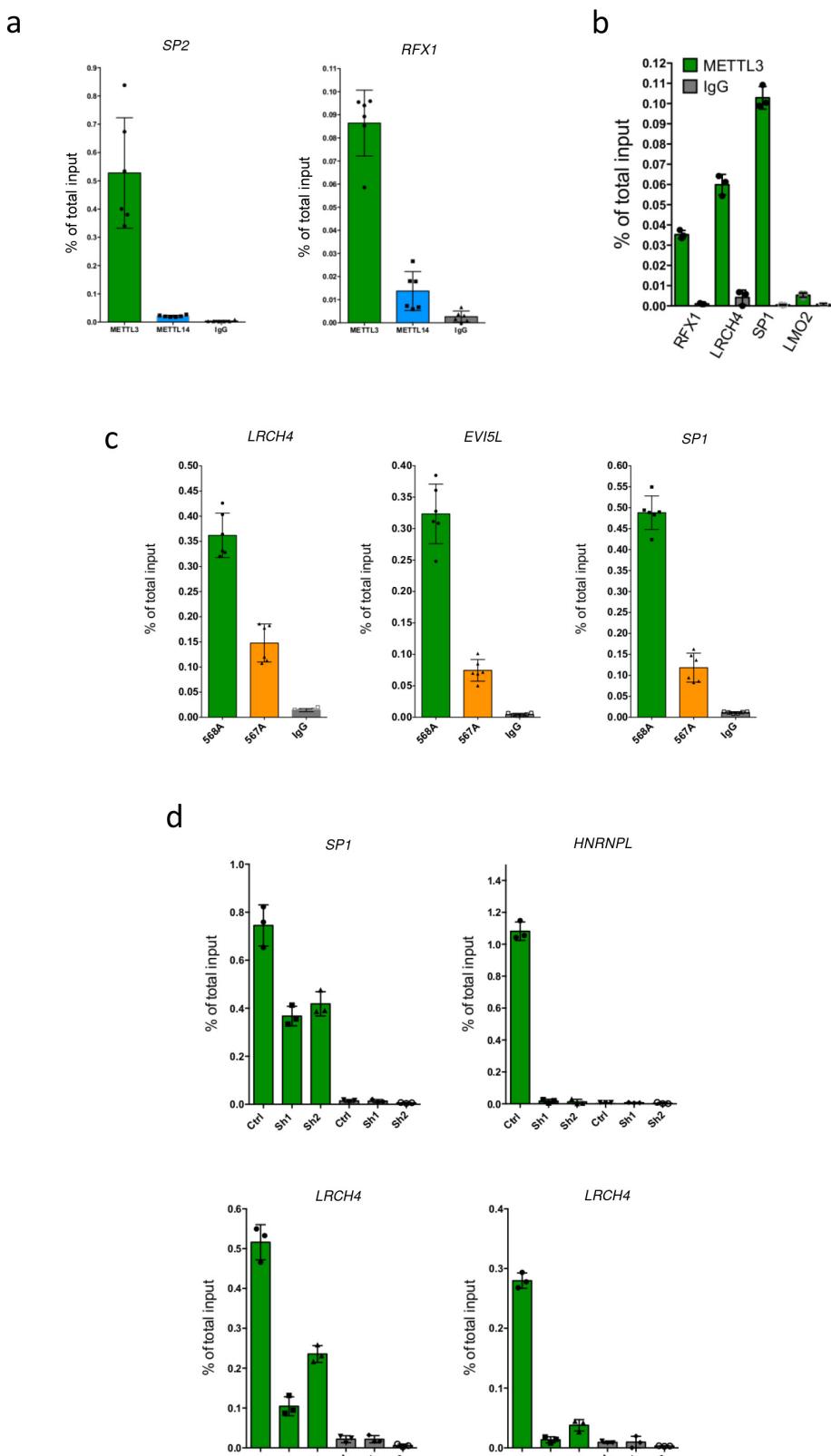
**Extended Data Figure 3 | METTL3 depletion in AML human cell lines leads to cell cycle arrest.** Related to Fig. 1. **a**, *METTL3* mRNA levels detected by RT-qPCR 4 days after shRNA induction with doxycycline in MOLM13 cells. The mean  $\pm$  s.e.m. of four independent cultures is shown. **b**, Western blot showing METTL3 and H3 levels in MOLM13 cells infected with specific or control TET-inducible shRNAs 5 days after doxycycline treatment. For gel source data see Supplementary Information. **c**, *METTL3* mRNA levels detected by RT-qPCR 4 days after shRNA induction with doxycycline in THP1 cells (left). The mean  $\pm$  s.e.m. of three independent cultures is shown. A proliferation assay of the cells was then performed with cell numbers measured between day 0 (4 d post doxycycline) and day 4 (8 d post doxycycline) (right). The mean  $\pm$  s.d. of two independent replicates is shown. **d**, Western blot for METTL3 and

actin in mouse AML cells. *Npm1c/Flt3<sup>ltd/+</sup>/Rosa26<sup>Cas9/+</sup>* mouse AML cells were transduced with gRNAs targeting the catalytic domain of Mettl3 and plasmids expressing either wild-type METTL3 or a catalytically inactive mutant (DW/AA). For gel source data see Supplementary Information. **e**, Volcano plots for *METTL3-KD* versus control samples, showing the significance *P* value ( $\log_{10}$ ) versus fold change ( $\log_2$ ) of gene expression. Significantly upregulated and downregulated transcripts are shown in red ( $|\log FC| > 1$ ,  $P < 0.001$ , FDR  $< 0.01$ ). **f**, Graphical representation of KEGG pathway regulation showing cell cycle downregulation (upper panel) and haematopoietic differentiation upregulation (lower panel) as obtained by comparing RNA-seq data from *METTL3-KD* and control MOLM13 cells (upregulated genes, red; downregulated genes, green).



**Extended Data Figure 4 | METTL3 is overexpressed in human AML and it is recruited on chromatin.** Related to Figs 1, 2. **a**, *METTL3* (top) and *METTL14* (bottom) mRNA expression levels across cancer types from the TCGA database. **b**, Proliferation assay of human AML cell lines upon transduction with a vector expressing *METTL3*. Cell numbers were measured between day 1 and day 3 after electroporation. The mean  $\pm$  s.d. of three independent replicates is shown. **c**, Western blot for *METTL3*, *METTL14*, GAPDH and histone H3 on cytoplasmic, nucleoplasmic and chromatin fractions from MOLM13 cells. For gel source data see Supplementary Information. **d**, Genomic browser screenshot of *METTL14*

and H3K4me3 normalized ChIP-seq datasets on the human SP2 gene locus from MOLM13 cells. **e**, Pie charts of genomic regions associated with *METTL14* (top) and *METTL3* (bottom) ChIP-seq peaks. **f**, Distribution of *METTL14* ChIP-seq reads centred on TSSs (upper) and histogram of ChIP-seq reads distribution relative to TSSs (lower). **g**, Top, Venn diagram showing the overlap between *METTL3* and *METTL14* peak datasets (statistical significance was evaluated by a  $\chi^2$  test). Bottom, distribution of *METTL3* and *METTL14* ChIP-seq reads centred on *METTL14* (left) or *METTL3* (right) peaks.



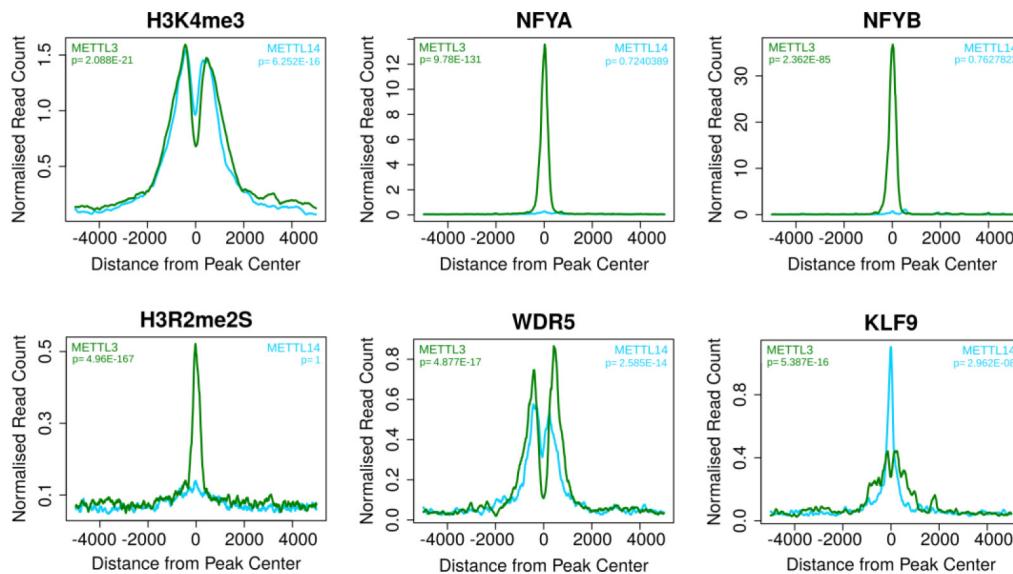
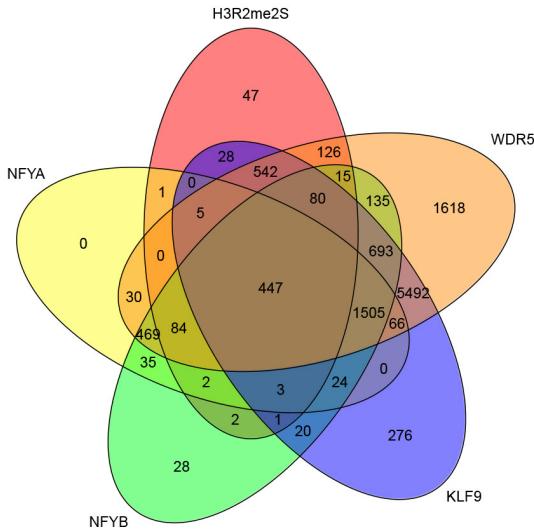
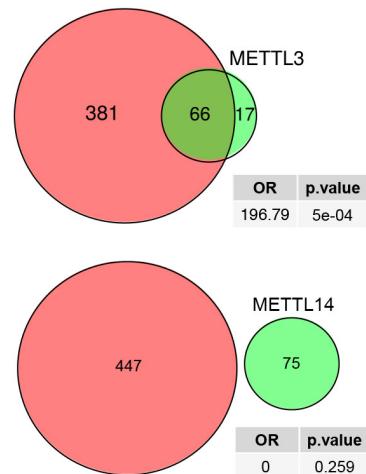
**Extended Data Figure 5 | Validation of METTL3 ChIP-seq.** Related to Fig. 2. **a**, ChIP-seq validation by ChIP-qPCR of METTL3 and METTL14 binding on the *SP2* and *RFX1* loci. The mean of six technical replicates  $\pm$  s.d. is shown. The experiment was performed independently three times. **b**, METTL3 ChIP-seq validation by ChIP-qPCR on the indicated loci. The *LMO2* promoter was used as a negative control. The mean of three technical replicates  $\pm$  s.d. is shown. The experiment was performed independently three times. **c**, METTL3 ChIP-seq validation by ChIP-qPCR on the indicated TSS in control or *METTL3*-KD MOLM13 cells, showing a specific reduction of METTL3 binding in *METTL3*-KD cells. The mean of three technical replicates  $\pm$  s.d. is shown. The experiment was performed independently three times.

by ChIP-qPCR on the indicated TSSs using two independent METTL3 antibodies in MOLM13 cells. The mean of six technical replicates  $\pm$  s.d. is shown. The experiment was performed independently three times.

**d**, METTL3 ChIP-seq validation by ChIP-qPCR on the indicated TSS in control or *METTL3*-KD MOLM13 cells, showing a specific reduction of METTL3 binding in *METTL3*-KD cells. The mean of three technical replicates  $\pm$  s.d. is shown. The experiment was performed independently three times.

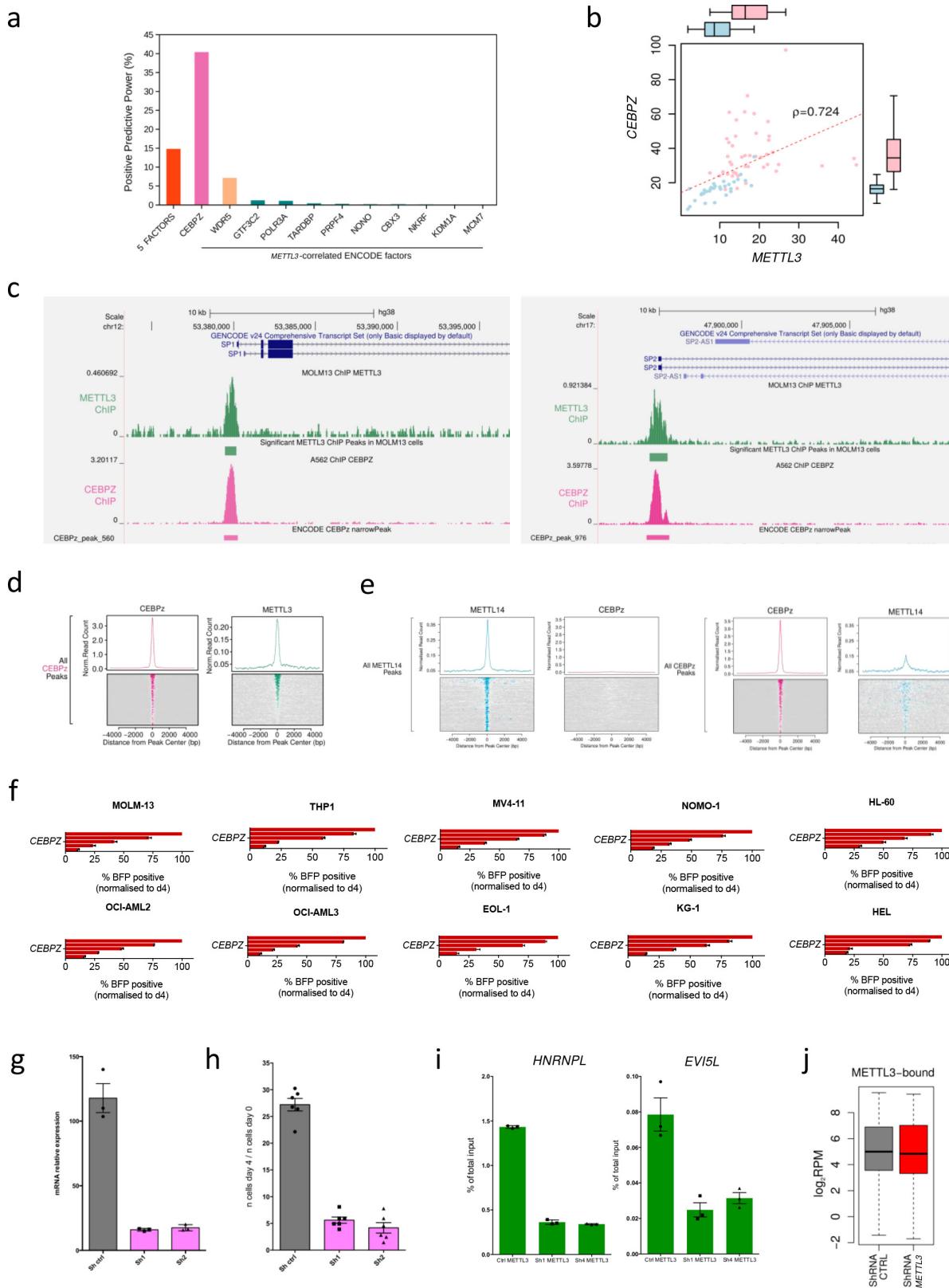
**a**

Rank	Motif	Name	P-value	log P-pvalue	q-value (Benjamini)	# Target Sequences with Motif	% of Targets Sequences with Motif	# Background Sequences with Motif	% of Background Sequences with Motif
1		NFY(CCAAT)/Promoter /Homer	1e-13	-3.050e+01	0.0000	20.0	58.82%	3508.6	8.03%
2		Klf9(Zf)/GBM-Klf9-ChIP-Seq(GSE62211)/Homer	1e-4	-9.345e+00	0.0139	10.0	29.41%	3104.8	7.11%
3		KLF14(Zf)/HEK293-KLF14 GFP-ChIP-Seq(GSE58341) /Homer	1e-3	-8.413e+00	0.0236	21.0	61.76%	13600.1	31.12%
4		Pbx3(Homeobox)/GM12878-PBX3-ChIP-Seq(GSE32465) /Homer	1e-2	-6.489e+00	0.1212	4.0	11.76%	643.0	1.47%
5		Pknox1(Homeobox)/ES-Prepl-ChIP-Seq(GSE63282) /Homer	1e-2	-6.222e+00	0.1267	4.0	11.76%	691.6	1.58%
6		Lhx3(Homeobox)/Neuron-Lhx3-ChIP-Seq(GSE31456) /Homer	1e-2	-5.144e+00	0.3103	10.0	29.41%	5318.1	12.17%

**b****c****d**

**Extended Data Figure 6 | METTL3 colocalizes with a defined set of chromatin factors.** Related to Fig. 2. a, Motif discovery analysis of the genomic sequences under METTL3 ChIP-seq peaks using HOMER. Significance was obtained using a hypergeometric test. b, Distribution of ChIP-seq reads for the indicated factors or histone modifications, centred on METTL3 (green) and METTL14 (blue) ChIP peaks. Statistical

significance of the binary overlap was evaluated by a  $\chi^2$  test. c, Venn diagram showing the overlap of H3R2me2s, WDR5, KLF9, NFYA and NFYB ChIP-seq peaks after filtering for H3K4me3 promoters. d, Venn diagram showing significant overlap between METTL3 peaks (but not METTL14 peaks) and the 447 loci carrying all five factors as in c. Statistical significance of the binary overlap was evaluated by a  $\chi^2$  test.

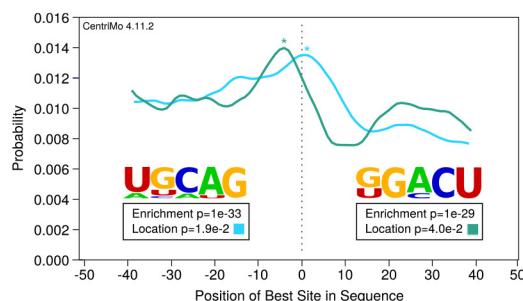
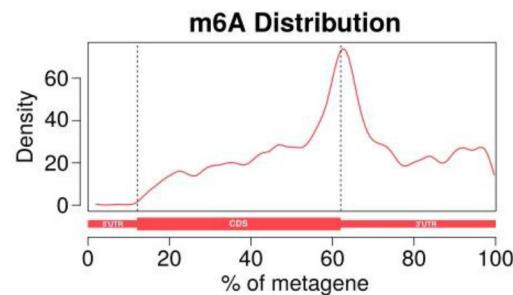
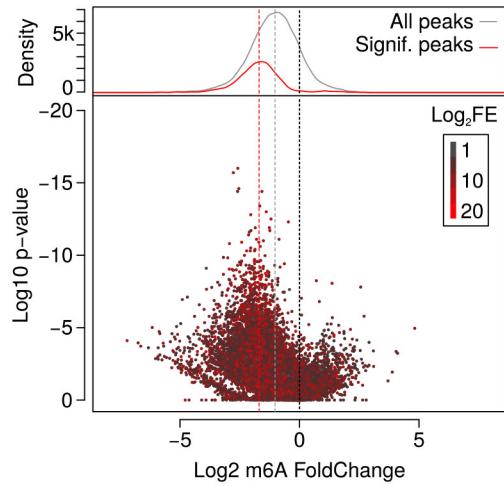
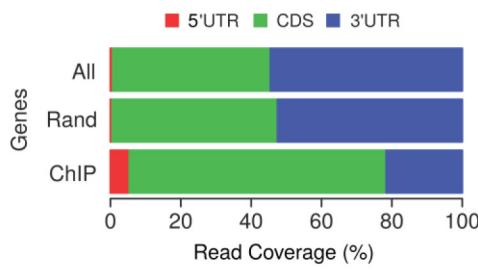
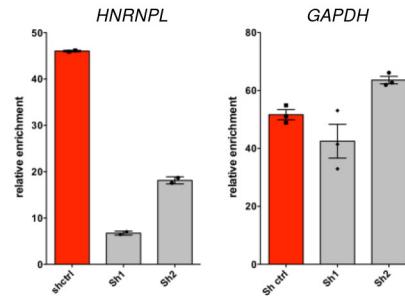
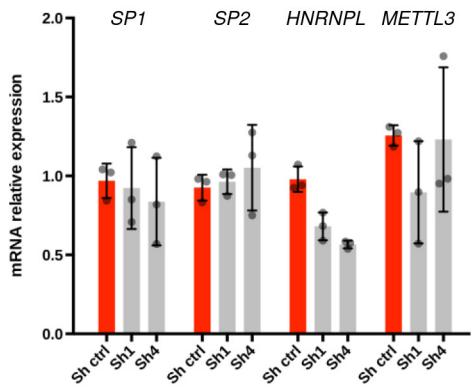
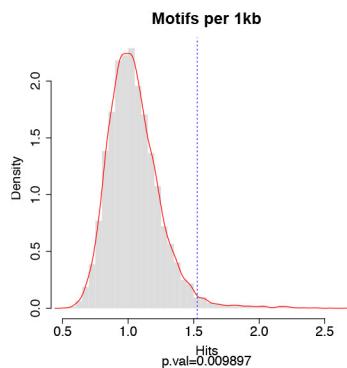


Extended Data Figure 7 | See next page for caption.

**Extended Data Figure 7 | CEBPZ recruits METTL3 on chromatin.**

Related to Fig. 2. **a**, Histogram representing the positive predictive power of the combined five factors compared with the predictive power of the ENCODE factors whose expression levels are tightly correlated with METTL3 expression. **b**, Correlation between *CEBPZ* and *METTL3* mRNA expression levels in the Human Protein Atlas RNA-seq datasets, including non-transformed (blue) and cancer (pink) cell lines. ( $\rho$ , Spearmann correlation coefficient). **c**, Genomic plot of METTL3 and CEBPZ normalized ChIP-seq datasets on human *SP1* and *SP2* gene loci in MOLM13 and K562 cells, respectively. **d**, Distribution and heatmaps of normalized ChIP-seq reads for METTL3 centred on CEBPZ peaks. **e**, Distribution and heat maps of normalized ChIP-seq reads of METTL14 and CEBPZ centred on METTL14 (left) and CEBPZ (right) peaks. **f**, Competitive co-culture assay showing negative selection of BFP<sup>+</sup> AML cell lines upon targeting of *CEBPZ* by CRISPR–Cas9 gRNAs. Cells were transduced with lentiviruses expressing a gRNA targeting the first exon

of *CEBPZ* and the BFP-positive fraction was compared with the non-transduced population. Results were normalized to those at day 4. The mean  $\pm$  s.d. of two independent infections is shown. **g**, *CEBPZ* mRNA levels detected by RT–qPCR 4 days after shRNA induction with doxycycline in MOLM13 cells. The mean  $\pm$  s.d. of three independent cultures is shown. **h**, Proliferation assay of control and *CEBPZ*-KD cells. Cell numbers were measured between day 0 (4 d post doxycycline) and day 4 (8 d post doxycycline). The mean  $\pm$  s.d. of six independent replicates is shown. **i**, ChIP–qPCR of METTL3 binding on target TSSs in MOLM13 cells expressing a control shRNA or two independent shRNAs against *CEBPZ*, showing a specific reduction of METTL3 binding in *CEBPZ*-KD cells. The mean of three technical replicates  $\pm$  s.d. is shown. The experiment was performed independently three times. **j**, Box plot representing the expression levels of METTL3 targets upon *METTL3*-KD from the dataset shown in Extended Data Fig. 3e.

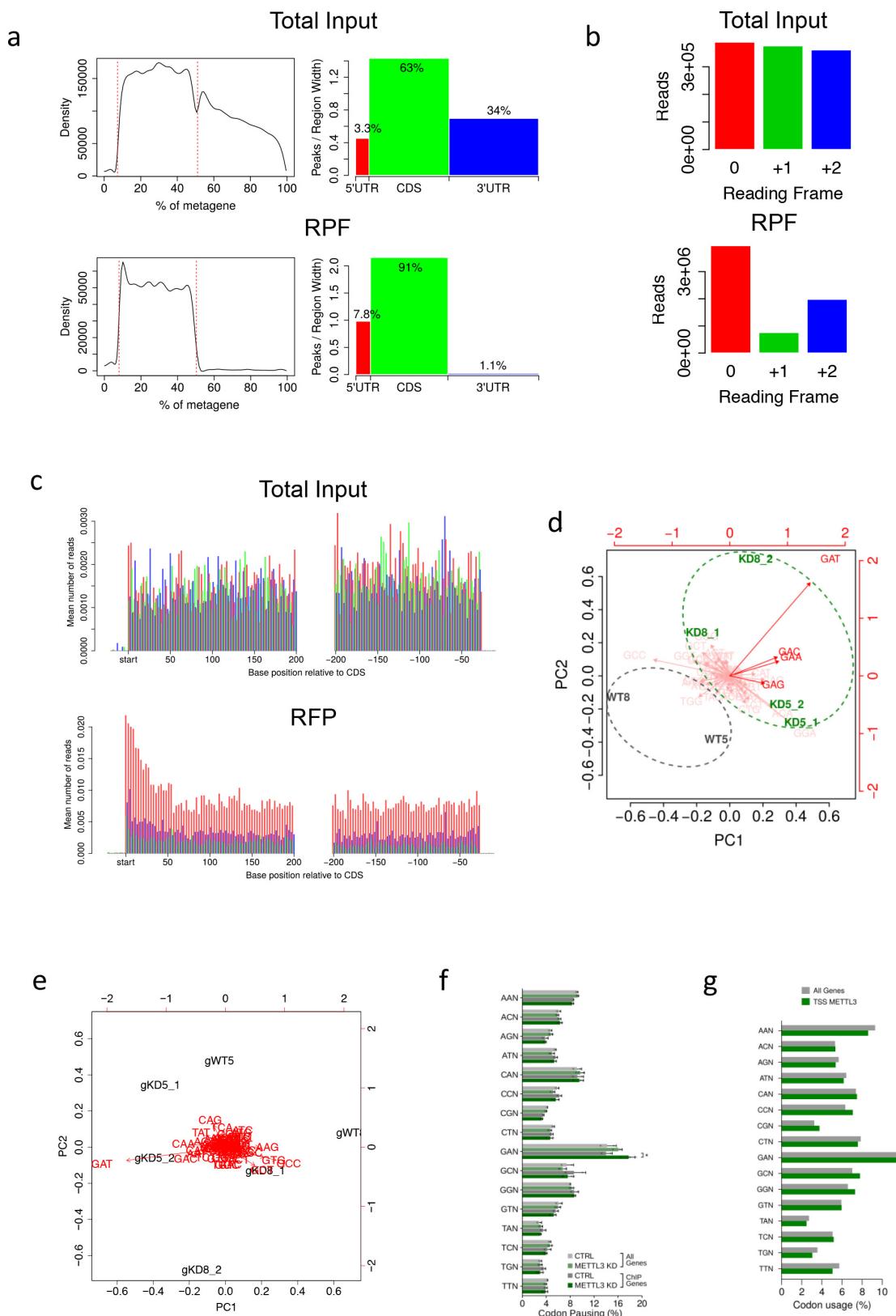
**a****b****c****d****e****f****g**

#### Extended Data Figure 8 | Validation of the m<sup>6</sup>A RNA

**immunoprecipitation upon METTL3 depletion.** Related to Fig. 3. **a**, Motif analysis under the identified m<sup>6</sup>A immunoprecipitation peaks showing enrichment of the expected UGCAG and GGACU sequences and their central distribution throughout the m<sup>6</sup>A immunoprecipitation peaks, as obtained by MEME and CentriMo. **b**, Distribution of m<sup>6</sup>A immunoprecipitation reads throughout the mRNA metatranscript, showing the expected enrichment around the STOP codon in MOLM13 cells. **c**, Scatter plots and density plot showing the general downregulation of m<sup>6</sup>A immunoprecipitation signal upon *METTL3*-KD in MOLM13 cells. **d**, Histogram showing METTL3-dependent m<sup>6</sup>A immunoprecipitation read coverage in mRNAs from METTL3-bound TSSs (ChIP), whole

transcriptome (All) or the permutation of random sets of genes (Rand).

**e**, m<sup>6</sup>A immunoprecipitation followed by qPCR for m<sup>6</sup>A peaks of *HNRNPL*, or *GAPDH* as a control. The plot shows the m<sup>6</sup>A immunoprecipitation signal over total input in MOLM13 cells expressing a control shRNA or shRNAs targeting *CEBPZ*. Mean  $\pm$  s.d. of three technical replicates is shown; experiment was performed independently twice. **f**, *SP1*, *SP2*, *HNRNPL* and *METTL3* mRNA levels detected by RT-qPCR 8 days after doxycycline induction in MOLM13 control or *CEBPZ*-KD cells. The mean  $\pm$  s.d. of three independent cultures is shown. **g**, Histogram showing the enrichment of the [GAG]<sub>n</sub> motif within the transcript sequences of METTL3 ChIP-targets compared with random permutations of genes.



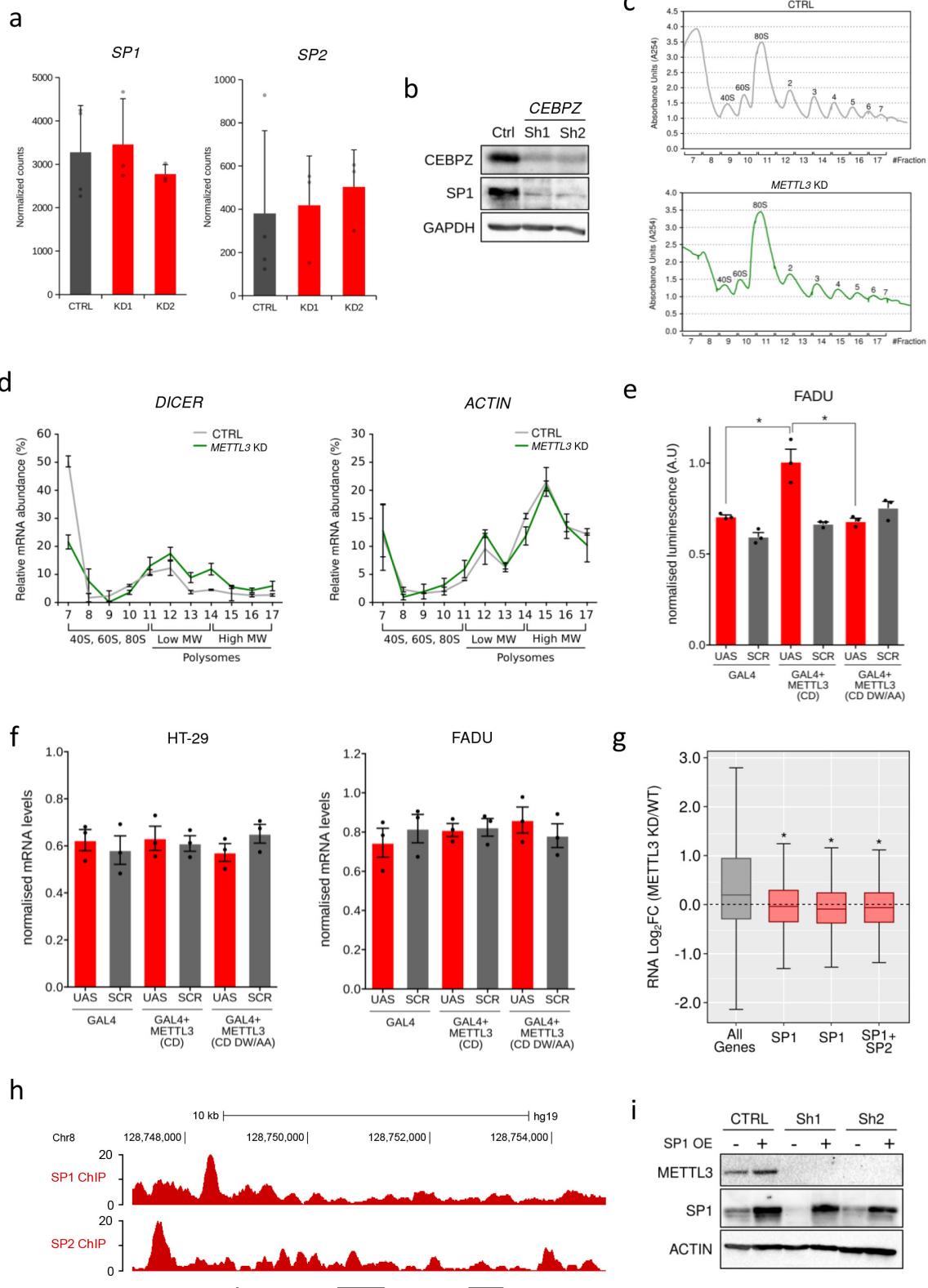
Extended Data Figure 9 | See next page for caption.

**Extended Data Figure 9 | Ribosome profiling analysis.** Related to Fig. 3. **a**, Distribution of ribosome profiling reads throughout the mRNA metatranscript from RNA inputs or ribosome-protected fragments (RPFs) showing absence of 3'UTR specifically in the RPF dataset.

**b**, Reading frame analysis of ribosome profiling reads from RNA inputs and RPFs in MOLM13 cells showing enrichment of the 0 reading frame specifically in the RPF reads. **c**, Average read alignments to 5' and 3' ends of coding sequences in RNA inputs (upper) or RPFs (lower) showing triplet periodicity and accumulation of reads on the start site typical of cycloheximide pre-treatment. **d**, Principal component analysis of P-site codon distribution on mRNAs from METTL3-bound TSSs obtained by

ribosome footprinting, 5 or 8 days after doxycycline administration, of METTL3-KD (KD5, KD8) or control (WT5, WT8) MOLM13 cells.

**e**, Principal component analysis of P-site codon distribution on all mRNAs, as in **d**. **f**, Frequency of P-site occupancy of codons in METTL3-KD or control MOLM13 cells for either all coding genes or genes harbouring a METTL3 ChIP peak on their promoter (\* $P < 0.05$ ; *t*-test). **g**, Frequency of codons within the coding sequence of METTL3 chromatin targets compared with the general frequency throughout the coding transcripts. The plot shows no significant overrepresentation of GAN codons in METTL3 chromatin targets.



Extended Data Figure 10 | See next page for caption.

**Extended Data Figure 10 | METTL3 controls the translation of SP1 and SP2.** Related to Fig. 4. **a**, RNA-seq normalized counts of *SP1* and *SP2* mRNAs from control or *METTL3*-KD MOLM13 cells at day 8 after doxycycline induction. Mean + s.d. of at least three biological replicates is shown. **b**, Western blot showing CEBPZ, SP1 and GAPDH levels in control and *CEBPZ*-KD cells. For gel source data, see Supplementary Information. **c**, Polysome fractionation analysis. Cell extracts from control or *METTL3*-KD cells were prepared and resolved in a 5–50% sucrose gradient. The absorbance at 254 was continuously measured. The peaks corresponding to free 40S and 60S subunits, 80S and polysomes are indicated. **d**, *DICER1* and *ACTB* mRNAs in each ribosome fraction were quantified through qPCR and plotted as a percentage of the total. Data are from two independent polysome-profiling experiments. Mean ± s.e.m. is shown. **e**, Firefly luciferase activity in FADU cell line from *UAS* or scrambled (SCR) sequence carrying plasmid in the presence of GAL4

either alone or fused with METTL3 wild-type (CD) or inactive (CD DW/AA) catalytic domain (\* $P < 0.05$ ; *t*-test). The mean + s.d. of three independent transfections is shown. **f**, Firefly luciferase mRNA from plasmids carrying *UAS* or scrambled sequence in the presence of GAL4 either alone or fused with METTL3 wild-type (CD) or inactive (CD DW/AA) catalytic domain, as evaluated by qPCR. The mean ± s.d. of three replicates is shown. **g**, Box plot showing transcriptional modulation of genes bound by SP1, SP2 or both between *METTL3*-KD and control MOLM13 cells (\* $P < 0.05$ ; Wilcoxon test). **h**, Genomic browser screenshot of SP1 and SP2 normalized ChIP-seq dataset on the human *MYC* gene locus in K562 cells (from ENCODE). **i**, Western blot showing METTL3, SP1 and ACTIN protein levels in MOLM13 cells infected with *METTL3*-specific or control TET-inducible shRNAs and with an *SP1* expression vector 5 days after doxycycline treatment. For gel source data see Supplementary Information.

## Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see [Reporting Life Sciences Research](#). For further information on Nature Research policies, including our [data availability policy](#), see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

### ► Experimental design

#### 1. Sample size

Describe how sample size was determined.

No sample size calculations were performed. Sample sizes were determined based on previous experience of the specific experimental setup.

#### 2. Data exclusions

Describe any data exclusions.

No Data were excluded

#### 3. Replication

Describe whether the experimental findings were reliably reproduced.

Experiments were successfully replicated the stated number of times in each figure legend.

#### 4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

No randomization

#### 5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

No blinding was used

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

#### 6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a Confirmed

- The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
- A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- A statement indicating how many times each experiment was replicated
- The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
- A description of any assumptions or corrections, such as an adjustment for multiple comparisons
- The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
- A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
- Clearly defined error bars

See the web collection on [statistics for biologists](#) for further resources and guidance.

## ► Software

Policy information about [availability of computer code](#)

### 7. Software

Describe the software used to analyze the data in this study.

The softwares used for the study are all publicly available.

>CRISPR Screens:

Enrichment and depletion of guides and genes were analyzed using MAGeCK statistical package; gRNAs were designed using WTSI Genome Editing website (<http://www.sanger.ac.uk/htgt/wge/>); Efficiency of genome editing in the pool of sgRNA-targeted cells was evaluated by Tracking of Indels by Decomposition (TIDE); Flow cytometry data were analysed with FlowJo.

>ChIP-seq:

Read quality was assessed using FastQC. Ribosomal contamination was removed by first mapping the reads to hg38 rDNA sequences using bwa and reads were mapped to the hg38 genome using bwa; Multiple reads mapping to a single genomic locus were removed using samtools rmdup. Reads were filtered to remove those with mapping quality less than 20 using samtools. Genome coverage bedgraph files were generated using genomeCoverageBed from the bedtools suite. Coverage files were converted to bigwig format using bedGraphToBigWig. Peaks were called against input sample using MACS2. Read count for each replicate was calculated using the GenomicRanges package in R. Genomic annotation of ChIP peaks and reads and gene set operations were performed taking advantage of the R packages ChIPseeker and VennDiagram, respectively. HOMER tool suite was used for DNA motif discovery.

>RNA sequencing experiments (M6A-IP, Ribosome Profiling, RNA-seq):

Read quality was assessed using FastQC. Ribosomal contamination was removed by first mapping the reads to hg38 rDNA sequences using bwa and reads were mapped to the hg38 genome using Tophat2; Reads were filtered to remove those with mapping quality less than 20 using samtools. Genome coverage bedgraph files were generated using genomeCoverageBed from the bedtools suite. Coverage files were converted to bigwig format using bedGraphToBigWig. Transcript assembly was performed using cufflinks and a single transcript database was generated using cuffmerge.

>m6A RNA-IP:

Statistical analysis of differentially methylated peaks was performed using the R package MeTDiff. Metagene plots were generated by RNAModR package (<https://github.com/mevers/RNAModR>). For evaluating statistical significance of [GAG]<sub>n</sub> motif enrichment, individual motif occurrences were searched throughout human transcriptome with FIMO program (MEME suite).

>Ribosome Profiling:

Statistical analysis of differentially translated genes in CTRL and METTL3 KD cells was performed using the R package xtail. In order to estimate the offset value relative to the read start required to localize the position of P-sites we employed the function “psite” of the plastid Python library.

>RNA-seq:

Gene counts were calculated at the transcript level for the combined transcript database from cuffmerge using summarizeOverlaps from the GenomicAlignments package in R. Differential gene expression analysis was conducted using DESeq2. Gene set enrichment analysis was performed using R package GAGE.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* [guidance for providing algorithms and software for publication](#) provides further information on this topic.

## ► Materials and reagents

Policy information about [availability of materials](#)

### 8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

All materials are readily available.

### 9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

For the ChIP experiments the following antibodies were used: anti-METTL3 from Bethyl Laboratories (A301-568A), anti-METTL3 from Bethyl Laboratories (A301-567A), rabbit polyclonal anti-METTL14 from Abcam (ab98166), anti H3k4me3 from Abcam (ab8580) and IgG Isotype Control (ab171870). Western blot experiment were performed using the following antibodies: anti-METTL3 from Bethyl Laboratories (A301-568A), anti-METTL14 from Abcam (ab98166), anti-Histone H3 from Active motif (39763), anti-WDR5 from Abcam (ab178410), anti-CEBPZ from Abcam (ab176579), anti-SP1 from Abcam (ab13370) and anti-SP2 from Abcam (ab137238), anti-ACTIN from Abcam (ab8227). For each genomic experiment the antibody lot number is provided on the GEO submission.

### 10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

All human cancer cell lines were obtained from the Sanger Institute Cancer Cell Collection.

b. Describe the method of cell line authentication used.

Cell lines were checked for morphology by microscope, as indicated by ATCC.

c. Report whether the cell lines were tested for mycoplasma contamination.

Cells were routinely tested for mycoplasma contamination by PCR.

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by [ICLAC](#), provide a scientific rationale for their use.

None of the cell lines used is listed within the ICLAC dataset.

## ► Animals and human research participants

Policy information about [studies involving animals](#); when reporting animal research, follow the [ARRIVE guidelines](#)

### 11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

$1 \times 10^5$  cells were transplanted into female 6-week old Rag2<sup>-/-</sup> IL2RG<sup>-/-</sup> mice by tail-vein injection. At day 14 post-transplant, the tumour burdens of the animals were detected using IVIS Lumina II (Caliper) with Living Image version 4.3.1 software (PerkinElmer). Animals were culled when the tumour burden was 108 photons per second or higher. Diseased mice welfare was assessed blindly by qualified animal technicians from the Sanger mouse facility. All animal studies were carried out in accordance with the Animals (Scientific Procedures) Act 1986, UK and approved by the Ethics Committee at the Sanger Institute.

Policy information about [studies involving human research participants](#)

### 12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

The study does not involve human research participants.

## ChIP-seq Reporting Summary

Form fields will expand as needed. Please do not leave fields blank.

### ► Data deposition

1. For all ChIP-seq data:

- a. Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
- b. Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

2. Provide all necessary reviewer access links.

*The entry may remain private before publication.*

<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE94613>

3. Provide a list of all files available in the database submission.

For each file we provide both raw sequencing data in SRA format and track-level information in bigwig format. Peak information is provided in the Supplementary Tables for easier access (equivalent to bed files).

4. If available, provide a link to an anonymized genome browser session (e.g. [UCSC](#)).

N/A

### ► Methodological details

5. Describe the experimental replicates.

For the METTL3, METTL14, H3K4me3 and IgG ChIP-Seq experiment 5, 5, 2 and 3 independent biological replicates were used, respectively.

6. Describe the sequencing depth for each experiment.

Total read number for each single replicate was on average around 30M, of which 85% unique mapping.

7. Describe the antibodies used for the ChIP-seq experiments.

IP antibody: H3K4me3; vendor: Abcam; catalog num: ab8580; lot num: GR240214-4; validation: <http://www.abcam.com/histone-h3-tri-methyl-k4-antibody-chip-grade-ab8580.html>

IP antibody: METTL3; vendor: Bethyl Laboratories; catalog num: A301-568A; lot num: 1;

IP antibody: METTL14; vendor: Abcam; catalog num: ab98166; lot num: GR277753-1;

8. Describe the peak calling parameters.

Peaks were called against input sample using MACS2 using default parameters. Peaks from all replicates were merged to give a master list of potential binding loci per condition, and read count (normalised by overall read depth of the library) for each replicate was calculated using the GenomicRanges package in R. Peaks were treated as potential binding loci if all replicates showed normalised score greater than 1 and did not overlap a peak called in the IgG. Genomic annotation of ChIP peaks and reads and gene set operations were performed taking advantage of the R packages ChIPseeker and VennDiagram, respectively.

9. Describe the methods used to ensure data quality.

In our analysis, METTL3 called peaks all had by definition FDR < 5% and 36% of them displayed fold enrichment greater than 5 (99% had FE greater than 2.5%)

10. Describe the software used to collect and analyze the ChIP-seq data.

Reads were trimmed to remove the TRUseq adapter using trim\_galore with parameters '-q 0 -a AGATCGGAAGAGCACACGTCTGAACTCAGTCAC --phred33 --fastqc'. Read quality was assessed using FastQC. Ribosomal

contamination was removed by first mapping the reads to hg38 rDNA sequences using bwa36 (default parameters), and removing all reads that mapped. Trimmed non-ribosomal reads were mapped to the hg38 genome using bwa with parameters '-n 3 -k 2 -R 300 -t 4'. Multiple reads mapping to a single genomic locus were treated as PCR duplicates, and were removed using samtools rmdup. Mapped reads were filtered to remove reads mapping to more than one unique genomic locus (multihits) by keeping only reads with flag XT:A:U in the output bam file from bwa. Reads were further filtered to remove those with mapping quality less than 20 using samtools. Genome coverage bedgraph files were generated using genomeCoverageBed from the bedtools suite of tools. Coverage files were converted to bigwig format using bedGraphToBigWig. HOMER tool suite was used for DNA motif discovery coupled with the hypergeometric enrichment calculations (or binomial) to determine motif enrichment.

## Flow Cytometry Reporting Summary

Form fields will expand as needed. Please do not leave fields blank.

### ► Data presentation

For all flow cytometry data, confirm that:

- 1. The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- 2. The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- 3. All plots are contour plots with outliers or pseudocolor plots.
- 4. A numerical value for number of cells or percentage (with statistics) is provided.

### ► Methodological details

5. Describe the sample preparation.

Cells were detached by trypsinization, washed and resuspended in PBS at RT. For immunodetection experiments, cells were stained with anti-mouse CD11b PE/Cy5 (Biolegend, cat. no. 101210) and anti-human CD11b PE (eBiosciences, cat. no. 9012-0118).

6. Identify the instrument used for data collection.

All flow cytometry data were collected on a LSRFortessa (BD) except for sorting experiment, which were performed on SH800S Cell Sorter (Sony).

7. Describe the software used to collect and analyze the flow cytometry data.

Data were acquired with the default control softwares of the machines employed, exported as .FCS files and analyzed with FlowJo software (FLOWJO, LLC).

8. Describe the abundance of the relevant cell populations within post-sort fractions.

The purity of the relevant cells was over 98% in the post-sort fraction as assessed by flow cytometry.

9. Describe the gating strategy used.

Cells were gated according to physical parameters in order to discard cell debris (FSC/SSC) and cell clumps (Width vs. Area). Dead cells were excluded by selecting the DAPI or PI- negative cell populations. Fluorescent cells were gated with a threshold capturing the 0.5% upper tail of a negative control population.

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.