



Clonal Evolution at First Sight: A Combined Visualization of Diverse Diagnostic Methods Improves Understanding of Leukemic Progression

Sarah Sandmann^{1†}, Yvonne Lisa Behrens^{2*†}, Claudia Davenport², Felicitas Thol³, Michael Heuser³, Daniela Dörfel⁴, Friederike Löhr⁵, Agnes Castrup⁶, Doris Steinemann², Julian Varghese¹, Brigitte Schlegelberger², Martin Dugas^{1,7} and Gudrun Göhring²

OPEN ACCESS

Edited by:

Robert Ohgami,
University of California, San Francisco,
United States

Reviewed by:

Babal Jha,
Cleveland Clinic, United States
Michael Diamantidis,
University Hospital of Larissa, Greece

*Correspondence:

Yvonne Lisa Behrens
Behrens.Yvonne@mh-hannover.de

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Hematologic Malignancies,
a section of the journal
Frontiers in Oncology

Received: 02 March 2022

Accepted: 10 June 2022

Published: 08 July 2022

Citation:

Sandmann S, Behrens YL,
Davenport C, Thol F, Heuser M,
Dörfel D, Löhr F, Castrup A,
Steinemann D, Varghese J,
Schlegelberger B, Dugas M
and Göhring G (2022) Clonal
Evolution at First Sight: A Combined
Visualization of Diverse Diagnostic
Methods Improves Understanding
of Leukemic Progression.
Front. Oncol. 12:888114.
doi: 10.3389/fonc.2022.888114

¹ Institute of Medical Informatics, University of Münster, Münster, Germany, ² Department of Human Genetics, Hannover Medical School, Hannover, Germany, ³ Department of Hematology, Hemostasis, Oncology and Stem Cell Transplantation, Hannover Medical School, Hannover, Germany, ⁴ Department of Hematology, Oncology and Immunology, Klinikum Region Hannover (KRH) Klinikum Siloah, Hannover, Germany, ⁵ Department of Hematology and Oncology, Klinikum Braunschweig, Braunschweig, Germany, ⁶ Hämato-Onkologische Praxis, Hämato-Onkologische Praxis im Medicum, Bremen, Germany, ⁷ Institute of Medical Informatics, Heidelberg University Hospital, Heidelberg, Germany

Patients with myeloid neoplasia are classified by the WHO classification systems. Besides clinical and hematological criteria, cytogenetic and molecular genetic alterations highly impact treatment stratification. In routine diagnostics, a combination of methods is used to decipher different types of genetic variants. Eight patients were comprehensively analyzed using karyotyping, fluorescence *in situ* hybridization, array-CGH and a custom NGS panel. Clonal evolution was reconstructed manually, integrating all mutational information on single nucleotide variants (SNVs), insertions and deletions (indels), structural variants and copy number variants (CNVs). To allow a correct integration, we differentiate between three scenarios: 1) CNV occurring prior to the SNV/indel, but in the same cells. 2) SNV/indel occurring prior to the CNV, but in the same cells. 3) SNV/indel and CNV existing in parallel, independent of each other. Applying this bioinformatics approach, we reconstructed clonal evolution for all patients. This generalizable approach offers the possibility to integrate various data to analyze identification of driver and passenger mutations as well as possible targets for personalized medicine approaches. Furthermore, this model can be used to identify markers to assess the minimal residual disease.

Keywords: clonal evolution, bioinformatics, single nucleotide variants, copy number variants, cancer cell fraction, leukemic progression

INTRODUCTION

Myeloid neoplasia, including acute myeloid leukemia (AML) and myelodysplastic syndrome (MDS) are heterogeneous hematopoietic stem cell disorders, which are marked by the acquisition of somatic alterations and clonal evolution (1–4). Besides clinical and hematological criteria, cytogenetic and molecular genetic alterations highly impact treatment stratification (2). In recent

years, high-throughput sequencing technologies have led to the identification of many driver and passenger alterations that enable precision medicine (e.g. *IDH1* inhibitor ivosidenib in AML patients) (2, 5–9). The different subgroups of myeloid neoplasia (e.g. MDS and AML) show a highly heterogeneous cytogenetic and molecular genetic profile. Multiple subclones may exist in one patient. Distribution of recurrently mutated genes and clonal architecture are different, e.g. in MDS/MPN subtypes (4, 10, 11). Consequently, leukemic progression may be influenced by several leukemic clones, and the clonal composition identified at diagnosis can differ from the composition identified at relapse or progression (12). Patients with myeloid neoplasia, especially patients with MDS, often progress to AML (sAML) through a process of clonal evolution (2, 11–13). Clonal evolution can be associated with poor prognosis, relapse and therapeutic resistance (14, 15). To this day different models of tumor evolution have been reported (e.g. linear, branching or neutral evolution) (16).

Recent publications have shown that reconstruction of clonal evolution and characterization of clonal architecture are important for the understanding of tumor development and treatment failure (14). Furthermore, they can help in finding new treatment strategies (e.g. combination of drugs, or new drugs for patients with relapse) (11, 12, 14). Different bioinformatic approaches already exist for the calculation of clonal evolution of bulk sequencing data. In addition, the single cell sequencing method has expanded the field of reconstructing clonal evolution (14, 17–19). A study of Morita et al. has shown both genetic and phenotypic evolution in AML by single cell sequencing and cell surface protein analyses (14). However, in routine diagnostics, a combination of methods (e.g. karyotyping and NGS panel) is used to identify alterations that are important for treatment stratification of patients with myeloid neoplasia, and which may not be detected using only a single method (20). These methods decipher different types of genetic variants, including single nucleotide variants (SNVs), insertions/deletions (indels), structural variants (SVs) and copy number variations (CNVs). The combination of a wide spectrum of methods could lead to the identification of many alterations, which, in turn, can make interpretation difficult. Unfortunately, an approach that combines all identified molecular genetic and cytogenetic aberrations and that reconstructs and visualizes the genetic architecture and clonal evolution, is still lacking.

In the present study, we performed karyotyping, fluorescence *in situ* hybridization, array-CGH and used a custom NGS panel in a cohort of eight patients with myeloid neoplasia. Here, we propose a bioinformatic approach as well as a visualization of occurrences of genetic alterations to improve our understanding of leukemic progression (driver and passenger alterations) and assist in finding the best treatment stratification paving the way for personalized medicine approaches. Clonal evolution and clonal architecture not only play a role in therapeutic resistance, relapse and poor outcomes in myeloid neoplasia, but also in other non-solid tumors (14). Our proposed bioinformatic model is a general approach, which may be used on all kinds on non-solid tumors showing clonal evolution.

MATERIALS AND METHODS

Study Population

A total of eight patients with myeloid neoplasia and clonal evolution were included in this study. The patients were analyzed between 2011 and 2021 at our department. The ethical review boards of Hannover Medical School approved this study and all patients gave their written consent. Our cohort was analyzed at one or more time points. All results of performed analyses as well as clinical data are shown in **Table 1** and in more detail in **Table S1**.

Karyotyping and Fluorescence *In-Situ* Hybridization (FISH) Analysis

Cytogenetic and molecular cytogenetic analyses were performed on bone marrow aspirates or peripheral blood cultures. Chromosome preparation and fluorescence R-banding were performed as described previously (21, 22). Altogether, whenever possible, 15 to 25 metaphases were examined per patient. The karyotype was described according to guidelines of the International System for Human Cytogenetic Nomenclature (23). Depending on the cytogenetic aberrations in each patient, FISH analyses on interphase nuclei were performed using: (1) break apart probes for the loci 3q26 (EVI1 (MECOM) Break Apart; Cytocell, Cambridge, UK), 11q23 and 12p13 (Vysis MLL and ETV6 FISH Probe Kit; Abbott, IL, USA); (2) dual color probes for the loci 5p15.2, 5q31, CEP7, 7q31, 9q34.1, 17p13.1, CEP17, 22q11.2 (Vysis EGR1/D5S23, D5S721, D7S486/CEP7, BCR/ABL and TP53/CEP 17 FISH Probe Kit; Abbott); and (3) single color probe for the locus 20q12 (Vysis D20S108 FISH Probe Kit; Abbott). At least 200 interphase nuclei were analyzed for each probe.

Array-CGH (Array-Based Comparative Genomic)

According to the manufacturer's protocol (e-Array design 84704, Agilent Technologies, Waldbronn, Germany), three patients (patients c1, 6 and 7) were screened for CNVs by high-resolution array-CGH. Microarray slides were scanned immediately using an Agilent microarray scanner at a resolution of 2 μ m. Fluorescence ratios were calculated using Feature Extraction Software and copy number states analyzed using the CGH data analysis software Genomic Workbench (Agilent Technologies).

NGS With IDT Custom Panel

DNA sequencing was performed for all patients using an IDT custom panel (Integrated DNA Technologies Inc, Iowa, USA) including 148 leukemia-associated genes. Sequencing was performed on an Illumina NextSeq sequencer. All samples were aligned against the reference genome GRCh37 using BWA mem (24). The pipeline uses eight open-source tools for independent variant calling. Subsequently, raw calls are automatically combined and characterized with respect to data quality. Default settings were used (minimum number of reads 50, minimum number of reads with the alternate allele 20, minimum variant allele frequency VAF 0.01). Excluding low-quality calls, the remaining variants are

TABLE 1 | Overview of analyzed patients (n=8) (All results of performed analyses as well as clinical data are shown in **Table S1**).

patient (#)	WHO diagnosis	status	cytogenetic aberrations	DNA analysis NGS (somatic)			treatment
				gene	variant	variant allele frequency (VAF)	
#1	AML*1	initial	del(5)(q14q33), -17, add(18)(q22),+mar1,+mar2,+mar3	<i>ASXL1</i>	c.2317G>T p.(Glu773*)	10.83%	not known
				<i>DNMT3A</i>	c.2645G>A p.(Arg882His)	18.69%	
				<i>IDH1</i>	c.394C>T p.(Arg132Cys)	13.85%	
				<i>TP53</i>	c.427G>A p.(Val143Met)	20.49%	
				<i>BCOR</i>	c.4639+1G>A	66.76%	
#2	MDS*2	initial	+8,+10	<i>DNMT3A</i>	c.2645G>A p.(Arg882His)	39.43%	Azacitidine plus Venetoclax
				<i>KRAS</i>	c.34G>C p.(Gly12Arg)	4.03%	
				<i>SF3B1</i>	c.2098A>G p.(Lys700Glu)	1.31%	
				<i>STAG2</i>	c.3362_3365dup p.(Ser1123Hisfs*14)	52.12%	
				<i>U2AF1</i>	c.101C>T p.(Ser34Phe)	31.50%	
				<i>NF1</i>	c.2033del p.(Pro678Arg*10)	49.52%	
#3	t-MN*6	initial	t(9;20)(q11;q11),t(12;22)(p13;q11)	<i>NF1</i>	c.2033del p.(Pro678Arg*10)	49.52%	CPY-351, allogenic stem cell planned
#4	t-MN	initial	del(5)(q14q34),+8,i(8)(q10)x2	<i>IDH2</i>	c.419G>A p.(Arg140Gln)	46.79%	Azacitidine plus Venetoclax
				<i>RUNX1</i>	c.420T>G p.(Ser140Arg)	43.00%	
				<i>SRSF2</i>	c.284C>T p.(Pro95Leu)	39.78%	
				<i>TET2</i>	c.1455del p.(Asn486Thr*11)	47.20%	
				<i>TET2</i>	c.3473del p.(Ala1158Glu*68)	41.43%	
				<i>TP53</i>	c.844C>T p.(Arg282Trp)	90.92%	
				<i>IDH2</i>	c.419G>A p.(Arg140Gln)	33.28%	
				<i>KRAS</i>	c.35G>C p.(Gly12Ala)	1.96%	
		progression	t(2;3)(p23;q27),del(5)(q14q34),del(7)(q21),+8,i(8)(q10)x2	<i>NRAS</i>	c.35G>C p.(Gly12Ala)	1.70%	
				<i>PTPN11</i>	c.1508G>C p.(Gly503Ala)	2.80%	
				<i>RUNX1</i>	c.420T>G p.(Ser140Arg)	38.29%	
				<i>SRSF2</i>	c.284C>T p.(Pro95Leu)	40.88%	
				<i>TET2</i>	c.1398_1402dup p.(His468Leu*20)	5.11%	
				<i>TET2</i>	c.1455del p.(Asn486Thr*11)	44.10%	
				<i>TET2</i>	c.3473del p.(Ala1158Glu*68)	36.71%	
				<i>TP53</i>	c.844C>T p.(Arg282Trp)	74.41%	
#5	AML	initial	t(4;14;11)(q22;q32;q23),add(10)(p14)	<i>ASXL1</i>	c.2077C>T p.(Arg693*)	0.68%	7+3
				<i>KRAS</i>	c.38G>A p.(Gly13Asp)	0.35%	
				<i>NRAS</i>		0.50%	

(Continued)

TABLE 1 | Continued

patient (#)	WHO diagnosis	status	cytogenetic aberrations	DNA analysis NGS (somatic)			treatment		
				gene	variant	variant allele frequency (VAF)			
#6	initial: MPN*3; 6-years later: suspicion on sAML*4	progression	t(4;14;11)(q22;q32;q23),+der(4)t(4;14;11),+8,+9,add(10)(p14),+19,+21	ASXL1	c.182A>G p.(Gln61Arg)	7.84%	Hydroxyurea & phlebotomy/blood letting; -6 years later: Jakavi		
		initial	-	JAK2	c.2077C>T p.(Arg693*)	7.59%			
					KRAS	c.38G>A p.(Gly13Asp)		5.34%	
		Progression	add(5)(q12),-7,-13,del(14)(q12q31),der(17)t(13;17)(q21;p12)	JAK2	c.182A>G p.(Gln61Arg)	73.76%			
					JAK2	c.1849G>T p.(Val617Phe)		97.53%	
		progression	add(2)(q37),add(5)(q12),-7,-13,del(14)(q12q31),der(17)t(13;17)(q21;p12),del(20)(q12q13)	TP53	c.1849G>T p.(Val617Phe)	19.56%			
					TP53	c.814G>A p.(Val272Met)		98.12%	
		#7	sAML from atypical CML*5	initial	ins(9;12)(q34;p12p13),+12	no variants detected			7+3 plus Dasatinib, allogenic Tx
				progression	+X,ins(9;12)(q34;p12p13),+11,+12,del(12)(p13),+19	no variants detected			
				remission	-	no variants detected			
relapse	der(7)t(7;9)(q35;q21),ins(9;12)(q34;p12p13),+12			no variants detected					
#8	AML without maturation	initial	del(9)(q21q31)	NRAS	c.37G>C p.(Gly13Arg)	12.69%	7+3, allogenic transplantation, after relapse: FLA-V-IDA (=FLAG-IDA with Venetoclax) plus donor lymphocytes		
		relapse	t(1;16)(p12;q21),del(9)(q21q31)	WT1	c.1136_1142dup p.(Ala382Thr*5)	23.26%			
					no variants detected				
		remission	-	no variants detected					
		progression	t(1;16)(p12;q21),?add(9)(p12),del(9)(q21q31),add(10)(p12),add(17)(q22)	NRAS	c.37G>C p.(Gly13Arg)	12.78%			
					NRAS	c.37G>T p.(Gly13Cys)		5.03%	
				WT1	c.1141_1144dup p.(Ala382Val*4)	4.51%			
					WT1	c.1136_1142dup p.(Ala382Thr*5)		11.35%	
				WT1	c.1128dup p.(Thr377Asp*8)	1.48%			
					WT1	c.1110dup p.(Val371Cys*14)		3.69%	
				NRAS	c.37G>C p.(Gly13Arg)	17.38%			
					NRAS	c.37G>T p.(Gly13Cys)		4.93%	
				WT1	c.1141_1144dup p.(Ala382Val*4)	4.03%			
					WT1	c.1136_1142dup p.(Ala382Thr*5)		17.18%	
		WT1	c.1128dup p.(Thr377Asp*8)	4.76%					
			WT1	c.1110dup p.(Val371Cys*14)	4.96%				
remission	-	no variants detected							

*1 AML, acute myeloid leukemia; *2 MDS, myelodysplastic syndrome; *3 MPN, myeloproliferative neoplasm; *4 sAML, secondary AML; *5 CML, chronic myeloid leukemia

further characterized (e.g. presence in common databases and in silico effect prediction). Combining all information, an artifact- and a polymorphism score is calculated for every SNV and small indel. The scores allow for automatic classification of the calls as true variants, polymorphisms and artifacts. Additionally, all detected variants were manually checked using the Integrative Genomics Viewer (IGV) to confirm automatic classification (25, 26). All detected variants were classified according to the standards and guidelines of the American College of Medical Genetics and Genomics (ACMG) (27). Here, we report only variants of unknown significance (VUS), likely pathogenic and pathogenic variants.

Bioinformatic Approach

SNVs and short indels were detected using targeted NGS and variant calling pipeline appreci8 (28). SVs and CNVs were detected performing karyotyping, FISH and array-CGH analysis. Clonal evolution was reconstructed manually, integrating all mutational information on SNVs, indels, SVs and CNVs as described in Reutter et al. (29). The results were visualized by fishplots (30). In short, percentages of cancer cell fractions (CCFs) for SNVs and indels were estimated based on VAFs, assuming heterozygous variants ($2 \cdot \text{VAF} = \text{CCF}$). Percentages of CCFs for SVs and CNVs were estimated based on cell counts reported for karyotyping and FISH analyses. For SVs and CNVs that were only detected in array-CGH percentages of CCF were estimated based on \log_2 Ratio of fluorescence intensities for reference versus tumor probes. In case of CNVs overlapping the position of an SNV or indel, the calculation of CCF is less straightforward. Altogether, we differentiate between three possible scenarios:

Scenario 1: The CNV occurred prior to the SNV/indel, but in the same cells.

Scenario 2: The SNV/indel occurred prior to the CNV, but in the same cells.

Scenario 3: SNV/indel and CNV exist in parallel, independent of each other.

We defined w as the ratio of cells featuring a CNV and an SNV/indel. Analogously, x , y and z were defined in the same manner. The sum over all cell ratio was always $w+x+y+z=1.0$. Furthermore, the known CCF for the CNV was defined as $\text{CCF}_{\text{CNV}}=w+x$. Additionally, two formulas were derived from the model (see **Figure 1A**):

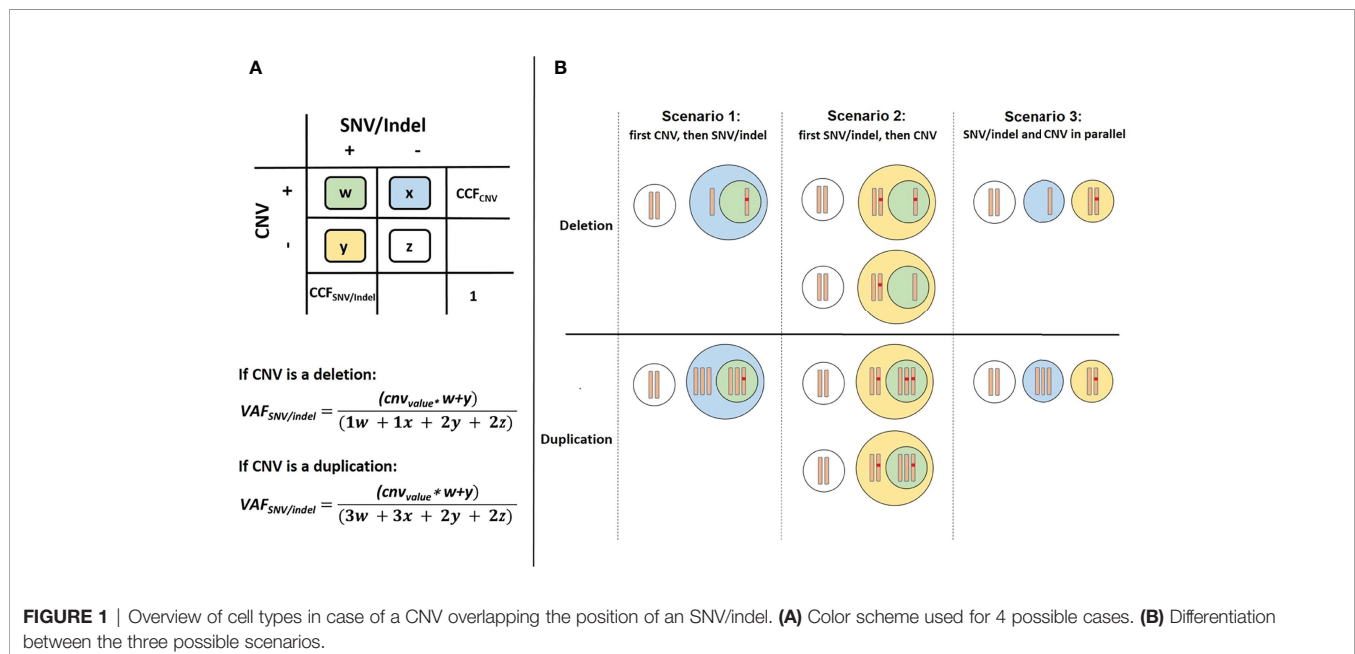
$$\text{VAF}_{\text{SNV/indel}} = (\text{cnv}_{\text{value}} \cdot w + y) / \quad (1)$$

$$(1 \cdot w + 1 \cdot x + 2 \cdot y + 2 \cdot z) \quad \text{if CNV is deletion}$$

$$\text{VAF}_{\text{SNV/indel}} = (\text{cnv}_{\text{value}} \cdot w + y) / \quad (2)$$

$$(3 \cdot w + 3 \cdot x + 2 \cdot y + 2 \cdot z) \quad \text{if CNV is duplication}$$

We defined $\text{cnv}_{\text{value}}$ as a parameter to model the number of alleles with the SNV/indel depending on the CNV. For scenarios 1 and 3 the $\text{cnv}_{\text{value}}=1$ and for scenario 2 the $\text{cnv}_{\text{value}}$ can be 0, 1 or 2. Dependent on the three scenarios we considered, the two equations may be further simplified: For scenario 1, the CNV occurring prior to the SNV/indel, $y=0$ (see **Figure 1B**). Thus, $\text{CCF}_{\text{CNV}}=w+x$ and $\text{CCF}_{\text{SNV/indel}}=w$. As $\text{VAF}_{\text{SNV/indel}}$ is also known, w and $\text{CCF}_{\text{SNV/indel}}$ can be easily determined. For scenario 2, a distinction of cases has to be made. Both, the allele with or without SNV/indel may be affected by CNV. Regarding deletions, $\text{cnv}_{\text{value}}=0$ if the allele with SNV/indel is deleted, and $\text{cnv}_{\text{value}}=1$ if the allele without SNV/indel is deleted. Regarding duplications, $\text{cnv}_{\text{value}}=2$ if the allele with SNV/indel is duplicated, and $\text{cnv}_{\text{value}}=1$ if the allele without SNV/indel is duplicated. Additionally, $x=0$, so $\text{CCF}_{\text{CNV}}=w$. The unknown



value of y and thus $CCF_{SNV/indel}=w+y$ can be determined using $VA_{SNV/indel}$ following the distinction of cases. For the third scenario of SNV/indel and CNV occurring in parallel, independent of each other, $w=0$. Thus, $CCF_{CNV}=x$ and $CCF_{SNV/indel}=y$. As $VA_{SNV/indel}$ is known, the unknown value of $CCF_{SNV/indel}$ can be determined. Although $CCF_{SNV/indel}$ can always be determined using the above-mentioned formulas, the result may not always be distinct. Information on whether the CNV occurred prior or after the SNV/indel is in general not available. Therefore, for application of our approach, all scenarios have to be considered and four possible values for $CCF_{SNV/indel}$ have to be calculated. Scenarios may be excluded if the basic assumption $w+x+y+z=1.0$ is violated.

RESULTS

Eight patients with myeloid neoplasia were comprehensively analyzed using karyotyping, fluorescence *in situ* hybridization (FISH), and a custom NGS panel. In addition, array-CGH analysis was performed whenever possible. All patients showed clonal evolution with many genomic variants, either at the initial time point and/or during the course of the disease.

Identification of Linear Clonal Evolution in Patients Analyzed at One Time Point

In order to show that clonal evolution can be reconstructed and visualized even when only one time point is available, we examined three patients (patient 1, 2 and 3) with the aforementioned bioinformatic approach. The total number of aberrations in a patient ranged from 3 (patient 3) to 11 (patient 1). Due to this approach it was possible to determine clones by clustering mutations. These clones were sorted according to CCF, in order of their appearance. Through reconstruction and visualization of clonal evolution it was possible to identify the initial clone (stemline) as well as the passenger aberrations for all patients. For example, patient 1 (AML) showed a deletion in 5q (del(5)(q14q33)) in the initial clone, patient 2 (MDS) showed a trisomy 8 as well as a pathogenic variant in *DNMT3A* in the initial clone, and patient 3 (therapy related myeloid neoplasm) showed a translocation t(12;22) in combination with a pathogenic variant in *NFI* in the initial clone (see **Figure 2** and **Table 1/S1**). The total number of clones in a patient ranged from 2 (patient 3) to 7 (patient 1). In summary, our approach was successful in identifying linear clonal evolution in all patients.

Identification of Different Clonal Evolution Models (Linear, Branching or Neutral) in Patients Analyzed at More Than One Time Point

Next, we analyzed five patients where data was available for more than one time point (2 to 5 time points of analysis) to show that the bioinformatic approach can reconstruct different models of clonal evolution. At each time point, all patients were analyzed by a combination of methods, including karyotyping, FISH and a custom NGS-Panel. The estimated therapy effect was included in the reconstruction of clonal evolution. The approach reconstructed

a distinct model of clonal evolution for patients 5, 7 and 8 (see **Figures 2C–E**). Patient 5 was analyzed at two time points – initial and progression. The bioinformatic approach reconstructed a linear clonal evolution pattern. The initial clone showed a translocation t(4;14;11) and a structural aberration in chromosome arm 10p (add(10)(p14)). At this time point, pathogenic variants in *KRAS*, *ASXL1* and *NRAS* were identified with a very low percentage of CCF (close to 1%). However, the pathogenic variants were only identified by deep sequencing. At the second time point (progression), additional chromosome aberrations (+der(4)(4;14;11), +8, +9, +19 and +21) were identified and this subclone became the mainline (see **Figure 2D**). Patient 7 was analyzed at four time points – initial, progression, remission and relapse. Stem cell transplantation was performed between the third and fourth time point of analysis. The initial clone showed an insertion of a segment of chromosome 12 into chromosome 9 (ins(9;12)) and a trisomy 12. At time point two, the percentage of cells only harboring the stemline's mutations were greatly reduced (~5%) and a subclone developed. This clone showed a deletion in 12p (del(12)(p13)) and a trisomy 11, 19 and X. For the subclone, the approach calculated linear clonal evolution. At time point three, the subclone was no longer detectable, but a new subclone was identified at time point four. This subclone presented a derivative chromosome 7 with a translocation t(7;9) (der(7)t(7;9)(q35;q21)). At this time point, the clonal evolution pattern of patient 7 changed from linear to branched (see **Figure 2E**). Patient 8 was analyzed at five time points – initial, remission, progression, relapse and remission. Stem cell transplantation took place between the first and second time point. The initial clone showed a deletion in 9q (del(9)(q21q31)) and a subclone with a likely pathogenic variant in *WT1* as well as a pathogenic variant in *NRAS*. At this time point, a linear evolution pattern was reconstructed. At time point two, after stem cell transplantation, no variants were detected – neither by karyotyping nor deep sequencing. However, the aforementioned variants were present at time point three (progression). At this time point, a linear clonal evolution pattern was reconstructed in one subclone (light red). Furthermore, four additional subclones were observed. All subclones developed independently of the initial clone with a deletion in 9q (del(9)(q21q31)). At the fourth time point of analysis, three additional subclones were detected. The clonal evolution pattern of patient 8 changed from linear to neutral. At the last time point of analysis (remission), no variants were detected (see **Figure 2F**). In summary, a linear evolution pattern was reconstructed for patient 5, a branched evolution pattern was reconstructed for patient 7 and a neutral evolution pattern was reconstructed for patient 8.

Reconstruction of clonal evolution in patients 4 and 6 showed more than one possible clonal evolution pattern. All possible versions of reconstruction are shown in **Figure 3**. Patient 4 was analyzed at two time points (initial and progression). The initial clone showed a deletion in 5q (del(5)(q14q34)). Additionally, four subclones developed. At this time point, patient 4 showed a linear clonal evolution pattern. At time point two, four additional subclones were detected and the evolution pattern changed from a linear to a branched evolution pattern. Although the model of clonal evolution might be clear, it was not possible to reconstruct

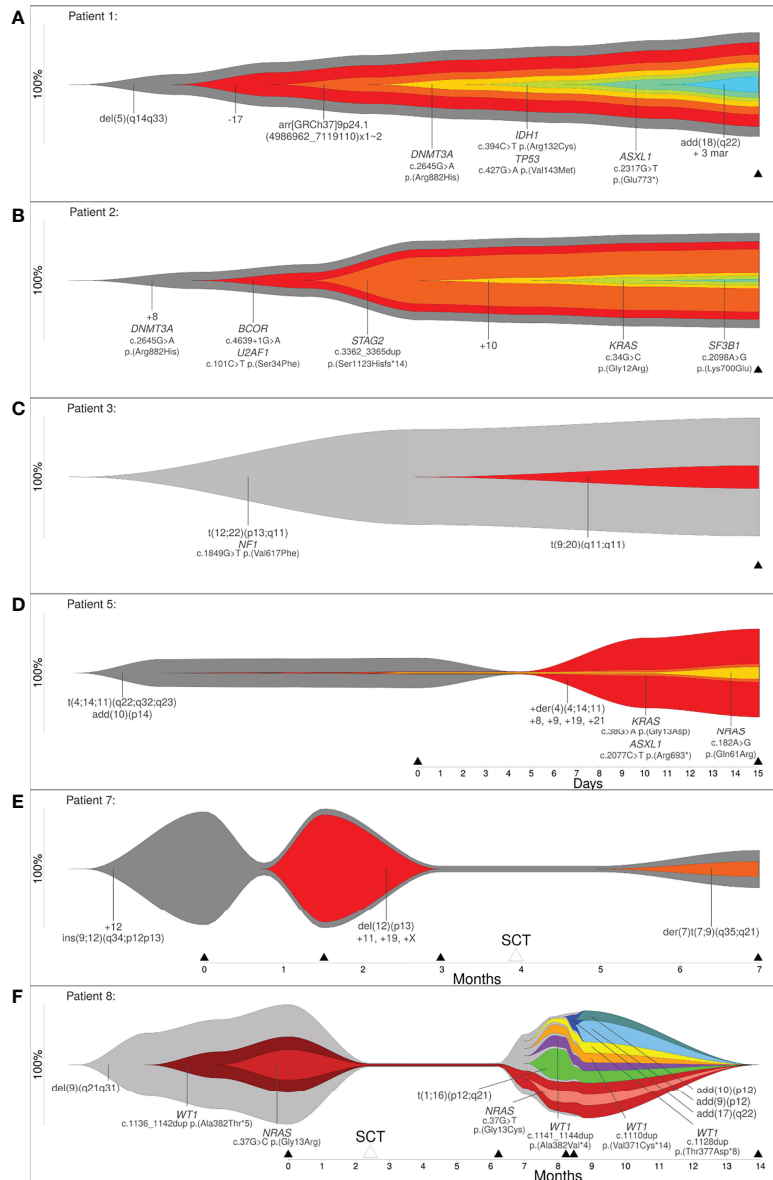


FIGURE 2 | (A–C) Reconstruction of clonal evolution of three patients analyzed at one time point. **(D–F)** Clonal evolution patterns of patient 5 (linear evolution), patient 7 (branched evolution) and patient 8 (neutral evolution). Y-axis represents the percentage of clone size, black triangles indicate the analyzed time point and white triangles indicate time point of stem cell transplantation.

the precise subclonal relationship. The smallest subclone with (likely) pathogenic variants in *KRAS*, *NRAS* and *PTPN11* showed a low percentage of CCF (5%). Based on the available data, the parent clone of the smallest subclone cannot be definitely determined and, therefore, patient 4 showed three possible versions of branched evolution patterns (see **Figure 3A**). Patient 6 was analyzed at three time points – initial and two progression time points. The initial clone showed a pathogenic variant in *JAK2*. At the second time point of analysis, we identified two events: (1) a likely pathogenic variant in *TP53* and (2) a derivative chromosome 17 resulting in loss of *TP53* (der

(17)t(13;17)(q21;p12)) with further aberrations. Based on the available data, it was not possible to reconstruct which of the two events took place first. Furthermore, it was not possible to calculate if both events affected the same allele or different alleles. As described in Materials and Methods, four different cases have to be considered (see **Figure 1A**). Two out of four cases were in line with the basic assumption $w+x+y+z=1.0$. Both cases were confirmed by the observed percentages of CCF and VAFs at the time points one and two of analysis. According to the first case, the CNV event took place first, followed by the likely pathogenic variant in *TP53*. Both aberrations affected different alleles. For

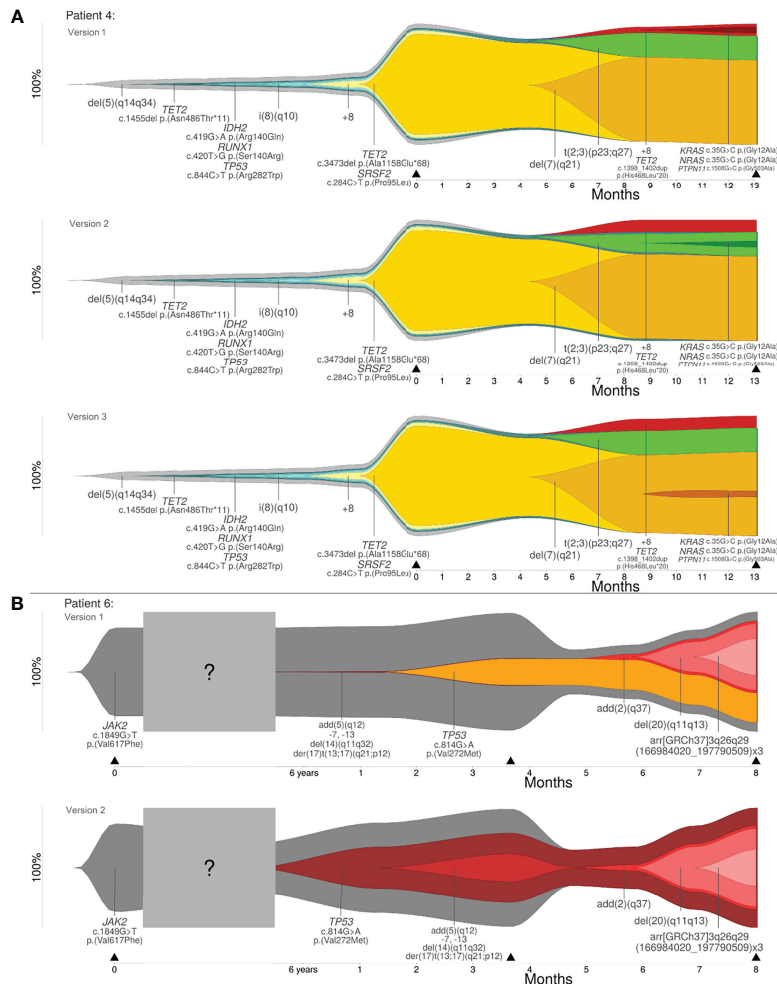


FIGURE 3 | (A) Clonal evolution patterns of patient 4: three versions of branched evolution. **(B)** Clonal evolution patterns of patient 6: linear and branched evolution. Y-axis represents the percentage of clone size and black triangles indicate the analyzed time point.

this case, a branched clonal evolution pattern was reconstructed (see **Figure 3B**). Mathematically, it is also possible that the point mutation in *TP53* took place first and subsequently the mutated allele was affected by the deletion. For this case, a linear clonal evolution pattern was reconstructed. However, a biallelic inactivation of *TP53* frequently occurs during disease progression of hematologic neoplasms (e.g. MDS) (31). In summary, three possible branched evolution patterns were reconstructed for patient 4, and two models of clonal evolution (linear and branching) were reconstructed for patient 6.

DISCUSSION

In this study we propose a bioinformatic approach to analyze clonal evolution and the genetic architecture comprising SNVs, indels and SVs, especially CNVs in patients with myeloid neoplasia. In order to show the possibility of reconstruction and

visualization of clonal evolution at one time point of analysis, we analyzed three patients with myeloid neoplasia. Here, all patients showed a linear clonal evolution pattern. Furthermore, with this bioinformatic approach, we addressed all alterations that may play a role in the pathogenesis of the disease (driver) and alterations, which may occur during disease development (passenger). For example, patient 1 and 2 showed a *DNMT3A*^{R882} mutation at the initial time point of analysis. Pathogenic variants in *DNMT3A* occur in ~30% of AML patients and studies of clonal architecture have shown that pathogenic variants in *DNMT3A* occur as an early event in leukemogenesis. For that reason, *DNMT3A* could be a possible therapeutic target in the future (32).

In recent years, several targeted therapies have been approved for the treatment of patients with myeloid neoplasia. For example, the development of genome sequencing has identified numerous somatic alterations in patients with AML. Some of these alterations (“actionable mutations”) can be targeted by specific drugs to improve the outcome of the patients (33, 34).

For example, patient 1 had an *IDH1*^{R132} mutation, which can be found in 5-10% of AML patients. Patients with an *IDH1*^{R132} mutation can be treated with ivosidenib. A recent study has shown that in these patients single-agent ivosidenib treatment leads to a complete remission (CR) or CR with incomplete haematological recovery rates of 42.4% (5). A phase trial study has analyzed a cohort of AML patients with a pathogenic variant in *IDH1*, which were treated with ivosidenib-and-azacitidine or with placebo-and-azacitidine. Patients treated with ivosidenib-and-azacitidine showed a significantly longer event-free survival and overall survival (24.0 month vs. 7.9 month) in comparison to patients treated with placebo-and-azacitidine (35). In addition, Bolton et al. have analyzed the clonal hematopoiesis in patients with therapy-related myeloid neoplasms. The study has shown that pathogenic variants show a different clonal evolution based on exposures (e.g. radiation, topoisomerase II inhibitors), for example, clones with pathogenic variants in DNA damage response genes outcompetes other clones (36). Thus, reconstruction of clonal evolution as well as identification of driver and passenger alterations allow a better understanding of the mechanisms of leukemogenesis and improvement of (future) treatment strategies (personalized therapy).

We analyzed five patients at more than one time point to show that the approach can reconstruct different models of clonal evolution (i.e. linear, branching and neutral evolution). For all of these patients the estimated therapy effect was included. The analyzed data of three out of these five patients allowed a distinct reconstruction of clonal evolution (patients 5, 7 and 8). Linear evolution is characterized by a dominant clone which overgrows the ancestral clone after acquisition of additional mutations in a stepwise manner (3, 4) - a linear mutation pattern was reconstructed for patient 5. Branching evolution is characterized by the occurrence of different subclones from one ancestral clone (3, 4) - as detected in patient 7. Neutral evolution is an extreme case of branching evolution, which is characterized by the accumulation of random mutations over time (16) - as was observed in patient 8. A possible advantage of reconstructing clonal evolution could be to show if both pathogenic variants co-occur in the same clone or in different clones. This could influence the treatment strategy. For example, patient 8 showed a pathogenic variant in *NRAS* and in *WT1*. Pathogenic variants in *NRAS* occur in ~12% of all AML cases and co-occur with pathogenic variants in epigenetic modifiers (*TET2/IDH/WT1*). The co-occurrence of pathogenic variants in *NRAS* and *WT1* showed a sensitivity to MAPK kinase inhibition, which has been shown in patient samples and mouse models providing a possible treatment strategy in the future (32). Especially patients with AML show a high number of heterogeneous pathogenic variants at diagnosis and relapse. These pathogenic variants are organized into a hierarchy of clones and they are able to adapt and evolve in response to therapeutic pressure (5). In the future, the clonal evolution pattern and genetic architecture of many patients must be analyzed in order to learn more about therapy resistance and new possibilities for therapy strategies to improve patient outcome.

For the remaining two patients (patient 4 and 6) the analyzed data showed more than one possible clonal evolution pattern.

The proposed bioinformatic approach can be used to analyze the order and occurrence of genetic alterations, which could be relevant for treatment decisions and for the understanding of leukemic progression. For example, patient 6 showed a *JAK2/TP53* clone at the transformation to AML, which replaced the dominant *JAK2* clone in the myeloproliferative neoplasia (initial diagnosis). This observation is in concordance with a recent publication, where a significant alteration in clonal architecture or a “clonal sweep” with emergence of new dominant clone(s) was observed with single cell sequencing (17). These observations can be used for therapy stratification during the disease course.

Altogether, we present a generalizable approach that enables the combination of a wide spectrum of methods and to analyze retrospective samples to visualize leukemic progression and genetic architecture. Recent publications have reconstructed clonal evolution with single cell sequencing where the available data has allowed definitive identification of clonal architecture (17, 19, 37). Single cell sequencing has many advantages: exact distinction of different clones, measurement of accurate clonal complexity or resolution of mutational order (17, 38). But there are also many disadvantages, including limited depth of sequencing, no possibility of analyzing retrospective samples (no suitable material), or missing information of chromosomal aberrations (important for treatment stratification). Even though the used methods in this project do not offer the possibility to analyze clonal evolution at a single cell level, the inclusion of cytogenetic and molecular genetic alterations allow the identification of driver and passenger alterations, which is currently impossible using only single cell sequencing. For future projects, a combination of methods (bulk and single cell), including karyotyping and single cell sequencing, could be helpful for the exact reconstruction of clonal evolution.

In conclusion, this bioinformatic approach offers the possibility of analyzing clonal evolution and the order and occurrence of many cytogenetic and molecular genetic alterations (genetic architecture) at one or more time points of analysis. Different models of clonal evolution (i.e. linear, branching and neutral) can be reconstructed with this approach. As the approach describes integration of data for reconstruction of clonal evolution in general, its application is not limited to myeloid neoplasms. Instead, it may be applied on all kinds of non-solid tumors showing clonal evolution. The visualization of the results in fishplots contributes to a better understanding of genetic architecture and leukemic progression. This approach helps to identify possible targets for the disease (personalized therapy) and can be used to identify markers in order to assess minimal residual disease.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: PRJNA847150 link to <https://www.ncbi.nlm.nih.gov/sra/PRJNA847150>.

ETHICS STATEMENT

The study was conducted according to the guidelines of the Declaration of Helsinki. The ethical review boards of Hannover Medical School approved this study (Nr. 8657_BO_K_2019) and all patients gave their written consent. Informed consent was obtained from all subjects involved in the study. Reported results can be found in **Supplementary Materials**.

AUTHOR CONTRIBUTIONS

The study was conceived by SS, YB, and GG. Data and material of the study was generated by YB, GG, SS, DS, FT, MH, AC, DD, FL, JV, MD, and BS. Data collection was performed by SS, YB, GG. Interpretation and analysis were conducted by YB, GG, SS, MD, and BS. SS and YB wrote the manuscript. The paper was edited by all authors.

REFERENCES

- Doulatov S, Papapetrou EP. Studying Clonal Evolution of Myeloid Malignancies Using Induced Pluripotent Stem Cells. *Curr Opin Hematol* (2021) 28(1):50–6. doi: 10.1097/MOH.0000000000000620
- E Swerdlow SH Campo, NL Harris, ES Jaffe, SA Pileri, H Stein and J Thiele eds. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. Revised 4th edition*. Lyon, France: International Agency for Research on Cancer (2017).
- da Silva-Coelho P, Kroeze LI, Yoshida K, Koorenhof-Scheele TN, Knops R, van de Locht LT, et al. Clonal Evolution in Myelodysplastic Syndromes. *Nat Commun* (2017) 8:15099. doi: 10.1038/ncomms15099
- Vosberg S, Greif PA. Clonal Evolution of Acute Myeloid Leukemia From Diagnosis to Relapse. *Genes Chromosomes Canc* (2019) 58(12):839–49. doi: 10.1002/gcc.22806
- Dohner H, Wei AH, Lowenberg B. Towards Precision Medicine for AML. *Nat Rev Clin Oncol* (2021) 18(9):577–90. doi: 10.1038/s41571-021-00509-w
- Papaemmanuil E, Gerstung M, Bullinger L, Gaidzik VI, Paschka P, Roberts ND, et al. Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N Engl J Med* (2016) 374(23):2209–21. doi: 10.1056/NEJMoa1516192
- Metzler KH, Herold T, Rothenberg-Thurley M, Amler S, Sauerland MC, Gorlich D, et al. Spectrum and Prognostic Relevance of Driver Gene Mutations in Acute Myeloid Leukemia. *Blood* (2016) 128(5):686–98. doi: 10.1182/blood-2016-01-693879
- Haferlach T, Nagata Y, Grossmann V, Okuno Y, Bacher U, Nagae G, et al. Landscape of Genetic Lesions in 944 Patients With Myelodysplastic Syndromes. *Leukemia* (2014) 28(2):241–7. doi: 10.1038/leu.2013.336
- Papaemmanuil E, Gerstung M, Malcovati L, Tauro S, Gundem G, Van Loo P, et al. Clinical and Biological Implications of Driver Mutations in Myelodysplastic Syndromes. *Blood* (2013) 122(22):3616–27; quiz 99. doi: 10.1182/blood-2013-08-518886
- Palomo L, Meggendorfer M, Hutter S, Twardziok S, Adema V, Fuhrmann I, et al. Molecular Landscape and Clonal Architecture of Adult Myelodysplastic/Myeloproliferative Neoplasms. *Blood* (2020) 136(16):1851–62. doi: 10.1182/blood.2019004229
- Makishima H, Yoshizato T, Yoshida K, Sekeres MA, Radivoyevitch T, Suzuki H, et al. Dynamics of Clonal Evolution in Myelodysplastic Syndromes. *Nat Genet* (2017) 49(2):204–12. doi: 10.1038/ng.3742
- Onecha E, Rapado I, Luz Morales M, Carreno-Tarragona G, Martinez-Sanchez P, Gutierrez X, et al. Monitoring of Clonal Evolution of Acute Myeloid Leukemia Identifies the Leukemia Subtype, Clinical Outcome and Potential New Drug Targets for Post-Remission Strategies or Relapse. *Haematologica* (2021) 106(9):2325–33. doi: 10.3324/haematol.2020.254623
- Menssen AJ, Walter MJ. Genetics of Progression From MDS to Secondary Leukemia. *Blood* (2020) 136(1):50–60. doi: 10.1182/blood.2019000942
- Morita K, Wang F, Jahn K, Hu T, Tanaka T, Sasaki Y, et al. Clonal Evolution of Acute Myeloid Leukemia Revealed by High-Throughput Single-Cell Genomics. *Nat Commun* (2020) 11(1):5327. doi: 10.1038/s41467-020-19119-8
- McGranahan N, Swanton C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. *Cell* (2017) 168(4):613–28. doi: 10.1016/j.cell.2017.01.018
- Davis A, Gao R, Navin N. Tumor Evolution: Linear, Branching, Neutral or Punctuated? *Biochim Biophys Acta Rev Cancer* (2017) 1867(2):151–61. doi: 10.1016/j.bbcan.2017.01.003
- Miles LA, Bowman RL, Merlinsky TR, Csete IS, Ooi AT, Durruthy-Durruthy R, et al. Single-Cell Mutation Analysis of Clonal Evolution in Myeloid Malignancies. *Nature* (2020) 587(7834):477–82. doi: 10.1038/s41586-020-2864-x
- Weber LL, El-Kebir M. Distinguishing Linear and Branched Evolution Given Single-Cell DNA Sequencing Data of Tumors. *Algorith Mol Biol* (2021) 16(1):14. doi: 10.1186/s13015-021-00194-5
- Potter N, Miraki-Moud F, Ermini L, Titley I, Vijayaraghavan G, Papaemmanuil E, et al. Single Cell Analysis of Clonal Architecture in Acute Myeloid Leukaemia. *Leukemia* (2019) 33(5):1113–23. doi: 10.1038/s41375-018-0319-2
- Akkari YMN, Baughn LB, Dubuc AM, Smith AC, Mallo M, Dal Cin P, et al. Guiding the Global Evolution of Cytogenetic Testing for Hematologic Malignancies. *Blood* (2022) 139(15):2273–84. doi: 10.1182/blood.2021014309
- Schlegelberger B, Metzke S, Harder S, Zühlke-Jenisch R, Zhang Y, Siebert R. Classical and Molecular Cytogenetics of Tumor Cells. Berlin, Heidelberg: Springer. In: Springer Lab Manual (1999). p. 151–85.
- Schlegelberger B, Zwingers T, Harder L, Nowotny H, Siebert R, Vesely M, et al. Clinicopathogenetic Significance of Chromosomal Abnormalities in Patients With Blastic Peripheral B-Cell Lymphoma. Kiel-Wien-Lymphoma Study Group. *Blood* (1999) 94(9):3114–20. doi: 10.1182/blood.V94.9.3114
- McGowan-Jordan J, Hastings R, Moore S. *An International System for Human Cytogenomic Nomenclature*. Basel: Karger. J McGowan-Jordan, R Hastings and S Moore, editors (2020).
- Li H. *Aligning Sequence Reads, Clone Sequences and Assembly Contigs With BWA-MEM*. Karger. Basel (2013).
- Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): High-Performance Genomics Data Visualization and Exploration. *Brief Bioinform* (2013) 14(2):178–92. doi: 10.1093/bib/bbs017
- Robinson JT, Thorvaldsdottir H, Wenger AM, Zehir A, Mesirov JP. Variant Review With the Integrative Genomics Viewer. *Cancer Res* (2017) 77(21):e31–e4. doi: 10.1158/0008-5472.CAN-17-0337

All authors contributed to the article and approved the submitted version.

FUNDING

GG and BS have been supported by the BMBF MyPred (01GM1911A).

ACKNOWLEDGMENTS

We would like to thank Dr. Marius Wöste for his input on analyzing and interpreting array-CGH data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2022.888114/full#supplementary-material>

27. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* (2015) 17(5):405–24. doi: 10.1038/gim.2015.30
28. Sandmann S, Karimi M, de Graaf AO, Rohde C, Gollner S, Varghese J, et al. Appreci8: A Pipeline for Precise Variant Calling Integrating 8 Tools. *Bioinformatics* (2018) 34(24):4205–12. doi: 10.1093/bioinformatics/bty518
29. Reutter K, Sandmann S, Rohde J, Muller S, Woste M, Khanam T, et al. Reconstructing Clonal Evolution in Relapsed and non-Relapsed Burkitt Lymphoma. *Leukemia* (2021) 35(2):639–43. doi: 10.1038/s41375-020-0862-5
30. Miller CA, McMichael J, Dang HX, Maher CA, Ding L, Ley TJ, et al. Visualizing Tumor Evolution With the Fishplot Package for R. *BMC Genomics* (2016) 17(1):880. doi: 10.1186/s12864-016-3195-z
31. Bernard E, Nannya Y, Hasserjian RP, Devlin SM, Tuechler H, Medina-Martinez JS, et al. Implications of TP53 Allelic State for Genome Stability, Clinical Presentation and Outcomes in Myelodysplastic Syndromes. *Nat Med* (2020) 26(10):1549–56. doi: 10.1038/s41591-020-1008-z
32. Kishtagari A, Levine RL, Viny AD. Driver Mutations in Acute Myeloid Leukemia. *Curr Opin Hematol* (2020) 27(2):49–57. doi: 10.1097/MOH.0000000000000567
33. Collignon A, Hospital MA, Montersino C, Courtier F, Charbonnier A, Saillard C, et al. A Chemogenomic Approach to Identify Personalized Therapy for Patients With Relapse or Refractory Acute Myeloid Leukemia: Results of a Prospective Feasibility Study. *Blood Cancer J* (2020) 10(6):64. doi: 10.1038/s41408-020-0330-5
34. Carr TH, McEwen R, Dougherty B, Johnson JH, Dry JR, Lai Z, et al. Defining Actionable Mutations for Oncology Therapeutic Development. *Nat Rev Canc* (2016) 16(5):319–29. doi: 10.1038/nrc.2016.35
35. Montesinos P, Recher C, Vives S, Zarzycka E, Wang J, Bertani G, et al. Ivosidenib and Azacitidine in IDH1-Mutated Acute Myeloid Leukemia. *N Engl J Med* (2022) 386(16):1519–31. doi: 10.1056/NEJMoa2117344
36. Bolton KL, Ptashkin RN, Gao T, Braunstein L, Devlin SM, Kelly D, et al. Cancer Therapy Shapes the Fitness Landscape of Clonal Hematopoiesis. *Nat Genet* (2020) 52(11):1219–26. doi: 10.1038/s41588-020-00710-0
37. Pellegrino M, Sciambi A, Treusch S, Durruthy-Durruthy R, Gokhale K, Jacob J, et al. High-Throughput Single-Cell DNA Sequencing of Acute Myeloid Leukemia Tumors With Droplet Microfluidics. *Genome Res* (2018) 28(9):1345–52. doi: 10.1101/gr.232272.117
38. Brierley CK, Mead AJ. Single-Cell Sequencing in Hematology. *Curr Opin Oncol* (2020) 32(2):139–45. doi: 10.1097/CCO.0000000000000613

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Sandmann, Behrens, Davenport, Thol, Heuser, Dörfel, Löhr, Castrup, Steinemann, Varghese, Schlegelberger, Dugas and Göhring. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.