Upregulated cholesterol biosynthesis facilitates the survival of methylation-retaining AML cells following decitabine treatment.

Danielle R Bond[1,2^*], Sean M Burnard[1,2^], Kumar Uddipto[1,2], Kooper V Hunt[1,2], Brooke M Harvey[1,2], Luiza Steffens Reinhardt[1,3], Charley Lawlor-O'Neill[1], Ellise A Roper[1], Sam Humphries[1,2], Heather C. Murray[1,2], Nikola A Bowden[4,5], Anoop K Enjeti[2,4,6,7], Nicole M Verrills[1,2], Carlos Riveros[8], Kim-Anh Lê Cao[9], Heather J Lee[1,2*]

[1]*School of Biomedical Science and Pharmacy, The University of Newcastle, Australia.*

[2]*Precision Medicine Research Program, Hunter Medical Research Institute, Australia.*

[3]*Cancer Detection and Therapy Research Program, Hunter Medical Research Institute, Australia.*

[4]*School of Medicine and Public Health, The University of Newcastle, Australia.*

[5]*Drug Repurposing and Medicines Program, Hunter Medical Research Institute, Australia.*

[6]*Department of Haematology, Calvary Mater Hospital, Australia.*

[7]*New South Wales Health Pathology, John Hunter Hospital, Australia.*

[8]*Data Science, Hunter Medical Research Institute, Australia.*

[9]*Melbourne Integrative Genomics, School of Mathematics and Statistics, The University of Melbourne, Australia.*

^ Joint lead authors

* Corresponding authors

Abstract

DNA hypomethylating agents (HMAs) are used to treat acute myeloid leukaemia (AML) and myelodysplasia patients who are unsuitable for intensive chemotherapy, but low response rates and therapy-resistant relapse remain significant challenges. To optimise HMA efficacy, we must understand how resistance and relapse arise from cells that survive treatment. Here we combine single-cell multi-omic analysis with parallel colony-forming assays to link HMA-induced molecular heterogeneity with functional consequences in AML cells. HMAs, azacytidine (AZA) and decitabine (DAC), induced global epigenetic heterogeneity associated with upregulation of inflammatory responses and cell death pathways in a subset of hypomethylated cells. Some AML cells maintained high DNA methylation during treatment, and these methylation-retaining cells had increased self-renewal capacity following DAC, but not AZA. Molecular profiling of individual colonies revealed upregulated cholesterol biosynthesis as an adaptation to HMA treatment, and inhibition by rosuvastatin enhanced DAC effects *in vitro* and *in vivo*. Thus, HMA-induced heterogeneity has important implications for AML cell growth and statins are a candidate co-treatment strategy to delay or prevent HMA-resistant relapse.

Introduction

Hypomethylating agents (HMAs) are DNA methyltransferase (DNMT) inhibitors that are used to treat patients with Acute Myeloid Leukaemia (AML) and the pre-leukaemic condition Myelodysplastic Syndrome (MDS). Two commonly prescribed HMAs are the cytidine analogues, azacytidine (AZA, 5-azacytidine) and decitabine (DAC, 2'-deoxy-5-azacytidine), which are incorporated into DNA during replication[1,2]. This leads to degradation of DNMTs[3] and loss of DNA methylation in subsequent cell divisions[2,4,5].

The relatively limited side effects of these epigenetic therapies make them useful alternatives to standard intensive chemotherapies, and they are routinely administered to older, or otherwise unfit, AML patients. While single agent HMA treatment extends survival in many patients[6-12], only 20% have a complete response to therapy[13]. Responses are also short-lived (e.g., 8-15 months[14]), with acquired resistance leading to relapse in most patients. To address these limitations, many clinical trials are testing co-treatment strategies with some studies delivering improved outcomes[15]. For example, the pro-apoptotic therapy, venetoclax, has increased response rates in elderly AML patients undergoing HMA treatment[16]. However, relapse remains a significant problem with the median duration of response being 11-18 months for patients treated with both venetoclax and HMA therapy[16,17]. To improve the long-term benefits of HMA therapy, we must understand the

molecular mechanisms responsible for HMA efficacy, as well as the processes by which relapse arises from cells that survive treatment.

While initial studies demonstrated that HMA treatment can eliminate cancer cells with self-renewal capacity[18], some leukaemic stem cell (LSC) populations were shown to survive AZA treatment in AML and MDS patients[19]. This may be facilitated by integrin signaling in the bone marrow microenvironment, which induces quiescence in blasts and predicts AZA response in MDS patients[20]. Another study demonstrated that LSCs increase CD70 expression following HMA treatment, and CD70 blockade by cusatuzumab was shown to reduce the self-renewal capacity of AML patient blasts[21]. A phase I/II clinical trial of this combination recently reported a marginal improvement in survival compared to historical data from a similar patient cohort[22,23].

Altered pyrimidine metabolism can also facilitate HMA resistance. Prior to DNA incorporation, both DAC and AZA must be converted into a deoxycytidine triphosphate analogue, Aza-dCTP[24]. Unlike DAC, AZA can also be converted to AZA-CTP and incorporated into RNA, where it influences transcript stability and translation[25,26]. Distinct enzymes catalyse these reactions and are dysregulated at relapse[24]. For example, Deoxycytidine Kinase (DCK) is required for the metabolism of DAC, but not AZA. Loss of *DCK* promotes resistance to DAC[27], and its expression is downregulated in patients who develop relapse on DAC, but not AZA[24].

In this study, we further characterise the cellular processes that facilitate AML cell survival and proliferation following HMA treatment. Our single-cell multi-omic analysis reveals global HMA-induced DNA methylation heterogeneity and methylation-retaining cells that appear to evade treatment. In parallel colony-forming assays, we show that methylation-retaining cells have a growth advantage following treatment with DAC, but not AZA, and reveal upregulation of cholesterol biosynthesis in cells surviving treatment. Co-treatment with statins enhances the effects of DAC in colony assays and extends the survival of leukaemia-bearing mice. Together, our work suggests that relapse may arise from metabolic adaptation in therapy-evading cells and identifies a candidate treatment strategy to enhance the long-term efficacy of HMA therapy.

Results

DNA methylation heterogeneity induced by HMA treatment

To characterise the responses of individual AML cells to HMA treatment, we performed single-cell analysis of DNA methylation. AML cell lines (HL-60, MOLM-13, MV-4-11) were treated with low doses of decitabine (DAC) or azacytidine (AZA) to induce maximal demethylation with minimal effects on cell growth and viability (Supplementary Fig. 1). After 3 days of HMA treatment, striking

86 heterogeneity in DNA methylation levels was observed (Fig. 1A, Supplementary Table 1). While
87 untreated cells had homogeneously high levels of DNA methylation (e.g., HL-60: 65-73%), the extent
88 of hypomethylation varied substantially among cells treated with DAC (e.g., HL-60: 17-69%) or AZA
89 (e.g., HL-60: 20-69%). Interestingly, a small proportion (1-5%) of methylation-retaining cells (Fig. 1A,
90 dashed boxes) displayed no evidence of HMA-induced hypomethylation, with DNA methylation
91 levels comparable to untreated cells. The extent of HMA-induced hypomethylation was related to
92 cell division rate, as indicated by positive correlations between DNA methylation and CellTrace
93 fluorescence (Fig. 1B). This is consistent with HMA incorporation during replication[2] and confirms
94 that slowly dividing cells can avoid the effects of HMA treatment.

95 We next examined DNA methylation in different genomic contexts to test whether heterogeneity is
96 observed across the genome. Contexts with high levels of DNA methylation in untreated HL-60 cells
97 (i.e., exons, introns, intergenic regions) showed the greatest reductions upon HMA treatment (Fig.
98 1C), and this loss of methylation was accompanied by increased cell-to-cell heterogeneity (Fig. 1D).
99 Active promoters marked by H3K4me3 were the only genomic features without increased DNA
100 methylation heterogeneity following HMA treatment.

101 To explore the consequences of DNA methylation heterogeneity on other layers of genetic
102 regulation, we used the multi-omic data collected from HL-60 cells. The scNMT-seq method profiles
103 DNA methylation, DNA accessibility and gene expression in parallel[28], allowing these molecular
104 modalities to be correlated with each other across the genome of individual cells (Fig. 1E). Untreated
105 cells showed the expected trends, with DNA methylation being negatively correlated to accessibility
106 and gene expression, while DNA accessibility and transcription were positively correlated. In HMA-
107 treated cells, the relationship between DNA methylation and accessibility was weakened in all
108 genomic contexts. HMA treatment also weakened the correlation between DNA methylation (in CpG
109 islands, promoters, exons, and introns) and expression of associated genes. In contrast, HMA
110 treatment had minimal impact on associations between DNA accessibility and gene expression.
111 These observations imply that loss of DNA methylation is not always accompanied by increased
112 accessibility and transcription, but may introduce epigenetic noise in an otherwise well-regulated
113 system.

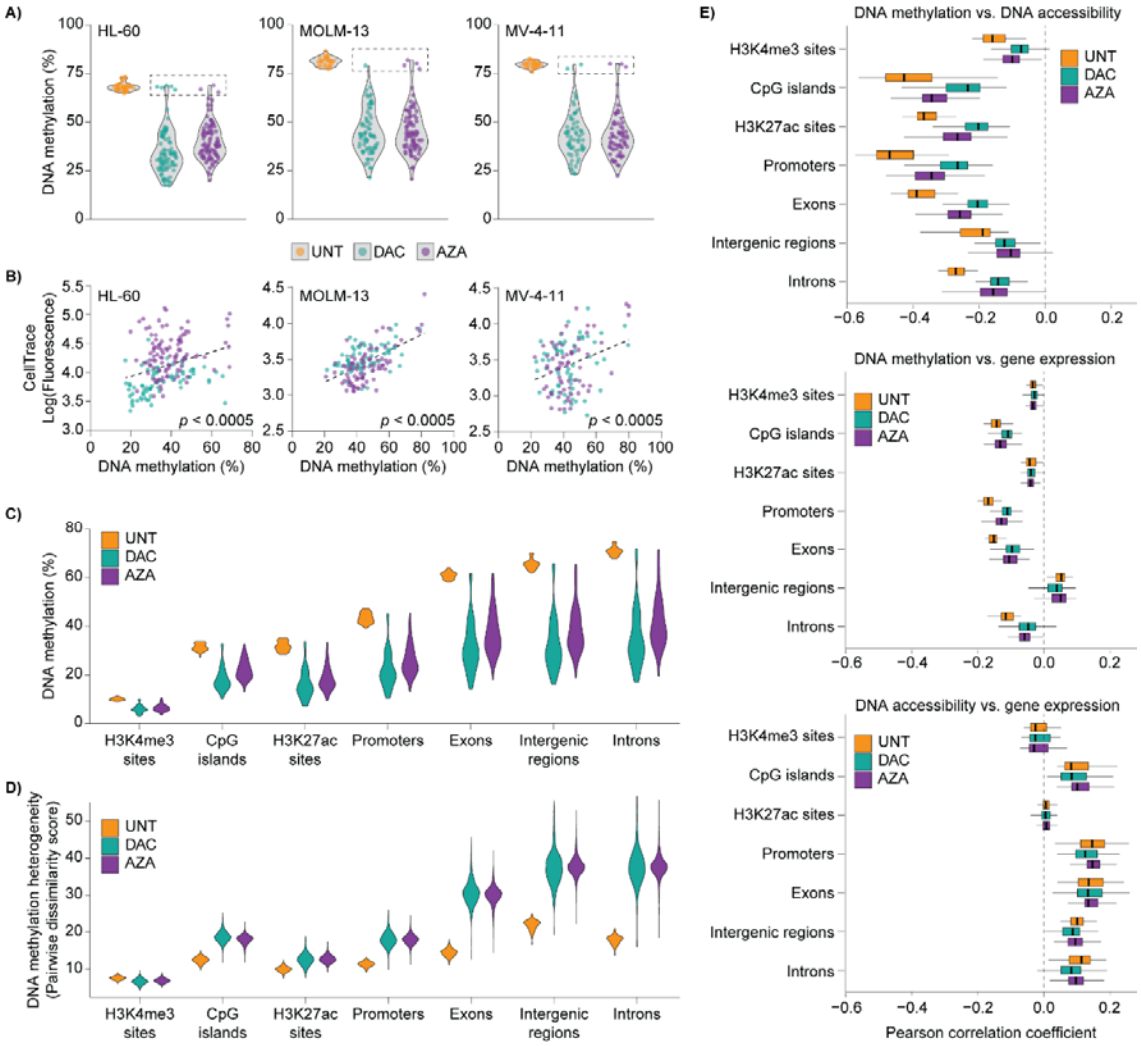
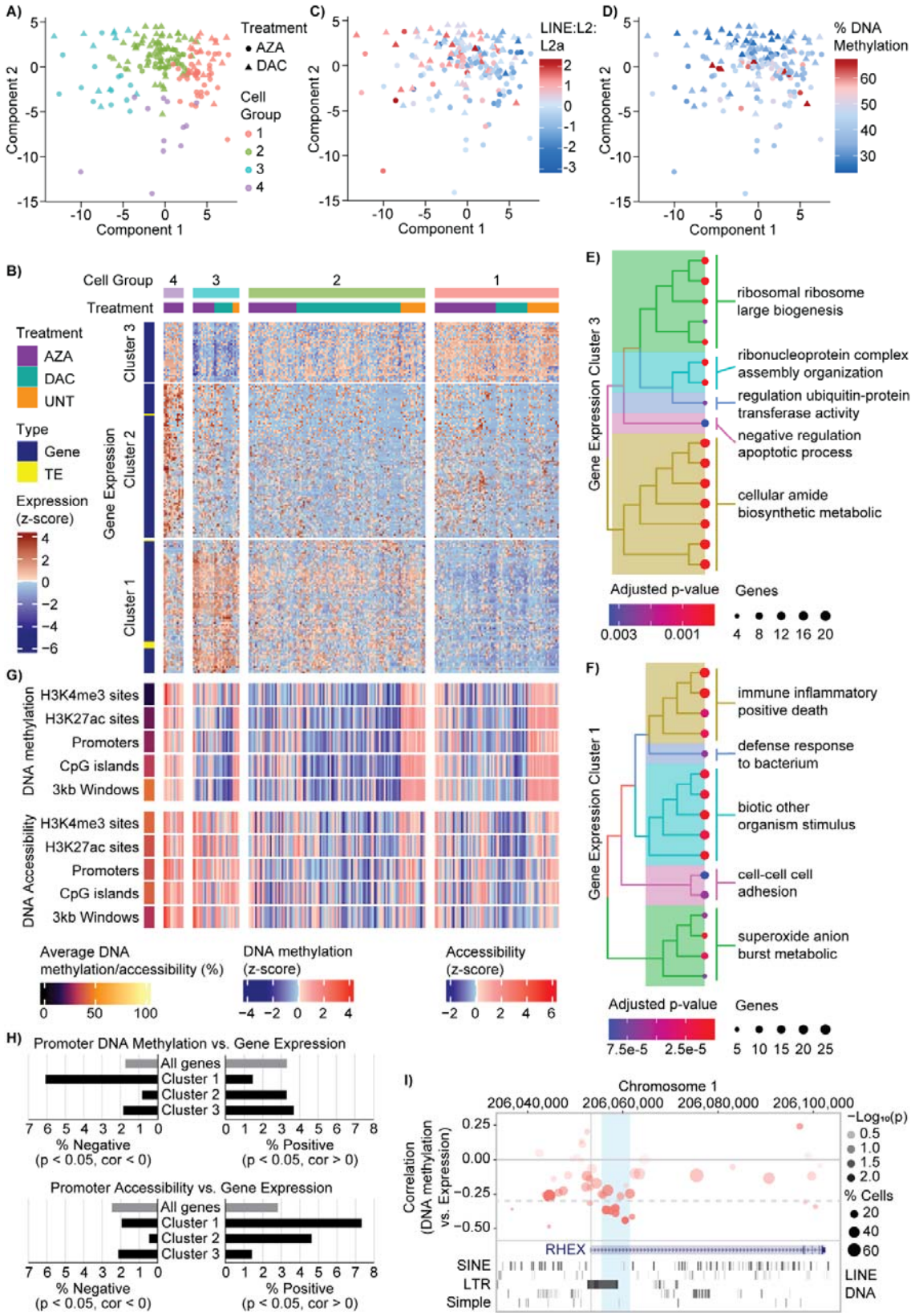


Figure 1. HMA treatment induces DNA methylation heterogeneity in AML cells. HL-60, MOLM-13 and MV-4-11 cells were labelled with CellTrace and treated with decitabine (DAC; 100nM) or azacytidine (AZA; HL-60: 2000nM, MOLM-13 and MV-4-11: 500nM) every 24h for 72h. Single cells collected by indexed FACS were subjected to scNMT-seq (HL-60) or scTEM-seq (MOLM-13, MV-4-11). **A)** Violin plots of DNA methylation levels in single HL-60 (left), MOLM-13 (centre) and MV-4-11 (right) cells. Superimposed points show single-cell values from untreated (UNT, orange), DAC (cyan) and AZA (purple) groups. Dashed boxes surround DAC and AZA cells with methylation levels within the range of UNT samples. Data are shown for 185-222 cells collected from 2-3 replicate experiments in each cell line (UNT, *n = 27-38*; DAC *n = 63-93*; AZA *n = 68-91*). **B)** Scatter plots comparing CellTrace fluorescence and DNA methylation in single cells, with linear regression analysis and F-test *p*-values. **C)** Violin plots of DNA methylation levels in different genomic contexts from HL-60 scNMT-seq data. **D)** Violin plots of DNA methylation heterogeneity, as determined by pairwise dissimilarity analysis, within different genomic contexts from HL-60 scNMT-seq data. **E)** Box and whisker plots of Pearson correlation coefficients computed within single cells from HL-60 scNMT-seq data. DNA methylation and DNA accessibility were considered in different genomic contexts, and individual loci were matched based on genomic co-ordinates. Correlations were performed between DNA methylation and DNA accessibility (top), DNA methylation and gene expression (middle), and DNA accessibility and gene expression (bottom). Boxes depict the interquartile range (IQR) with median. Whiskers extend to the highest and lowest data points within 1.5 x IQR of the first and third quartile. Outlying data points are not shown.

Transcriptional programs linked to HMA-induced epigenetic heterogeneity

HMA mechanisms of action have been difficult to precisely define because genome-wide loss of DNA methylation is associated with pleiotropic transcriptional changes[15]. HMA-induced promoter hypomethylation is thought to allow re-expression of tumour suppressor genes[29], re-activation of DNA repair pathways[30] and upregulation of differentiation markers[31,32]. HMA treatment can also increase expression of cancer testis antigens[33] and enhance antigen presentation on cancer cells[34]. Furthermore, genome-wide de-repression of transposable elements (TEs) has been shown to trigger a viral mimicry response in which dsRNA stimulate interferon signalling and apoptosis[35-37]. To clarify how HMA-induced epigenetic heterogeneity influences transcriptional responses we performed a multivariate analysis to integrate the three molecular layers: DNA methylation, DNA accessibility, and transcription. We excluded untreated cells to focus on variability among HMA-treated cells and applied an unsupervised sparse Partial Least Squares (sPLS) method[38]. This method performed feature selection to identify variably expressed transcripts that are highly correlated to changes in DNA methylation and accessibility in regulatory regions (CpG islands, promoters, H3K4me3 sites and H3K27ac sites) and 3kb genomic windows.

The 200 transcript features selected by sPLS were further examined in both treated and untreated cells, revealing 4 groups of cells with distinct transcriptional profiles (Fig. 2A, B). Consistent with the induction of viral mimicry, a gradient of TE expression was observed across component 1. For example, cell group 3 had high expression of LINE: L2a and low global DNA methylation levels (Fig. 2A, C and D). Gene ontology (GO) over-representation analysis (ORA) was performed for each of the 3 clusters of transcript features (Fig. 2B; Supplementary Table 2). Cell group 3 had low expression of genes in cluster 3, which were related to translation and inhibition of cell death (Fig. 2E, Supplementary Table 3). Simultaneously, cell group 3 had high expression of genes in cluster 1, which were enriched in terms related to immune inflammatory response and positive regulation of cell death (Fig. 2F, Supplementary Table 3). This transcriptional profile is consistent with the expected effects of HMA treatment[1], and 29 of the 78 genes in expression cluster 1 were significantly upregulated by DAC and/or AZA in matched bulk RNA sequencing (RNA-seq) data (Supplementary Table 2). In contrast, cell group 1 displayed an inverted gene expression pattern which was shared with many untreated cells (observed:expected ratio = 1.48). This suggests that cell group 1 did not activate transcriptional pathways commonly associated with HMA treatment, despite low methylation levels in most cells (Fig. 2A, D). The expression profile of cell group 2 was intermediate between groups 1 and 3, whereas cell group 4 showed distinctive upregulation of gene expression cluster 2. No ontology terms from gene cluster 2 retained significance after multiple testing correction.


A)
Component 2
Component 1
Treatment
AZA
DAC
Cell Group
1
2
3
4
C)
LINE:L2: L2a
D)
% DNA Methylation
B)
Cell Group
Treatment
Treatment
AZA
DAC
UNT
Type
Gene
TE
Expression (z-score)
Cluster 3
Cluster 2
Cluster 1
Gene Expression
E)
Gene Expression Cluster 3
ribosomal ribosome large biogenesis
ribonucleoprotein complex assembly organization
regulation ubiquitin-protein transferase activity
negative regulation apoptotic process
cellular amide biosynthetic metabolic
Adjusted p-value
Genes
F)
Gene Expression Cluster 1
immune inflammatory positive death
defense response to bacterium
biotic other organism stimulus
cell-cell cell adhesion
superoxide anion burst metabolic
G)
DNA methylation
H3K4me3 sites
H3K27ac sites
Promoters
CpG islands
3kb Windows
DNA Accessibility
Average DNA methylation/accessibility (%)
DNA methylation (z-score)
Accessibility (z-score)
H)
Promoter DNA Methylation vs. Gene Expression
All genes
Cluster 1
Cluster 2
Cluster 3
% Negative (p < 0.05, cor < 0)
% Positive (p < 0.05, cor > 0)
Promoter Accessibility vs. Gene Expression
I)
Chromosome 1
Correlation (DNA methylation vs. Expression)
−Log10(p)
% Cells
RHEX
SINE
LTR
Simple
LINE
DNA

169

170

171 **Figure 2. HMA-induced epigenetic heterogeneity influences transcriptional programs in AML cells.**
172 HMA treated (AZA and DAC only) HL-60 scNMT-seq data underwent multivariate feature selection by
173 sparse least squares (sPLS) using an unsupervised model. **A)** sPLS projection of cells based on
174 transcript features coloured by cell group 1-4 (from B). **B)** Heatmap of transcript features selected by
175 sPLS displaying all samples (treated and untreated) as columns, are split by k-means clustering and
176 grouped by treatment. Individual gene and TE expression levels (rows) are z-score normalised and
177 split by k-means clustering with internal hierarchical clustering. **C-D)** sPLS projections of cells based
178 on transcript features are coloured by **C)** LINE:L2:L2a expression, and **D)** global methylation level. **E-**
179 **F)** Tree plots of the gene ontology (GO) over-representation analysis (ORA) for gene expression
180 clusters 3 and 1. **G)** Heatmaps summarising the DNA methylation and DNA accessibility features
181 selected by sPLS. Samples (columns) are ordered according to the heatmap in B. The average
182 methylation or accessibility of all sPLS selected features across all samples (treated and untreated) is
183 displayed on the left for each genomic context. The two heatmaps show z-score normalised
184 averages of DNA methylation and accessibility for all sPLS selected features in each genomic context.
185 **H)** Pearson correlations were computed between gene expression and DNA methylation (top) or
186 accessibility (bottom) of associated promoters. Bar graphs show the percentage of correlations ($p <$
187 0.05) with negative (left) and positive (right) coefficients for all genes and filtered by gene expression
188 cluster (identified in D). **I)** Correlations between *RHEX* expression and DNA methylation of adjacent
189 loci (top) are displayed together with annotated short and long interspersed nuclear elements (SINE,
190 LINE), long terminal repeat (LTR), DNA, and Simple repeat sequences (bottom). The blue shading
191 highlights a promoter-proximal region of intron 1 that included 6 regions with negative correlations
192 (cor < -0.3, $p < 0.05$) between DNA methylation and *RHEX* expression.

193

194 When considering the epigenetic features selected by sPLS (Supplementary Tables 4 and 5), we
195 noted that cell group 3 was the only group to have relatively low DNA methylation and high
196 chromatin accessibility across several genomic contexts (Fig. 2G). This suggests that the
197 transcriptional response to HMA treatment depends on both reductions in DNA methylation and
198 gains in accessibility. Consistently, transcript features from expression cluster 1 had predominantly
199 positive correlations with accessibility features, and many negative correlations to methylation
200 features, especially in CpG islands and 3kb genomic windows (Supplementary Fig. 2).

201 To test whether epigenetic alterations in cis-regulatory elements could influence transcriptional
202 responses to HMA treatment, we next correlated gene expression to DNA methylation and
203 accessibility in nearby genomic loci. Genes from expression cluster 1 showed a significant shift
204 toward negative correlations with promoter methylation ($p = 2.15\text{x}10^{-11}$, χ^2 test) and positive
205 correlations with promoter accessibility ($p = 2.26\text{x}10^{-8}$, χ^2 test) (Fig. 2H, Supplementary Table 6). An
206 interesting example is *RHEX* (regulator of hemoglobinization and erythroid cell expansion), which is
207 highly expressed in AML[39]. Negative correlations between *RHEX* expression and DNA methylation
208 were concentrated in a promoter-proximal region of intron 1 that contained several conserved long
209 terminal repeat (LTR) TEs (Fig. 2I). A previous study has identified an AML-specific *RHEX* transcript
210 resulting from onco-exaptation of an upstream LTR2B element[39], so HMA-induced hypomethylation
211 could induce other non-canonical transcripts by activating additional TEs.

In summary, our single-cell multi-omic analysis has identified transcriptional changes linked to the patterns of epigenetic heterogeneity in HMA-treated AML cells. Importantly, activation of genes involved in inflammatory response and cell death pathways (expression cluster 1) with simultaneous down-regulation of inhibitory genes involved in these pathways (expression cluster 3) is observed in only a sub-set of hypomethylated cells (cell group 3) (Fig. 2A, C, D).

Functional consequences of HMA-induced heterogeneity

To determine the functional consequences of epigenetic and transcriptional heterogeneity in HMA-treated cells, we next performed colony-forming assays. DAC and AZA significantly decreased colony counts (Fig. 3A), confirming that HMA treatment decreases the self-renewal capacity of AML cells[18]. DAC treatment also significantly increased the size of MOLM-13 and MV-4-11 colonies, indicative of increased proliferation during colony formation (Supplementary Fig. 3A). To characterise the molecular profiles of colonies formed after HMA treatment, we performed single-colony analysis of DNA methylation (Fig. 3B, C). Following DAC treatment (Fig. 3B, solid fill), we observed a high percentage of colonies (34-65%) with DNA methylation levels comparable to untreated colonies (≥75%). This was in stark contrast to the low percentage of methylation-retaining cells (1-5%) observed in single-cell data after 72h DAC treatment (Fig. 1A, dashed box; Fig. 3B, dashed line). Surprisingly, this shift towards higher DNA methylation levels was far less pronounced in colonies established following AZA treatment (Fig. 3C).

To assess recovery of DNA methylation following HMA treatment, a time-course analysis was performed during the colony-forming assay (Fig. 3D). DNA methylation levels in HMA-treated HL-60 colonies were low throughout the time-course (Fig. 3D, top), mirroring an analysis performed in suspension culture (Supplementary Fig. 1E). Similar results were obtained for MOLM-13 and MV-4-11 colonies derived after AZA treatment (Fig. 3D, centre and bottom). In contrast, high levels of DNA methylation were observed at early stages of colony formation (experiment day 6) following DAC treatment of MOLM-13 and MV-4-11 cells (Fig. 3D, centre and bottom). This suggests that the shift toward high DNA methylation observed in these colonies (Fig. 3B) is not due to a gradual recovery of methylation. Rather, our data is consistent with the selection of highly methylated cells in colony-forming assays performed after DAC treatment. We deduce that methylation-retaining cells have increased self-renewal and proliferative capacity relative to hypomethylated cells, after treatment with DAC, but not AZA.

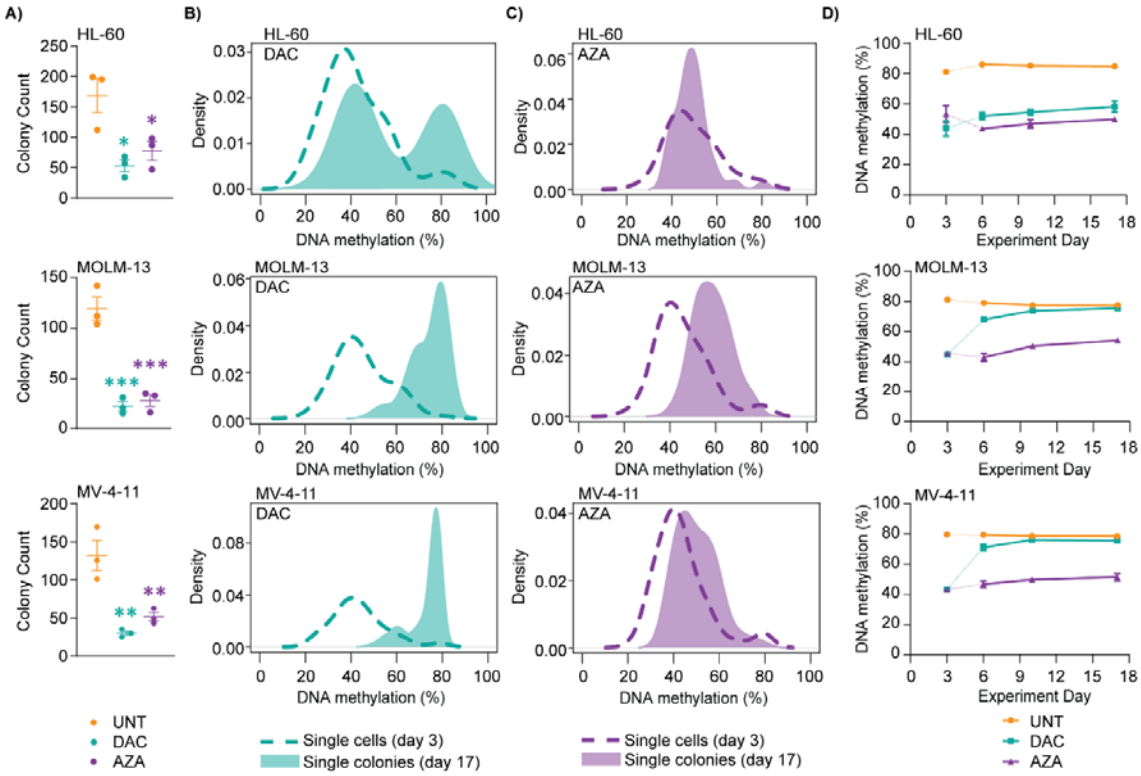


Figure 3. Highly methylated AML cells display a growth and survival advantage following treatment with DAC, but not AZA. A) HL-60, MOLM-13, and MV-4-11 colony counts following treatment with DAC (cyan) or AZA (purple) vs. untreated cells (UNT, orange). **B-C)** Density plots show the average DNA methylation levels for single cells collected on experiment day 3 (dashed line) and individual colonies collected on experiment day 17 (solid fill) following treatment with DAC or AZA. HL-60 scNMT-seq data were filtered for cytosines within SINE Alu sites for direct comparison to scTEM-seq data from colonies. Data are shown for 288 colonies collected from triplicate experiments in each cell line (*n = 96* per treatment). **D)** Time-course experiment showing changes in average DNA methylation of cells collected at different time points throughout the colony-forming assay (experiment days 6, 10 and 17). Values for experiment day 3 were obtained from single-cell data (Fig. 1A). Data in A and D are expressed as mean +/- standard error of the mean (SEM). Statistical analysis of colony counts (A) was performed using ordinary one-way ANOVA with Dunnett's multiple comparisons test with a $p < 0.05$ cut-off for significance ($p < 0.03$*, $p < 0.006$**, $p < 0.0002$***).

Transcriptional programs associated with recovery following HMA exposure

To identify cellular processes active during recovery from HMA exposure, we next generated matched transcriptomes from the same set of colony samples. Single-colony RNA-seq and principal component analysis (PCA) showed that DAC and AZA samples were generally distinct from untreated samples, regardless of their global DNA methylation levels, in all cell lines (Fig. 4A). This implies that HMA exposure has substantial effects on the transcriptome, even in highly methylated cells.

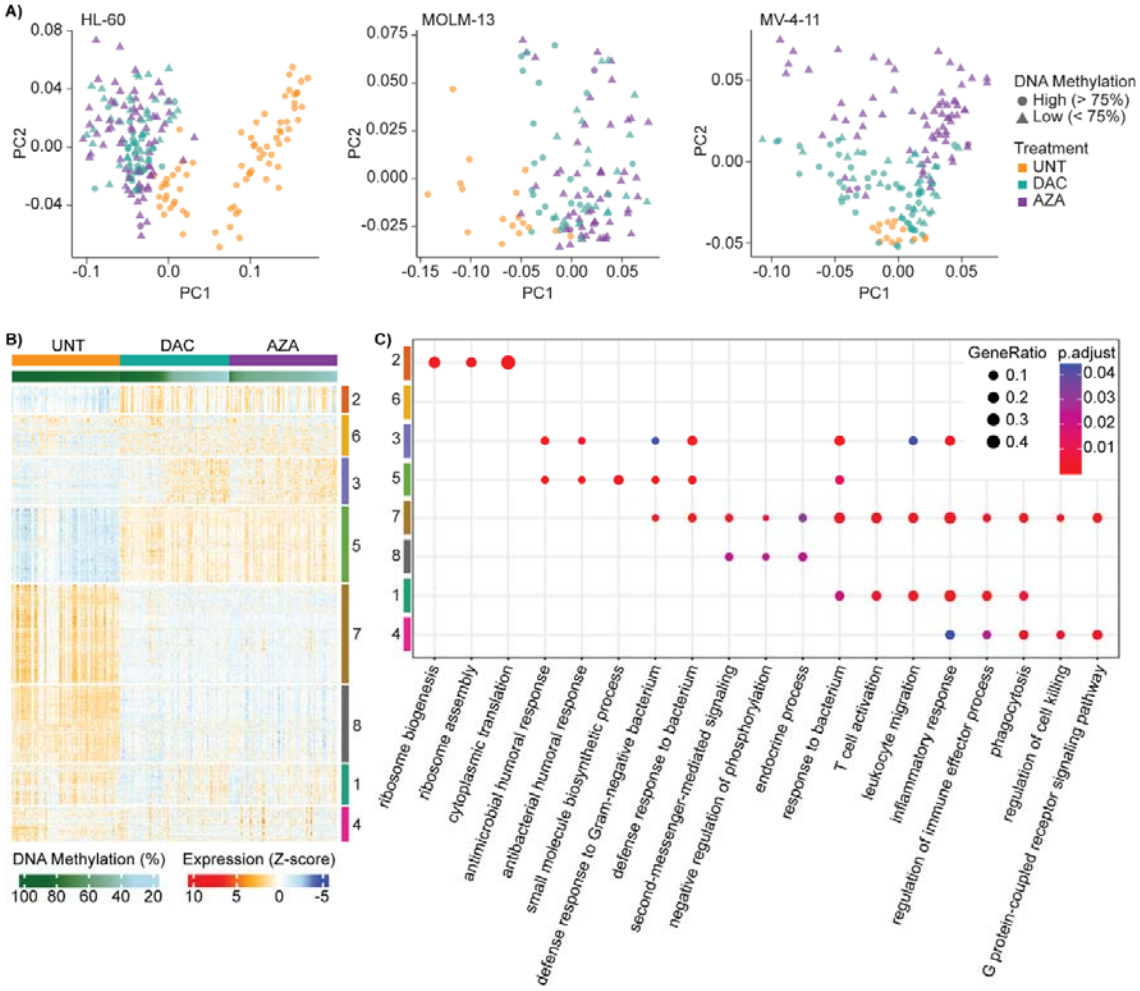


Figure 4. HMA treatment has sustained transcriptional effects in both methylation-retaining and hypomethylated cells. A) Principal Component Analysis (PCA) plots of single-colony RNA-seq data from AML cell lines, highlighting treatment groups (UNT = orange; DAC = cyan; AZA = purple) and matched mean DNA methylation levels (circle: high > 75%; triangle: low < 75%). Data shown for 119 - 220 colonies collected from 3 replicate experiments in each cell line (UNT, *n = 14-73*; DAC, *n = 46-78*; AZA, *n = 56-73*). **B)** Heatmap of the top 2000 highly variable genes from colony RNA-seq data. Samples are ordered by decreasing global methylation levels (green gradient) within each treatment group. Rows are grouped by K-means clusters based on gene expression, with hierarchical clustering by Euclidean distance within each cluster. **C)** GO analysis of the clusters from the top 2000 highly variable genes from B. The size of the circles indicates the gene ratio (number of genes from the input list annotated to the GO term divided by the total number of genes in the input list), and the colour represents the significance of the adjusted *p*-value. Gene clusters are colour-coded on the y-axis and GO processes are shown on the x-axis.

Of the 2000 most variably expressed genes among HL-60 samples, only 215 had increased expression specific to hypomethylated colonies (Fig. 4B, cluster 3; Supplementary Table 8). Many of these genes (42.3%) were upregulated following 72h treatment with either DAC or AZA in bulk HL-60 RNA-seq data, and several were associated with activation of inflammatory responses within the

283 sPLS model (e.g., *S100A8*, *S100A9*; Supplementary Table 8; Fig. 2B, cluster 1). In contrast, genes in
284 cluster 5 were upregulated following HMA exposure in both hypomethylated and highly methylated
285 colonies (Fig. 4B). Only 3.5% of these genes were induced by HMA treatment in bulk RNA-seq data,
286 but several were included in gene expression cluster 3 from the sPLS model (e.g., *MYC*, *MPO*;
287 Supplementary Table 8; Fig. 2B). Thus, some transcriptional changes that occur immediately after
288 HMA treatment are maintained in hypomethylated colonies, while other genes are upregulated
289 during colony formation and are independent of HMA-induced global hypomethylation. GO ORA
290 revealed an enrichment of anti-microbial and immune-related processes among both
291 hypomethylation-dependent and -independent gene sets, whereas 'small molecule biosynthetic
292 process' was uniquely over-represented among the hypomethylation-independent cluster 5 genes
293 (Fig. 4C; Supplementary Table 9).

294 To identify cellular processes that could favour the growth of methylation-retaining cells (Fig. 3), we
295 next focused on the HL-60 colonies derived following DAC treatment. Among these samples, the
296 broad range of DNA methylation levels (Fig. 3B) allowed us to compute correlations to gene and TE
297 expression (Fig. 5A). We found an enrichment for negative correlations with TE expression,
298 consistent with their upregulation in hypomethylated cells[35-37]. We also identified transcripts with
299 both significant negative (cor < -0.4, adjusted *p*-value < 0.05, *n* = 722) and positive (cor > 0.4,
300 adjusted *p*-value < 0.05, *n* = 345) correlations to global DNA methylation levels and divided these
301 genes into 4 clusters based on their expression patterns across treatment groups (Fig. 5B, C;
302 Supplementary Table 10). Genes with increased expression in hypomethylated colonies (Fig. 5B,
303 cluster 2) were enriched for GO terms related to defense responses, cell motility and chemotaxis
304 (Supplementary Table 11). Interestingly, we also identified 233 genes with specific upregulation in
305 highly methylated HL-60 colonies derived following DAC treatment (Fig. 5B, cluster 1). These genes
306 displayed significant enrichment of cholesterol-related ontologies, which included many enzymes
307 required for *de novo* cholesterol biosynthesis downstream of mevalonate[40] (Figure 5C,
308 Supplementary Table 11; Supplementary Fig. 4). Other genes involved in the mevalonate pathway
309 (e.g., *MVD, MVK,* and *PMVK*) were increased by HMA treatment in most HL-60 colonies, regardless
310 of their global DNA methylation level. MOLM-13 and MV-4-11 colonies also displayed increased
311 expression of cholesterol biosynthesis genes after HMA treatment (Supplementary Fig. 5).
312 Interestingly, these changes were not observed immediately following HMA treatment, nor after
313 long-term culture in suspension (i.e., in day 3 and day 21 HL-60 bulk RNA-seq data). On the contrary,
314 several members of the cholesterol biosynthesis pathway were significantly decreased (FDR < 0.05,
315 Log2(fold change) < -1) following 72h DAC and/or AZA treatment of HL-60 cells (*LSS*, *DHCR7,*
316 Supplementary Table 8; *HMGCS1*, *HMGCR,* data not shown). These observations suggest that

317 increased cholesterol production is a delayed response to HMA treatment that occurs independently
318 of methylation changes and facilitates the self-renewal and proliferation of AML cells.

319 Inhibition of cholesterol biosynthesis following HMA exposure

320 To test whether inhibition of cholesterol biosynthesis could enhance HMA efficacy, we performed
321 co-treatments with rosuvastatin in colony-forming assays. Rosuvastatin is a potent inhibitor of the
322 rate limiting enzyme of the cholesterol biosynthetic pathway, HMG-CoA reductase (HMGCR), and is
323 frequently prescribed to reduce the risk of cardiovascular disease associated with
324 hypercholesterolaemia[41]. In all cell lines, rosuvastatin caused a trend towards decreased colony
325 numbers following DAC treatment, with a significant decrease in MOLM-13 colonies, whereas co-
326 treatment with AZA had varied effects (Fig. 5D). Following DAC and rosuvastatin co-treatment,
327 colonies also displayed significantly reduced size (Supplementary Fig. 3B) and DNA methylation
328 levels (Fig. 5E) in MOLM-13 and MV-4-11 cell lines, where strong selection for highly methylated
329 cells was previously noted (Fig. 3B). This suggests that inhibition of cholesterol biosynthesis
330 specifically inhibits the self-renewal capacity of methylation-retaining cells and identifies statins as a
331 candidate co-treatment strategy to target HMA-evading AML blasts.

332 The efficacy and potential clinical utility of DAC and rosuvastatin co-treatment was then tested in
333 immunocompromised mice engrafted with luciferase-tagged MOLM-13 cells (Fig. 5F). In a dose
334 optimisation experiment, combining DAC and rosuvastatin led to a significant increase in median
335 survival for all doses tested (Fig. 5F, left). The 1mg/kg/day dose of rosuvastatin was chosen for
336 further validation in this model, as it is a low dose that is commonly prescribed for preventing
337 cardiovascular disease (equivalent to approximately 5-10mg/day oral dose in humans). DAC and
338 rosuvastatin co-treatment once again led to a significant increase in median survival compared to
339 DAC treatment alone (24 vs. 21 days) in this aggressive *in vivo* AML model (Fig. 5F, right).
340 Rosuvastatin alone provided no benefit, with survival times comparable to vehicle treated mice. DAC
341 and rosuvastatin co-treated mice also exhibited a trend towards decreased tumour burden, as
342 measured by BLI, between days 17 – 24 of the experiment, compared to mice treated with DAC
343 alone (Supplementary Fig. 6). These *in vivo* results provide evidence that HMAs combined with
344 statins have the potential to decrease tumour burden and improve AML survival.

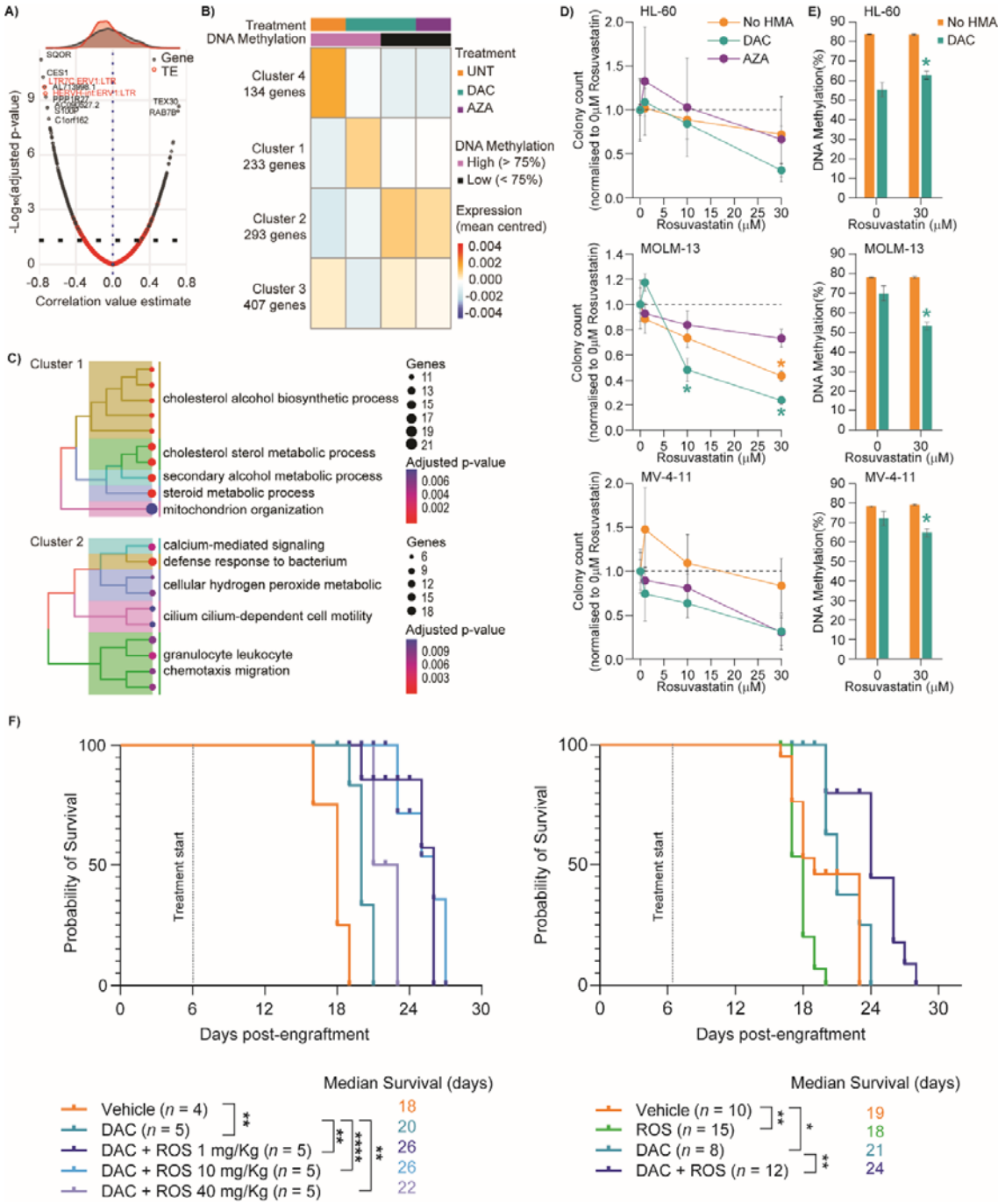


Figure 5. Upregulation of cholesterol biosynthesis in cells surviving HMA therapy can be targeted to enhance treatment efficacy. A) Volcano plot showing the Pearson correlation coefficient and adjusted *p*-value for correlations between gene or transposable element (TE) expression and global DNA methylation levels from HL-60 colonies derived following DAC treatment. The upper density plot shows a bias toward negative correlations, especially between global DNA methylation level and TE expression (red). **B)** Simplified heatmap of k-means clustering for the 1,067 genes with significant correlations to global DNA methylation level (adjusted *p*-value ≤ 0.05 and 0.4 ≤ correlation estimate ≤ 0.4, from A). **C)** Summarised tree plots displaying GO terms with significant (adjusted *p*-value < 0.05) over-representation in clusters 1 and 2 (from B). **D)** Colony counts for HL-60 (top), MOLM-13 (middle) and MV-4-11 (bottom) cell lines obtained following HMA and rosuvastatin co-treatments.

Data from No HMA (orange), DAC (cyan) and AZA (purple) groups are normalised to the corresponding 0 µM rosuvastatin control. Means ± SEM for *n = 3* experiments. Significance determined by two-way ordinary ANOVA with Dunnett's multiple comparisons test, $p < 0.05$* vs. corresponding 0 µM rosuvastatin control. **E)** DNA methylation of colonies formed following DAC and rosuvastatin co-treatments. Means ± SEM for *n = 3* experiments. Significance determined by one-way ANOVA with Tukey's multiple comparisons test, $p < 0.05$* vs. corresponding 0 µM rosuvastatin control. **F)** Left: rosuvastatin dose optimisation experiment showing median survival of mice engrafted with MOLM-13-luc cells following treatment with DAC (0.2mg/kg/day) +/- rosuvastatin (1, 10, 40mg/kg/day) on a treatment schedule of '5 days on, 2 days off' for 3 cycles. Right: Validation of survival benefit when DAC (0.2mg/kg/day) is combined with rosuvastatin (1mg/kg/day) in mice engrafted with MOLM-13-luc AML cells. Survival analyses were performed using Kaplan-Meier analysis followed by the Log-rank (Mantel-Cox) test and an adjusted *p*-value of < 0.05 was considered statistically significant. Left: $p < 0.002$**, $p < 0.0001$****; Right: $p < 0.05$*, $p < 0.005$**.

Discussion

The clinical benefits of HMA therapy are limited by the rapid development of treatment-resistant relapse. We have characterised the heterogeneous responses of AML cells to HMA treatment, revealing new insights into how cells survive and adapt to treatment.

Our single-cell multi-omic analyses revealed global DNA methylation heterogeneity induced by treatment, and transcriptional responses linked to epigenetic changes. We observed activation of inflammatory response and cell death pathways in only a minor subset of hypomethylated cells (Fig. 2B), consistent with our previous report of heterogeneous TE expression following DAC treatment[42] and scRNA-seq data from a colorectal cancer cell line[43]. These observations suggest that additional epigenetic modifications, transcription factors, or other mechanisms, can suppress transcriptional responses in hypomethylated cells (e.g., H3K9me3[44,45]). Alternatively, loss of methylation at specific loci may be required for HMAs to exert their effects. Our observations are consistent with the lack of correlation between HMA-induced hypomethylation and clinical response[46-48], and support the use of locus-specific methylation changes to build a predictor of patient response[49].

In contrast to our single-cell analyses, Li *et al.* reported reduced epigenetic and transcriptional variance in AML cells collected after 12 weeks of AZA treatment in a transgenic mouse model[50]. We suggest that HMAs initially increase epigenetic and transcriptional diversity, allowing some cells to gain a relative growth or survival advantage. Subsequent expansion of those clones would lead to the reduced heterogeneity reported by Li *et al.*[50]. Consistently, we found that some AML cells retain high levels of DNA methylation during HMA treatment (Fig. 1A) and have a relative growth advantage following drug withdrawal (Fig. 3). Methylation-retaining cells tended to divide less frequently during treatment (Fig. 1B) but had higher self-renewal and proliferative capacity than hypomethylated cells after DAC treatment (Fig. 3). Interestingly, this selection for methylation-

393 retaining cells was not observed following AZA treatment, reflecting previous reports of distinct
394 effects of these HMAs[51-53]. Since AZA is incorporated into both DNA and RNA[54], while DAC is
395 restricted to DNA, we speculate that RNA-mediated toxicities (such as translational inhibition[25,26])
396 could prevent the growth of highly methylated cells following AZA treatment. Together, our results
397 suggest that methylation-retaining cells are a likely source of AML relapse, especially following DAC
398 therapy.

399 We observed upregulated cholesterol biosynthesis, particularly in colonies that had high DNA
400 methylation after HMA treatment (Fig. 5 and Supplementary Fig. 4). This indicates that HMA
401 exposure causes transcriptional changes and metabolic alterations, even in highly methylated cells.
402 Previous studies have shown that HMAs perturb the homeostasis of pyrimidine metabolism
403 independently of DNA methylation changes[24], and similar effects have been linked to altered
404 cholesterol and lipid metabolism in AZA-treated liver cell lines[55]. Therefore, we speculate that
405 upregulation of cholesterol biosynthesis is a delayed response to HMA treatment that occurs in
406 highly methylated cells, potentially via altered pyrimidine metabolism.

407 Increased cholesterol demand is an established feature of AML cells[56], and upregulation of
408 cholesterol biosynthesis genes has shown prognostic value in AML[57-59]. While the precise
409 mechanisms by which cholesterol confers a survival advantage remain unclear, inhibiting cholesterol
410 biosynthesis has been shown to sensitize AML cells to radiation and chemotherapy[60,61]. In addition,
411 we have demonstrated that inhibition of cholesterol biosynthesis by rosuvastatin can decrease the
412 self-renewal capacity and global DNA methylation levels of MOLM-13 colonies when combined with
413 DAC treatment (Fig. 5D, E). This suggests that the upregulation of cholesterol biosynthesis facilitates
414 the self-renewal and proliferation of cells that retain DNA methylation during HMA treatment.
415 Recently, cholesterol metabolism was also linked to DAC resistance in AML cell lines, with statin co-
416 treatment showing synergistic inhibition of *in vitro* AML cell growth[62]. *In vivo*, we observed
417 significantly improved survival of leukaemia-bearing mice treated with DAC and rosuvastatin (Fig.
418 5F), suggesting that co-treatment may increase the duration of HMA response in some AML and
419 MDS patients.

420 Encouragingly, a recent retrospective analysis of MDS patients (including some HMA-treated high-
421 risk cases) reported improved survival and reduced progression to AML for patients who
422 commenced statin treatments three months before or after MDS diagnosis[63]. Current clinical trials
423 are testing the safety of a statin (pitavastatin) in combination with AZA and venetoclax in AML
424 patients[64] and our results imply that DAC and statin co-treatments should also be assessed. This
425 therapeutic avenue is of particular interest since statins are commonly prescribed, well-tolerated,

oral medications, which could be rapidly repositioned for use in AML and MDS patients receiving HMA therapy.

Methods

Cell lines and culture

AML cell lines, HL-60 (ATCC #CCL-240), MOLM-13 (DSMZ #ACC-554), and MV-4-11 (ATCC #CRL-9591) were maintained in tissue culture flasks (Greiner Bio-One) at 37°C, 5% CO_2, and sub-cultured at 500,000 cells/mL every 2-3 days with fresh medium. HL-60 cells were maintained in Iscove's Modified Dulbecco's media (IMDM; Sigma-Aldrich) supplemented with 4mM GlutaMAX (Thermo Fisher Scientific) and 10% Fetal Bovine Serum (FBS; Sigma-Aldrich). MOLM-13 and MV-4-11 cells were maintained in Roswell Park Memorial Institute 1640 media (RPMI; Sigma-Aldrich) supplemented with 2mM GlutaMAX and 10% FBS. All cell lines were mycoplasma negative based on routine testing using MycoAlert Mycoplasma Detection Kit (Lonza). Cell line authentication was routinely performed by the Australian Genome Research Facility (AGRF).

CellTrace staining

Prior to drug treatment (Day 0), AML cells ($2x10^6$ cells/mL) were stained with 1µM (MOLM-13 and MV-4-11) or 3µM (HL-60) CellTrace Far Red (Thermo Fisher Scientific) to monitor cell divisions, according to manufacturer's instructions. Cells with uniformly high CellTrace fluorescence underwent fluorescence activated cell sorting (FACS) prior to treatment with hypomethylating agents (HMAs) in MOLM-13 and MV-4-11 cell lines, whereas all cells were used for HL-60 treatments.

HMA treatments

AML cell lines were treated with HMAs at various doses in suspension culture, every 24h for 72h total. All cell lines were treated with 100nM decitabine (DAC; Selleckchem #S1200), HL-60 cells were treated with 2000nM azacytidine (AZA; Selleckchem # S1782), and MOLM-13 and MV-4-11 cells were treated with 500nM AZA. Untreated cells (UNT) were given an equal volume of 0.1% DMSO in UltraPure™ DNase/RNase-Free Distilled Water (Thermo Fisher Scientific) and served as a negative control. After HMA treatment, cells were prepared for FACS or colony forming assays, as described below.

Fluorescence activated cell sorting (FACS)

HMA treated cells were stained with propidium iodide (PI, 1.5µg/mL) and prepared for FACS. Viable (PI^-) single cells were sorted into 2.5µL of RLT PLUS buffer (Qiagen) containing 2.5U SUPERas-In (Thermo Fisher Scientific) in LoBind 96-well full skirted plates (Eppendorf) using indexed sorting on a

FACS Aria II (BD Biosciences). Plates were sealed and briefly centrifuged before storage at -80°C for sequencing analyses.

Colony-forming assays

HMA treated cells were seeded in MethoCult Optimum (H4034; STEMCELL Technologies Inc.) at 500 cells/well in 6-well plates, with rosuvastatin (Selleckchem # S2169) added to the MethoCult at various doses (0, 1, 10, 30µM). Cells were then cultured at 37°C, 5% CO_2 for 14 days to allow colony formation. Wells containing colonies were imaged using Cytation3 (Biotek). Colony counts and sizes were analysed using ImageJ software. Individual colonies were manually plucked using a 20µL pipette tip into 100µL of media, centrifuged at 200xg for 5 mins, and then resuspended in 20µL of RLT PLUS buffer before storage at -80°C. Alternatively, all colonies in each well were collected by resuspending the MethoCult Optimum media (and colonies) in 3mL of standard culture media (IMDM or RPMI), centrifuging at 200xg for 5 mins, and resuspending the cell pellet in 20-50µL RLT PLUS buffer, before storage at -80°C.

Library preparation and sequencing

scNMT-seq library preparation and sequencing

For scNMT-seq, matched scNOMe-seq and scRNA-seq libraries were prepared from sorted HL-60 single cells, as previously described[28]. Minor modifications to the published protocol are in the Supplementary Methods.

For scNOMe-seq libraries, paired-end 150bp sequencing was performed on the NovaSeq (Illumina) platform. For scRNA-seq libraries, paired-end 75bp sequencing was performed on the NovaSeq or NextSeq (Illumina) platform.

scTEM-seq library preparation and sequencing

For scTEM-seq analysis of global DNA methylation levels in single MOLM-13 and MV-4-11 cells, library preparation was performed as previously described[65].

Paired-end 150bp sequencing was performed on the MiSeq (Illumina) platform.

Colony TEM-seq and RNA-seq library preparation and sequencing

Single colony TEM-seq (Fig. 3B, C) and parallel RNA-seq analysis was performed as described[65,66] with minor modifications. Lysates from single colonies (HL-60: 7.5µL; MOLM-13 and MV-4-11: 2.5µL) were used to separate genomic DNA and mRNA. During single colony TEM-seq library preparation, the number of SINE Alu amplification cycles was reduced to 29. For RNA-seq analysis, 15 cycles of cDNA amplification were used.

TEM-seq analysis of pooled colonies (Fig. 3D, 5E) was performed as described[65,66] using 5-10µL of cell lysate as input for bisulphite conversion, and 29 cycles for SINE Alu amplification.

All TEM-seq libraries were sequenced using 150bp paired-end sequencing on the MiSeq (Illumina) platform. For RNA-seq libraries, paired-end 75bp sequencing was performed on the NextSeq or NovaSeq platform.

Data processing

Sequencing data were processed and aligned as described in the Supplementary Methods.

scNMT-seq data analysis

Quality control

For scNMT-seq data, cells were required to pass both scNOMe-seq and scRNA-seq quality control (QC). Cells that had less than 500,000 CpG sites covered, less than 5,000,000 GpC sites covered, greater than 15% CHH methylation rate, or less than 2% GpC methylation, failed scNOMe-seq QC. For scRNA-seq, QC was performed using bam files from hisat2, and the SeqMonk[67] (v1.47.1) 'RNA-seq QC Plot'. Cells that had less than 70% reads in exons, or less than 15% genes measured, failed scRNA-seq QC. In total, 222 scNMT-seq samples passed QC (Supplementary Table 1).

scNOMe-seq normalisation and batch correction

scNOMe-seq libraries provide information on both DNA methylation (CpG sites) and DNA accessibility (GpC sites). For both CpG (methylation) and GpC (accessibility) datasets, several genomic annotation contexts were considered: introns, exons, intergenic regions, CpG islands, promoters (-1500 to +500 bp of transcription start sites), H3K4me3 sites (ENCODE[68-70] accession ID: ENCFF021JBH, experiment: ENCSR000DUO) and H3K27ac sites (ENCODE[68-70] accession ID: ENCFF763UAG, experiment: ENCSR919WLM). In addition, unbiased 3kb windows of the whole genome were generated with a step size of 1.5kb.

For DNA methylation, the CpG methylation rate was estimated within each annotation window using the Bayes binomial approximation as in Smallwood *et. al.*[71].

The GpC methylation, which marks accessible DNA in scNOMe-seq libraries, is introduced *in vitro* using a bacterial GpC methyltransferase enzyme (Supplementary Methods). To remove batch effects resulting from differences in enzymatic activity, data normalisation and batch correction were performed as follows. GpC data for the whole genome was aggregated in windows of 500kb in length with 250kb overlap separately for methylated and unmethylated GpC counts for each cell. Per-cell pooled size factors were computed from these 500kb windows using the method of Lun *et. al.*[72] scaling by total library size, as implemented in the single-cell R package "scuttle" (v1.8.4)[73].

Batch scaling factors were estimated from corrected methylated and unmethylated window log counts using the rescaleBatches method from the R package "batchelor" (v1.14.1)[74]. Per cell methylated and unmethylated cell scaling factors were calculated as the ratio of the batch-corrected sum of counts to the mean sum of counts across cells. Finally, unscaled methylated and unmethylated counts in each cell were independently scaled by the product of the cell pooled size factor and the methylated/unmethylated count batch correction factor, respectively. GpC methylation rate for each annotation window was then computed using the normalised batch corrected counts by the Bayesian binomial approximation.

From the overall distribution of counts across in the annotation layer for CpG and GpC methylation data, minimum total count thresholds per window of 5 counts (CpG) and 20 counts (GpC) were established and applied to discard windows with unreliable methylation rate estimation.

Pairwise distance analysis of DNA methylation heterogeneity

To assess the DNA methylation heterogeneity per treatment group and genomic context (Fig. 1D), pairwise CpG methylation distance analysis was performed. The mean absolute methylation difference was computed for each cell pair (A, B) as the mean of the absolute difference in methylation rate at each common cytosine position in the relevant genomic annotation. To make the comparison of methylation patterns meaningful, only cytosine loci with data in both cells in the pair were used. These mean absolute methylation differences were grouped by the treatment combination of the cell pair. The global summaries shown in Fig. 1D corresponds to the groups where both cells in the pair had the same treatment. Higher values indicate a more heterogeneous methylation pattern when cells in the same treatment group are compared vis-a-vis.

Cell-wise correlation analysis

To assess the relationships between DNA methylation, DNA accessibility and gene expression within individual cells (Fig. 1E), Pearson correlations were computed using HL-60 scNMT-seq data. For this analysis, RNA-seq data was normalised and log transformed per batch using ‘scuttle::logNormCounts()’ (v1.6.2)[73] without batch correction or prior count filtering. DNA methylation was correlated to DNA accessibility at matched loci based on genomic coordinates. For correlations with gene expression, methylation and accessibility measurements at promoters, introns and exons were matched to the corresponding transcript. For CpG islands, H3K27ac sites and H3K4me3 sites, methylation and accessibility measurements were matched to all transcripts within 10kb. For intergenic regions and 3kb genomic windows, methylation and accessibility measurements were matched to every transcript within 1bp to assess the expression of immediately adjacent

genes. For each cell, Pearson correlation estimates were then computed using all matched values and the cor.test function in R.

scRNAseq normalisation and batch correction

scRNA-seq libraries from HL-60 scNMT-seq data were filtered to remove lowly expressed genes, requiring at least 5 counts in 10% of cells. Normalisation and variance stabilisation were performed by scTransform[75] and batch correction by anchor-based integration using the R package "Seurat" (v4.2.0)[76]. First, batches were independently normalised by scTransform. The top 5,000 most variable features that were in common across batches were identified to determine anchors for integration and batch correction of the data (using default parameter and k.weigh=50), applied to and retaining those 5,000 commonly variable features. Finally, a sparse RNA-seq matrix was utilised, whereby gene imputation calculations were ignored and removed by reintroducing 'NAs' in place of genes with originally 'missing data' (zeros). Downstream analyses considered only autosomal genes (Chr1-22).

Integrative sparse partial least squares (sPLS) analysis

Mixomics[38] (v6.20.0) was used to perform a multivariate integrative analysis of HL-60 scNMT-seq data (Fig. 2A-D). Feature selection was performed to identify variably expressed transcripts that are highly correlated to changes in DNA Methylation and accessibility following HMA treatment. We performed an unsupervised sparse Partial Least Squares (sPLS) analysis using the function 'mixOmics::mint.block.spls()' which combines a multivariate integrative (MINT) method and a multiblock sPLS integrative analysis. MINT[77] accounts for multiple batches (Supplementary Table 1) measured on the same variables, while the multiblock sPLS seeks for correlated patterns between DNA methylation and DNA accessibility rates that are split into multiple genomic regions ('blocks') and explain (correlated to) the predictor (transcriptome).

To focus on transcriptomic and epigenetic changes resulting from HMA treatment only treated cells (AZA and DAC) were included in the sPLS model. The genomic regions included in this analysis for both DNA methylation and DNA accessibility were CpG islands, promoters, H3K27ac sites, H3K4me3 sites and 3kb windows. The rates from these genomic regions were filtered to retain only those detected in greater than 10% of cells. The sPLS model was implemented assessing 2 components, selecting 50 features per component and per block (genomic region) in the DNA methylation and DNA accessibility data sets, and 100 genes per component in the transcriptome data set. More details are provided in our GitHub page.

Heatmap visualisation and identification of cell and expression clusters from sPLS selected features was performed using ComplexHeatmap[78] (v2.12.1). sPLS selected features for components 1 and 2

were extracted using the function ‘mixOmics::selectVar()’. Heatmap visualisation was performed on sPLS selected features using transcriptomic (converted to z-score) and epigenetic rates (mean of features in genomic regions and converted to z-score) that were entered into the model and included both treated and untreated cells. K-means clustering was performed on sPLS selected transcriptomic features, first on Gene Expression (row_km=3) followed by Cell Group (column_km=4).

sPLS sample projections (Fig. 2A, C and D) were plotted using ggplot2[79] (v3.3.6) by extracting the sPLS components 1 and 2 for a given block (RNA or epigenetic genomic region) and overlayed with relevant information i.e. cell group identified from k-means clustering and treated cell type or average DNA methylation.

Gene Ontology (GO) Over Representation Analysis (ORA) was calculated using clusterProfiler[80] (v4.4.4) for ‘biological process’ and displayed using enrichplot[81] V1.16.1 (Fig. 2E, F). Gene Expression k-means clusters (Fig. 2D) and sPLS selected features per component (1-2) were assessed by ‘enrichGO(p.adj=0.05, p.adj.method = "fdr", q.val.threshold = 0.4)’ and the list of genes from the batch corrected transcriptome dataset (entered into the sPLS model) as the background. Results were displayed as treeplots using default settings for pairwise ’termsim()’ and ‘treeplot(nCluster=5, showCategory = 10)’.

The correlation of sPLS features (Supplementary Fig. 2) was calculated as a similarity matrix using ‘mixOmics::circosPlot()’ on the sPLS model. The results were displayed using ComplexHeatmap showing DNA methylation and DNA accessibility features related to transcript features split by the previously identified Gene Expression k-means clusters.

Locus-specific correlation analysis

To compare gene expression to adjacent epigenetic features, locus-specific correlations were performed using HL-60 scNMT-seq data (Fig. 2H and I). DNA methylation and DNA accessibility measurements were paired to genes based on annotation (Promoters) or by strand-aware position within 10kbp of the gene transcription start site (CpG islands, H3K27ac sites, H3K4me3 sites and 3kb windows). For paired sites, Pearson correlations were computed between CpG or GpC methylation rate and log gene expression values. All cells with data (i.e. UNT, DAC and AZA groups) were combined in these correlations, and a minimum of 22 cells with paired data (i.e. both gene expression and DNA methylation/accessibility measurements) were required for the correlation to be performed.

Colony sequencing analysis

RNA-seq quality control, normalisation and batch correction

Samples were excluded if they had less than 35% genes measured, or less than 70% reads in exons for HL-60 and MOLM-13 samples, or less than 65% reads in exons for MV-4-11 samples. RNA-seq data from single colonies were filtered to remove lowly expressed genes, requiring at least 5 counts in 3 samples. For each cell line, normalisation was performed by 'scuttle::logNormCounts()' (v1.6.2) and batch corrected using mutual nearest neighbours method by 'batchelor::mnnCorrect()' (v1.12.3) with default parameters. Downstream analyses considered only autosomal genes (Chr1-22).

Highly variable gene analysis

For Figure 4, highly variable genes (HVGs) were identified from colony RNA-seq data and PCA was performed using 'scater::calculatePCA(ntop = 2,000)'. K-means clusters of HVGs was determined using the r stats package (v4.2.1) with 'kmeans(centers = 8, iter.max = kmeans.iter, nstart = 50)'. Heatmapping of HVGs and k-means cluster was performed using 'ComplexHeatmap::pheatmap()' with z-scored values and Euclidean distance hierarchical clustering within row clusters (k-means groups) and columns (samples) ordered by treatment and descending average global methylation level.

GO ORA of k-means clusters was compared using clusterProfiler for 'biological process' by 'compareCluster(pAdjustMethod = "fdr", p.adj.threshold = 0.05, qvalueCutoff=0.4)' and the full list of genes from the batch corrected dataset (for each cell type) as the background list. Plots were created using 'clusterProfiler::dotplot(showCategory = 3) + coord_flip()'.

Correlation analysis

Pearson correlations comparing gene expression to mean global methylation in DAC HL-60 colonies (Fig. 5A) were performed using 'cor.test()' and underwent Benjamini–Hochberg false discovery rate adjustment using 'p.adjust(method="BH")'. Gene clustering and heatmap visualisation was performed on significantly correlated genes (p.adj ≤ 0.05 & cor.value.estimate ≤ -0.4 or cor.value.estimate ≥ 0.4). The average expression of each gene was calculated per treatment group with DAC split into high (≥75%) and low (<75%) global methylation groups. The R package "pheatmap"[82] (v1.0.12) was used to plot the mean centred treatment group average expression levels with rows aggregated into 4 'kmeans_k' clusters. The genes from each 'Kmeans_k' cluster was extracted and underwent GO ORA for biological process individually using 'enrichGO()' with fdr adjustment and results displayed as treeplots.

AML cell-line xenograft model

All experimental procedures were reviewed, approved, and carried out according to the Animal Care and Ethics Committee of the University of Newcastle (approval number: A-2023-303), and with consideration of the ARRIVE guidelines (Supplementary Methods).

MOLM-13 cells transduced with firefly luciferase (MOLM-13-luc) were a kind gift from Dr Charles de Bock (Children's Cancer Institute, UNSW Sydney). Five-week-old female NOD.Cg-Prkdc scid Il2rg tm1Wjl /SzJ (NSG) mice were obtained from the Australian Bioresources (ABR, Moss Vale, NSW, Australia) and were acclimatised for one week prior to any experimental procedure. The NSG mice were inoculated with MOLM-13-luc cells ($5x10^5$ cells suspended in 100 µL of PBS) by injection into the lateral tail vein. Tumour burden was assessed by bioluminescence imaging (BLI) twice a week using an IVIS Spectrum *in vivo* imaging system (PerkinElmer, Waltham, MA, USA), following intraperitoneal injections of 150 mg/kg D-luciferin (Promega, Alexandria, NSW, Australia) and under anaesthesia with isoflurane. Treatments commenced on day 6 post-engraftment, when a positive luminescence signal was detected.

In a pilot experiment, three different doses of rosuvastatin were co-administered with DAC. Mice ($n = 5$ per group) were treated by intraperitoneal injection of either vehicle (2% DMSO, 30% PEG300 in water), DAC (0.2 mg/Kg in saline), or rosuvastatin (1 mg/kg, 10 mg/kg, or 40 mg/kg in 30% PEG300 in water) combined with DAC (0.2 mg/Kg) once a day (5 days on, 2 days off) for up to 3 weeks. The animals were monitored until they reached ethical endpoint.

In a second experiment, mice ($n= 15$ per group) received vehicle (2% DMSO, 30% PEG300 in water), DAC (0.2 mg/Kg), rosuvastatin (1 mg/Kg), or DAC + rosuvastatin treatments via intraperitoneal injections once a day (5 days on, 2 days off) for up to 3 weeks, and mice were monitored until ethical endpoint.

Survival analyses were performed using Kaplan-Meier analysis followed by the Log-rank (Mantel-Cox) test. All statistical analyses were performed using GraphPad Prism v. 9.0 (GraphPad Software, La Jolla, CA, USA). An adjusted p-value of < 0.05 was considered statistically significant.

Data and code availability

Sequencing data and analysis code will be made available upon reasonable request and published following peer review.

References

1 Stomper, J., Rotondo, J. C., Greve, G. & Lübbert, M. Hypomethylating agents (HMA) for the treatment of acute myeloid leukemia and myelodysplastic syndromes: mechanisms of resistance and novel HMA-based therapies. *Leukemia* (2021). https://doi.org:10.1038/s41375-021-01218-0

2 Jones, P. A. & Taylor, S. M. Cellular differentiation, cytidine analogs and DNA methylation. *Cell* **20**, 85-93 (1980). https://doi.org:http://dx.doi.org/10.1016/0092-8674(80)90237-8

3 Patel, K. *et al.* Targeting of 5-aza-2-deoxycytidine residues by chromatin-associated DNMT1 induces proteasomal degradation of the free enzyme. *Nucleic Acids Research* **38**, 4313-4324 (2010). https://doi.org:10.1093/nar/gkq187

4 Mayyas, I. M. *et al.* Hairpin-bisulfite sequencing of cells exposed to decitabine documents the process of DNA demethylation. *Epigenetics*, 1-9 (2020). https://doi.org:10.1080/15592294.2020.1861169

5 Issa, J. P. *et al.* Safety and tolerability of guadecitabine (SGI-110) in patients with myelodysplastic syndrome and acute myeloid leukaemia: a multicentre, randomised, dose-escalation phase 1 study. *Lancet Oncol* **16**, 1099-1110 (2015). https://doi.org:10.1016/S1470-2045(15)00038-8

6 Fenaux, P. *et al.* Efficacy of azacitidine compared with that of conventional care regimens in the treatment of higher-risk myelodysplastic syndromes: a randomised, open-label, phase III study. *Lancet Oncol* **10**, 223-232 (2009). https://doi.org:10.1016/s1470-2045(09)70003-8

7 Jabbour, E. *et al.* Randomized phase 2 study of low-dose decitabine vs low-dose azacitidine in lower-risk MDS and MDS/MPN. *Blood* **130**, 1514-1522 (2017). https://doi.org:10.1182/blood-2017-06-788497

8 Fenaux, P. *et al.* Azacitidine prolongs overall survival compared with conventional care regimens in elderly patients with low bone marrow blast count acute myeloid leukemia. *J Clin Oncol* **28**, 562-569 (2010). https://doi.org:10.1200/jco.2009.23.8329

9 Kantarjian, H. M. *et al.* Survival advantage with decitabine versus intensive chemotherapy in patients with higher risk myelodysplastic syndrome: comparison with historical experience. *Cancer* **109**, 1133-1137 (2007). https://doi.org:10.1002/cncr.22508

10 Kantarjian, H. M. *et al.* Multicenter, randomized, open-label, phase III trial of decitabine versus patient choice, with physician advice, of either supportive care or low-dose cytarabine for the treatment of older patients with newly diagnosed acute myeloid leukemia. *J Clin Oncol* **30**, 2670-2677 (2012). https://doi.org:10.1200/jco.2011.38.9429

11 Lee, J. H. *et al.* Comparison of 7-day azacitidine and 5-day decitabine for treating myelodysplastic syndrome. *Ann Hematol* **92**, 889-897 (2013). https://doi.org:10.1007/s00277-013-1702-8

12 Dombret, H. *et al.* International phase 3 study of azacitidine vs conventional care regimens in older patients with newly diagnosed AML with >30% blasts. *Blood* **126**, 291-299 (2015). https://doi.org:10.1182/blood-2015-01-621664

13 Helbig, G. *et al.* Real Life Data on Efficacy and Safety of Azacitidine Therapy for Myelodysplastic Syndrome, Chronic Myelomonocytic Leukemia and Acute Myeloid Leukemia. *Pathology & Oncology Research* **25**, 1175-1180 (2019). https://doi.org:10.1007/s12253-018-00574-0

14 Stahl, M. *et al.* Hypomethylating agents in relapsed and refractory AML: outcomes and their predictors in a large international patient cohort. *Blood Adv* **2**, 923-932 (2018). https://doi.org:10.1182/bloodadvances.2018016121

15 Fennell, K. A., Bell, C. C. & Dawson, M. A. Epigenetic therapies in acute myeloid leukemia: where to from here? *Blood* **134**, 1891-1901 (2019). https://doi.org:10.1182/blood.2019003262

726 16 DiNardo, C. D. *et al.* Venetoclax combined with decitabine or azacitidine in treatment-naive,
727 elderly patients with acute myeloid leukemia. *Blood* **133**, 7-17 (2019).
728 https://doi.org:10.1182/blood-2018-08-868752
729 17 DiNardo, C. D. *et al.* Azacitidine and Venetoclax in Previously Untreated Acute Myeloid
730 Leukemia. *New England Journal of Medicine* **383**, 617-629 (2020).
731 https://doi.org:10.1056/NEJMoa2012971
732 18 Tsai, H. C. *et al.* Transient low doses of DNA-demethylating agents exert durable antitumor
733 effects on hematological and epithelial tumor cells. *Cancer Cell* **21**, 430-446 (2012).
734 https://doi.org:10.1016/j.ccr.2011.12.029
735 19 Craddock, C. *et al.* Azacitidine fails to eradicate leukemic stem/progenitor cell populations in
736 patients with acute myeloid leukemia and myelodysplasia. *Leukemia* **27**, 1028-1036 (2013).
737 https://doi.org:10.1038/leu.2012.312
738 20 Unnikrishnan, A. *et al.* Integrative Genomics Identifies the Molecular Basis of Resistance to
739 Azacitidine Therapy in Myelodysplastic Syndromes. *Cell Rep* **20**, 572-585 (2017).
740 https://doi.org:10.1016/j.celrep.2017.06.067
741 21 Riether, C. *et al.* Targeting CD70 with cusatuzumab eliminates acute myeloid leukemia stem
742 cells in patients treated with hypomethylating agents. *Nature Medicine* **26**, 1459-1467
743 (2020). https://doi.org:10.1038/s41591-020-0910-8
744 22 Pabst, T. *et al.* Results from a phase I/II trial of cusatuzumab combined with azacitidine in
745 patients with newly diagnosed acute myeloid leukemia who are ineligible for intensive
746 chemotherapy. *Haematologica* **108**, 1793-1802 (2023).
747 https://doi.org:10.3324/haematol.2022.281563
748 23 Zeidan, A. M. *et al.* Clinical outcomes of older patients with AML receiving hypomethylating
749 agents: a large population-based study in the United States. *Blood Adv* **4**, 2192-2201 (2020).
750 https://doi.org:10.1182/bloodadvances.2020001779
751 24 Gu, X. *et al.* Decitabine- and 5-azacytidine resistance emerges from adaptive responses of
752 the pyrimidine metabolism network. *Leukemia* **35**, 1023-1036 (2021).
753 https://doi.org:10.1038/s41375-020-1003-x
754 25 Aimiuwu, J. *et al.* RNA-dependent inhibition of ribonucleotide reductase is a major pathway
755 for 5-azacytidine activity in acute myeloid leukemia. *Blood* **119**, 5229-5238 (2012).
756 https://doi.org:10.1182/blood-2011-11-382226
757 26 Diesch, J. *et al.* Inhibition of CBP synergizes with the RNA-dependent mechanisms of
758 Azacitidine by limiting protein synthesis. *Nature Communications* **12**, 6060 (2021).
759 https://doi.org:10.1038/s41467-021-26258-z
760 27 Gruber, E., Franich, R. L., Shortt, J., Johnstone, R. W. & Kats, L. M. Distinct and overlapping
761 mechanisms of resistance to azacytidine and guadecitabine in acute myeloid leukemia.
762 *Leukemia* **34**, 3388-3392 (2020). https://doi.org:10.1038/s41375-020-0973-z
763 28 Clark, S. J. *et al.* scNMT-seq enables joint profiling of chromatin accessibility DNA
764 methylation and transcription in single cells. *Nature Communications* **9**, 781 (2018).
765 https://doi.org:10.1038/s41467-018-03149-4
766 29 Daskalakis, M. *et al.* Demethylation of a hypermethylated P15/INK4B gene in patients with
767 myelodysplastic syndrome by 5-Aza-2'-deoxycytidine (decitabine) treatment. *Blood* **100**,
768 2957-2964 (2002). https://doi.org:10.1182/blood.V100.8.2957
769 30 Kiziltepe, T. *et al.* 5-Azacytidine, a DNA methyltransferase inhibitor, induces ATR-mediated
770 DNA double-strand break responses, apoptosis, and synergistic cytotoxicity with doxorubicin
771 and bortezomib against multiple myeloma cells. *Molecular Cancer Therapeutics* **6**, 1718-
772 1727 (2007). https://doi.org:10.1158/1535-7163.Mct-07-0010
773 31 Dunshee, D. R. *et al.* CC-486 Mechanism Imparted By Extended Exposure of Azacitidine
774 Upregulates Myeloid Differentiation Markers and Induces Cell Death in AML Cells. *Blood*
775 **136**, 33-34 (2020). https://doi.org:https://doi.org/10.1182/blood-2020-137606

776 32 Momparler, R. L., Bouchard, J. & Samson, J. Induction of differentiation and inhibition of
777 DNA methylation in HL-60 myeloid leukemic cells by 5-AZA-2-deoxycytidine. *Leukemia*
778 *Research* **9**, 1361-1366 (1985). https://doi.org:https://doi.org/10.1016/0145-
779 2126(85)90123-7
780 33 Goodyear, O. *et al.* Induction of a CD8+ T-cell response to the MAGE cancer testis antigen by
781 combined treatment with azacitidine and sodium valproate in patients with acute myeloid
782 leukemia and myelodysplasia. *Blood* **116**, 1908-1918 (2010). https://doi.org:10.1182/blood-
783 2009-11-249474
784 34 Nahas, M. R. *et al.* Hypomethylating agent alters the immune microenvironment in acute
785 myeloid leukaemia (AML) and enhances the immunogenicity of a dendritic cell/AML vaccine.
786 *Br J Haematol* **185**, 679-690 (2019). https://doi.org:10.1111/bjh.15818
787 35 Roulois, D. *et al.* DNA-Demethylating Agents Target Colorectal Cancer Cells by Inducing Viral
788 Mimicry by Endogenous Transcripts. *Cell* **162**, 961-973 (2015).
789 https://doi.org:10.1016/j.cell.2015.07.056
790 36 Liu, M. *et al.* Vitamin C increases viral mimicry induced by 5-aza-2'-deoxycytidine. *Proc Natl*
791 *Acad Sci U S A* **113**, 10238-10244 (2016). https://doi.org:10.1073/pnas.1612262113
792 37 Chiappinelli, Katherine B. *et al.* Inhibiting DNA Methylation Causes an Interferon Response in
793 Cancer via dsRNA Including Endogenous Retroviruses. *Cell* **162**, 974-986 (2015).
794 https://doi.org:10.1016/j.cell.2015.07.011
795 38 Rohart, F., Gautier, B., Singh, A. & Le Cao, K. A. mixOmics: An R package for 'omics feature
796 selection and multiple data integration. *PLoS Comput Biol* **13**, e1005752 (2017).
797 https://doi.org:10.1371/journal.pcbi.1005752
798 39 Deniz, Ö. *et al.* Endogenous retroviruses are a source of enhancers with oncogenic potential
799 in acute myeloid leukaemia. *Nature Communications* **11**, 3506 (2020).
800 https://doi.org:10.1038/s41467-020-17206-4
801 40 Huang, B., Song, B.-l. & Xu, C. Cholesterol metabolism in cancer: mechanisms and
802 therapeutic opportunities. *Nature Metabolism* **2**, 132-141 (2020).
803 https://doi.org:10.1038/s42255-020-0174-0
804 41 Teramoto, T. & Watkins, C. Review of efficacy of rosuvastatin 5 mg. *Int J Clin Pract* **59**, 92-
805 101 (2005). https://doi.org:10.1111/j.1742-1241.2005.00346.x
806 42 Hunt, K. V. *et al.* scTEM-seq: Single-cell analysis of transposable element methylation to link
807 global epigenetic heterogeneity with transcriptional programs. *Scientific Reports* **12**, 5776
808 (2022). https://doi.org:10.1038/s41598-022-09765-x
809 43 Takeshima, H. *et al.* Low-dose DNA demethylating therapy induces reprogramming of
810 diverse cancer-related pathways at the single-cell level. *Clinical Epigenetics* **12**, 142 (2020).
811 https://doi.org:10.1186/s13148-020-00937-y
812 44 Griffin, G. K. *et al.* Epigenetic silencing by SETDB1 suppresses tumour intrinsic
813 immunogenicity. *Nature* (2021). https://doi.org:10.1038/s41586-021-03520-4
814 45 Irwin, R. *et al.* UHRF1 suppresses viral mimicry through both DNA methylation-dependent
815 and -independent mechanisms. *bioRxiv*, 2020.2008.2031.274894 (2020).
816 https://doi.org:10.1101/2020.08.31.274894
817 46 Maria Teresa, V. *et al.* Why methylation is not a marker predictive of response to
818 hypomethylating agents. *Haematologica* **99**, 613-619 (2014).
819 https://doi.org:10.3324/haematol.2013.099549
820 47 Unnikrishnan, A. *et al.* AZA-MS: a novel multiparameter mass spectrometry method to
821 determine the intracellular dynamics of azacitidine therapy in vivo. *Leukemia* **32**, 900-910
822 (2018). https://doi.org:10.1038/leu.2017.340
823 48 Yang, A. S. *et al.* DNA methylation changes after 5-aza-2'-deoxycytidine therapy in patients
824 with leukemia. *Cancer Res* **66**, 5495-5503 (2006). https://doi.org:10.1158/0008-5472.CAN-
825 05-2385

826 49 Greve, G. *et al.* In vivo kinetics of early, non-random methylome and transcriptome changes
827 induced by DNA-hypomethylating treatment in primary AML blasts. *Leukemia* **37**, 1018-1027
828 (2023). https://doi.org:10.1038/s41375-023-01876-2
829 50 Li, S. *et al.* Somatic Mutations Drive Specific, but Reversible, Epigenetic Heterogeneity States
830 in AML. *Cancer Discovery* **10**, 1934-1949 (2020). https://doi.org:10.1158/2159-8290.Cd-19-
831 0897
832 51 Hollenbach, P. W. *et al.* A Comparison of Azacitidine and Decitabine Activities in Acute
833 Myeloid Leukemia Cell Lines. *PLOS ONE* **5**, e9001 (2010).
834 https://doi.org:10.1371/journal.pone.0009001
835 52 Aumer, T. *et al.* Comprehensive comparison between azacytidine and decitabine treatment
836 in an acute myeloid leukemia cell line. *Clinical Epigenetics* **14**, 113 (2022).
837 https://doi.org:10.1186/s13148-022-01329-0
838 53 Flotho, C. *et al.* The DNA methyltransferase inhibitors azacitidine, decitabine and zebularine
839 exert differential effects on cancer gene expression in acute myeloid leukemia cells.
840 *Leukemia* **23**, 1019-1028 (2009). https://doi.org:10.1038/leu.2008.397
841 54 Li, L. H., Olin, E. J., Buskirk, H. H. & Reineke, L. M. Cytotoxicity and mode of action of 5-
842 azacytidine on L1210 leukemia. *Cancer Res* **30**, 2760-2769 (1970).
843 55 Poirier, S. *et al.* The epigenetic drug 5-azacytidine interferes with cholesterol and lipid
844 metabolism. *J Biol Chem* **289**, 18736-18751 (2014).
845 https://doi.org:10.1074/jbc.M114.563650
846 56 Vitols, S., Norgren, S., Juliusson, G., Tatidis, L. & Luthman, H. Multilevel regulation of low-
847 density lipoprotein receptor and 3-hydroxy- 3-methylglutaryl coenzyme A reductase gene
848 expression in normal and leukemic cells. *Blood* **84**, 2689-2698 (1994).
849 https://doi.org:10.1182/blood.v84.8.2689.2689
850 57 Banker, D. E. *et al.* Cholesterol synthesis and import contribute to protective cholesterol
851 increments in acute myeloid leukemia cells. *Blood* **104**, 1816-1824 (2004).
852 https://doi.org:10.1182/blood-2004-01-0395
853 58 Bai, S. *et al.* Lipid profile as a novel prognostic predictor for patients with acute myeloid
854 leukemia. *Frontiers in Oncology* **13** (2023). https://doi.org:10.3389/fonc.2023.950732
855 59 Kuliszkiewicz-Janus, M., Małecki, R. & Mohamed, A. Lipid changes occuring in the course of
856 hematological cancers. *Cellular and Molecular Biology Letters* **13** (2008).
857 https://doi.org:10.2478/s11658-008-0014-9
858 60 Dimitroulakos, J. *et al.* Increased Sensitivity of Acute Myeloid Leukemias to Lovastatin-
859 Induced Apoptosis: A Potential Therapeutic Approach. *Blood* **93**, 1308-1318 (1999).
860 https://doi.org:10.1182/BLOOD.V93.4.1308
861 61 Li, H. Y., Appelbaum, F. R., Willman, C. L., Zager, R. A. & Banker, D. E. Cholesterol-modulating
862 agents kill acute myeloid leukemia cells and sensitize them to therapeutics by blocking
863 adaptive cholesterol responses. *Blood* **101**, 3628-3634 (2003).
864 https://doi.org:10.1182/blood-2002-07-2283
865 62 Yabushita, T. *et al.* Mitotic perturbation is a key mechanism of action of decitabine in
866 myeloid tumor treatment. *Cell Reports* **42**, 113098 (2023).
867 https://doi.org:https://doi.org/10.1016/j.celrep.2023.113098
868 63 Afzal, A., Fiala, M. A., Jacoby, M. A. & Walter, M. J. Statin use in myelodysplastic syndromes
869 is associated with a better survival and decreased progression to leukemia. *Blood Advances*
870 **7**, 3838-3841 (2023). https://doi.org:10.1182/bloodadvances.2023009818
871 64 Brem, E. A. *et al.* A Phase 1 Study of Adding Pitavastatin to Venetoclax-Based Therapy in AML
872 and CLL/SLL. *Blood* **142**, 5940-5940 (2023). https://doi.org:10.1182/blood-2023-188086
873 65 Hunt, K. V. *et al.* scTEM-seq: Single-cell analysis of transposable element methylation to link
874 global epigenetic heterogeneity with transcriptional programs. *Sci Rep* **12**, 5776 (2022).
875 https://doi.org:10.1038/s41598-022-09765-x

66 Hunt KV, B. S., Bond DR, Lee HJ. Protocol for targeted analysis of transposable element methylation levels and transcriptome in single cells using scTEM-seq. *ProtocolExchange* (2022). https://doi.org:10.21203/rs.3.pex-2075/v1
67 SeqMonk (Babraham Bioinformatics).
68 Benjamin, C. H. *et al.* The ENCODE Uniform Analysis Pipelines. *bioRxiv*, 2023.2004.2004.535623 (2023). https://doi.org:10.1101/2023.04.04.535623
69 An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74 (2012). https://doi.org:10.1038/nature11247
70 Luo, Y. *et al.* New developments on the Encyclopedia of DNA Elements (ENCODE) data portal. *Nucleic Acids Res* **48**, D882-d889 (2020). https://doi.org:10.1093/nar/gkz1062
71 Smallwood, S. A. *et al.* Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nature Methods* **11**, 817-820 (2014). https://doi.org:10.1038/nmeth.3035
72 L. Lun, A. T., Bach, K. & Marioni, J. C. Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. *Genome Biology* **17**, 75 (2016). https://doi.org:10.1186/s13059-016-0947-7
73 McCarthy, D. J., Campbell, K. R., Lun, A. T. L. & Wills, Q. F. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* **33**, 1179-1186 (2017). https://doi.org:10.1093/bioinformatics/btw777
74 Haghverdi, L., Lun, A. T. L., Morgan, M. D. & Marioni, J. C. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nature Biotechnology* **36**, 421-427 (2018). https://doi.org:10.1038/nbt.4091
75 Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biology* **20**, 296 (2019). https://doi.org:10.1186/s13059-019-1874-1
76 Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587.e3529 (2021). https://doi.org:10.1016/j.cell.2021.04.048
77 Rohart, F., Eslami, A., Matigian, N., Bougeard, S. & Lê Cao, K.-A. MINT: a multivariate integrative method to identify reproducible molecular signatures across independent experiments and platforms. *BMC Bioinformatics* **18**, 128 (2017). https://doi.org:10.1186/s12859-017-1553-8
78 Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847-2849 (2016). https://doi.org:10.1093/bioinformatics/btw313
79 Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag New York, 2016).
80 Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation* **2**, 100141 (2021). https://doi.org:https://doi.org/10.1016/j.xinn.2021.100141
81 Yu, G. enrichplot: Visualization of Functional Enrichment Result. R package. (2022).
82 Kolde, R. & Kolde, M. R. Package 'pheatmap'. *R package* **1**, 790 (2015).

Acknowledgements

Nicole Cole (University of Newcastle) provided technical support for fluorescence activated cell sorting. Al J. Abadi (Melbourne Integrative Genomics, University of Melbourne) assisted with the implementation of sPLS analysis. The authors thank Professor Tri Phan (Garvan Institute) and Professor Xu Dong Zhang (University of Newcastle) for critical review of the manuscript.

H.J.L. and C.R. received funding from the National Health and Medical Research Council of Australia (NHMRC; GNT1143614, GNT1173892 and GNT2016283). K.A.L.C. and H.J.L. received funding from the Australian Research Council (ARC; DP200102903). H.J.L. and H.C.M. received funding from the Cancer Institute NSW (ECF171145 and ECF1299). D.R.B received funding from Cure Cancer Australia Foundation (CCAF2023-Bond) and NSW Health Pathology North. N.A.B received funding from the McGuigan family through the Hunter Medical Research Institute (HMRI778). The contents of the published material are solely the responsibility of the research institutions involved or individual authors and do not reflect the views of funding agencies.

Author Contributions

H.J.L. conceived and oversaw the project. D.R.B., A.K.E., C.R., K.A.L.C. and H.J.L. acquired funding. D.R.B., K.U., K.V.H., B.H., C.L.O., E.A.R., S.H., L.S.R., H.M. and H.J.L. performed experiments. D.R.B., S.M.B., K.U., K.V.H., B.H., C.L.O., L.S.R., C.R. and H.J.L. performed data analysis. S.M.B. and C.R. developed analytical approaches. S.M.B. managed sequencing data and analysis code. D.R.B., N.A.B., A.K.E, N.M.V., K.A.L.C. and H.J.L. supervised work. D.R.B., S.M.B., K.U., B.H., L.S.R., C.R. and H.J.L. prepared figures and wrote the manuscript. All authors read and approved the manuscript.

Competing Interests

A.K.E. declares the following competing interests: Advisory board/ Honoraria from AbbVie, Astellas, Gilead, Servier, Jazz, Otsuka/Astex and RACE Oncology; Speakers fees from AbbVie, Otsuka/Astex, Astellas and Jazz; Research Funding from RACE oncology and Otsuka/Astex.

Correspondence

Correspondence should be directed to Dr Heather Lee and Dr Danielle Bond.