



Article

# Single-Nucleotide Variations, Insertions/Deletions and Copy Number Variations in Myelodysplastic Syndrome during Disease Progression Revealed by a Single-Cell DNA Sequencing Platform

Paul Lee <sup>1,†</sup> , Rita Yim <sup>1,†</sup> , Sin-Hang Fung <sup>2,†</sup> , Kai-Kei Miu <sup>2,†</sup> , Zhangting Wang <sup>2</sup> , Ka-Chun Wu <sup>3</sup>, Lester Au <sup>1</sup>, Garret Man-Kit Leung <sup>1</sup>, Victor Ho-Fun Lee <sup>3</sup> and Harinder Gill <sup>1,\*</sup>

<sup>1</sup> Department of Medicine, School of Clinical Medicine, LKS Faculty of Medicine, University of Hong Kong, Hong Kong, China; pl85@hku.hk (P.L.); ritayim@hku.hk (R.Y.); lesterau@hku.hk (L.A.); garretleung@gmail.com (G.M.-K.L.)

<sup>2</sup> School of Biomedical Sciences, Faculty of Medicine, Chinese University of Hong Kong, Hong Kong, China; shfung@link.cuhk.edu.hk (S.-H.F.); kelvinmiu@cuhk.edu.hk (K.-K.M.); wangzhangting@cuhk.edu.hk (Z.W.)

<sup>3</sup> Department of Clinical Oncology, School of Clinical Medicine, LKS Faculty of Medicine, University of Hong Kong, Hong Kong, China; eymonwu@hku.hk (K.-C.W.); vhflee@hku.hk (V.H.-F.L.)

\* Correspondence: gillhsh@hku.hk; Tel.: +852-2255-4542; Fax: +852-2816-2863

† These authors contributed equally to this work.



**Citation:** Lee, P.; Yim, R.; Fung, S.-H.; Miu, K.-K.; Wang, Z.; Wu, K.-C.; Au, L.; Leung, G.M.-K.; Lee, V.H.-F.; Gill, H. Single-Nucleotide Variations, Insertions/Deletions and Copy Number Variations in Myelodysplastic Syndrome during Disease Progression Revealed by a Single-Cell DNA Sequencing Platform. *Int. J. Mol. Sci.* **2022**, *23*, 4647. <https://doi.org/10.3390/ijms23094647>

Academic Editor: James K. Bashkin

Received: 25 March 2022

Accepted: 21 April 2022

Published: 22 April 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Simple Summary:** Myelodysplastic syndrome (MDS) is a myeloid neoplasm associated with complex clonal architecture. The application of single-cell sequencing is capable of revealing the clonal dynamics of MDS during disease progression and treatment resistance. This has advantages over bulk-tumor sequencing which is limited by its resolution. In this study, we evaluated two patients with MDS for the clonal dynamics of pathogenic mutations at the single-cell level of disease progression and resistance to hypomethylating agents (HMAs). There were two key observations. First, changes in the clonal heterogeneity of the pathogenic *FLT3-ITD*, *IDH2*, *EZH2*, or *GATA2* mutations was associated with disease progression and resistance to HMA. Secondly, disease progression and resistance to HMA was accompanied by the acquisition of copy number variations of *DNMT3A*, *TET2*, and *GATA2*.

**Abstract:** Myelodysplastic syndrome (MDS) is a clonal myeloid neoplasm characterized by ineffective hematopoiesis, cytopenia, dysplasia, and clonal instability, leading to leukemic transformation. Hypomethylating agents are the mainstay of treatment in higher-risk MDS. However, treatment resistance and disease transformation into acute myeloid leukemia (AML) is observed in the majority of patients and is indicative of a dismal outcome. The residual cell clones resistant to therapy or cell clones acquiring new genetic aberrations are two of the key events responsible for drug resistance. Bulk tumor sequencing often fails to detect these rare subclones that confer resistance to therapy. In this study, we employed a single-cell DNA (sc-DNA) sequencing approach to study the clonal heterogeneity and clonal evolution in two MDS patients refractory to HMA. In both patients, different single nucleotide variations (SNVs) or insertions and deletions (INDELs) were detected with bulk tumor sequencing. Rare cell clones with mutations that are undetectable by bulk tumor sequencing were detected by sc-DNA sequencing. In addition to SNVs and short INDELs, this study also revealed the presence of a clonal copy number loss of *DNMT3A*, *TET2*, and *GATA2* as standalone events or in association with the small SNVs or INDELs detected during HMA resistance and disease progression.

**Keywords:** myelodysplastic syndrome; hypomethylating agents; single-cell sequencing; clonal evolution

## 1. Introduction

Myelodysplastic syndrome (MDS) is a myeloid neoplasm associated with complex clonal architecture and is characterized by ineffective hematopoiesis, cytopenia, dysplasia, and clonal instability [1,2]. Common chromosomal aberrations include loss of chromosomes 5q and 7 and gains of chromosomes 8, 19, and 21. Mutations of *SF3B1*, *TET2*, *ASXL1*, *SRSF2*, and *DNMT3A* are reported in more than 10% of all MDS [3,4]. These genetic lesions contribute to the clinical presentations, risk of progression to acute myeloid leukemia (AML), treatment response, and survival [4–7]. Prognosis is conventionally determined by the International Prognostic Scoring System-Revised (IPSS-R) and more recently by personalized prognostic models incorporating somatic mutations [4,6–8]. Resistance to the hypomethylating agents (HMAs) azacitidine (AZA) and decitabine (DEC) in higher-risked MDS (HR-MDS) represents a major unmet need [3]. Resistance to HMA indicates a poor prognosis with a median overall survival (OS) of only 4.3 months in HR-MDS [9].

Somatic mutations and epigenetic alterations contribute to HMA resistance and clonal evolution in response to treatment remains a challenge in the management of MDS [10–12]. The use of next-generation sequencing (NGS) may help to predict the response to HMA and guide the clinical use of novel targeted therapy such as *FLT3*, *BCL2*, *IDH1*, and *IDH2* inhibitors [13,14]. While bulk tumor sequencing may help to detect somatic mutations that predict treatment outcomes, the clonal heterogeneity characteristic of MDS remains a challenge for the detection of somatic mutations in small subclones that may confer treatment resistance. In small studies, the single-cell RNA (sc-RNA) sequencing of CD34+ cell has identified sub-populations with distinct gene expression profiles and demonstrated myeloid-biased hematopoiesis in patients with MDS [15,16]. This single cell-based approach increases the sensitivity and facilitates sub-clonal analysis. Nevertheless, CNV detection at the RNA level is still challenging as the bioinformatic inference of CNV lacks accuracy.

To study clonal evolution and changes in clonal architecture during the resistance to HMA at the single-cell level, we performed DNA sequencing in two patients with MDS who acquired resistance to HMA. In addition to studying SNV, INDEL, and CNV serially at single-cell levels, we also evaluated the sensitivity of single-cell sequencing technology in detecting rare sub-population of cells harboring pathogenic mutations.

## 2. Materials and Methods

### 2.1. Patients

Bone marrow samples of two patients with MDS treated with HMA were collected serially during response and subsequent resistance. Mononuclear cells were isolated by density separation prior to sequencing. This study was approved by the Institutional Review Board (IRB) of the University of Hong Kong and Hong Kong West Cluster and written informed consent was obtained.

Patient 1 was a 53-year-old man diagnosed with MDS with excess blasts-2 (MDS-EB-2). This patient was treated with AZA (100 mg/m<sup>2</sup>/day on days 1–7 per cycle) every 28 days. After 14 cycles of AZA, the patient with progressed secondary AML was treated with induction with daunorubicin (90 mg/m<sup>2</sup>/d on days 1–3), cytarabine (100 mg/m<sup>2</sup>/day on days 1–7), and the FLT3 inhibitor midostaurin (50 mg twice per day on days 8–21). He subsequently received one cycle of a high dose of cytarabine (3 g/m<sup>2</sup>/dose every 12 h for 4 doses) and midostaurin (50 mg twice per day on days 8–21). He relapsed again and was refractory to treatment with venetoclax (400 mg/day on days 1–28)—DEC (20 mg/m<sup>2</sup>/day on days 1–5) and Gilteritinib (120 mg/day) monotherapy. He succumbed to secondary refractory AML. Patient 2 was a 51-year-old woman with MDS-EB-2, who refused treatment at initial presentation, received AZA (100 mg/m<sup>2</sup>/day on days 1–7) every 28 days for 7 months after her initial diagnosis. This patient subsequently refused induction chemotherapy or allogeneic hematopoietic stem cell transplantation (HSCT) and continued six further cycles of AZA, after which this patient progressed to secondary AML.

## 2.2. NGS of the Bulk Tumor Sample

DNA was extracted from bone marrow samples serially using a DNA blood mini extraction kit. NGS was performed serially using a 69-gene customized myeloid panel as previously described in the Supplementary Methods [17,18].

## 2.3. Targeted sc-DNA Sequencing

Sc-DNA sequencing was performed using the Tapestry Single-cell DNA AML Panel Kit (Mission Bio, San Francisco, CA, USA) according to the manufacturer's protocol. Sequencing data were processed using Mission Bio's Tapestry Pipeline V2.0. This pipeline comprised the following key steps: (1) adapter-trimming using Cutadapt, (2) reference genome alignment to hg19, (3) cellular barcode demultiplexing, (4) cell-based genotype calling using GATK/Haplotypecaller. Detection of *FLT3*-ITD was performed using additional scripts provided by Mission Bio and details were as described in the Supplementary Methods.

## 2.4. Data Sharing

All high-throughput sequencing data supporting the findings of this study were deposited in the Sequence Read Archive (SRA) with the BioProject number PRJNA748569. The sc-DNA raw reads were deposited to SRA under the accession codes SRR15209068, SRR15209069, SRR15209070, SRR15209071, SRR15209072, SRR15209073, SRR15209076, SRR15209077, and SRR15209078. The bulk tumor sequencing raw data were deposited under the accession codes SRR15209074, SRR15209075, SRR15209079, and SRR15209080, respectively.

## 3. Results

### 3.1. Mutation Dynamics on Bulk Tumor Sequencing and Cytogenetics

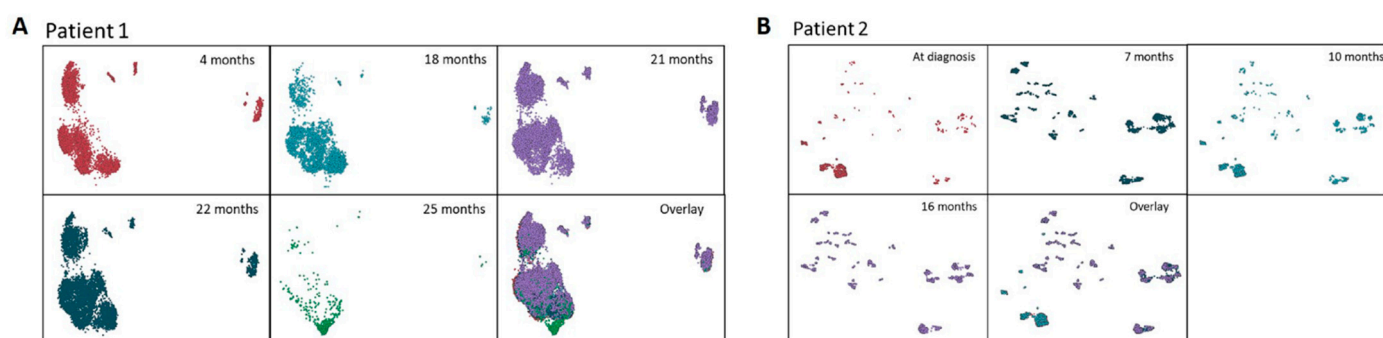
Two patients with MDS were first studied by bulk sequencing of the diagnostic bone marrow and a serial reassessment of bone marrow samples. Patient 1 with MDS-EB-2 had a 12% blast at diagnosis, and the karyotype was normal. Bulk sequencing using a 69-gene custom panel detected 37 variants at diagnosis with the major pathogenic mutation being *FLT3*-ITD with a VAF of 50% (Supplemental File S1). After 14 cycles of AZA (18 months), Patient 1 displayed a reduction in the mutation signature of *FLT3*-ITD VAF to 15.18% and a change in the list of variants was detected. Among the variants, the pathogenic *IDH2* R140H mutation was detected with a VAF of 30.64%. The patient subsequently progressed to secondary AML and was treated with induction with daunorubicin, cytarabine, and midostaurin. He achieved complete remission with incomplete hematologic recovery (CRi). He subsequently received one cycle of high-dose cytarabine and midostaurin. He relapsed again on the 21 month and was refractory to treatment with venetoclax—DEC and Gilteritinib monotherapy while the bulk tumor sequencing no longer detect any *FLT3*-ITD or *IDH2* mutation. He succumbed secondary to refractory AML.

Patient 2 with MDS-EB-2 was presented with 13% bone marrow blasts at diagnosis and a complex karyotype as following: 46,XX, add (1) (p11), add (5) (q11.2), add (6) (p23), -8, + mar [11]/46, XX, add (1) (p11), add (5) (q11.2), add (6) (p23), -8,+ r [2]/46, XX [6]. Bulk sequencing showed two frameshift mutations of *KMT2A* (M2894fs or G2895fs) and *KMT2D* (G2346S) at a VAF of 17% and 45.92%, respectively (Supplementary File S1). These mutations were predicted to be functionally disruptive. This patient, who refused treatment at initial presentation, started treatment with AZA 7 months after diagnosis. Bulk tumor sequencing performed at this time point showed the negative *KMT2A* mutation while the *KMT2D* mutation persisted. The bone marrow examination at the start of treatment showed a bone marrow blast percentage of 17%. After three cycles of AZA, the bone marrow blast percentage was reduced to 13% but bulk sequencing was not performed due to limited sample availability. After six further cycles of AZA, this patient progressed to secondary AML with a bone marrow blast percentage of 21%. At progression to secondary AML, the

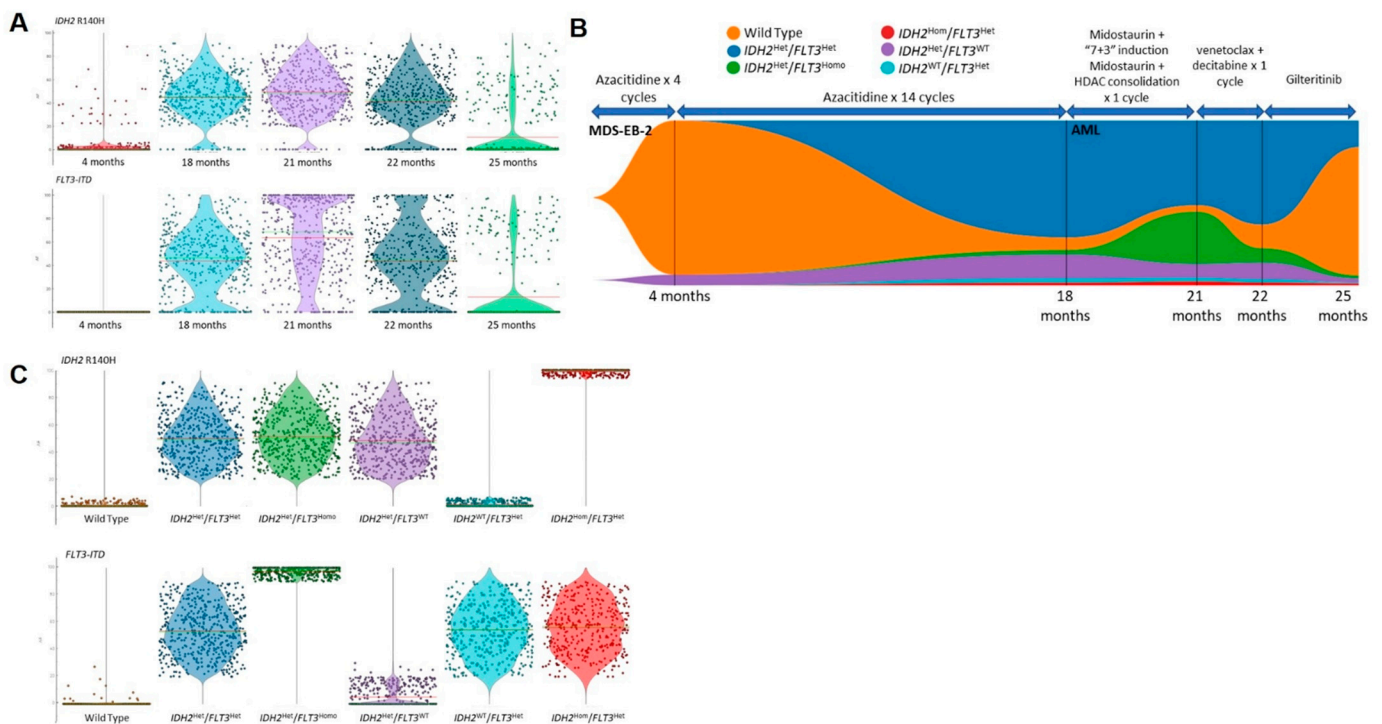
karyotype showed the following: 46, XX, add (1) (p11), der (5) t (1;5) (p31;q11.2), ad1d (6) (p23), -8, +mar [13]/46,XX.

### 3.2. Clonal Architecture of SNVs and INDELS on sc-DNA Sequencing

For Patient 1, a total number of 34,931 cells from five time points (8573, 3751, 7533, 14080, and 994 cells for each time point, respectively) were sequenced using sc-DNA sequencing platform. There were 98 variants detected across the five time points with good panel uniformity and an overall coverage of  $31 \times -600 \times$  per cell per amplicon (Supplemental File S2). Due to poor cell viability, the diagnostic sample of Patient 1 did not complete sc-DNA sequencing and sc-DNA analysis started after four cycles of AZA. The clustering analysis of pooled variants showed multiple cell clusters with a gradual reduction in clonal heterogeneity and dominance by 1 to 2 clusters during the treatment course. Despite having fewer cells successfully sequenced at 25 months from diagnosis, the dominant cell clone showed unique features compared with previous time points (Figure 1A). Among the list of variants, two pathogenic mutations were identified with high genotype quality (GQ) (Supplemental Files S3 and S4), and a dynamic mutation burden was observed as follows: (1) chr15:90631934:C/T (IDH2:p.R140H) and (2) chr13:28608262: /CTGAAATCAACGTAGAAG (FLT3-ITD) (Figure 2A). The average VAF of IDH2 remained stable at approximately 40% and decreased significantly at 25 months with a minor fraction of cells retaining the mutation. On the other hand, the VAF of FLT3-ITD mutation fluctuated between 18–22 months, when Patient 1 received midostaurin with induction/consolidation chemotherapy, and venetoclax-DEC. The VAF of FLT3-ITD decreased significantly when Patient 1 received Gilteritinib. At the clonal level, 94% of cells were wild type (WT) during initial treatment with AZA, with a minor clone carrying a heterozygous IDH2 mutation (Figure 2B,C). During transformation into AML at 18 months from diagnosis (after 14 cycles of AZA), this clone expanded together with another clone carrying the double heterozygous for IDH2 and FLT3-ITD mutations. Other minor subclones with a homozygous mutation of either or both of the two genes were also detected, and the homozygous FLT3 mutant clones expanded despite treatment with midostaurin and induction/consolidation chemotherapy. When gilteritinib was started at 22 months, all mutant clones were suppressed and the WT clone became the major clone again. Nevertheless, none of the mutant clones were completely eradicated.



**Figure 1.** Clonal heterogeneity during treatment by density-based spatial clustering of cells with noise (DBSCAN) based on the allelic burden of SNV and INDELS allele burden in 2 patients with MDS. (A): Patient 1 at 4 months (after 4 cycles of AZA), at 18 months (at progression to secondary AML after 18 cycles of AZA), at 21 months (at relapse after the first cycle of midostaurin and high-dose cytarabine consolidation), at 22 months (non-remission to one cycle of venetoclax-DEC) and at 25 months (non-remission after 3 months of Gilteritinib); (B): Patient 2 at diagnosis, at 7 months (before the first cycle of AZA), at 10 months (after 3 cycles of AZA), and at 16 months (at progression to secondary AML after 9 cycles of AZA).

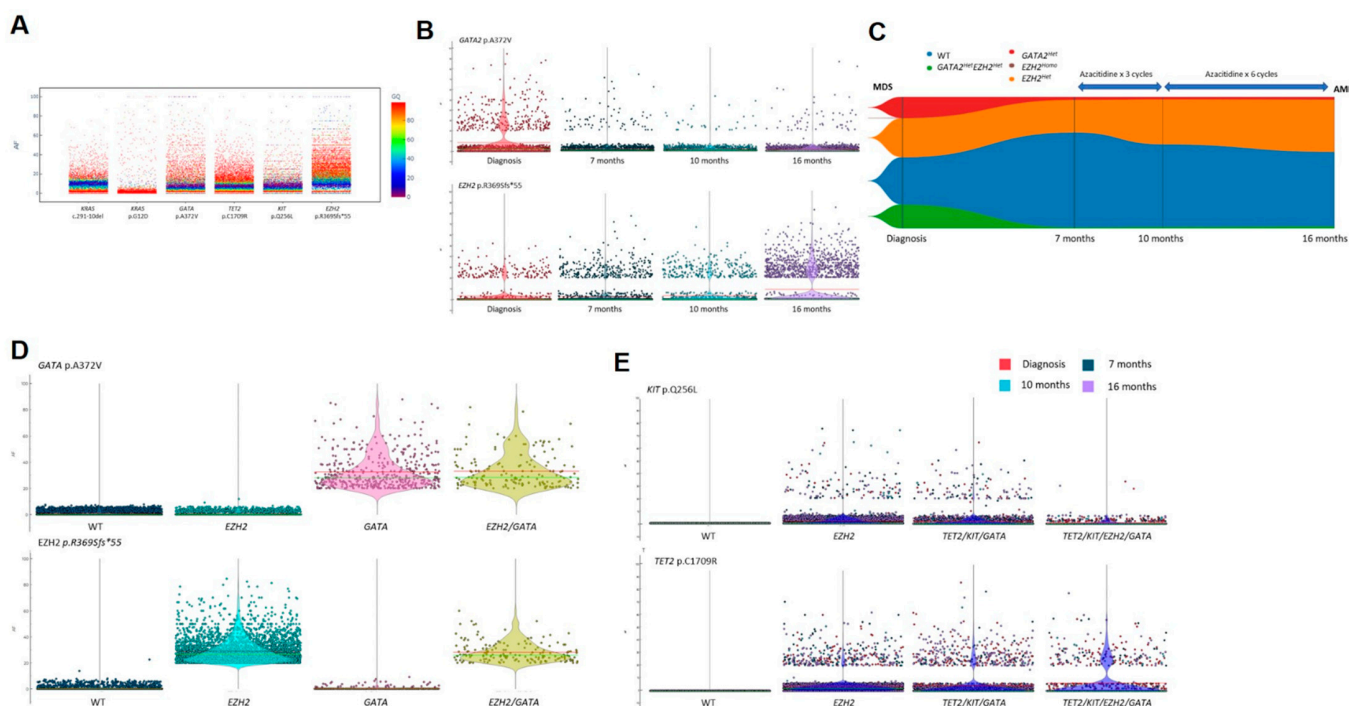


**Figure 2.** Clonal evolution of *IDH2* R140H and *FLT3*-ITD mutations in Patient 1. (A): Overall VAFs for *IDH2* R140H and *FLT3*-ITD across different time points evaluated; (B): Fish plot showing the evolution of WT, *FLT3*-ITD, and *IDH2* R140H clones serially with treatment; (C): VAFs of *IDH2* R140H and *FLT3*-ITD mutations at the sub-clonal level. Red and green lines on the violin plots represent median and mean VAFs, respectively.

### 3.3. Rare Cell Clones with SNVs and INDELS Undetectable with Bulk Tumor Sequencing

For Patient 2, there were 130 variants detected in 37,710 cells sequenced across four time points (4692, 6092, 11,438, and 15,488 for each time point, respectively) with good panel uniformity and an overall coverage of  $30\times$ – $76\times$  per cell per amplicon (Supplemental File S5). Total variant clustering showed a distinct profile across the four time points (at diagnosis, 7, 10, and 16 months from diagnosis). Clustering homologies were observed between the diagnostic time point and at 10 months (after 3 cycles of AZA) where the bone marrow blast percentage was the same at 13% (Figure 1B). On the other hand, the clustering profile at 7 months (at the start of AZA) reassembled at 16 months post-diagnosis when she completed nine cycles of AZA. Among the variants evaluated for pathogenicity, the following four missense and two frameshift variants were short-listed: (1) chr12-25398284-C-T (*KRAS*:p.G12D), (2) chr4-106196792-T-C (*TET2*:p.C1709R), (3) chr3-128200690-G-A (*GATA2*:p.A372V), (4) chr4-55569900-A-T (*KIT*:p.Q256L), (5) chr7-148515102-C- (*EZH2*:p.R369Sfs\*55) and (6) chr12-25378716-GA g (*KRAS*: c.291-10del). The five variants of *KRAS*: p.G12D, *TET2*: p.C1709R, *GATA2*: p.A372V, *EZH2*: p.R369Sfs \*55, and *KRAS*: c.291-10del were known pathogenic mutations, while *KIT*: p.Q256L was predicted to be tentatively pathogenic based on Varsome. A generally low VAF was observed for most variants but a low fraction of cells (<1% of total) showed high VAFs in *KRAS*, *GATA*, and *KIT* with high GQ (Figure 3A). For the *GATA2*, *TET2*, and *EZH2* variants, a more diverse VAF heterogeneity was observed (Figure 3A). A longitudinal analysis of the *GATA2* and *EZH2* mutations showed that the VAF of the *GATA2* mutation was highest at diagnosis and decreased after three cycles of AZA while the VAF of *EZH2* mutation remained constant initially and increased after nine cycles of AZA (Figure 3B). A downstream clonal analysis of *GATA2* and *EZH2* showed that these mutations were derived from different heterozygous cell clones harboring either one or both of the mutations (Figure 3C). The two *GATA2* mutant clones (*GATA2*<sup>Het</sup> single mutant and *GATA2*<sup>Het</sup>/*EZH2*<sup>Het</sup> double

mutant clones) diminished at disease progression but remained as minor clones throughout the remaining time points while the smallest EZH2 homozygous clone was no longer detectable at disease progression (Figure 3C). When reviewing the rare mutations, there were minor fractions of GATA2 and EZH2 mutant cells that harbored the KIT or TET2 mutation. Most of these triple mutant cells persisted across all time points with no specific time points being enriched for these minor subclones (Figure 3E).

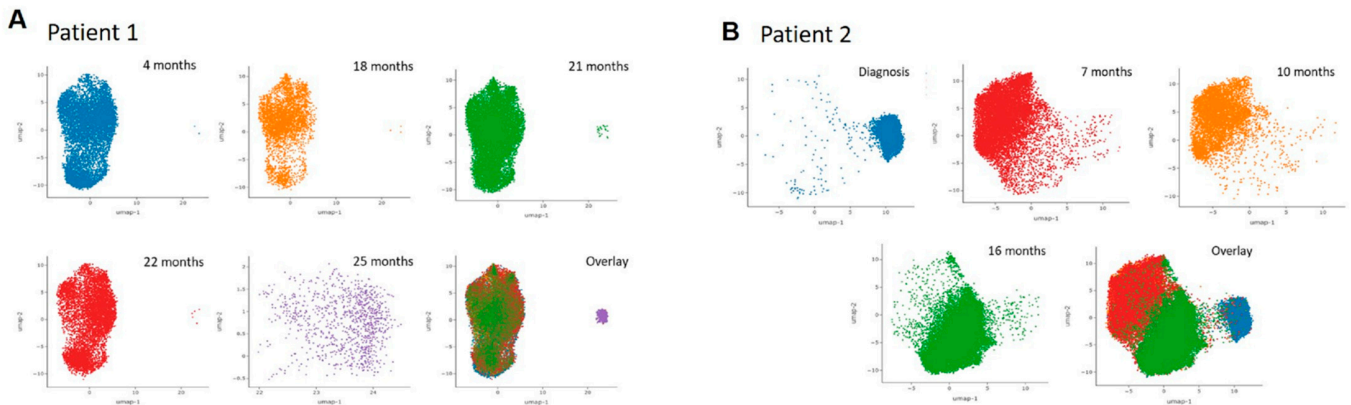


**Figure 3.** Detection and clonal evolutions of mutations in Patient 2. (A): VAFs and Quality of genotypes of rare pathogenic SNVs and INDELS which were present in <1% of total cells sequenced; (B): VAFs of major GATA2 and EZH2 mutations across different time points evaluated; (C): Fish plot evolution of WT, GATA2, and EZH2 clones serially with treatment; (D): VAFs of EZH2 and GATA2 mutations at sub-clonal level; (E): VAFs of rare KIT and TET2 mutations in association with GATA2 and EZH2 mutations across different time points. Red and green lines on violin plots represented median and mean VAF, respectively while the color of each cell represented corresponding time points.

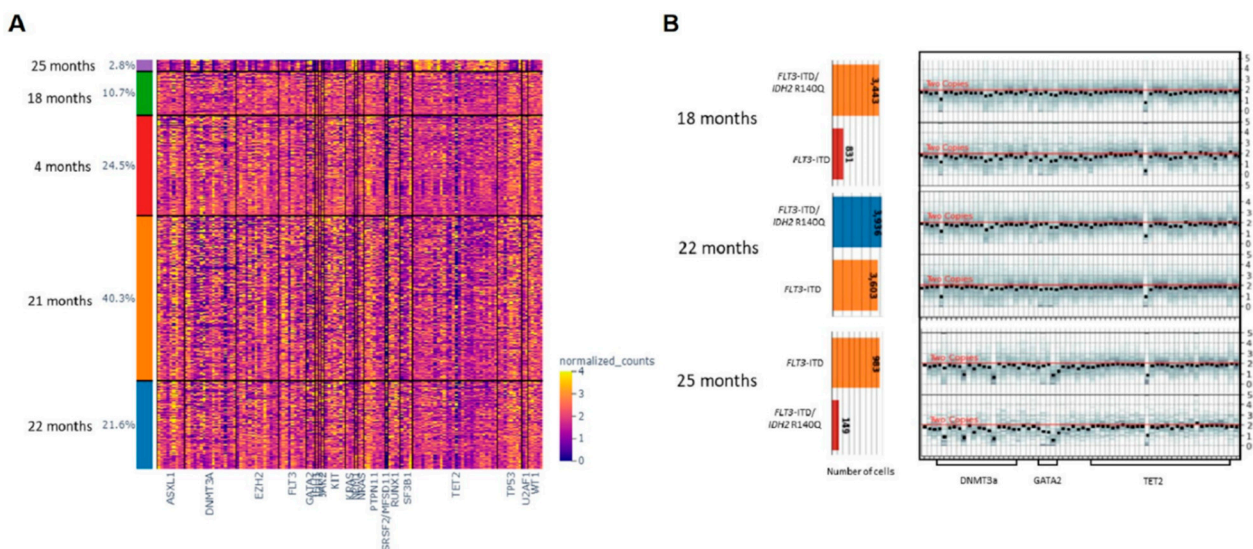
### 3.4. Clonal CNVs Associated with Pathogenic SNVs or INDELS during HMA Resistance

In addition to studying the dynamics of clonal SNV and INDEL, the CNVs of all 20 genes (127 amplicons) were studied and correlated with the IDH2 and FLT3-ITD mutations detected in Patient 1. In Patient 1, a cluster analysis of all amplicons pooled from all time points showed no significant differences during treatment with AZA and at progression to secondary AML (Figure 4A). An unbiased clustering analysis showed several CNVs in ASXL1, GATA2, DNMT3A, IDH1/2, KIT, and TET2. These CNVs were observed serially but ASXL1, DNMT3A, IDH1/2, and KIT deletion were enriched at 21 and 25 months when the patient failed treatment with midostaurin-based induction/consolidation, venetoclax-DEC and gilteritinib (Figure 5A). Despite observing a population of WT clone, loss of DNMT3A, GATA2, and TET2 were enriched in clones harboring FLT3-ITD and/or IDH2 mutations (Supplemental File S6). The magnitude of DNMT3A and GATA2 loss in FLT3 and/or IDH2 mutant clones also showed sequential changes as treatment resistance developed while TET2 CNV remained stable (Figure 5B). For Patient 2 where rare mutant clones were detected, a CNV analysis was performed to detect copy number abnormalities independent of the rare SNVs and INDELS to maximize sensitivity. Overall clustering in patient 2 showed distinct profiles at diagnosis compared with that at subsequent disease stages and homology was observed between 7 months (pre-treatment) and 10 months (after

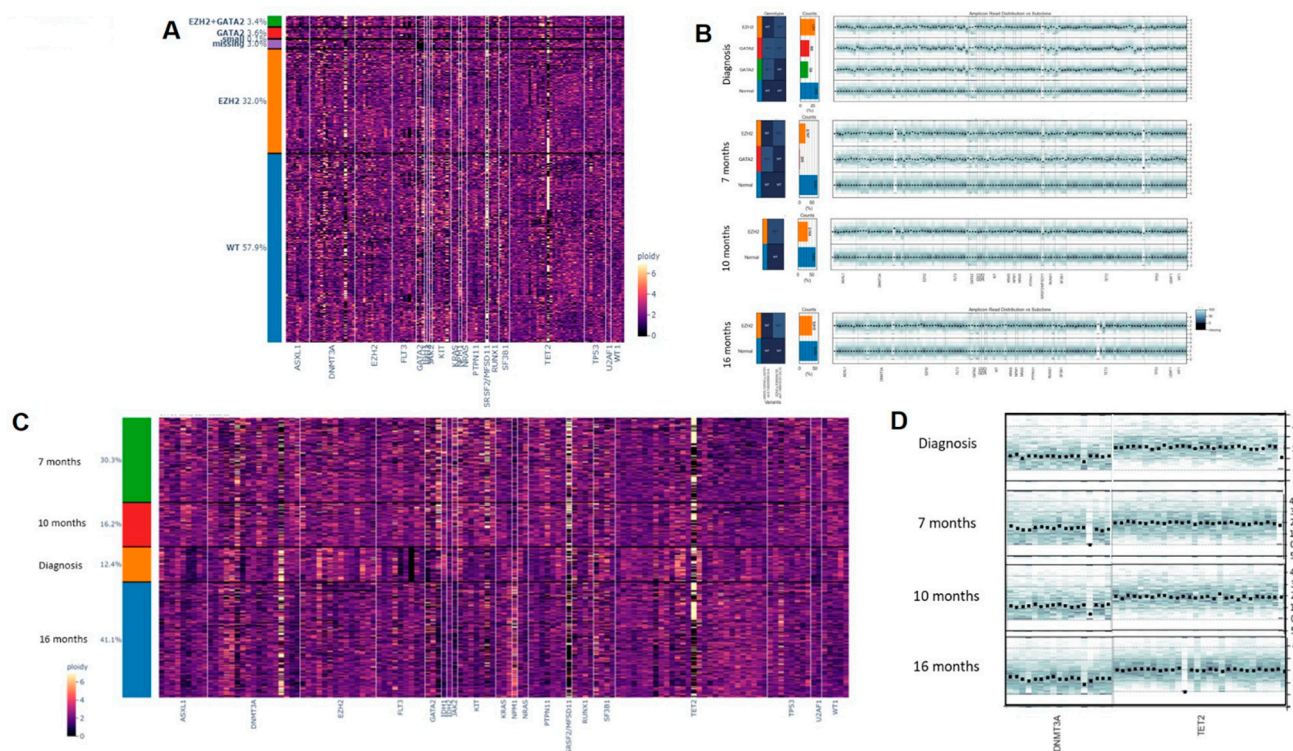
3 cycles of AZA) from diagnosis. The profile at 16 months (at progression to secondary AML after 9 cycles of AZA) was somewhat intermediate between 7 and 10 months (Figure 4B). An analysis of all time points in association with GATA2 and EZH2 mutations did not lead to the detection of any large clonal specific CNVs compared with WT cells, except for the amplicon loss of TET2 exon 3 (106158314) in the GATA2<sup>Het</sup> single mutant clone at 7 months diagnosis (Figure 6A,B). Consistent with the longitudinal analysis of GATA2 mutation, this clone was undetectable during treatment with AZA and at subsequent progression to secondary AML.



**Figure 4.** Heterogeneity of overall copy numbers during treatment by Principal Component Analysis (PCA) and Uniform Manifold Approximation Projection (UMAP) analysis. **(A):** Patient 1 at 4 months (after 4 cycles of AZA), at 18 months (at progression to secondary AML after 18 cycles of AZA), at 21 months (at relapse after the first cycle of midostaurin and high-dose cytarabine consolidation), at 22 months (non-remission to one cycle of venetoclax-DEC) and at 25 months (non-remission after 3 months of Gilteritinib); **(B):** Patient 2 at diagnosis, at 7 months (before the first cycle of AZA), at 10 months (after 3 cycles of AZA), and at 16 months (at progression to secondary AML leukemia after 9 cycles of AZA).



**Figure 5.** Co-occurrence of SNVs and short INDELS with CNVs in Patient 1. **(A):** merged CNV analysis of all time points; **(B):** Gradual copy number loss of DNMT3A, GATA2, and TET2 at 18 months (at progression to secondary AML after 18 cycles of AZA), at 22 months (non-remission to one cycle of venetoclax-DEC) and at 25 months (non-remission after 3 months of Gilteritinib).



**Figure 6.** Co-occurrence of SNVs and short INDELS with CNVs in Patient 2 as independent cell clones. (A): merged CNV analysis of all time points; (B): CNV analysis at individual time points in association with *GATA2* p.A372V and *EZH2* p.R369Sfs\*55 mutations; (C): Unbiased clustering of CNV analysis of all time points; (D): Gradual copy number loss of *DNMT3A* and *TET2* at 7 months (before the first cycle of AZA), at 10 months (after 3 cycles of AZA), and at 16 months (at progression to secondary AML after 9 cycles of AZA).

### 3.5. Clonal CNVs as Independent Events during HMA Resistance

To further detect clonal CNVs that were independent of *GATA2* and *EZH2* mutations in patient 2, unbiased clustering with pooled time points was performed. Putative deletions of *TET2*, *ASXL1*, *DNMT3A*, and potential artifacts of *FLT3* and *SRSF2* deletions were detected across multiple time points (Figure 6C). Both *FLT3* and *SRSF2* were later confirmed to be allele-dropouts due to poor PCR efficiency of amplicons and they were discarded from subsequent analysis (Figure 6C). To confirm the putative CNVs, repeated analyses normalized to cells with neutral polymorphisms were performed (Figure 6D). Among the three candidate CNVs, only *DNMT3A* and *TET2* deletions were confirmed. Consistent drops of all *DNMT3A* amplicons were observed at 7 months (at the start of AZA) and at 16 months (at progression to secondary AML) together with a homozygous loss of *TET2*.

## 4. Discussion

In this study, we used a single-cell-based approach to detect mutations that were associated with disease progression in MDS and secondary AML. Using this high-resolution and sensitive method, the complex clonal heterogeneity of MDS cells during treatment resistance and progression was deciphered, demonstrating a distinct clonal architecture involving both SNVs, INDELS, and CNVs. In addition to the good correlation between the VAF on bulk tumor sequencing and the VAF of the sub-clones harboring *FLT3*-ITD and *IDH2* mutations, rare mutations that were below the detection limit of bulk tumor sequencing were detectable with sc-DNA sequencing. When gilteritinib was administered at 22 months to Patient 1, all mutant clones were suppressed; however, this may result in possible resistance to *FLT3* inhibitors in several ways such as through the activation of survival pathways, cell adhesion to increase proliferation and adaptation of leukemic



cells in bone marrow microenvironment by attenuating the expression of *CYP3A4* in bone marrow stromal cells resulting in activation to FLT3 inhibitor [19–21].

Although under-sequencing was observed at the last time point in Patient 1 and at diagnosis in Patient 2 compared with other time points, the high sensitivity to detect rare mutations of *KIT* was achieved when integrated analyses of all serial time points were performed. *KIT* and *TET2* mutations have been implicated in cellular differentiation and DNA methylation in MDS (1). We also demonstrated that this sc-DNA platform allows for the definitive determination of co-occurring mutations within the same cell.

Excluding the low fraction of allele-dropout events that led to false positive CNVs of *SRSF2* and *FLT3*, pathogenic losses of *DNMT3A* and *TET2* were detected in both patients. With *DNMT3A* showing a progressive loss in both patients, this suggested that continuous treatment with AZA or other DNMT inhibitors created a selection pressure for cell clones with *DNMT3A* loss or promoted the clonal acquisition of de novo *DNMT3A* mutations. While both *TET2* and *DNMT3A* were tumor suppressive in MDS, this study showed that *DNMT3A* loss is the primary or causative genomic abnormality to disease progression and treatment resistance [22,23]. In Patient 1, while *FLT3*-ITD clones were suppressed during potent FLT3 inhibition with Gilteritinib, disease remission was not achieved. In addition to the persistence of *IDH2* R140H clones and other rare cell clones, our study showed that copy number loss of *DNMT3A*, *TET2* and *GATA2* may play a role in resistance to FLT3 inhibitors.

## 5. Conclusions

In conclusion, this study demonstrated the clonal evolution of MDS at single-cell level treatment resistance to HMA. sc-DNA sequencing can uncover genetic mosaicism, and detect cells that are implicated in transformation into AML cell states which are crucial in underlining AML development and providing insights into treatment responses towards resistance development. SNVs and CNVs detected in this study clearly demonstrate their implication in cancer heterogeneity, and minor pathogenic clones that cannot be resolved by bulk tumor sequencing could be selected during prolonged treatments leading to drug resistance. The underlying evolutionary process involves not only SNVs and short INDELS but also acquired pathogenic CNVs of *GATA2*, *DNMT3A*, and *TET2* that are associated with HMA resistance. These CNVs could be coupled with SNVs or small INDELS of *FLT3* and *IDH2*. Nevertheless, the effect of these mutations on cellular differentiation and proliferation, and the expression of cell surface markers will require a multi-omic approach to simultaneously detect mutations and phenotypic changes at the protein level. Accurately identified clonal populations and reconstructing clonal phylogenies could be more effectively achieved with single cell analysis. The resolution of clonal heterogeneity in MDS may allow for better disease monitoring and the early detection of resistant clones. Resistant clones prior to therapy may also be detected. This will facilitate the design of combinatorial treatment approaches in patients with MDS. The low mutation burden of common mutations such as *FLT3*-ITD in AML could also be better resolved to determine if monotherapy would be a good option when sc-DNA sequencing is applied to detect the proportional of homozygous clones.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms23094647/s1>. References [24–36] are cited in the Supplementary Materials.

**Author Contributions:** Conceptualization, H.G.; methodology, H.G.; formal analysis, P.L., R.Y., S.-H.F., K.-K.M., Z.W., K.-C.W. and V.H.-F.L.; data curation, P.L., R.Y., S.-H.F., K.-K.M., Z.W., K.-C.W., L.A. and G.M.-K.L.; writing—original draft preparation, P.L., R.Y. and H.G.; writing—H.G.; supervision, H.G.; project administration, L.A., G.M.-K.L. and H.G.; funding acquisition, H.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** The research was funded by the General Research Fund (GRF), Research Grant Council (RGC), University Grants Committee (UGC), Hong Kong, China (funding number: 17118914) and the Hong Kong Anti-Cancer Society Grant (project code: AR190027).

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of the University of Hong Kong (protocol code: UW 17-138 and date of approval: 14 June 2019).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** All high-throughput sequencing data supporting the findings of this study were deposited in the Sequence Read Archive (SRA) with the BioProject number PRJNA748569. The sc-DNA raw reads were deposited to SRA under the accession codes SRR15209068, SRR15209069, SRR15209070, SRR15209071, SRR15209072, SRR15209073, SRR15209076, SRR15209077, and SRR15209078. The bulk tumor sequencing raw data were deposited under the accession codes SRR15209074, SRR15209075, SRR15209079, and SRR15209080.

**Acknowledgments:** We would like to acknowledge the Genomic Core at the Center for PanorOmic Sciences of Hong Kong University for performing NovaSeq sequencing of the single-cell DNA libraries.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Gill, H.; Leung, A.Y.H.; Kwong, Y.-L. Molecular and Cellular Mechanisms of Myelodysplastic Syndrome: Implications on Targeted Therapy. *Int. J. Mol. Sci.* **2016**, *17*, 440. [[CrossRef](#)] [[PubMed](#)]
- Walter, M.J.; Shen, D.; Ding, L.; Shao, J.; Koboldt, D.C.; Chen, K.; Larson, D.E.; McLellan, M.D.; Dooling, D.; Abbott, R.; et al. Clonal Architecture of Secondary Acute Myeloid Leukemia. *N. Engl. J. Med.* **2012**, *366*, 1090–1098. [[CrossRef](#)] [[PubMed](#)]
- Garcia-Manero, G.; Chien, K.S.; Montalban-Bravo, G. Myelodysplastic syndromes: 2021 update on diagnosis, risk stratification and management. *Am. J. Hematol.* **2020**, *95*, 1399–1420. [[CrossRef](#)] [[PubMed](#)]
- Traina, F.; Visconte, V.; Elson, P.; Tabarrokhi, A.; Jankowska, A.M.; Hasrouni, E.; Sugimoto, Y.; Szpurka, H.; Makishima, H.; O'Keefe, C.L.; et al. Impact of molecular mutations on treatment response to DNMT inhibitors in myelodysplasia and related neoplasms. *Leukemia* **2014**, *28*, 78–87. [[CrossRef](#)]
- Woll, P.S.; Kjällquist, U.; Chowdhury, O.; Doolittle, H.; Wedge, D.; Thongjuea, S.; Erlandsson, R.; Ngara, M.; Anderson, K.; Deng, Q.; et al. Myelodysplastic Syndromes Are Propagated by Rare and Distinct Human Cancer Stem Cells In Vivo. *Cancer Cell* **2014**, *25*, 794–808. [[CrossRef](#)]
- Bersanelli, M.; Travaglino, E.; Meggendorfer, M.; Matteuzzi, T.; Sala, C.; Mosca, E.; Chiereghin, C.; Di Nanni, N.; Gnocchi, M.; Zampini, M.; et al. Classification and Personalized Prognostic Assessment on the Basis of Clinical and Genomic Features in Myelodysplastic Syndromes. *J. Clin. Oncol.* **2021**, *39*, 1223–1233. [[CrossRef](#)]
- Nazha, A.; Komrokji, R.; Meggendorfer, M.; Jia, X.; Radakovich, N.; Shreve, J.; Hilton, C.B.; Nagata, Y.; Hamilton, B.K.; Mukherjee, S.; et al. Personalized Prediction Model to Risk Stratify Patients With Myelodysplastic Syndromes. *J. Clin. Oncol.* **2021**, *39*, 3737–3746. [[CrossRef](#)]
- Platzbecker, U. Treatment of MDS. *Blood* **2019**, *133*, 1096–1107. [[CrossRef](#)]
- Jabbour, E.; Garcia-Manero, G.; Batty, N.; Shan, J.; O'Brien, S.; Cortes, J.; Ravandi, F.; Issa, J.-P.; Kantarjian, H. Outcome of patients with myelodysplastic syndrome after failure of decitabine therapy. *Cancer* **2010**, *116*, 3830–3834. [[CrossRef](#)]
- Santini, V. How I treat MDS after hypomethylating agent failure. *Blood* **2019**, *133*, 521–529. [[CrossRef](#)]
- Meldi, K.; Qin, T.; Buchi, F.; Droin, N.; Sotzen, J.; Micol, J.-B.; Selimoglu-Buet, D.; Masala, E.; Allione, B.; Gioia, D.; et al. Specific molecular signatures predict decitabine response in chronic myelomonocytic leukemia. *J. Clin. Investig.* **2015**, *125*, 1857–1872. [[CrossRef](#)] [[PubMed](#)]
- Tobiasson, M.; Abdulkadir, H.; Lennartsson, A.; Katayama, S.; Marabita, F.; De Paepe, A.; Karimi, M.; Krjtskov, K.; Einarsdottir, E.; Grövdal, M.; et al. Comprehensive mapping of the effects of azacitidine on DNA methylation, repressive/permissive histone marks and gene expression in primary cells from patients with MDS and MDS-related disease. *Oncotarget* **2017**, *8*, 28812–28825. [[CrossRef](#)] [[PubMed](#)]
- Nazha, A.; Sekeres, M.A.; Bejar, R.; Rauh, M.J.; Othus, M.; Komrokji, R.S.; Barnard, J.; Hilton, C.B.; Kerr, C.M.; Steensma, D.P.; et al. Genomic Biomarkers to Predict Resistance to Hypomethylating Agents in Patients With Myelodysplastic Syndromes Using Artificial Intelligence. *JCO Precis. Oncol.* **2019**, *3*, 1–11. [[CrossRef](#)] [[PubMed](#)]
- Bewersdorf, J.P.; Zeidan, A.M. Management of patients with higher-risk myelodysplastic syndromes after failure of hypomethylating agents: What is on the horizon? *Best Pract. Res. Clin. Haematol.* **2021**, *34*, 101245. [[CrossRef](#)]
- Zhao, X.; Gao, S.; Wu, Z.; Kajigaya, S.; Feng, X.; Liu, Q.; Townsley, D.M.; Cooper, J.; Chen, J.; Keyvanfar, K.; et al. Single-cell RNA-seq reveals a distinct transcriptome signature of aneuploid hematopoietic cells. *Blood* **2017**, *130*, 2762–2773. [[CrossRef](#)]
- Ganan-Gomez, I.; Yang, H.; Ma, F.; Pellegrini, M.; Clise-Dwyer, K.; Garcia-Manero, G.; Colla, S. Single-Cell RNA Sequencing Reveals Distinct Hematopoietic Stem Cell Hierarchies in MDS. *Blood* **2019**, *134* (Suppl. S1), 771. [[CrossRef](#)]

17. Gill, H.; Leung, G.M.; Yim, R.; Lee, P.; Pang, H.H.; Ip, H.-W.; Leung, R.Y.; Li, J.; Panagiotou, G.; Ma, E.S.; et al. Myeloproliferative neoplasms treated with hydroxyurea, pegylated interferon alpha-2A or ruxolitinib: Clinicohematologic responses, quality-of-life changes and safety in the real-world setting. *Hematology* **2020**, *25*, 247–257. [[CrossRef](#)]
18. Gill, H.; Ip, H.W.; Yim, R.; Tang, W.F.; Pang, H.H.; Lee, P.; Leung, G.M.K.; Li, J.; Tang, K.; So, J.C.C.; et al. Next-generation sequencing with a 54-gene panel identified unique mutational profile and prognostic markers in Chinese patients with myelofibrosis. *Ann. Hematol.* **2019**, *98*, 869–879. [[CrossRef](#)]
19. Traer, E.; Martinez, J.; Javidi-Sharifi, N.; Agarwal, A.; Dunlap, J.; English, I.; Kovacs, T.; Tyner, J.W.; Wong, M.; Druker, B. FGF2 from Marrow Microenvironment Promotes Resistance to FLT3 Inhibitors in Acute Myeloid Leukemia. *Cancer Res.* **2016**, *76*, 6471–6482. [[CrossRef](#)]
20. Chang, Y.T.; Hernandez, D.; Alonso, S.; Gao, M.; Su, M.; Ghiaur, G.; Levis, M.J.; Jones, R.J. Role of CYP3A4 in bone marrow microenvironment-mediated protection of FLT3/ITD AML from tyrosine kinase inhibitors. *Blood Adv.* **2019**, *3*, 908–916. [[CrossRef](#)]
21. Eguchi, M.; Minami, Y.; Kuzume, A.; Chi, S. Mechanisms Underlying Resistance to FLT3 Inhibitors in Acute Myeloid Leukemia. *Biomedicines* **2020**, *8*, 245. [[CrossRef](#)] [[PubMed](#)]
22. Feng, Y.; Li, X.; Cassady, K.; Zou, Z.; Zhang, X. TET2 Function in Hematopoietic Malignancies, Immune Regulation, and DNA Repair. *Front. Oncol.* **2019**, *9*, 210. [[CrossRef](#)] [[PubMed](#)]
23. Saliba, A.N.; John, A.J.; Kaufmann, S.H. Resistance to venetoclax and hypomethylating agents in acute myeloid leukemia. *Cancer Drug Resist* **2020**, *4*, 125–142. [[CrossRef](#)] [[PubMed](#)]
24. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [[CrossRef](#)]
25. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)]
26. Van der Auwera, G.A.; Carneiro, M.O.; Hartl, C.; Poplin, R.; Del Angel, G.; Levy-Moonshine, A.; Jordan, T.; Shakir, K.; Roazen, D.; Thibault, J.; et al. From FastQ data to high confidence variant calls: The Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinform.* **2013**, *43*, 11.10.1–11.10.33.
27. Koboldt, D.C.; Zhang, Q.; Larson, D.E.; Shen, D.; McLellan, M.D.; Lin, L.; Miller, C.A.; Mardis, E.R.; Ding, L.; Wilson, R.K. VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **2012**, *22*, 568–576. [[CrossRef](#)]
28. Ye, K.; Schulz, M.H.; Long, Q.; Apweiler, R.; Ning, Z. Pindel: A pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics* **2009**, *25*, 2865–2871. [[CrossRef](#)]
29. Cingolani, P.; Platts, A.; Wang, L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* **2012**, *6*, 80–92. [[CrossRef](#)]
30. Wang, K.; Li, M.; Hakonarson, H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **2010**, *38*, e164. [[CrossRef](#)]
31. Kopanos, C.; Tsiolkas, V.; Kouris, A.; Chapple, C.E.; Albarca Aguilera, M.; Meyer, R.; Massouras, A. VarSome: The human genomic variant search engine. *Bioinformatics* **2019**, *35*, 1978–1980. [[CrossRef](#)] [[PubMed](#)]
32. Landrum, M.J.; Lee, J.M.; Benson, M.; Brown, G.R.; Chao, C.; Chitipiralla, S.; Gu, B.; Hart, J.; Hoffman, D.; Jang, W.; et al. ClinVar: Improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* **2018**, *46*, D1062–D1067. [[CrossRef](#)] [[PubMed](#)]
33. Sherry, S.T.; Ward, M.; Sirotkin, K. dbSNP-database for single nucleotide polymorphisms and other classes of minor genetic variation. *Genome Res.* **1999**, *9*, 677–679. [[CrossRef](#)]
34. Smigielski, E.M.; Sirotkin, K.; Ward, M.; Sherry, S.T. dbSNP: A database of single nucleotide polymorphisms. *Nucleic Acids Res.* **2000**, *28*, 352–355. [[CrossRef](#)]
35. Karczewski, K.J.; Francioli, L.C.; Tiao, G.; Cummings, B.B.; Alfoldi, J.; Wang, Q.; Collins, R.L.; Laricchia, K.M.; Ganna, A.; Birnbaum, D.P.; et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **2020**, *581*, 434–443. [[CrossRef](#)]
36. Quang, D.; Chen, Y.; Xie, X. DANN: A deep learning approach for annotating the pathogenicity of genetic variants. *Bioinformatics* **2015**, *31*, 761–763. [[CrossRef](#)]