# Myelodysplastic syndrome can propagate from the multipotent progenitor compartment

*Kevin Rouault-Pierre,[1*] Alexander E. Smith,[2,3*] Syed A. Mian,[1,3] Irene Pizzitola,[1] Austin G. Kulasekararaj,[2,3] Ghulam J. Mufti[2,3**] and Dominique Bonnet[1**]*

*[1]Haematopoietic Stem Cell Laboratory, The Francis Crick Institute and [2]Kings College Hospital, Department of Haematology and [3]King's College London School of Medicine, Department of Haematological Medicine, London, UK*

**Authors contributed equally; **corresponding authors contributed equally.*

*Correspondence: dominique.bonnet@crick.ac.uk or ghulam.mufti@kcl.ac.uk*
*doi:10.3324/haematol.2016.152520*

**Isolation of $CD34^+$ cells, HSCs, MPPs, MLPs, CMPs, GMPs and MEPs**

Mononuclear cells (MNCs) from one MDS patient sample at diagnostic were received from King's College London Haemato-Oncology Tissue Bank after obtaining written informed consent for the samples to be used for all the experiments in accordance to King's College London Haemato-Oncology Tissue Bank research ethics protocol (08/H0906/94). Mononuclear cells (MNCs) were isolated from the bone marrow cells by centrifugation using Ficoll-Paque$^{TM}$ PLUS (GE Healthcare Life Sciences, Buckinghamshire, UK). The cells were processed within 24 hours following collection using an Easysep Human CD34 positive selection kit and Easysep magnet (StemCell Technologies, Vancouver, Canada) according to the manufacturer's instructions to enrich $CD34^+$ cells from patient samples. Mesenchymal Stem Cells (MSCs) were isolated from the $CD34^-$ fraction during the $CD34^+$ cell selection. Cells were seeded at a concentration of $1.10^6/cm^2$, in DMEM low glucose (Dulbecco's Modified Eagle Medium, Life Technologies, Paisley, UK) supplemented with FBS (Fetal Bovine Serum, Mesenchymal Stem Cell Qualified, life technologies, Paisley, UK). Culture media was replaced three days later, and cells were frozen at passage 2 maximum.

$CD34^+$ cells were stained with the following antibodies: Lineage Cocktail FiTc, CD123 PE, CD34 PerCP (BD Biosciences, Oxford, UK) and CD38 APC-Alexa eFluor 780, CD45RA PE-Cy7, CD90 APC (eBioscience, Hatfield, UK) (Supplementary Figure 1). DAPI was added to the cell suspension prior sorting to exclude dead cells. Cells were sorted on a BD FACS Aria SORP (San Jose, CA) operating in 4-way purity sort mode, and collected into 1.5 ml microfuge tubes.

**Colony-forming Cell and Long term culture (LTC)**

Five hundred cells from the bone marrow $CD34^+$ enriched fraction were plated in triplicate in 0.5 mL of MethoCult H4434 (StemCell Technologies, Vancouver, Canada) and cultured at 3% oxygen for 2 weeks. At day 14 of culture, the numbers of colonies were scored. Single cell colony were picked, harvested and washed twice with PBS.

LTC assay was performed by plating $10^3$ $CD34^+$ bone marrow cells in quadruplets on irradiated MS-5 murine stromal cells and cultured in Myelocult H5100 (StemCell Technologies, Vancouver, Canada) in the presence of cytokines (20 ng/ml G-CSF, 20ng/ml IL-3, and 20 ng/ml TPO from PeproTech, London, UK). After 5 weeks, live cells were stained and sorted on human $CD45^+$ phenotype using the FACS Aria SORP (BD Biosciences, Oxford, UK). Sorted cells were washed in PBS, harvested and stored as cell pellet and/or resuspended in Karnoy's solution in order to later perform FISH or sequencing analysis.

**Xenotransplant Assays**

NOD/SCID/IL2r$\gamma^{-/-}$ (NSG) mice and NOD/SCID/IL2r$\gamma^{-/-}$-3/GM/SF (NSG-S) were a kind gift of Dr Leonard Shultz (The Jackson Laboratory). All animal experiments were performed in accordance to Home Office and CRUK guidelines. Female NSG and NSG-S mice used in this study were 8 to 12 weeks old at start of experiments. Prior to transplantation, mice received a sub-lethal dose of radiation (375 cGy) from a cesium-137 source. Direct intra–bone marrow injection was performed in the tibia with 2 to $4x10^5$ bone marrow $CD34^+$ cells from patients or with cell numbers described Figure 2A when fractions were injected.

Engraftment was assessed over time by intra-tibia aspiration under isoflurane anaesthesia and the bone marrow was immuno-phenotyped by the presence of $mCD45^-$ PerCP, $hCD45^+$ PE-Cy7, $CD33^+$ PE, $CD19^+$ FITC and $CD3^+$ APC (using anti-mouse and anti-human antibodies

from BD Biosciences, Oxford, UK) cell populations. After 12-32 weeks, mice were terminated and cells were harvested from pooled bone marrow (femurs, tibias, pelvis). Live cells were stained and sorted on $hCD45^+$ phenotype using FACS Aria SORP (BD Biosciences, Oxford, UK). Sorted cells were washed in PBS and harvested in order to later perform genomic analysis and secondary transplantation.

**Mutation screening overview**

Whole exome and/or targeted amplicon sequencing was performed on samples as indicated in Supplementary Table 3. Where whole genome amplified (WGA) material was used, identified mutations and their allele burdens were confirmed, the remaining non-WGA material were confirmed when possible (Supplementary Table 3). Mutant allele burdens measurements presented in this paper are representative of at least two independent experiments and combine of both WGA and non-WGA samples where relevant.

**Whole-genome Amplification**

Whole Genome amplification (WGA) was performed on samples as specified in Supplementary Table 3 prior to exome or targeted amplicon sequencing. Reagents used were GenomePlex Single Cell Whole genome amplification (WGA4, Sigma-Aldrich Co, LLC), following manufacturers instructions, along with a no cell reaction as a negative control and a reaction of human tissue genomic DNA as positive control.

**Whole-exome Sequencing and Data Analysis**

Whole exome sequencing (WES) was performed on primary $CD34^+$ bone marrow cells and compared to paired skin samples to get a baseline readout of acquired mutations for

individual patients. WES was also used subsequently on selected samples following culturing and mouse engraftment as indicated in Supplementary Table 3.

gDNA (non-WGA, 100-500ng) or WGA DNA (500ng) was processed for exome sequencing (Agilent V4) and sequenced on the Illumina HiSeq2000 (Paired end V3 chemistry) according to manufacturer's instructions. Base calling was generated by the Illumina RTA software. Demultiplexing and conversion of basecalls to fastq was performed by Casava version 1.8.2, filtering out poor quality reads. Alignment, realignment and recalibration was performed using Burrows-Wheelers aligner (BWA) (1) and GATK (2) respectively, according to Broad Institute best practices for individual samples. VarScan 1.3.4 (3, 4) was subsequently used to call somatic variants on pileup files (Samtools) (5) for paired tumour and normal (skin or $CD3^+$ T-cells) tissue. Reads with alignment score of <10, base quality score <15 and fold strand bias >10 were excluded. All resulting variants were passed through ANOVAR (6) using refseq annotation and variants deemed to cause protein changes and not found in dbSNP135, esp5400 and 6500, and 1000 genomes databases at >0.001 population allele frequency were passed. Genomic duplicated regions were also filtered out unless associated with known mutations. Somatic mutations were subsequently passed when somatic p-value (VarScan, Fisher's exact T-test) was <0.01, had ≥3 reads supporting mutation, were present in paired normal tissue at less than 20% of the paired tumour tissue and had a allele burden of >10%. For isolated engrafted $CD45^+CD33^+$ cells where WGA material was used, candidate mutations were only passed if they were present in more than 1 engrafted sample (>1 mouse engraftment experiment) or found concurrently in engraftment experiments and LTC derived sample or primary $CD34^+$ cells for the same patient.

**Mapping Loss of Heterozygosity (LOH) Across Chromosomes by WES**

LOH was mapped from WES data according to the following criteria: Variants were called as described above. Variants found in 1000 genome project database with a population allele frequency less than 0.001 were selected and filtered out for known genomic duplicated regions, read depth <20 and somatic status according to VarScan. The filtered sets of variants were then plotted graphically for the required chromosome.

**Sequencing validation and follow-up targeted amplicon screening**

All somatic mutations were confirmed by sequencing from independent PCR reactions. PCR amplicons were generated using primers for candidate gene variants that designed by using Primer 3 program (7) with default settings. The acquired nature of the mutations was also confirmed by their absence in paired constitutional DNA from appropriate amplicon sequencing.

Here, PCR amplicons were normalized, mixed in batches and converted into sequencing libraries utilizing transposon-based Illumina NexteraXT reagents. These libraries were sequenced on the Illumina MiSeq utilizing version 2 chemistry, (150-250 paired-end reads with dual indexes). The same methodology was used for follow-up targeted screening as indicated in Supplementary Table 3. For all targeted sequencing experiments, >500x depth (Q-score > q25) was achieved and for many this was much higher (avg across all amplicons > 1500x). For single cell colony experiments, mutations were called when the mutant allele burden was greater than 25%. Reads were only considered for mutation calling when the q-score of the base in question was >Q25.

## References


1. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-60.
2. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research. 2010;20(9):1297-303.
3. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome research. 2012;22(3):568-76.
4. Koboldt DC, Chen K, Wylie T, Larson DE, McLellan MD, Mardis ER, et al. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. Bioinformatics. 2009;25(17):2283-5.
5. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9.
6. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic acids research. 2010;38(16):e164.
7. Koressaar T, Remm M. Enhancements and modifications of primer design program Primer3. Bioinformatics. 2007;23(10):1289-91.

## SUPPLEMENTARY TABLES & FIGURES LEGEND

### Supplementary table 1

Clinical data of the patient are presented in the table.

| Sample | WHO | Cytogenetics | Hb/Ne/Plt | BM blasts | % BM $CD34^+$ cells | IPSS | Therapy | Progression to AML | Mutations |
|---|---|---|---|---|---|---|---|---|---|
| **MDS 46year Male** | MDS/MPN-U | 47,XY,+8 [30] | 13.2/6.34/4.5 | 8% | 5.7 | Int-1 | 5-Azacytidine DA X2 HSCT | Yes (post HSCT) | EZH2, PHACTR3, RUNX1, WDR64, ANGPTL5, ASXL1 |

### Supplementary Table 2

Genetics of the patient at diagnosis are presented in the table.

| Gene | Class | Locus | Description | $CD34^+$ MAB |
|---|---|---|---|---|
| **EZH2** | Missense | chr7:148526910 | G to T : P93T | 97% (LOH chr 7) |
| **PHACTR3** | Missense | chr20:58330363 | C to T: T159M | 49.00% |
| **RUNX1** | Frameshift | chr21:36171652 | Ins T: P277fs | 44.00% |
| **WDR64** | Missense | chr1:241901671 | C to T: R391W | 51.30% |
| **ANGPTL5** | Missense | chr11:101775568 | G to T: P139L | 54.00% |
| **ASXL1** | Frameshift | chr 20:31022438 | Del GGAGG : G642fsX13 | 47.00% |

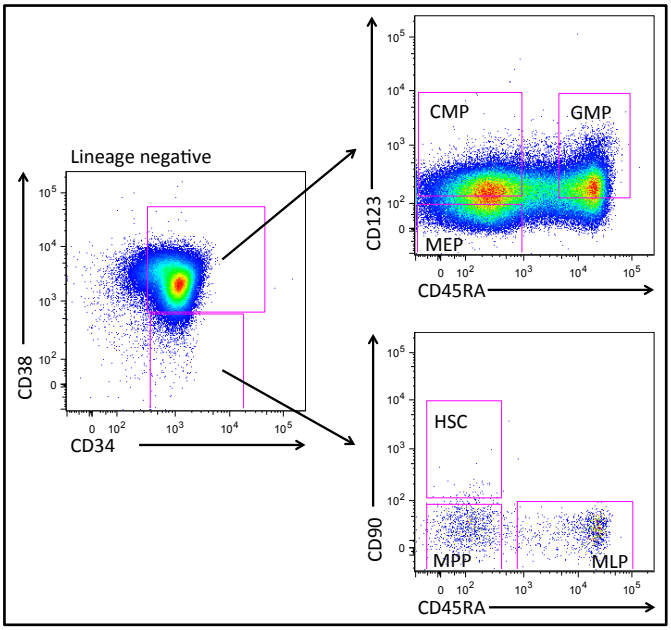
**Supplementary Table 3**

Inventory of genomic analysis performed on the different samples.

| Sample | | Exome (genomic DNA) | Exome (WGA DNA) | Targeted mutation screen/confirmation (genomic DNA) | Targeted mutation screen/confirmation (WGA DNA) |
|---|---|---|---|---|---|
| $CD34^+$ primary cells | | ✓ | – | ✓ | – |
| CFU pool | | – | – | ✓ | – |
| LTC pool | | ✓ | – | ✓ | – |
| Individual CFUs from day0 sample | | – | – | ✓ | – |
| Individual CFUs from $I^{IY}$ mouse | | – | – | – | ✓ |
| NSG mice injected with MDS-$CD34^+$ (sorted on $CD33^+$ at sacrifice) | Mouse1 | – | ✓ | ✓ | ✓ |
| | Mouse2 | – | ✓ | ✓ | ✓ |
| | Mouse3 | – | ✓ | ✓ | ✓ |
| | Mouse4 | – | ✓ | – | ✓ |
| NSG mice Co-injected with MDS-$CD34^+$/ MSCs (sorted on $CD33^+$ at sacrifice) | Mouse1 | – | ✓ | – | ✓ |
| | Mouse2 | – | ✓ | – | ✓ |
| NSG-S mice injected with MDS-$CD34^+$ (sorted on $CD33^+$ at sacrifice) | Mouse1 | – | – | ✓ | – |
| | Mouse2 | – | – | ✓ | – |
| | Mouse3 | – | – | ✓ | – |
| | Mouse4* | – | – | – | – |
| NSG mouse injected with MDS-MPP | | – | – | ✓ | – |
| NSG-S mouse injected with MDS-MPP | | – | – | ✓ | – |
| HSC and progenitor cell pools | | – | – | ✓ | ✓ |

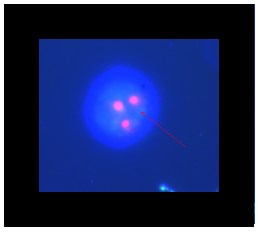
**Supplementary Figure 1**

A) Facs plot showing the sorting strategy to isolate cells from different bone marrow compartments at day 0.

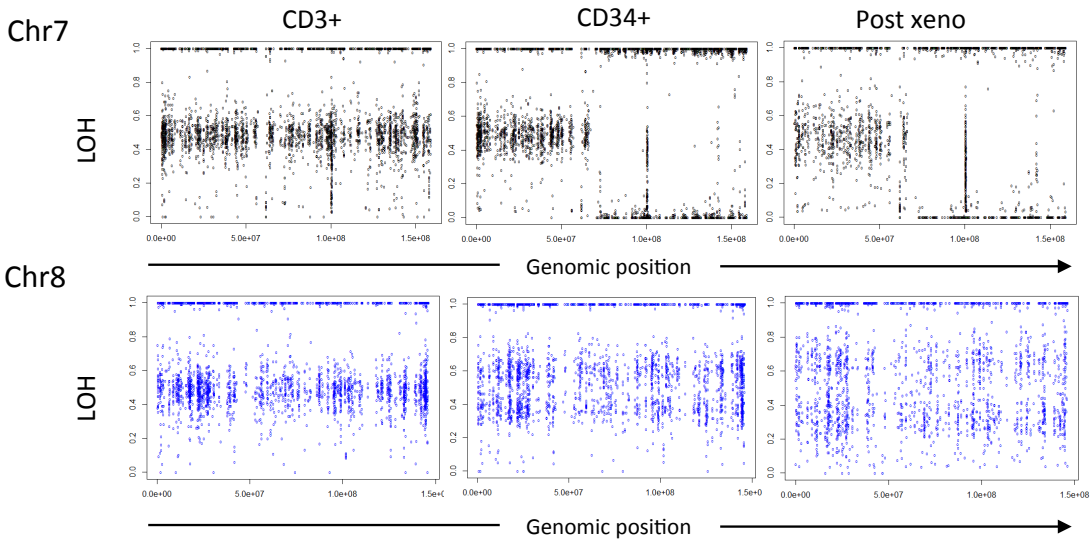


**Supplementary Figure 2**

(A) Top left hand side panel showing the FISH analysis of the trisomy 8 present in cells obtained from patient bone marrow cells expanded in LTC. Bottom panel showing the WES SNP data depicting the LOH at chromosome 7 and chromosome 8 in day 0 $CD3^+$ cells, day 0 $CD34^+$ cells and post xenograft cells.



| Sample | Number of CD34+ cells | 5wks LTC fold expansion | 5wks LTC +3GT fold expansion | FISH LTC +3GT |
|---|---|---|---|---|
| **MDS** | 4000 | 7.5 | 55 | 70% (130T8/56normal) |



**Supplementary Figure 3**

A) Kinetic of engraftment in NSG, NSG co-injected with MSC or in NSG-SGM3 mice, injected with CD34 from MDS patient. B) Tracking of the malignant clone in NSG-SGM3 mice engrafted with CD34 from MDS patient. C) Left panel: Multilineage engraftment and kinetic of engraftment in an NSG-SGM3 mouse. Mutations were assessed at 8 weeks by bone marrow puncture and at sacrifice. Mutations were not detectable at sacrifice. Right panel: Multilineage engraftment and kinetic of engraftment in an NSG-SGM3 mouse. Mutations were not detectable at sacrifice. D) Multilineage engraftment and kinetic of engraftment in an NSG-SGM3 mouse. Mutations were not detectable at sacrifice.