Supplemental Information

Direct Modulation of the Bone Marrow Mesenchymal Stromal Cell Compartment by Azacitidine Enhances Healthy Hematopoiesis

Catharina Wenk[1*], Anne-Kathrin Garz[1*], Sonja Grath[2], Christina Huberle[1], Denis Witham[1], Marie Weickert[1], Roberto Malinverni[3,4], Julia Niggemeyer[5], Michele Kyncl[1], Judith Hecker[1], Charlotta Pagel[1], Christopher B. Mulholland[2], Catharina Müller-Thomas[1], Heinrich Leonhardt[2], Florian Bassermann[1,6], Robert A.J. Oostendorp[1], Klaus H. Metzeler[5,6], Marcus Buschbeck[3,4] and Katharina S. Götze[1,6]

1. Department of Medicine III, Technische Universität München (TUM), Munich, Germany
2. Department of Biology II, Ludwig-Maximilians-Universität München (LMU), Munich, Germany
3. Campus ICO – Germany Trias I Pujol – Universidad Autonoma de Barcelona, Josep Carreras Leukaemia Research Institute, Badalona, Spain
4. Program for Predictive and Personalized Medicine of Cancer, Germans Trias i Pujol Research Institute (PMPPC-IGTP), Badalona, Spain
5. Laboratory for Leukemia Diagnostics, Department of Medicine III, Ludwig-Maximilians-Universität München (LMU), Munich, Germany
6. German Cancer Consortium (DKTK) and German Cancer Center (DKFZ), Heidelberg, Germany

* These authors contributed equally to this work.

Scientific Category: Myeloid Neoplasia

Running Title: AZA modulates mesenchymal stromal cells in MDS

Key words: azacitidine, MSC, HSPC, MDS, epigenetic drug

Word count: 4000

Figures/Tables: 7/1

References: 40

Corresponding author:

Katharina S. Götze, MD

Department of Medicine III

Technische Universität München

Ismaninger Str. 15

D-81675 Munich, Germany

Tel. +49 89 4140 5618

Fax: + 49 89 4140 4879

Email: katharina.goetze@tum.de

Conflict of Interest Statement: KSG has received honoraria from Celgene. KFB and HM have received research support and honoraria from Celgene. All other authors report no conflict of interest.

Supplemental Materials and Methods

Cell lines and reagents

EL08-1D2 stromal cells were used up to passage 12 and cultured as described.[1] AZA (Celgene Corp., Munich, Germany) was freshly prepared directly before use following manufacturer's instructions.

Colony forming unit (CFU) genotyping

For genotyping of CFU, individual colonies were plucked after scoring by visual identification under an inverted microscope using a micropipette. Genomic DNA was extracted from ten colonies for each condition and patient using the viral DNA kit (Zymo Research). Each colony was genotyped for one or two known MDS-associated mutations which had been previously determined to be present in the samples at high VAF by targeted next-generation sequencing at diagnosis. Genotyping of individual colonies was performed using PCR and Sanger sequencing.

Competitive co-culture of healthy and MDS HSPC on healthy MSC

Primary healthy MSC were treated once with AZA (10 µM) or left untreated for 48 hours. In the meantime, MDS and healthy CD34+ HSPC were prepared for competitive co-culture experiments. BM samples were thawed and subjected to Ficoll gradient centrifugation to remove dead cells. Viable cells from each sample were equally divided and then stained either with CellTrace Violet (CTV, 0.8 uM) or carboxyfluorescein succinmidyl ester (CFSE, 1 uM) for 20 min at 37°C followed by washing in IMDM. In addition, cells were stained with PI and fluorescence coupled antibodies against CD34-PE/Cy7 (clone 561, BioLegend) and CD45-APC/Cy7 (clone H130, BioLegend). Stained BM samples were sorted on a MoFlo XDP sorter (Beckman Coulter) for PI-/CD34+/CD45+/CTV+ or PI-/CD34+/CD45+/CFSE+ cells, respectively. Sorted MDS and healthy HSPC were deposited directly onto prepared confluent healthy MSC -/+ AZA treated layers in a 1:1 ratio in serum-free medium with 5 growth factors.[2] In total, the following cultures were set up on untreated as well as AZA treated MSC: I) CFSE+ MDS HSPC + CTV+ healthy HSPC, II) CTV+ MDS HSPC + CFSE+ healthy HSPC and as a control III) CFSE+ healthy HSPC + CTV+ healthy HSPC. An MDS control culture not possible due to the very small cell number after sorting. After 4 days, hematopoietic cells were harvested from the co-culture by repeated washing, stained for CD34-PE/Cy7, CD45-APC/Cy7 and PI to exclude stromal cells as well as dead cells and analyzed by flow cytometry using Flow-Count fluorosphere beads (Beckman Coulter) for quantification. Data were analyzed with FlowJo software (FlowJo, LLC).

RNA Sequencing of EL08-1D2 cells

Total RNA was isolated from EL08-1D2 cells after 4 days of culture with or without AZA (10 µM) added on day 0, using the RNeasy Mini Kit (Qiagen) as per manufacturer's recommendations.

RNA sequencing (RNA-seq) library preparation and data processing was performed according to the Unique Molecular Identifiers (UMI) sequencing protocol, a modified bulk single-cell RNA barcoding and sequencing (SCRB)-seq protocol.[3] Briefly, cDNA libraries were prepared from 5 ng RNA per sample. Unique molecular identifiers (UMIs) and sample barcodes were introduced during reverse transcription. Next, cDNA of all samples was pooled, unincorporated primers were digested with Exonuclease I and pre-amplified by single-primer PCR (12 cycles). After quality control of amplified cDNA using a Bioanalyzer (Agilent Technologies), the sequencing library was constructed from 0.7 ng of full-length cDNA using the Nextera XT Kit (Illumina). To enrich for barcoded 3' ends in the Nextera XT library, a custom i5 primer was used. Library pools were sequenced on a rapid flow-cell with paired-end layout using an Illumina HiSeq1500. The first read contained the sample barcode and the UMI sequences (16 cycles), while the second read contained the cDNA fragment (50 cycles). Note that UMI sequencing is a 3' mRNA sequencing approach used to count the number of unique transcripts. This procedure has the advantage that fewer reads are required compared to full-length RNA sequencing methods, as only the 3' transcript ends need to be covered.[3,4] All samples were mapped against the mouse (mm10) genome and gene features were quantified using Ensembl gene models (v75).

Differential Gene Expression Analysis

Gene expression analysis and visualization was done with R version 3.2.0 on a x86_64-pc-linux-gnu (64-bit) platform under Ubuntu 14.04.2 LTS. Differential gene expression analysis was performed using DESeq2.[5] DESeq2 uses library-size-corrected read count data to find differentially expressed genes and is an error model based on the negative binomial distribution for the read counts. For gene expression analysis, the fits of the negative binomial with a generalized linear model were analyzed. Coefficients (interpreted as the log2-fold changes) were tested using the Wald test. The false discovery rate (FDR) according to Benjamini-Hochberg was used for multiple testing corrections.[6] All treated samples were compared to untreated samples and a log2-fold change for each gene and comparison was calculated (Supplemental Table 2). Hierarchical clustering was applied gene- and sample-wise with single linkage based on Euclidian distances of variance-stabilized counts of the 500 most differentially expressed genes (500 genes with lowest adjusted p-value, false-discovery-rate adjustment based on.[6] The result was visualized as heat map with R package heatmap3 version 1.1.1. (relates to Figure 5B).

Gene Ontology Analysis

We performed gene ontology (GO) enrichment analysis on all significantly differentially expressed genes with GOrilla.[7] Two unranked lists of genes (target: up- or down-regulated genes, background: all genes) were committed to the web tool and enrichment was determined for Biological Process, Molecular Function and Cellular Component. Enrichment is defined by four parameters (N: total

number of genes; B: total number of genes that are associated with a specific GO term; n: number of genes at the top of the target list; b: number of genes in the intersection), such that Enrichment = (b/n)/(B/N). Enrichment p-value is corrected for multiple testing using Benjamini and Hochberg and stated as 'FDR q-value' (Supplemental Table 3).[6]

Fixed network enrichment analysis (FNEA)

Over-representation of significantly differentially expressed genes in Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways and Gene Ontology (GO) categories were tested by a fixed network enrichment analysis (FNEA) implemented in the neaGUI R package[8] (Supplemental Table 4; relates to Figure 5C).

Gene set enrichment analysis (GSEA)

Gene set enrichment analysis (GSEA) based on gene set 1[9] was done with the GSEA Desktop Application v. 3.0 (http://www.broad.mit.edu/gsea/)[8] using the tool GSEAPreranked. For ranking all genes, a metric score was calculated by multiplying the signed log2 fold changes by the inverse of the p-values. Statistical significance was determined by 1000 gene set permutations (Supplemental Table 5; relates to Figure 6A).

Weighted gene co-expression analysis (WGCNA)

We used weighted gene co-expression analysis (WGCNA)[10] to identify networks in all genes and restricted to HSPC supportive genes (gene set 1 in[9]). The analysis was done with R version 3.2.0 on a x86_64-pc-linux-gnu (64-bit) platform under Ubuntu 14.04.2 LTS. We used the R package WGCNA version 1.63. For WGCNA, outlier samples should be removed as well as genes with too few reads or zero-variance. Further, the basis in WGCNA is a matrix of Pearson's correlation coefficients for all pair-wise comparisons of genes. For Pearson's correlation the data should be approximately normal meaning we need some variance stabilization. To this end, we applied the DESeq2 function varianceStabilizingTransformation on the count matrix. In addition, we only included genes with enough information so that the variances can be reliably estimated without spurious correlations. This was achieved by excluding genes with p-values set to NA in the DESeq2 analysis. To measure co-expression of genes, an adjacency matrix was calculated using absolute Pearson correlation coefficients between two genes for all pair-wise comparisons. The soft-thresholding power parameter ß that resulted in approximately scale-free network topologies, defined as $R^2 > 0.9$, was 2 for all genes and 3 for HSPC supportive genes. Subsequently, the adjacency matrix was transformed into a topological overlap matrix (TOM), a similarity measure found to result in biologically meaningful modules.[11,12] Modules were constructed with average linkage clustering and a minimum module size

of 30. Each module is given an arbitrary color where genes that do not fall into any particular co-expression module are gathered in the grey module. For each module, we defined hub genes as the 10% most intra-connected genes. The eigengene of each module, i.e. the first principal component, was calculated and associations between treatment and module eigengenes were determined by a Kruskal-Wallis test (Figure 6B). Pathway enrichment analysis for each module was done with the R package ReactomePA version 1.12.3 (Supplemental Figures 5-7). Networks were visualized with the function cnetplot (Figure 6C, Supplemental Figures 8-9).

Meta-Analysis

We compared our list of AZA-sensitive genes in EL08-1D2 cells to available published RNA-seq datasets derived from primary untreated MDS MSC in comparison to healthy controls (Figure 6D and Supplemental Figure 10).[13,14] To compare mouse with human genes, we used human orthologous genes obtained using biomaRt package (Bioconductor).[15] RNA-seq data from Medyouf et al[13] and Chen et al[14] were analyzed using the same pipeline. Fastq files were cleaned using Cutadapt software version 1.16 and subsequently aligned to ensembl.GRCh37_67 genome assembly using Tophat2 (sensitive parameters), reads count was done using featureCounts software ver. 1.6.0 from subread package. The statistical analysis was performed using DESeq2 (package from Bioconductor release 3.6). Genes considered differentially expressed showed an adjusted p-value < 0.05 (relates to Figure 6D).

Supplemental Results

AZA treatment conditions for stromal cells *in vitro*

To establish adequate AZA treatment conditions for stroma *in vitro*, AZA was freshly prepared in aqueous solution and directly diluted once (day 0) into EL08-1D2 cultures at increasing concentrations. AZA had no effect on EL08-1D2 cell viability and did not induce cell death at any concentration over a period of 6 days (Supplemental Figure 1).

At the standard treatment schedule of 75 mg/m^2, the maximal plasma concentrations of AZA and its active metabolites achieved in MDS patients range between 3 to 11 µM.[16] At 37°C *in vitro*, AZA is stable in solution and induces methylation changes in human cancer cell lines 48-72 hours after its one-time addition to culture, even if the drug is washed out before this time-point, suggesting a prolonged effect.[17,18] A 24-hour treatment with 3 µM AZA has been shown to be sufficient to induce differentiation of mouse BMSC into functional cardiomyocytes.[19] Thus, we were confident to elucidate prolonged AZA specific effects in stromal cells - if there were any - using a dose of 10 µM in further experiments. Doses above 20 µM are not reported as doses primarily associated with epigenetic effects which is why we stayed below this dose in our experiments.

CFU genotyping reveals presence of MDS-associated mutations.

Colonies scored after co-culture of HSPC with untreated or AZA-treated MSC (Figure 3, main manuscript) were additionally genotyped for at least one MDS-associated mutation known to be present at high VAF in the BM sample (Supplemental Figure 2). This analysis showed that with the exception of one colony (colony #5 from untreated MSC, patient sample #32) all genotyped colonies contained the known mutations, indicating they belonged to the dominant MDS clone. Of note, in patient sample #38, not all colonies contained the SRSF2 mutation compared to the TET2 mutation. This is most likely a reflection of the clonal hierarchy in this BM sample (SRSF2 VAF 30%, TET2 VAF 50%, see Table 1). One non-clonal (i.e. healthy) HSPC clone was detected (patient sample #32, colony #5 from untreated MSC).

Competitive co-culture of healthy and MDS HSPC.

To examine the competitive behavior of healthy and MDS HSPC in the same culture and determine the influence of AZA-treated MSC on their proliferation, we performed co-culture experiments over a 4-day period. To distinguish healthy from MDS HSPC from separate BM samples, CD34+ HSPC were stained either with CellTrace Violet or CFSE. CFSE proved to be toxic to MDS HSPC whereas healthy HSPC remained viable during both staining conditions. Thus, Supplemental Figure 3 depicts the results of CD34+ healthy HSPC (stained with CFSE) and CD34+ MDS HSPC (stained with CellTrace Violet) cultured on healthy MSC layers, either untreated or previously treated with AZA. Although sorted in a

1:1 ratio, FACS analysis after sorting revealed a ratio in favor of MDS HSPC over healthy HSPC (63.6 vs 32.3%, respectively). However, after 4 days of culture healthy HSPC had expanded while MDS HSPC were diminished. In line with our CFU results (Figure 3, main manuscript) proliferation of healthy HSPC was enhanced on AZA treated MSC. Survival of MDS HSPC over the 4-day period was severely reduced, which may have been due to toxicity of the cell staining with CellTraceViolet. Although MDS HSPC survived the initial sort and dye concentrations had been previously tested, we must acknowledge that MDS HSPC are more sensitive to such procedures than healthy HSPC. Thus, a definite conclusion regarding proliferation potential of MDS HSPC in competitive co-culture cannot be drawn from these results.

RNA Sequencing

Differential gene expression

Principal component analysis showed that the samples separate according to treatment (data not shown). There are 21,822 genes with non-zero total read count. Out of these genes, there are 1,183 genes significantly differentially expressed (i.e., adjusted p-value < 0.1) between (+)AZA and (-)AZA samples. 362 genes are significantly down-regulated in AZA-treated samples compared to untreated samples, 821 genes are significantly up-regulated (Supplemental Figure 4). The full summary statistics for all genes are provided in Supplemental Table 2 (genes are ordered according to adjusted p-value, significantly differentially expressed genes are marked in bold and yellow). The hierarchical clustering of the 500 most differentially expressed genes shows distinct expression patterns between AZA-treated and untreated samples (relates to Figure 5B).

Weighted gene co-expression analysis (WGNCA)

We identified four distinct gene co-expression modules where each module was given an arbitrary color (grey, blue, brown and turquoise). Based on Reactome terms, 3, 18 and 5 biological pathway terms were enriched in the blue, brown and turquoise module, respectively (p-value < 0.05, adjusted to Benjamini-Hochberg; Supplemental Figures 5-7). Further, we identified the 10% most intra-connected members (hub genes) of each module. The blue module contained 11 hub genes (H2-K1, Glipr1, Ccl5, Ifi27, Rtp4, Oasl2, Ifit1, Gatsl3, Ly6a, Stat1, Oasl1), all of which are significantly regulated under AZA. The brown module contained 10 hub genes (Emb, Crip2, Aig1, Gdf15, Lgmn Col3a1, Nrep, Col1a2, Col1a1, Loxl1), all significantly regulated under AZA. None of the 22 hub genes of the turquoise module were significantly differentially expressed between AZA and controls.

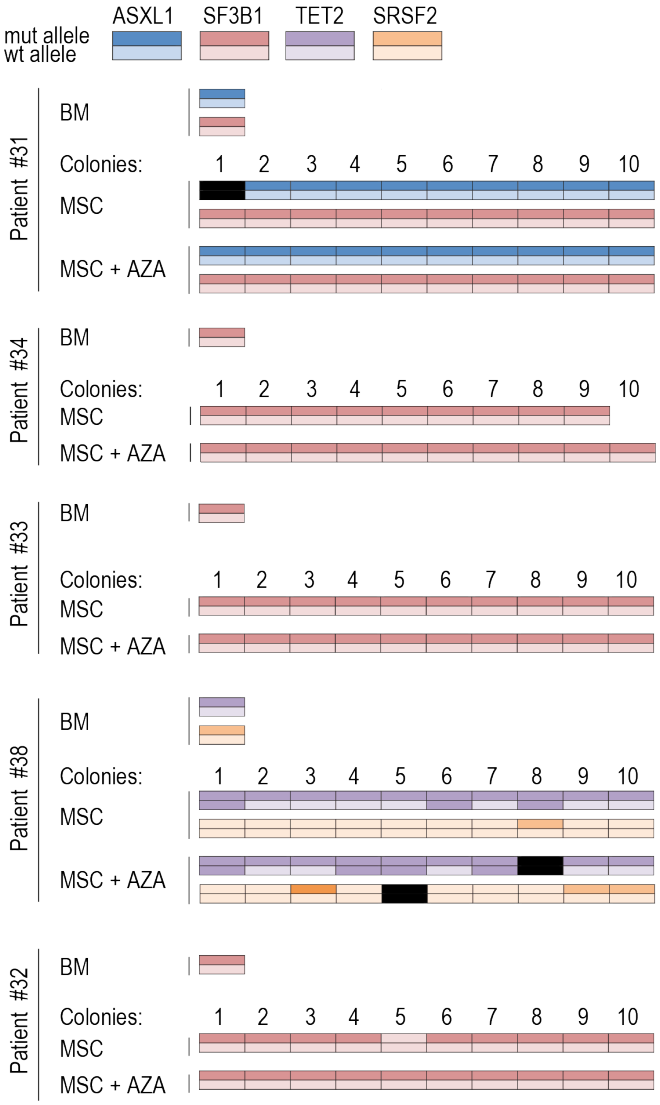
Meta-analysis of MDS specific genes regulated by AZA.

To test in how far AZA differentially affects gene expression in healthy and MDS MSC we compared our list of AZA-sensitive EL08-1D2 genes to available published RNA-seq datasets derived from primary MDS MSC in comparison to healthy controls from Medyouf et al.[13] and Chen et al.[14]. Of note, only 68 genes in total overlapped between the two published datasets. The high variance may be in part explained by differences in isolation of MSC (*ex vivo* expansion by plastic adherence and MSC selective growth media[13] versus prospective FACS sort from fresh BM samples[14]). However, by taking both datasets together we identified a total of 107 MDS MSC specific AZA-sensitive genes (Supplemental Figure 10). GO annotation of the MDS-specific AZA-sensitive genes revealed that genes involved in ECM interaction (especially collagens) and cell adhesion were again most significantly overrepresented (Supplemental Figure 10).

Supplemental Figures

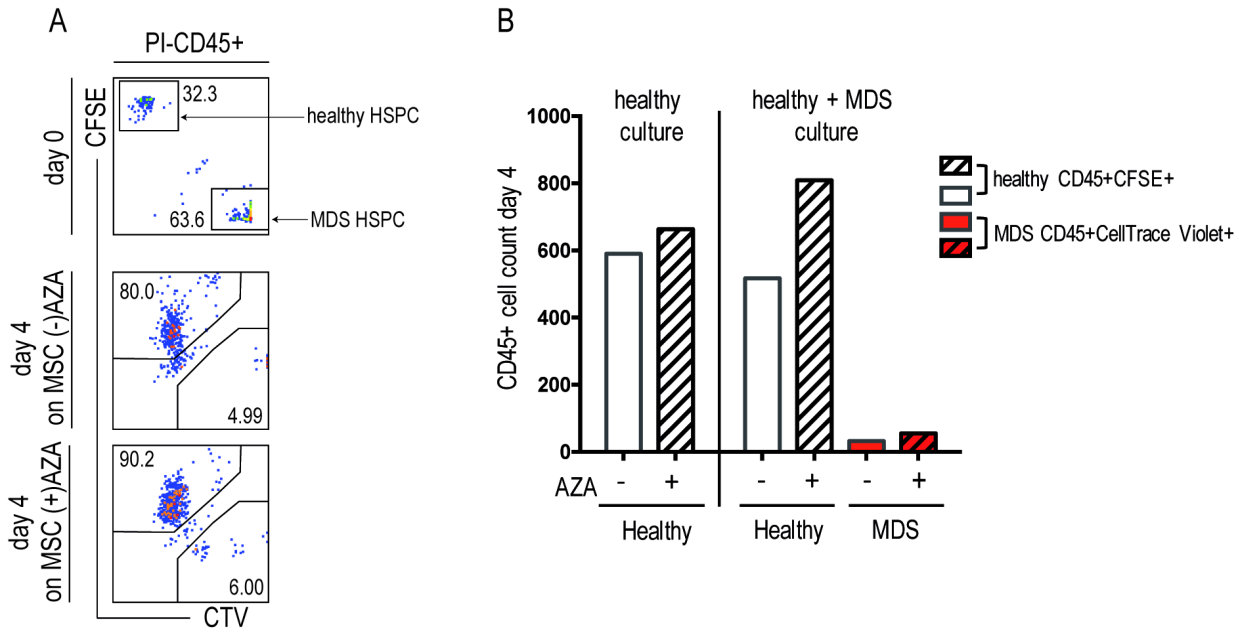
Supplemental Figure 1: Establishment of adequate AZA treatment conditions for *in vitro* stroma cultures

The stromal cell line EL08-1D2 was grown to 80% confluency and treated once on day (D) 0 with indicated AZA concentrations. After 48 hours stromal layers were washed twice with PBS and fresh medium was added. Untreated EL08-1D2 cells were used as controls. Cell viability was assessed by Annexin V/PI flow cytometry after trypsinization at indicated time points. Shown is the mean frequency +/- SEM of viable (Annexin V-/PI-) cells of two independent experiments.



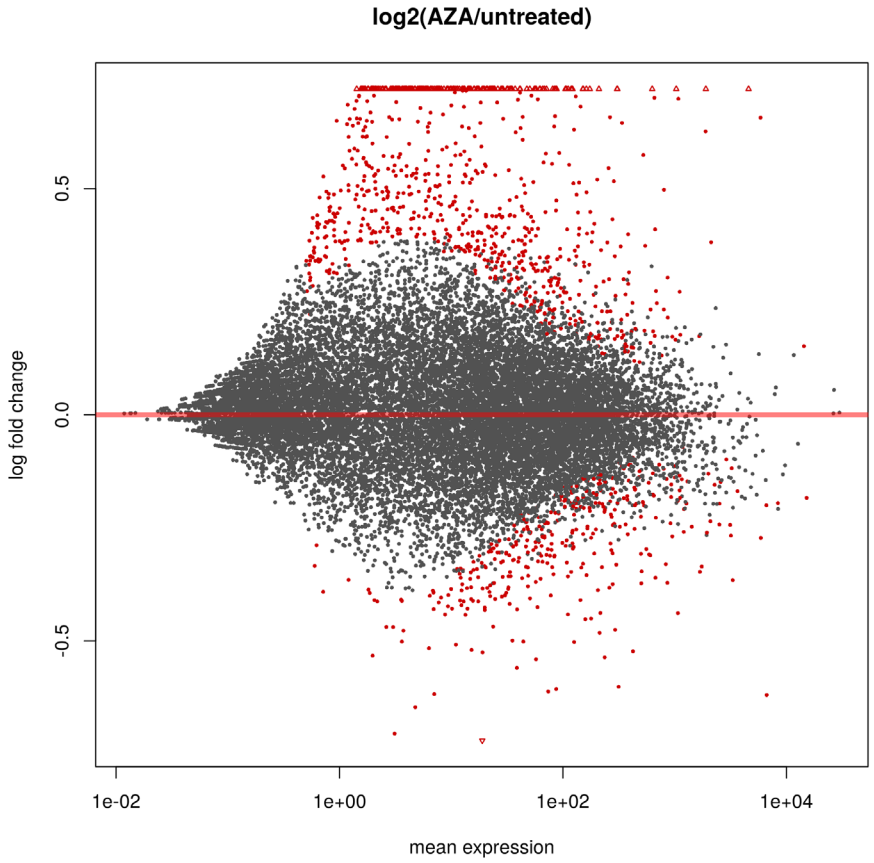
Supplemental Figure 2: CFU genotyping reveals presence of MDS-associated mutations.

After scoring, individual colonies were picked and genotyped for at least one mutation known to be present at high VAF at diagnosis (ASXL1, SF3B1, TET2, SRSF2). Ten colonies were picked and analyzed for each sample condition. Dark colors represent the mutant allele, light colors represent wild-type alleles, black represents technical failure. BM, bone marrow sample at diagnosis; MSC, mesenchymal stromal cell; AZA, azacitidine. Patient sample numbers correspond to Table 1.

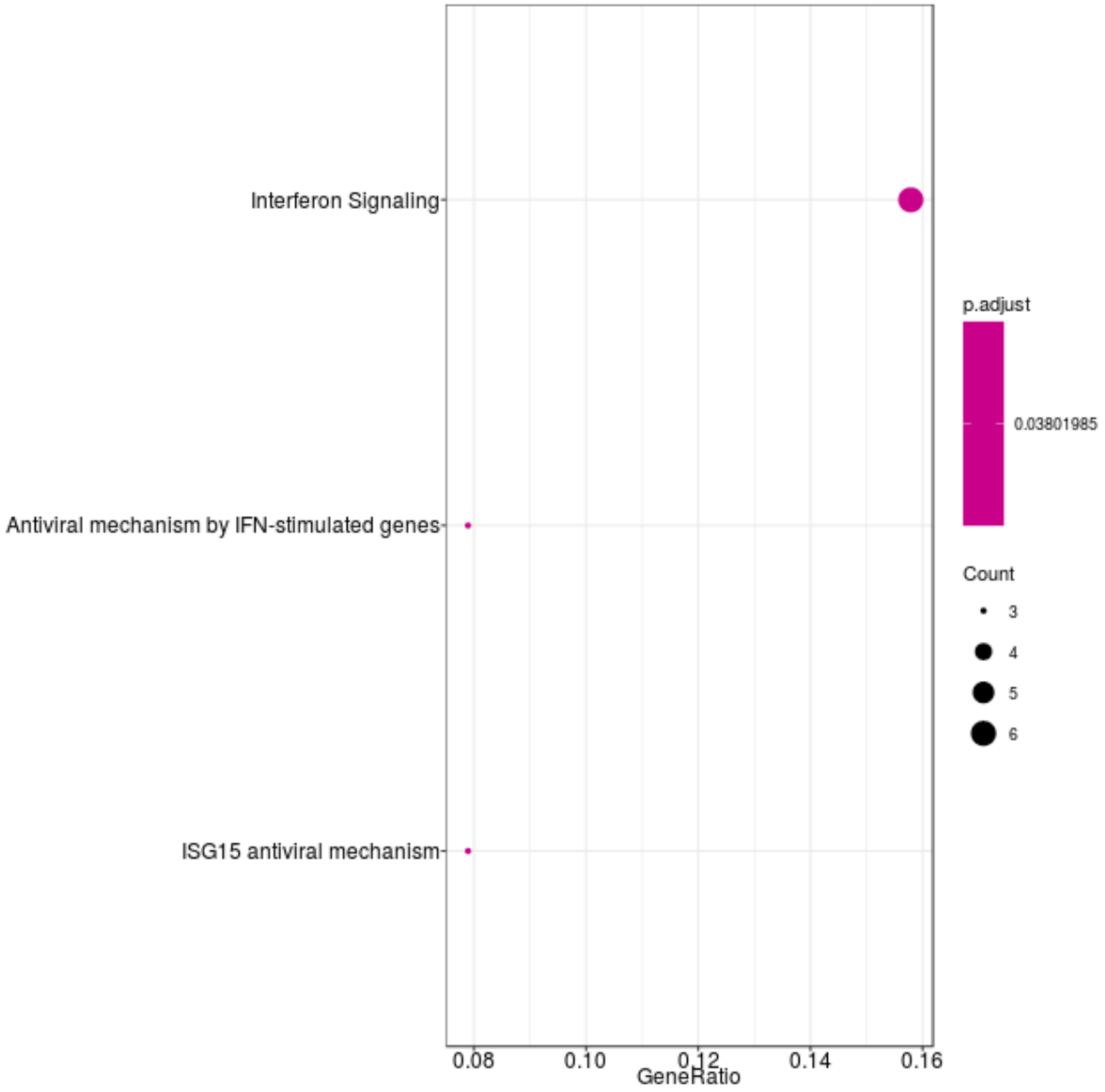


Supplemental Figure 3. Competetive co-culture of healthy and MDS HSPC on MSC tracked by intracellular cell staining.

(A) Cell staining with CFSE and CTV discriminates between healthy and MDS HSPC in the same culture before and after 4 day co-culture on healthy MSC +/- AZA treatment. Percentages of recovered cells are shown (B) Cell numbers of healthy HPSC (CFSE+) cultured alone (left panel) or together with MDS HSPC (CTV+, right panel) after 4 days on untreated or AZA-treated healthy MSC.

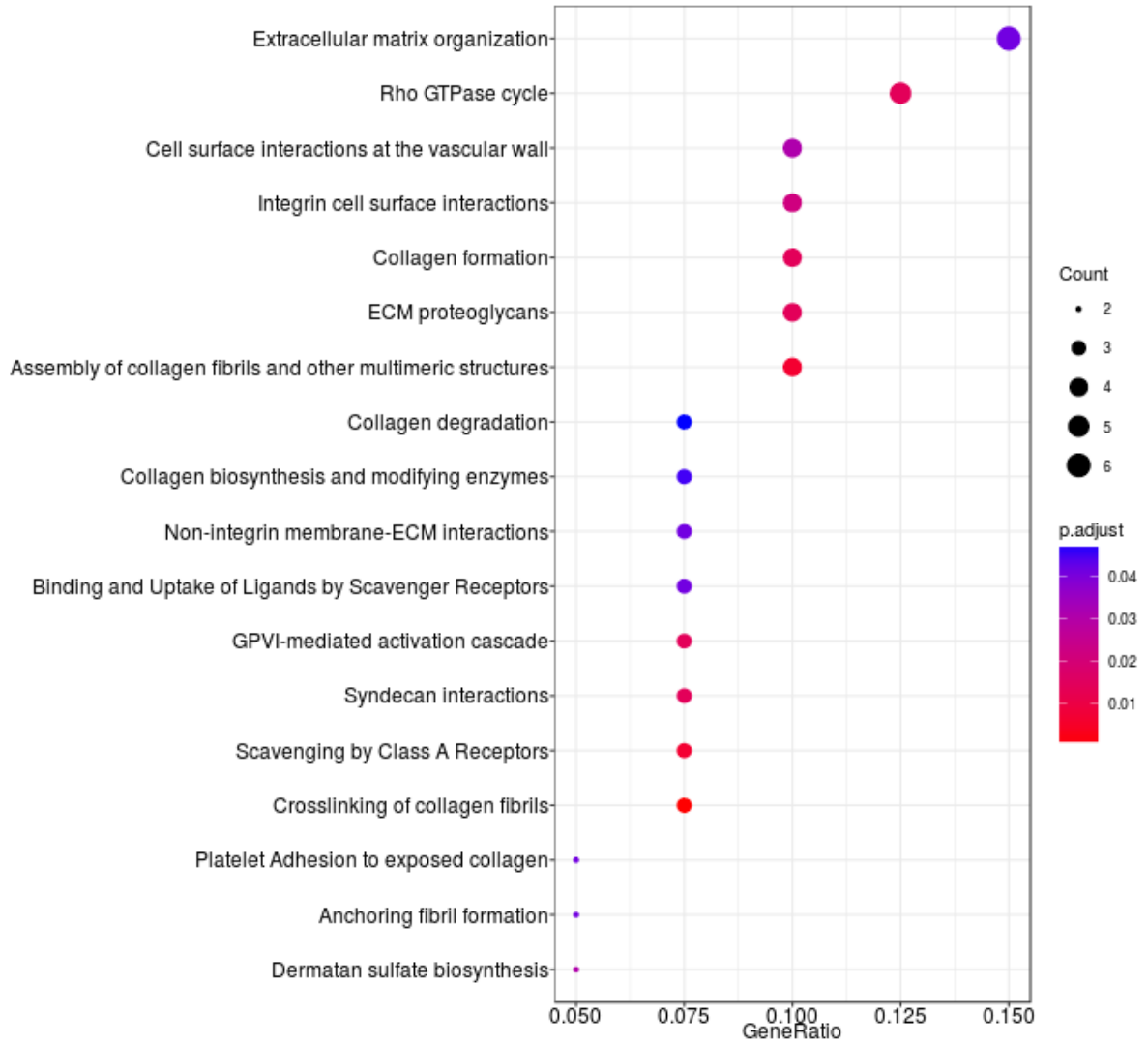


Supplemental Figure 4: Mean expression plotted against log2-fold change (AZA-treated versus untreated samples). Genes where the adjusted p-value (FDR) < 0.1 are marked in red. Genes that have a log2 fold-change higher than 4 are displayed as red triangles. Overall, 821 genes are significantly up-regulated (below red line), 362 genes are significantly down-regulated (above red line) in AZA-treated versus untreated samples (relates to Figure 5)



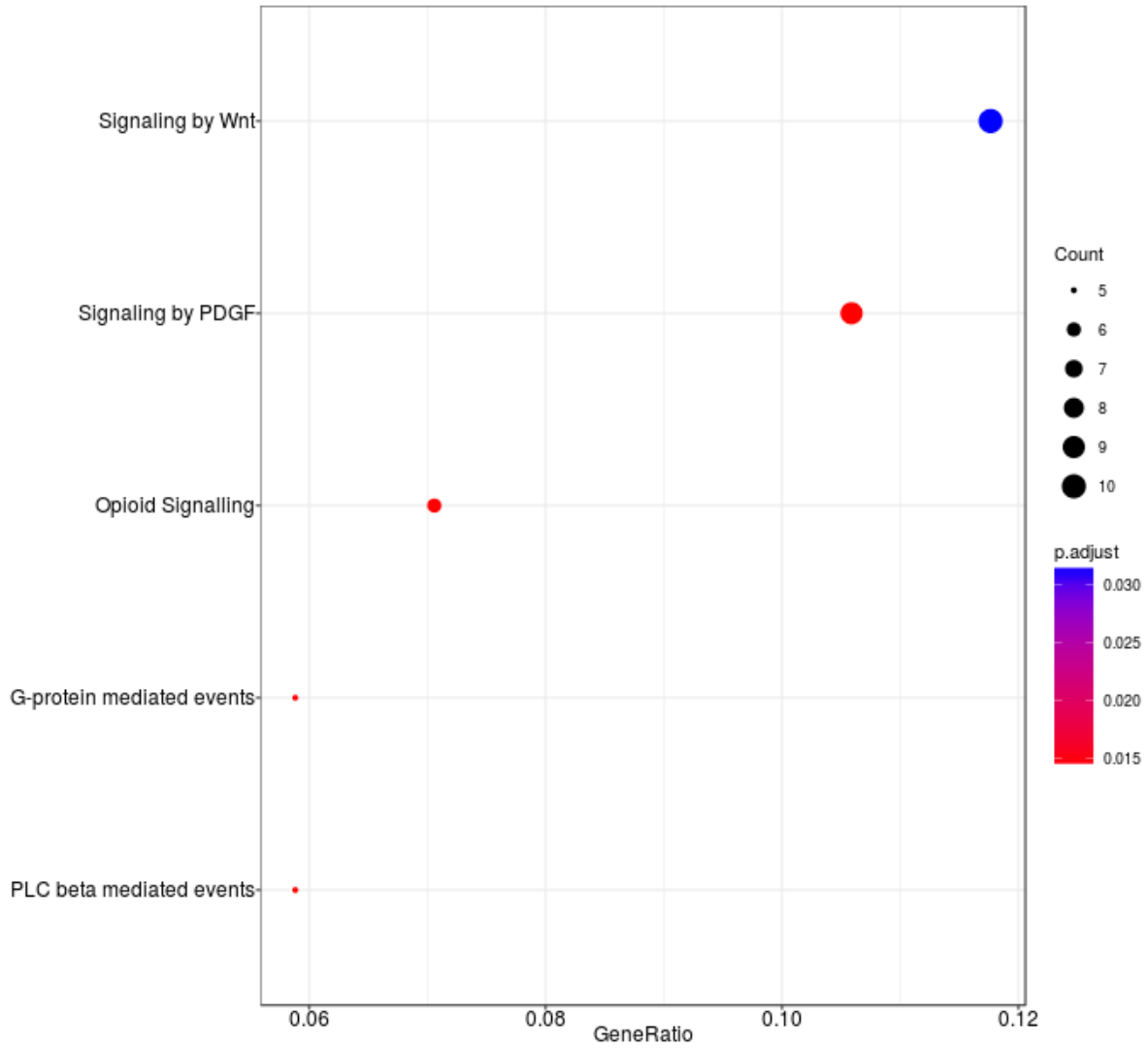
Supplemental Figure 5: Reactome dotplot of the categories that have the highest enrichment in the blue module (IFN signaling pathways).

Pathway enrichment analysis was performed based on the REACTOME pathway database using the R package ReactomePA version 1.12.3. Shown are enriched pathways at a significance level smaller than 0.05, p-value adjusted to Benjamini-Hochberg. The x-axis (GeneRatio) indicates the percentage of genes affected, the point size (Count) indicates their absolute number.



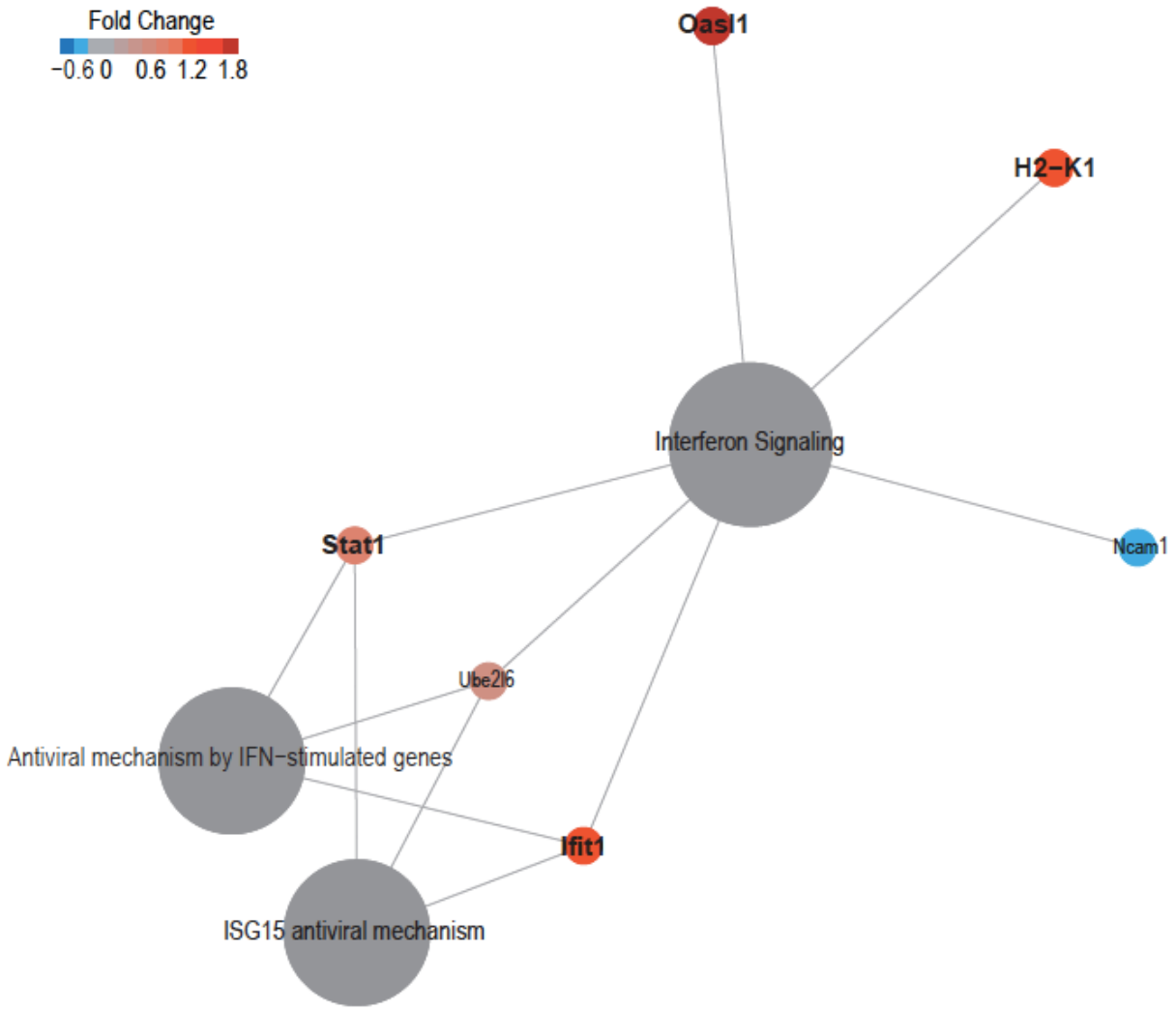
Supplemental Figure 6: Reactome dotplot of the categories that have the highest enrichment in the brown module (ECM receptor interaction pathways).

Pathway enrichment analysis was performed based on the REACTOME pathway database using the R package ReactomePA version 1.12.3. Shown are enriched pathways at a significance level smaller than 0.05, p-value adjusted to Benjamini-Hochberg. The x-axis (GeneRatio) indicates the percentage of genes affected, the point size (Count) indicates their absolute number.



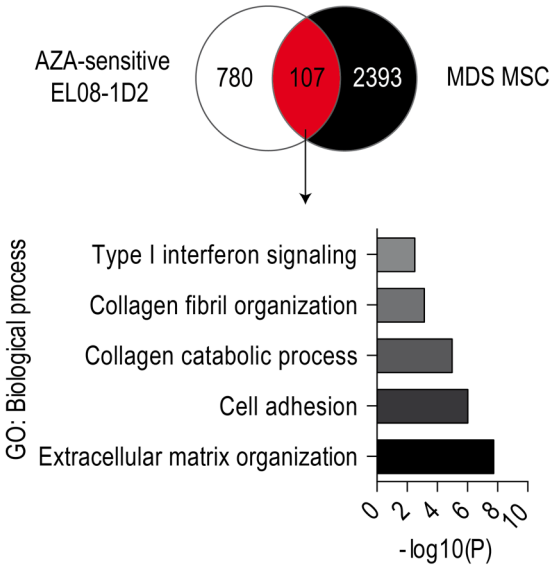
Supplemental Figure 7: Reactome dotplot of the categories that have the highest enrichment in the turquoise module (WNT and PDGF signaling pathways).

Pathway enrichment analysis was performed based on the REACTOME pathway database using the R package ReactomePA version 1.12.3. Shown are enriched pathways at a significance level smaller than 0.05, p-value adjusted to Benjamini-Hochberg. The x-axis (GeneRatio) indicates the percentage of genes affected, the point size (Count) indicates their absolute number.



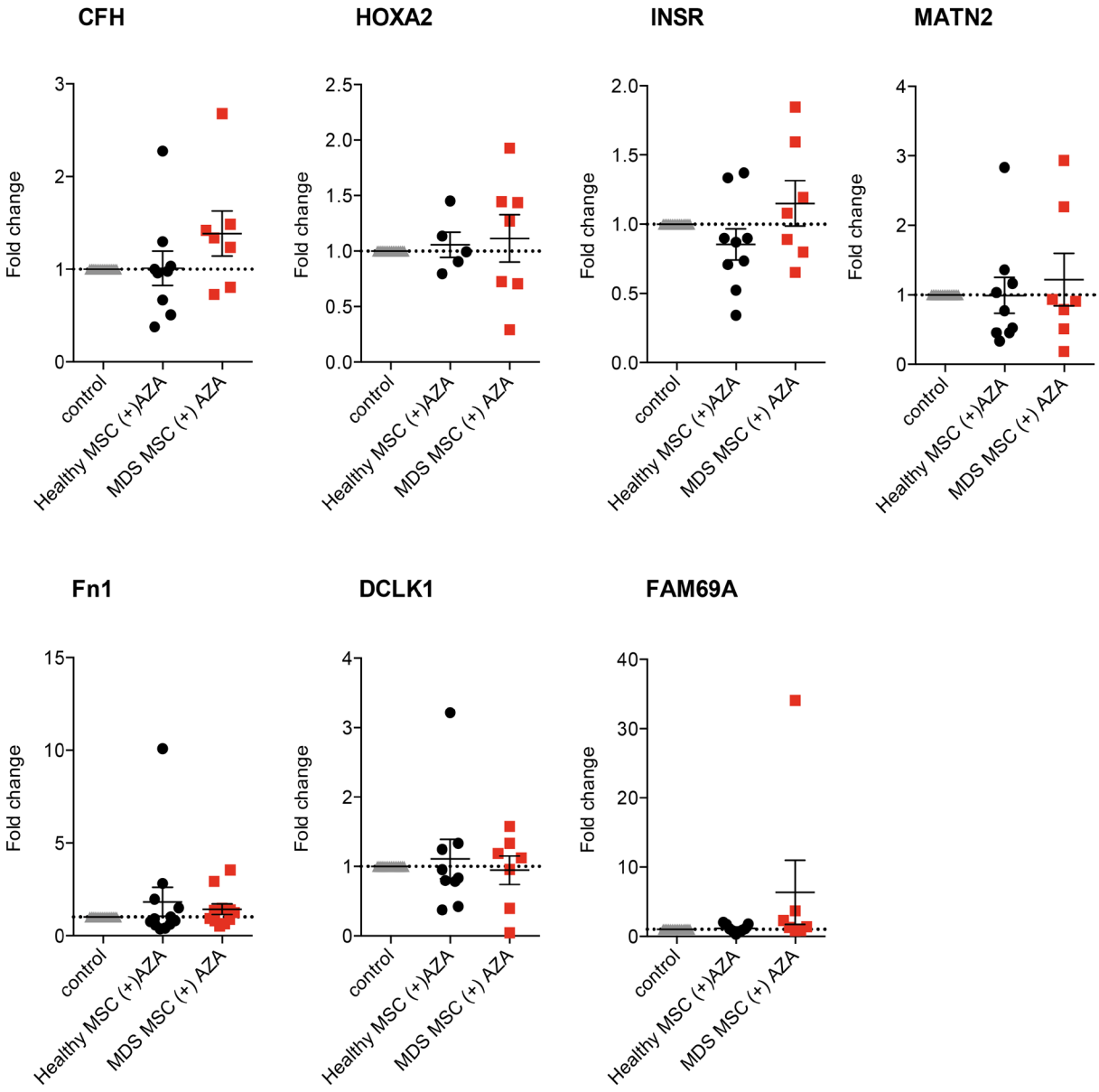
Supplemental Figure 8: Network diagram of the blue module (IFN signaling pathways) showing the 3 significant categories (p-value < 0.05, adjusted to Benjamini-Hochberg) and their associated genes. Color indicates log2fold change in gene expression (AZA vs. untreated).

Supplemental Figure 9: Network diagram of the turquoise module (WNT and PDGF signaling pathways) showing the 5 significant categories (p-value < 0.05, adjusted to Benjamini-Hochberg) and their associated genes. Color indicates log2fold change in gene expression (AZA vs. untreated).



Supplemental Figure 10: Data mining reveals MDS-specific AZA-sensitive genes.

Venn diagram of meta-analysis showing overlap of 107 genes (in red) between AZA-sensitive EL08-1D2 genes (in white) and genes that are "MDS specific", i.e. differentially expressed between MDS and healthy MSC (in black[13,14]). In addition, the five most significant GO terms for MDS-specific AZA-sensitive genes are shown below.



Supplemental Figure 11: Expression of selected HSPC supportive stromal genes from EL08-1D2 cells not found to be differentially regulated in primary MSC following AZA treatment.

Primary healthy and MDS MSC were treated with AZA (10 µM) once or left untreated. After 4 days, MSC were harvested and expression of selected stromal HSPC supportive genes was validated by comparative RT-PCR (from Figure 6D, in red).

Supplemental Tables

Supplemental Table 1: Primary healthy bone marrow samples.

#	Sex	Age (y)	Sample
1	f	41	MSC
2	m	59	MSC
3	m	41	MSC
4	m	83	MSC
5	f	52	MSC
5	m	29	MSC
6	m	67	MSC
7	f	48	MSC
8	m	77	MSC
9	f	18	MSC
10	m	44	MSC
11	f	71	MSC
12	m	75	MSC
13	m	61	MSC
14	m	61	MSC
15	m	49	MSC
16	m	70	MSC
17	f	55	MSC
18	f	62	MSC
19	f	54	MSC
20	f	78	MSC
21	m	57	MSC
22	f	80	MSC
23	f	78	MSC
24	f	60	MSC
25	m	65	MSC
26	m	60	CD34
27	m	45	CD34
28	w	63	CD34
29	m	41	CD34
30	f	30	CD34
31	m	65	CD34
32	f	66	CD34
33	m	67	CD34
34	f	55	CD34
35	f	33	CD34
36	m	65	CD34

Supplemental Table 2: Differentially expressed genes in (+) AZA vs (-) AZA.

See attached Excel file.

Supplemental Table 3: GO_BP_up and downregulated in (+) AZA vs (-) AZA.

See attached Excel file.

Supplemental Table 4: FNEA_up and downregulated genes in (+) AZA vs (-) AZA.

See attached Excel file.

Supplemental Table 5: GSEA for stromal HSPC supportive genes.

See attached Excel file.

Supplemental Table 6. Primer sequences used for RT-PCR.

Gene symbol	Gene name	PrimerBank ID (or literature reference)	Amplicon size	Sequence (5' -> 3')	
Col1a1	collagen, type I, alpha 1	110349771c2	119	FW	GTGCGATGACGTGATCTGTGA
				REV	CGGTGGTTTCTTGGTCGGT
Col1a2	collagen, type I, alpha 2	48762933c3	92	FW	GAGCGGTAACAAGGGTGAGC
				REV	CTTCCCCATTAGGGCCTCTC
Col4a1	collagen, type IV, alpha 1	148536824c3	203	FW	CCAGGGGTCGGAGAGAAAG
				REV	GGTCCTGTGCCTATAACAATTCC
Col4a2	collagen, type IV, alpha 2	116256353c3	218	FW	AAGGGCTTCATCGGAGACC
				REV	CCAGCGTCACCTTTCCACC
Col5a1	collagen, type V, alpha 1	89276750c2	97	FW	TACCCTGCGTCTGCATTTCC
				REV	GCTCGTTGTAGATGGAGACCA
CFH	Homo sapiens complement factor H (CFH), nuclear gene encoding mitochondrial protein, transcript	184172391c1	98	FW	GTGAAGTGTTTACCAGTGACAGC

	variant 2, mRNA				
				REV	AACCGTACTGCTTGTCCAAAA
PRSS3	Homo sapiens protease, serine, 3 (PRSS3), transcript variant 4, mRNA.		174	FW	CCACCCTAAATACAACAGGGAC
				REV	TCAGCACCAAAGCTCAGAGTG
HOXA2	Homo sapiens homeobox A2 (HOXA2), mRNA	37596298c1	90	FW	CCCCTGTCGCTGATACATTTC
				REV	TGGTCTGCTCAAAAGGAGGAG
NID1	homo sapiens nidogen 1 (NID1), mRNA.	115298673c1	90	FW	CGGGGATGACTTCGTCTCTC
				REV	GTGGTGACGTAGACTGCGT
DCLK1	doublecortin and CaM kinase-like 1; doublecortin-like kinase [Homo sapiens].	4758128a3	116	FW	CCAACTGGTCGGTGAACGTC
				REV	CTTGGGCCGAATGAAATCCTTA
FAM69A	Homo sapiens family with sequence similarity 69, member A (FAM69A), mRNA.	165932390b1	115	FW	GCCTGCATGTAACAGCCTTTG
				REV	CACAACACCTGGTAGATTATCCC
INSR	Homo sapiens insulin	119395735c1	90	FW	AAAACGAGGCCCGAAGATTTC

	receptor (INSR), transcript variant 1, mRNA.				
				REV	GAGCCCATAGACCCGGAAG
COL16A1	Homo sapiens collagen, type XVI, alpha 1 (COL16A1), mRNA.	100913219c2	132	FW	CCACCAGAAGACGTGGTATCT
				REV	CAGGACACAAAGTCGCCATC
F3	Homo sapiens coagulation factor III (thromboplastin, tissue factor) (F3), transcript variant 1, mRNA.	296010910c1	102	FW	GGCGCTTCAGGCACTACAA
				REV	TTGATTGACGGGTTTGGGTTC
COX4I1	Homo sapiens cytochrome c oxidase subunit IV isoform 1 (COX4I1), nuclear gene encoding mitochondrial protein, mRNA.	325301085c1	113	FW	CAGGGTATTTAGCCTAGTTGGC
				REV	GCCGATCCATATAAGCTGGGA
MATN2	Homo sapiens matrilin 2 (MATN2), transcript variant 2, mRNA.	62548861c2	78	FW	GTGTCAACACCCATGACTATGC

				REV	CATCAGGACCAATGTCCAAGAA
LOXL1	Homo sapiens lysyl oxidase-like 1 (LOXL1), mRNA.	67782345c2	97	FW	CCACTACGACCTACTGGATGC
				REV	GTTGCCGAAGTCACAGGTG
Col3a1	collagen, type III, alpha 1	110224482c2	83	FW	TTGAAGGAGGATGTTCCCATCT
				REV	ACAGACACATATTTGGCATGGTT
Col5a3	collagen, type V, alpha 3	110735434c1	165	FW	TGACCGGGCATTCAGAATTGG
				REV	CGGGCACCCCTTTCATCAT
Col6a3	collagen, type VI, alpha 3	240255534c1	76	FW	ATGAGGAAACATCGGCACTTG
				REV	GGGCATGAGTTGTAGGAAAGC
Zyx	zyxin	58530844c1	102	FW	TCTCCCGCGATCTCCGTTT
				REV	CCGGAAGGGATTCACTTTGGG
Met	met proto-oncogene	188595715c1	198	FW	AGCAATGGGGAGTGTAAAGAGG
				REV	CCCAGTCTTGTACTCAGCAAC
Fn1	fibronectin 1	47132556c2	106	FW	AGGAAGCCGAGGTTTTAACTG
				REV	AGGACGCTCATAAGTGTCACC
TGFB1	transforming growth factor beta 1	Xie et al[20]	201	FW	CCCAGCATCTGCAAAGCTC
				REV	GTCAATGTACAGCTGCCGCA

References

1. Oostendorp RAJ, Harvey KN, Kusadasi N, et al. Stromal cell lines from mouse aorta-gonads-mesonephros subregions are potent supporters of hematopoietic stem cell activity. *Blood*. 2002;99(4):1183–1189.
2. Parmar A, Marz S, Rushton S, et al. Stromal niche cells protect early leukemic FLT3-ITD+ progenitor cells against first-generation FLT3 tyrosine kinase inhibitors. *Cancer Research*. 2011;71(13):4696–4706.
3. Parekh S, Ziegenhain C, Vieth B, Enard W, Hellmann I. The impact of amplification on differential expression analyses by RNA-seq. *Sci Rep*. 2016;6:1–11.
4. Liu Y, Zhou J, White K. RNA-seq differential expression studies: more sequence or more replication? *Bioinformatics*. 2014;1–4.
5. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.
6. Benjamini Y, Yosef H. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J.R. Statist. Soc. B*. 1995;289–300.
7. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics*. 2009;10:48.
8. Alexeyenko A, Lee W, Pernemalm M, et al. Network enrichment analysis: extension of gene-set enrichment analysis to gene networks. *BMC Bioinformatics*. 2012;13(1):226.
9. Charbord P, Pouget C, Binder H, et al. A systems biology approach for defining the molecular framework of the hematopoietic stem cell niche. *Cell Stem Cell*. 2014;15(3):376–391.
10. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9(1):559.
11. Ravasz E, Somera AL, Mongru DA, Oltvai ZN, Barabási AL. Hierarchical organization of modularity in metabolic networks. *Science*. 2002;297(5586):1551–1555.
12. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. - PubMed - NCBI. *Statistical Applications in Genetics and Molecular Biology*. 2005;4(1):Article17.
13. Medyouf H, Mossner M, Jann J-C, et al. Myelodysplastic cells in patients reprogram mesenchymal stromal cells to establish a transplantable stem cell niche disease unit. *Cell Stem Cell*. 2014;14(6):824–837.
14. Chen S, Zambetti NA, Bindels EMJ, et al. Massive parallel RNA sequencing of highly purified mesenchymal elements in low-risk MDS reveals tissue-context-dependent activation of inflammatory programs. *Leukemia*. 2016;30(9):1938–1942.
15. Durinck S, Moreau Y, Kasprzyk A, et al. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics*. 2005;21(16):3439–3440.
16. Marcucci G, Silverman L, Eller M, Lintz L, Beach CL. Bioavailability of azacitidine subcutaneous versus intravenous in patients with the myelodysplastic syndromes. *J Clin Pharmacol*. 2005;45(5):597–602.
17. Stresemann C, Lyko F. Modes of action of the DNA methyltransferase inhibitors azacytidine and decitabine. *International Journal of Cancer*. 2008;123(1):8–13.
18. Aimiuwu J, Wang H, Chen P, et al. RNA-dependent inhibition of ribonucleotide reductase is a major pathway for 5-azacytidine activity in acute myeloid leukemia. *Blood*. 2012;119(22):5229–5238.
19. Makino S, Fukuda K, Miyoshi S, et al. Cardiomyocytes can be generated from marrow stromal cells in vitro. *J. Clin. Invest.* 1999;103(5):697–705.
20. Xie S, Macedo P, Hew M, et al. Expression of transforming growth factor-β (TGF-β) in chronic idiopathic cough. *Respiratory Research 2009 10:1*. 2009;10(1):40.