Integrative Genomics Identifies the Molecular Basis of Resistance to Azacitidine Therapy in Myelodysplastic Syndromes

**Highlights**
- AZA non-responder hematopoietic progenitor cells are relatively quiescent
- AZA responsiveness can be induced in non-responders
- AZA alters the sub-clonal architecture of MDS stem/progenitor cells
- Hematopoiesis in AZA responders remains clonal with founder mutations

**Authors**
Ashwin Unnikrishnan, Elli Papaemmanuil, Dominik Beck, ..., Jason W.H. Wong, Peter J. Campbell, John E. Pimanda

**Correspondence**
ashwin.unnikrishnan@unsw.edu.au (A.U.),
pc8@sanger.ac.uk (P.J.C.),
jpimanda@unsw.edu.au (J.E.P.)

**In Brief**
Unnikrishnan et al. discover that patients who fail to respond to frontline therapy with 5-azacitidine (AZA) have a higher baseline proportion of quiescent hematopoietic progenitor cells. Longitudinal sampling reveals that, although AZA response fails to eradicate clonal hematopoiesis, it restores functional hematopoiesis from progenitors with a lower mutational burden.

**Accession Numbers**
GSE76203
SUMMARY

Myelodysplastic syndromes and chronic myelomonocytic leukemia are blood disorders characterized by ineffective hematopoiesis and progressive marrow failure that can transform into acute leukemia. The DNA methyltransferase inhibitor 5-azacytidine (AZA) is the most effective pharmacological option, but only ~50% of patients respond. A response only manifests after many months of treatment and is transient. The reasons underlying AZA resistance are unknown, and few alternatives exist for non-responders. Here, we show that AZA responders have more hematopoietic progenitor cells (HPCs) in the cell cycle. Non-responder HPC quiescence is mediated by integrin α5 (ITGA5) signaling and their...
hematopoietic potential improved by combining AZA with an ITGAV5 inhibitor. AZA response is associated with the induction of an inflammatory response in HPCs in vivo. By molecular bar coding and tracking individual clones, we found that, although AZA alters the sub-clonal contribution to different lineages, founder clones are not eliminated and continue to drive hematopoiesis even in complete responders.

INTRODUCTION

Myelodysplastic syndrome (MDS) is a diverse group of hematological disorders characterized by impaired peripheral blood cell production and abnormal bone marrow (BM) that affects 3.5–12.6 per 100,000 people per year. The incidence increases with age and is approximately five times greater in people over the age of 70. Patients with MDS die as a consequence of marrow failure or progression to acute myeloid leukemia (Nimer, 2008). Chronic myelomonocytic leukemia (CMML) is a related hematological neoplasm with features of both MDS and the myeloproliferative neoplasms. Cytogenetic defects and mutations are common in both MDS and CMML, with recurrent mutations observed in the RNA splicing machinery (Yoshida et al., 2011), cohesion complex (Kon et al., 2013), transcription factors, signal transduction machinery, and epigenetic modifiers (Cazzola et al., 2013). Although some mutations have clear clinical and prognostic implications (Bejar et al., 2011, 2014a; Itzykson et al., 2011; Traina et al., 2014), their etiological roles are still being elucidated.

Azacitidine (AZA) is a cytidine analog that is incorporated into DNA and RNA (Issa and Kantarjian, 2009). In the process of DNA methylation, DNA methyltransferases (DNMTs) become covalently linked to DNA-incorporated AZA and degraded, leading to global DNA demethylation (Issa and Kantarjian, 2009). Approximately half of all AZA-treated MDS or CMML patients respond to treatment, and response is associated with improved survival outcomes and decreased likelihood of leukemic transformation (Fenaux et al., 2009; Silverman et al., 2002). The efficacy of AZA compared with supportive care alone has been shown in MDS (Fenaux et al., 2009) and CMML (Adès et al., 2013; Costa et al., 2011). However, the dynamics of AZA response are slow, requiring 4–6 months of treatment to become apparent. With few alternative therapeutic options for AZA non-responders, the prognosis for such patients is consequently very poor (Prébet et al., 2011). Although some clinical parameters (Adès et al., 2013; Itzykson et al., 2013a) and genetic mutations (Bejar et al., 2014a; Itzykson et al., 2011) have weak correlations with favorable AZA response, the molecular mechanisms underlying primary AZA resistance are poorly understood. Furthermore, AZA response is rarely sustained, and a significant fraction of patients who initially respond will eventually relapse within a 2-year period, with very poor subsequent prognosis (Prébet et al., 2011). The reasons for secondary AZA resistance are also poorly understood.

Here we describe investigations that revealed that cell cycle quiescence, primarily of hematopoietic progenitor cells (HPCs), underpins primary AZA resistance. We show that AZA responsiveness is improved in vitro by combinatorially blocking integrin α5 signaling, known to mediate cell cycle quiescence. Our discoveries, therefore, not just offer a clinically useful method to prospectively identify AZA non-responders early, thus averting months of potentially futile therapy, but also provide one rational avenue for developing future combination therapies to benefit such patients. We have also identified that AZA treatment induces an inflammatory response in vivo. Last, to shed light on the in vivo fates of dysplastic cells upon AZA response, we “molecularly bar-coded” clones using somatic mutations personal to each patient, identified through exome sequencing. We have discovered that AZA response does not alter clonal hematopoiesis or reduce the abundance of clones bearing driver mutations. However, using single-colony genotyping, we have uncovered that AZA alters the sub-clonal contribution to distinct hematopoietic lineages.

RESULTS

Identification of Pre-treatment Gene Expression Differences between AZA Responders and Non-responders

Subsidized AZA access in Australia and Europe for MDS or CMML patients is limited to those with poor prognostic scores and high risk of transformation to acute myeloid leukemia (AML). Prior to the availability of AZA on the Pharmaceuticals Benefit Scheme (PBS) in Australia in 2011, we enrolled 19 patients (nine with CMML, ten with MDS; Table S1) on a compassionate access program. Inclusion criteria were in line with those stipulated for MDS or CMML patients to access AZA on the PBS (Table S1). AZA was administered at a dose of 75 mg/m2/day for 7 days on a 4-week cycle. The dose was increased to 100 mg/m2/day when no beneficial effect and no significant toxicity were seen after two treatment cycles. Dose reductions were not required. Patients who had received any prior disease-modifying agents were excluded. Blood products were allowed, but the use of hematopoietic growth factors during the study was prohibited.

We collected 100 mL of BM and isolated CD34+ hematopoietic/stem progenitor cells (HSPCs) with high purity at distinct stages during treatment for transcriptomic, genomic, and matching functional analyses (Figure 1A). Twelve patients were AZA responders, and seven were non-responders (Table S1), according to the International Working Group criteria (Cheson et al., 2006). We hypothesized that pre-existing molecular differences at baseline in HSPCs might drive primary AZA resistance.

To identify differentially expressed genes at pre-treatment, we performed RNA sequencing (RNA-seq) on pre-treatment HSPCs in a discovery cohort of 13 patients (seven responders: MDS, n = 4 and CMML, n = 3; six non-responders: MDS, n = 4 and CMML, n = 2; Table S2). Tissue samples from the remaining six responders (MDS, n = 3; CMML, n = 3) were saved for independent validation, along with samples sourced from two independent cohorts from the United Kingdom and Sweden (as described later). Bioinformatics analyses identified a set of genes (n = 731; Table S3) that were differentially expressed at false discovery rate (FDR) ≤ 0.05 between responders and non-responders (Figure 1B; Figure S1A). There was little overlap
Figure 1. Identification of Pre-treatment Gene Expression Differences between AZA Responders and Non-responders

(A) Schematic depicting cycles of AZA treatment and time points of BM collection per individual. Right: the different BM sub-fractions that were collected per time point and the analyses performed on them.

Legend continued on next page.
between our gene set and two recently published prognostic marker gene sets (Pellagatti et al., 2013; Figure S1B), nor any correlation with clinical parameters (Table S4). To conclusively rule out the possibility that gene expression differences were simply due to some other unappreciated difference in disease severity, we utilized a cohort of AZA-untreated MDS/RAEB-2 patients (n = 43, Oxford, United Kingdom) where transcriptome data from HSPCs (Pellagatti et al., 2010) and survival data were available. We dichotomized patients into putative AZA responders and non-responders based on the relative expression of our gene set but found no differences in survival between these two groups (Figure 1C). Comparable matched HSPC transcriptome data and long-term survival data in AZA-untreated CMML patients were not available, precluding identical analysis in CMML patients. However, from available evidence, these analyses support the alternative hypothesis that the differential gene expression in our cohort was related to a predisposition to respond to AZA therapy.

Pathway analyses of the upregulated genes in responders showed significant enrichment for pathways related to cell cycle progression (Ingenuity Pathway Analysis [IPA], see Figure 1D; Metacore, see Figure S1C), covering all stages of the cell cycle (Figure S2A), and DNA damage response pathways that are inextricably linked to cell cycle passage (Branzei and Foiani, 2008; Figure 1D; Figures S1C–S2C). Gene set enrichment analysis (GSEA) also uncovered significant enrichment for a number of cell cycle-related gene sets in responders (Figure 1E). All of these data indicated downregulation of cell-cycle-related pathways in the BM CD34+ HSPCs of primary AZA non-responders prior to treatment.

Cell Cycle Quiescence of Hematopoietic Progenitors Marks AZA Non-responders

We performed independent validation of the cell cycle differences using pre-treatment BM CD34+ HSPCs from a cohort of 54 patients from the United Kingdom (n = 14), Sweden (n = 34), and Australia (n = 6), consisting of 33 AZA responders (MDS, n = 28; CMML, n = 5; Table S5) and 21 non-responders (MDS, n = 19; CMML, n = 2; Table S5). We identified a set of 20 cell cycle-related genes that had significant differential expression between AZA responders and non-responders at pre-treatment in our discovery cohort (median FDR = 0.0002; Table S3) and that were consistently present among the top deregulated cell cycle pathways in our IPAs (Figure 1D). Furthermore, the differential expression of just these genes was sufficient to efficiently separate responders from non-responders in our discovery cohort (Figure S3A). Using custom-designed real-time PCR probes against these genes, we confirmed that the differential expression of these genes could also identify responders from non-responders in the validation cohort (Figure 2A) with a sensitivity of 78.8% and a positive predictive value of 76.5%. Our findings are consistent with the recent observation of increased cell cycle quiescence in CMML non-responders to decitabine therapy (Meldi et al., 2015). However, we were unable to observe similar gene expression differences in cell cycle genes using matched bulk unfractionated mononuclear cells (MNCs) collected from the patients in our cohort (data not shown), highlighting that cell type is an important parameter that warrants addressing in cross-cohort validation and limiting the utility of a gene expression-based assay as a routine diagnostic tool.

To further pursue the link between cell cycle quiescence and AZA resistance with an assay that can be performed on unfractionated cell populations, we adapted a flow cytometry-based method (Tehranchi et al., 2010) to measure cell cycle parameters in matched pre-treatment MNCs from 12 patients (six responders and six non-responders) in whom we performed gene expression analyses as well as BM MNCs from six healthy adult volunteers (Experimental Procedures; Figure S3B). Healthy Lin− CD34+ CD38low−/C0 CD90+ hematopoietic stem cells (HSCs) were mostly quiescent (the median G0, G1, and S/G2/M proportions were 79.7%, 20%, and 0.7%, respectively; Figures 2B and 2C), whereas Lin− CD34+ CD38− HPCs were cycling (median G0, G1, and S/G2/M: 20.1%, 71.8%, and 6.6%, respectively; Figure 2C). HSCs (Lin− CD34+ CD38−) of non-responders were markedly quiescent, with few replicating cells (median G0, non-responders = 81.75%, non-responders = 77.3%, p = 0.76). However, the HPCs (Lin− CD34+ CD38−) of non-responders were mostly quiescent, with few replicating cells (median G0, G1, non-responders = 65.2%, responders = 46.3%, p = 0.019; median S/G2/M, non-responders = 1.24%, responders = 7.68%, p = 0.0023; Figures 2B and 2C). These data were not only consistent with the gene expression-based identification that CD34+ HSPCs in non-responders were quiescent but, additionally, revealed that it is specifically the population of normally cycling CD34+ CD38+ HPCs in the BM that is quiescent in AZA non-responders.

Targeting Cell Cycle Quiescence to Overcome AZA Resistance

We hypothesized that disruption of pathways that maintain quiescence might restore AZA sensitivity. Within the genes that were upregulated in AZA non-responders, we observed enrichment for a number of signaling pathways, chiefly integrin-linked signaling through integrin α5/ITGA5 (Figure 1D; Figure S1C). In our RNA-seq data from pre-treatment BM CD34+ cells, we observed that transcription of ITGA5 itself, the receptor for the extracellular matrix protein fibronectin and a protein known to be required for maintaining quiescence in HSCs (van der Loo et al., 1998), was upregulated in AZA non-responders (median

(B) Heatmap indicating the relative differential expression of the 731 genes (in rows) from RNA-seq at pre-treatment between AZA responders (green) and non-responders (orange). The expression values for each gene have been standardized to have a mean of 0 and an SD of ± 1 across all samples.

(C) Kaplan-Meier survival analysis of an independent cohort of AZA-untreated MDS patients, clustered by k-means based on the differential expression of the 731 genes into putative responders or non-responders, illustrating no significant difference in survival.

(D) IPA of genes upregulated in responders (left) and of genes upregulated in non-responders (right), with the identities of enriched pathways and associated p values.

(E) Gene set enrichment analysis (GSEA) indicating that cell cycle-related gene sets are enriched in responders compared with non-responders. Normalized enrichment scores (NESs) and FDR q values are given for the specific gene sets.
counts per million, non-responders = 1942.8495, responders = 1089.364, q = 2.09E-07; Figure S3C; Table S3). We independently validated these findings in pre-treatment CD34+ cells at both the transcriptional level by qRT-PCR and at the protein level using flow cytometry (Figure 3A).

To test the effect of blocking ITGA5 in combination with AZA, we developed a stromal co-culture assay wherein pre-treatment BM HSPCs from patients were cultured with MS5 murine stromal cells and treated daily with combinations of drugs (as outlined in Figure 3B). In vivo, response to AZA manifests as improved hematopoiesis. On this basis, we utilized the capacity to form colony forming units (CFUs) as an in vitro measure of improved hematopoiesis. In AZA responders, however, there was no further improvement of CFU-Cs with combination treatment compared with AZA alone (PD7153 and PD7155, CMML; PD7168, MDS; Figure 3C), with no difference in cell viability (Figure S3C).

In AZA responders, we observed a discernible difference in CFU-Cs between any of the treatment conditions (Figure 3E). The increase in CFU-Cs in AZA responders mirrored what occurs in vivo in AZA responders, with significantly greater numbers of colonies after six cycles of AZA treatment (C6d28) compared with pre-treatment (a representative non-responder, PD7166, Figure 3F, and two additional non-responders, Figure S3F).

Combination therapy, utilizing a well-characterized anti-ITGA5-specific antibody (Nam et al., 2010; Sawada et al., 2008) together with AZA, specifically led to an increase in the total numbers of CFU cells (CFU-Cs) in three different non-responders (PD7158 and PD7170, CMML; PD7169, MDS; Figure 3C), with no difference in cell viability (Figure S3C). A receiver operator characteristic (ROC) curve based on the expression of the 23 genes per individual is shown on the right, along with the area under the curve (AUC) measure. A boxplot summarizing the proportions of BM HSCs (top) and HPCs (bottom) in different stages of the cell cycle in six healthy individuals, responders and non-responders each. The whiskers extend to the maximum and minimum values, and the median is depicted as a horizontal bar. Median values for percentages of HSCs in S/G2/M are indicated above their respective bars. *p < 0.05 from a two-tailed t test.

Figure 2. Increased Cell Cycle Quiescence Characterizes HPCs of AZA Non-responders
(A) Hierarchical clustering heatmap of Fluidigm-based qRT-PCR measurement of the expression of 20 cell cycle-related genes (rows) at pre-treatment in an independent cohort of AZA-treated patients (columns). The top bar indicates the patients’ predicted AZA response based on the hierarchical clustering (predicted responders in brown, non-responders in light blue), whereas the bottom bar indicates the clinically observed response based on International Working Group (IWG) criteria (AZA responders in green, non-responders in orange). Adult BM CD34+ from a healthy individual (yellow in the bottom bar) is also included as a reference. A boxplot summarizing the proportions of BM HSCs (top) and HPCs (bottom) in different stages of the cell cycle in six healthy individuals, responders and non-responders each. The whiskers extend to the maximum and minimum values, and the median is depicted as a horizontal bar. Median values for percentages of HSCs in S/G2/M are indicated above their respective bars. *p < 0.05 from a two-tailed t test.
**Figure 3. Targeting Cell Cycle Quiescence by Blocking ITGA5 in CD34+ HSPCs Overcomes AZA Resistance**

(A) Upregulation of ITGA5 transcripts in BM CD34+ HSPCs of AZA non-responders compared with responders at pre-treatment, detected at the transcriptional level by qRT-PCR (left) and at protein level by flow cytometry (right). The mean is depicted as a horizontal bar. *p < 0.05, two-tailed t test.

(B) Schematic depicting the stromal co-culture drug testing setup.

(C) CFU colony counts for pre-treatment BM CD34+ HSPCs from three AZA non-responders (PD7170, PD7158, and PD7169) treated with vehicle control (control), 500 nM AZA only (AZA), 10 μg/mL anti-ITGA5 antibody only (ITGA5i), or a combination of 500 nM AZA + 10 μg/mL anti-ITGA5 antibody (AZA + ITGA5i) for three consecutive days in a stromal co-culture. The legend indicates the three different possible types of myeloid colonies (CFU-granulocyte, erythrocyte, monocyte, megakaryocyte, CFU-GEMM; CFU-granulocyte, monocyte, CFU-GM; BFU, Burst-forming unit) that were observed. *p < 0.05, two-tailed t test.

(D) AZA Responders

(E) Healthy

(F) In vivo

(legend continued on next page)
AZA Therapy Induces Pro-inflammatory Pathways
In Vivo in Responders

The mechanisms that determine in vivo effects of response to AZA therapy are poorly understood. Although a number of previous studies have assessed the transcriptional response of in vitro AZA-treated cells (Chiappinelli et al., 2015; Hollenbach et al., 2010; Karpf et al., 1999; Li et al., 2014; Roulois et al., 2015), there is a paucity of data available for in vivo response, especially following the long period (4–6 months) of treatment required to see clinical improvements. To address this, we performed RNA-seq analyses on HSPCs from nine responders and five non-responders at cycle 6 day 28 (C6d28) and compared gene expression to pre-treatment. GSEA in responders identified an upregulation of inflammation-related genes (Figure 4A; Table S6) at C6d28 compared with pre-treatment. Recently, low-dose AZA treatment in vitro was shown to induce a set of AZA immune genes (Li et al., 2014). We also identified a strong enrichment for these genes at C6d28 in vivo following AZA response (Figure 4B). Orthogonal analyses of the upregulated genes by IPA also implicated upregulation of inflammation and immune response pathways following response to AZA treatment (Figure 4C). In non-responders, however, we did not observe any alteration of inflammatory pathways or of any other specific pathways in general. These findings provide in vivo evidence to corroborate recent reports that in vitro AZA treatment of colorectal and ovarian cancer cells induces inflammatory and immune-related responses through the induction of intracellular double-stranded RNA (Chiappinelli et al., 2015; Roulois et al., 2015). Our data suggest that one of the mechanisms through which AZA therapy effects in vivo improvement in MDS and CMML patients is through the modulation of inflammatory response pathways.

AZA Alters the Sub-clonal Contribution to Distinct Hematopoietic Lineages

Although approximately half of all patients will respond to AZA initially, a response is rarely sustained long-term. A significant fraction of responders will relapse within a 2-year period of beginning treatment, with a markedly poor prognosis thereafter (Prébet et al., 2011). The reasons for secondary AZA resistance are poorly understood and further undermined by a lack of clarity over the in vivo fate of dysplastic cells upon response to AZA.

By molecularly bar-coding clones using somatic mutations, we decided to track their fate upon AZA treatment. By exome sequencing of pre-treatment and C6d28 samples in 16 patients (Figure S4A), followed by high-depth targeted resequencing at each time point, including in long-term follow-up samples (where available), we obtained high-confidence variant allele frequencies (VAFs). Clustering of the VAFs enabled us to distinguish different clones. We identified an average of eight mutations per patient (Table S7). Hematopoiesis was clonal in all patients at pre-treatment, with a single major clone and a mode of two minor sub-clones per patient (range, one to three sub-clones), and remained clonal through treatment. In all non-responders (n = 6), we did not observe any significant changes in clonal abundance over the course of AZA treatment (Figure 5A; Figure S4B). These findings are consistent with previous reports showing that the subclonal structure is unaltered in MDS and CMML AZA non-responders (Craddock et al., 2013; Itzykson et al., 2013b). However, even in the majority (n = 8, 80%) of AZA responders, the clonal mix of MDS cells remained unchanged even upon complete response, and hematopoiesis remained fully clonal (Figures 5B–5D; Figure S4C). Only in a minority of responders (n = 2, 20%) did we observe a diminution of mutated clones (Figure S4B).

TET2 mutations have been previously correlated with favorable AZA response (Bejar et al., 2014a; Itzykson et al., 2011; Traina et al., 2014), but the fate of mutant clones is unknown. Four of the six patients with TET2 mutations were responders (PD7151, PD7153, PD7154, and PD7161). In all patients, the abundance of the TET2 mutant clones was unchanged with treatment, indicating that these clones were not eliminated by AZA treatment (Figure 5C). In PD7153, >85% of the cells in the BM arose from a clone with two somatic TET2 mutations (TET2 L1065* and TET2 Q685*; Figure 5C). The abundance of the TET2 clone remained virtually unchanged, even after 2 years of treatment (Figure 5D). Our data reveal that AZA therapy does not change the major clonal structure of hematopoiesis in vivo, even in the event of complete hematologic response. Our findings are consistent with a recent report of the effects of hypomethylating therapy on mutant clones in a cohort of CMML patients (Merlevede et al., 2016).

However, BM VAFs are not informative with respect to hematopoietic functionality of resident clones and sub-clones. To reconcile the persistence of mutant clones with the clinically observed normalization of hematopoiesis upon AZA response, we complemented our sequencing analysis of bulk cells with genotyping of individual CFU colonies to determine clonal contribution to different myeloid lineages (Figure 6A). We designed custom real-time PCR probes against all identified somatic mutations in two AZA responders (PD7153 and PD7165) and two non-responders (PD7166 and PD7158) and performed CFU assays with pre-treatment and C6d28 HSPCs. In PD7153, at pre-treatment, the colony genotyping data largely reflected the patterns observed by bulk genotyping; every single colony contained TET2 driver mutations, and the majority of the colonies derived from the major clone, with smaller contributions from the minor subclones (Figure 6B). However, at C6d28, we observed that a number of colonies that grew up did not contain the full complement of mutations, but, instead, the clonal contribution to distinct sub-lineages was varied (Figure 6B). Although all colonies from more primitive granulocyte, erythrocyte,
megakaryocyte, macrophage CFU (CFU-GEMM) possessed the full complement of founder mutations, including both TET2 mutations (Figure 6D), granulocyte macrophage CFU (CFU-GM) and erythroid burst-forming unit (BFU) colonies were increasingly derived from cells containing fewer sets of mutations, including just a single TET2 mutation (TET2 p.L1065fs*1; n = 14; 15%) or two mutations (TET2 p.L1065fs*1, RBPM2 p.V126L; n = 26; 28%; Figure 6D). Similarly, in another AZA responder (PD7165), we observed that a number of C6d28 CFU-GM colonies were derived from cells lacking any of the founder mutations (n = 18, 19%), whereas all pre-treatment colonies contained a full complement of mutations (Figures S5A–S5E). No sub-clonal variation was observed in AZA non-responders (Figure 6C; Figure S5E). Our data indicated that, although AZA therapy did not completely eliminate dysplastic clones upon response, it altered the clonal contribution to hematopoiesis, enabling previously dormant clones with a lower mutational burden to contribute to hematopoiesis.

**DISCUSSION**

Although AZA is the most efficacious pharmaceutical option for MDS and CMML, only about half of the patients will respond. Given the lack of understanding of the molecular basis of primary resistance and few therapeutic alternatives, the outcome for AZA non-responders is poor (Prébet et al., 2011). Even among the responders, the response is rarely sustained, with a poor prognosis thereafter (Fenaux et al., 2009). Although mutation profiling has recently identified mutations in TET2 and DNMT3A as predictors of good AZA response (Bejar et al., 2014a; Itzykson et al., 2011; Traina et al., 2014), they have limited clinical utility unless better alternative treatment options can be developed for non-responders.

The present study shows that primary AZA resistance is intrinsically linked to cell cycle quiescence of HPCs in non-responders before treatment. Active DNA replication is required for incorporation of AZA into DNA, enabling the subsequent trapping and degradation of DNMTs (Stresemann and Lyko, 2008), explaining why the quiescent cells of non-responders are refractory to treatment. Our findings are consistent with a recent observation of increased cell cycle quiescence in CMML non-responders (Meldi et al., 2015). Of the 61 genes common between our gene signature and that of Meldi et al. (2015), a large fraction is related to the cell cycle. By directly measuring cell cycle parameters in HSCs and HPCs using flow cytometry, we show that cell cycle differences between AZA responders and non-responders are limited to the HPC fraction. Importantly, given that flow cytometry-based cell cycle assessment is performed on unfractionated freeze-thawed BM MNCs, it is more amenable as a diagnostic tool than measuring gene expression in fractionated BM
Figure 5. Clonal Evolution in the BM of Patients in Response to AZA Therapy

(A) Clonal evolution in the BM of two representative AZA non-responders over the course of treatment. The heatmaps depict the variant allele frequencies for different somatic mutations (columns; the numbers at the bottom correspond to mutations listed in Table S7) over the entire treatment period (rows). Known driver mutations previously implicated in cancer or myeloid leukemias are marked in orange in the top bar, whereas mutations of unknown significance are marked in teal. The coxcomb plots below the heatmaps depict the variant allele frequencies (VAFs) of all mutations at pre-treatment, C6d28, and a long-term follow up (LTFU) where available. The schematic at the bottom of the coxcomb indicates the axes of the different somatic mutations plotted.

(B) Clonal evolution in the BM of two representative AZA responders over the course of treatment. The heatmaps and coxcomb plots are depicted as in (A) at each of the major treatment time points (including an LTFU sample where available).

(legend continued on next page)
CD34+ cells. However, our initial discoveries will need to be further ratified in larger cohorts of patient samples.

Predicting AZA resistance, however, is of limited benefit in the absence of therapeutic alternatives. An important outcome of this study is that we provide experimental evidence that AZA resistance can be overcome by targeting a pathway that is upregulated in AZA non-responders. We have identified that the increased cell cycle quiescence in AZA non-responders is correlated with the upregulation of ITGA5 and that blocking integrin signaling in combination with AZA improves the CFU capability of HSPCs in these patients. Conversely, our results suggest that combining AZA with drugs that inhibit the cell cycle could be counterproductive. Indeed, combining histone deacetylase inhibitors, which are known to inhibit the cell cycle, with AZA has been the subject of over a decade of clinical testing but has either failed to show improved efficacy or has been inferior to AZA alone (Prebet et al., 2014), highlighting the need for rational data-driven approaches to combination drug therapy.

By bar-coding clones by their mutations, we have identified that AZA does not alter the abundance of early-arising driver mutations or clonal hematopoiesis, even upon response. Our findings are consistent with those of a recent report of the effects of hypomethylating therapy on mutant clones in a cohort of CMML patients (Merlevede et al., 2016). However, in the absence of a functional assessment of persisting clones following treatment, the basis for restored hematopoiesis despite the lack of bulk clonal variation in AZA responders remained unresolved in the latter study. By interrogating clonal function using CFU assays and by performing single-colony genotyping, we show that, although hematopoiesis remains clonal in responders, AZA therapy significantly changed the sub-clonal contribution to distinct hematopoietic lineages. Although clones with a high mutation burden were present at all stages of myeloid differentiation (CFU-GEMM, CFU-GM, and BFU) before treatment, AZA treatment appears to specifically blunt the ability of these clones to become more lineage-committed cells (CFU-GM and BFU; Figure 6E). Instead, cells bearing fewer mutations are more abundantly detected in these populations. We speculate that the more mutated cells have a clonal advantage, enabling them to populate more of the BM, preferentially at pre-treatment, but are also less capable of differentiating normally. AZA prevents clones with a high mutational burden from crowding out less mutated cells that are more capable of differentiating normally, thereby improving hematopoiesis (Figure 6E).

However, there are caveats to our findings that necessitate further study; First, given the small number of patients in our cohort, although extensively interrogated molecularly, our discoveries warrant validation in larger sets of patients. Also, because we performed exome sequencing using DNA from BM MNCs with some differentiation potential, we cannot exclude the possibility that there are other clones bearing further mutations that are completely incapable of differentiation. Additionally, given that clonal hematopoiesis is frequently observed in aging (Busque et al., 2012), albeit with an increased risk of acquiring hematological malignancy (Genovese et al., 2014; Jaiswal et al., 2014; Xie et al., 2014), it is possible that some of the clonal hematopoiesis observed upon AZA response may not have accompanying dysplasia. Indeed, the association between specific mutations and marrow dysplasia warrants further investigation, given the current uncertainty over the distinction between clonal hematopoiesis of indeterminate potential and myelodysplastic syndromes (Steensma et al., 2015).

Last, we find in vivo evidence to support that AZA therapy induces inflammatory- and immune-related responses in HSPCs of AZA responders, consistent with recent work in solid tumor cell lines (Chiappinelli et al., 2015; Roulois et al., 2015). Taken together with our clonal evolution and functional data, these findings suggest that the induction of pro-inflammatory pathways in persisting dysplastic clones by AZA facilitates their contribution toward functional hematopoiesis in vivo, a clinical hallmark of AZA response.

Therefore by utilizing primary patient material from well annotated treatment cohorts and by integrating transcriptomics, genomics, and clonal function, we provide insights into the molecular basis of primary AZA resistance, mechanisms to predict and overcome AZA resistance, and evidence for clonal hematopoiesis in the presence of persistent, highly mutated MDS/CMML HSPCs in complete responders. These findings should serve as a platform for more effective utilization of AZA and development of more durable therapies for treating MDS and CMML.

**EXPERIMENTAL PROCEDURES**

**Patients and Sample Collections**

Nineteen individuals (seven females and 12 males) aged 59–88 with MDS or CMML (Table S1) were recruited from New South Wales, Australia on a compassionate access basis for AZA. All samples were obtained with written informed consent in accordance with the Declaration of Helsinki and approval of the human research ethics committee of the South Eastern Sydney Local Health District. BM samples were collected from each individual at defined time points. Immediately upon sample collection, MNCs were isolated from the BM by density centrifugation using Lymphoprep (STEMCELL Technologies). Approximately 2 × 10⁶ MNCs per sample were then incubated with CD34+ magnetic beads (Miltenyi Biotec) and separated using an AutoMACS Pro machine (Miltenyi Biotec) exactly as recommended by the manufacturer. Approximately 1 × 10⁶ CD34+ and CD34− cells were used to prepare RNA (CD34+) and DNA (CD34+, CD34−). All remaining cells (MNCs, CD34+, and CD34−) were frozen viably and stored to be used later in functional assays or for flow cytometry as necessary. Details are described in the Supplemental Experimental Procedures.

**Transcriptomic Analyses**

Total RNA was amplified using the Ovation RNA-seq System v2 kit (NuGEN) prior to RNA-seq. Approximately 1 μg of the resulting, amplified, double-stranded cDNA was fragmented and used to prepare 90-bp paired-end sequencing libraries using the TruSeq RNA library preparation kit (Illumina) according to the manufacturer’s recommendations. All sequencing was
**Figure 6. AZA Alters the Sub-clonal Contribution to Distinct Hematopoietic Lineages**

(A) Schematic of CFU genotyping.

(B) Heatmap representation of Fluidigm-based genotyping of individual CFU-Cs (columns) at pre-treatment and C6d28 from AZA responder PD7153. The homozygous wild-type allele at any locus is shaded in blue, the heterozygous allele is shaded in orange, and the homozygous mutant allele is shaded in red. The number of colonies is indicated below each heatmap.

(C) Fluidigm-based genotyping of individual CFU-Cs from an AZA non-responder (PD7166).

(D) Genotypes of different colony types from PD7153 based on the cell of origin.

(legend continued on next page)
performed on a HiSeq2000 sequencer (Illumina) at BGI. Cleaned raw sequencing reads were aligned individually against the human genome reference (hg19) using TopHat (version 2.0.4) (Kim et al., 2013) with default parameters. Gene level expression was quantified using the DESeq2 (version 1.10.1) (Love et al., 2014) package in R. The quartiles of genes with the lowest counts were filtered, followed by standard normalization procedures for the read sets across the samples. A full description of the methods is available in the Supplemental Experimental Procedures. Fluidigm analyses for gene expression were performed using dynamic arrays on the BioMark HD (Fluidigm). Details are described in the Supplemental Experimental Procedures.

Statistical Analyses
All statistical analyses on clinical parameters were performed in R. Fisher’s exact t test (for disease type, gender, and cytogenetics), log rank (Mantel-Cox) test for survival analyses, two-tailed, unpaired Student’s t test (for normally distributed data), and Wilcoxon signed-rank test (for non-normally distributed data) were performed where appropriate. Statistical correlations between the gene signature and clinical parameters were done by logistic regression using the glm package in R. Significance was considered for p < 0.05. FDRs for RNA-seq analyses were generated using DESeq2 (version 1.10.1) (Love et al., 2014). Receiver operator characteristic (ROC) analyses were performed in GraphPad Prism v6.0 (GraphPad). Details of the respective tests are described in full in the Supplemental Experimental Procedures.

ACCESSION NUMBERS
The accession number for the sequencing data reported in this paper is GEO: GSE76203.

SUPPLEMENTAL INFORMATION
Supplemental Information includes Supplemental Experimental Procedures, five figures, and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2017.06.067.

AUTHOR CONTRIBUTIONS
Most of the study design, experiments, and data analyses were performed, guided, and supervised by A.U., P.J.C., and J.P. A.U., E.P., D.B., and N.P.D. performed data analyses, supervised by M.R.W., J.W.H.W., P.J.C., and J.P. A.V., A.K., L.A.R., K.K., V.C., J.A.I.T., M.L.T., Y.H., Z.A., J.O., S.G., S.R.W., and K.L.M. helped perform experiments, data acquisition, and data analyses. P.S.W. and S.E.W.J. guided and advised on flow experiments. Clinical samples and associated data were generously donated by A.G.K., M.T., M.K., A.P., J.B., G.J.M., and E.H.L. A.G.K. and M.T. further assisted with the analysis of clinical data. Clinical trials were conducted and patient samples were collected by R.L., B.Y., R.R., C.A., R.S., P.C., J.C., P.W., F.R., and J.E.P. The clinical trial was designed by J.E.P., A.U., A.V., and J.A.I.T. assisted with the isolation of patient samples. K.L. assisted with the implementation of the clinical trial and facilitated the coordination of various aspects of experimental work. The manuscript was written by A.U. and J.E.P. and reviewed and agreed upon by all coauthors.

ACKNOWLEDGMENTS
The authors would like to thank Drs. Orly Lavee, Anna Ostberg, Greg Corboy, Nicole Wong Doo, Victoria Potter, Brendan Beaton, Carrie Vanderweyden, and Natalie Meyer and the nursing staff at the Haematology Unit, Prince of Wales Hospital for assistance with BM collection. We also thank Mr. Chris Brownlee (flow cytometry facility, Biological Resources Imaging Laboratory, UNSW), Drs. Kristin Miller and Helen Speirs (Ramaciotti Centre, UNSW), and Dr. Ling Zhong and A/Prof. Mark Raftery (BMSF, UNSW) for technical assistance rendered. The authors acknowledge funding from the National Health and Medical Research Council (NHMRC), Leukaemia Foundation, Anthony Rothe Foundation, Cancer Institute for New South Wales, South Eastern Area Laboratory Services (SEALS), Wellcome Trust, Leukemia and Lymphoma Society, Medical Research Council (UK), Swedish Cancer Society, Cancer Society in Stockholm, Swedish Research Council, Bloodwise UK, and the NIHR Biomedical Research Centre, Oxford.

(E) Schematic illustrating the effect of AZA therapy on clonal contribution to different hematopoietic lineages. The widths of the arrows indicate the functional capacity of the different clones to differentiate, and the lengths of the arrows indicate their relative contribution to terminal hematopoiesis. The shading of the cells indicates the number of mutations they possess, from less mutated (light red) to most mutated (dark red), with internal colored circles representing clusters of mutations. Grey cells represent possible other clones in the BM that were never able to differentiate into CD34+ cells and, therefore, were never sampled by exome sequencing.


