haematologica

Journal of the Ferrata Storti Foundation

Long-read sequencing resolves *TP53* allelic status in myeloid neoplasms

by Mette Zeuthen, Louise Serup Christoffersen, Mathis Hildonen, Ulf Birkedal, Lykke Grubach, Lasse Kjær, Vibe Skov, Jack Bernard Cowland and Mette Klarskov Andersen

Received: December 30, 2025.
Accepted: April 31, 2026.

Citation: Mette Zeuthen, Louise Serup Christoffersen, Mathis Hildonen, Ulf Birkedal, Lykke Grubach, Lasse Kjær, Vibe Skov, Jack Bernard Cowland and Mette Klarskov Andersen. Long-read sequencing resolves TP53 allelic status in myeloid neoplasms.
Haematologica. 2026 Apr 16. doi: 10.3324/haematol.2025.300463 [Epub ahead of print]

Long-read sequencing resolves *TP53* allelic status in myeloid neoplasms

*Mette Zeuthen[1], *Louise Serup Christoffersen[2], Mathis Hildonen[3], Ulf Birkedal[3,4], Lykke Grubach[2], Lasse Kjær[5], Vibe Skov[5], Jack Bernard Cowland[3] and Mette Klarskov Andersen[3]

*Authors contributed equally and shared first authorship.

[1]Department of Technology, Biomedical Laboratory Science, Faculty of Health, University College Copenhagen, Denmark.

[2]Department of Pathology, Aalborg University Hospital, Aalborg, Denmark.

[3]Department of Clinical Genetics, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark.

[4]Centre for Physical Activity, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark.

[5]Department of Hematology, Zealand University Hospital, Roskilde, Denmark.

Correspondence: M. Zeuthen – meze@kp.dk or LS. Christoffersen – loch@rn.dk

Disclosures: The authors have no conflicts of interest to disclose.

Contributions: MA and JC conceived and designed the study. MZ and LC developed the study concept, conducted the experiments, analysed the data, and drafted the initial manuscript. MZ and LC contributed equally as co-first authors. LK, VS and LG provided essential biological samples and scientific input to the study. UB and MH carried out bioinformatic data processing. All authors contributed to the revision and editing of the manuscript and approved the final submitted version.

Microsoft Copilot was employed for grammar and style checks; all scientific content was verified by the authors.

Data Availability Statement: The data supporting the findings of this study are available from the corresponding authors upon request, subject to ethical and legal restrictions.

Acknowledgements: The authors thank biomedical laboratory scientists Lone Sandbjerg Hindbæk and Anne Marie Høgh Lauridsen at The Department of Clinical Genetics, Rigshospitalet, Copenhagen University Hospital for their laboratory assistance.

Determining *TP53* allelic status is essential for risk stratification in myelodysplastic neoplasms (MDS). In a cohort of 29 patients with multiple *TP53* mutations, long-read sequencing (LRS) demonstrated that monoallelic *TP53* inactivation is rare, whereas multi-hit patterns dominate. These findings support the continued use of conventional short-read sequencing (SRS) together with variant allele frequencies (VAFs) and cytogenetics as reliable tools in routine diagnostics.

The 5th edition of the World Health Organization (WHO) classification of myeloid neoplasms refined the terminology of MDS and emphasized the integration of genetic and morphologic criteria.[1,2] Among its newly defined entities is *MDS with biallelic TP53 inactivation,* which is also incorporated into the International Consensus Classification (ICC).[3]

TP53 mutations - hereafter referring to pathogenic/likely pathogenic variants - are among the most clinically significant genetic alterations in MDS. Although historically considered uniformly high-risk, recent evidence clearly distinguishes between *monoallelic* and *biallelic TP53* inactivation. Monoallelic *TP53* mutations - defined as a single mutation without concurrent deletion or copy-neutral loss of heterozygosity (cnLOH) - have been shown to confer outcomes comparable to *TP53* wild-type.[4] In contrast, biallelic *TP53* inactivation, typically resulting from *multi-hit TP53* alterations, confers adverse prognosis and poor response to allogeneic hematopoietic stem cell transplantation (allo-HSCT), the only potentially curative therapy for patients with MDS.[4]

Accordingly, identification of multiple *TP53* alterations has gained diagnostic and prognostic importance.[4–6] However, comprehensive assessment of the *TP53* allelic status remains challenging in routine laboratories, particularly in cases without deletion of 17p or cnLOH of 17p.[5,7] Standard SRS accurately detects individual mutations but cannot reliably phase variants separated by more than ~200 bp, thereby limiting its ability to distinguish monoallelic from biallelic *TP53* configurations.[8] In contrast, LRS enables phasing of variants across large genomic distances, overcoming this major limitation of conventional SRS.[9]

In this context, we applied LRS to a cohort of MDS patients harboring multiple *TP53* mutations, aiming to determine whether these mutations resided on the same allele (monoallelic) or on different alleles, indicative of multi-hit *TP53* inactivation.

The study cohort comprised 29 patients diagnosed with MDS, each carrying ≥2 *TP53* mutations identified by SRS (total mutations: n = 62), with VAFs ranging from 1% to 53% (Figure 1, Figure S1). SRS panels used in the study provide full coverage of all coding exons, including ±3 bp of the flanking intronic regions of *TP53*. Patients were recruited from three hematology units in Denmark, and the study has been approved by the Regional Committee on Health Research Ethics (Ethics Commission: H-24080399.

Genomic DNA was obtained from two possible sources: mononuclear cells isolated from bone marrow or peripheral blood by density gradient centrifugation, or from whole bone marrow or peripheral blood samples. For each patient, DNA was extracted from only one source using standard protocols. DNA input for PCR amplification was 100 ng (Figure 2, Table S1).

PCR amplification of the full *TP53* coding region (exons 2-11; 7640 bp) was performed using LongAmp® Taq 2X Master Mix. PCR products were verified by gel electrophoresis prior to library preparation according to the Native Barcoding Kit 96 V14 (SQKNBD114.96) protocol. Sequencing was conducted on Oxford Nanopore platforms. MinKNOW software (version 23.04.5) served as the device control interface, and a Phred quality threshold of 10 was applied to ensure data integrity.

Bioinformatic processing included alignment of reads to the hg38 reference genome, sorting and indexing with Samtools, and retention of full-length reads defined by coverage across both ends of the target region. Reads shorter than 7300 nt were filtered out to remove PCR artefacts. Allele-specific analysis was performed using heterozygous SNPs to phase mutations and sort sequencing reads by individual variants, enabling distinction between monoallelic and multi-hit inactivation events. Variant interpretation utilized Integrative Genomics Viewer (IGV), with NM_000546.6 as the reference transcript.

All 62 *TP53* mutations initially detected by SRS were identified by LRS, demonstrating concordance between the two methods. Phasing of long reads enabled direct determination of allelic configuration. Among the 29 samples, only one (3.4%) exhibited monoallelic *TP53* inactivation, whereas the remaining 28 (96.6%) showed multi-hit configurations with mutations located on different alleles.

According to Bernard et al.[4] and the 5th WHO classification[1], cumulative SRS-derived VAFs exceeding 50% are consistent with biallelic *TP53* inactivation. Fourteen of the 28 multi-hit cases exhibited cumulative VAFs above 50%, strongly suggesting that mutations co-occur within the same cellular population, despite inherent statistical uncertainty in VAF-based estimations. In the remaining 14 multi-hit cases, cumulative VAFs remained below 50%, meaning that biallelic inactivation could not be definitively confirmed. These samples may therefore represent either true biallelic events or independent subclones carrying distinct *TP53* mutations, a distinction not resolvable using bulk sequencing alone. Although lower cumulative VAFs reduce - but do not eliminate - the likelihood of subclonal mosaicism, definitive allelic resolution would require single-cell sequencing, which was beyond the scope of this study.[10–12] However, given that multi-hit *TP53* cases with cumulative VAFs above 50% imply involvement of both alleles within the same clone, the predominance of transallelic configurations suggests that similar allelic patterns may extend to cases with lower VAFs, even when formal WHO criteria for biallelic inactivation are not met.

Figure 3 provides an overview of *TP53* allelic configurations in the cohort. The upper panel summarizes allelic configurations among the 29 patients, while the lower panels show representative IGV screenshots illustrating two patterns of *TP53* inactivation: monoallelic mutations at low VAF and multi-hit mutations at higher VAF.

Using long-read phasing across the full *TP53* coding region, we demonstrate that *TP53* mutations in MDS are rarely located on the same allele, with multi-hit configurations representing the predominant allelic architecture. By enabling phasing across the entire *TP53* coding region, LRS provides insight into allelic configurations that cannot be resolved by conventional short-read sequencing. This is particularly relevant for *TP53*, where pathogenic variants may be distributed across distant exons, that lie beyond the phasing capability of standard diagnostic approaches.

While *TP53* mutations predominantly cluster within exons 5-8, corresponding to the DNA-binding domain, mutations outside this region have also been reported.[7,13] Unlike previous studies that focused on short genomic intervals (e.g., 5-76 bp between mutations) or restricted analysis to the DNA-binding domain[8,14], our approach encompassed the entire *TP53* coding sequence (exon 2-11), enabling a more comprehensive assessment. Of the 62 mutations identified, 12 were located outside the DNA-binding domain; notably, 8 of these would have been misclassified as single-hit cases if only exons 5-8 been analyzed and one sample would have been classified as wild-type since both mutations were allocated in exon 4 (Id 29).

Given the poor prognosis associated with biallelic *TP53* inactivation, particularly in patients undergoing allo-HSCT, accurate assessment of allelic status is essential for risk stratification and treatment planning. As cytogenetic data were not included in this study, we cannot exclude the possibility that the one monoallelic case could represent biallelic inactivation if accompanied by del(17p) or cnLOH. However, as the primary aim of the study was to resolve allelic configuration by phasing, the absence of cytogenetic profiling does not affect our main findings.

Our findings indicate that *TP53* mutations in MDS are rarely located on the same allele - even when mutations are distributed across distant exons. The ability to phase mutations across long genomic distances provides novel insights into *TP53* biology and underscores the clinical relevance of allelic status.

MDS cases with two or more *TP53* mutations are rare, which makes our cohort size limited. Nevertheless, a consistent pattern emerges, as 28 out of the included 29 MDS cases exhibited multi-hit *TP53* mutations. Additional studies with larger cohorts will be required to support our findings.

Although current guidelines emphasize the importance of assessing *TP53* allelic status, this is, to our knowledge, not routinely implemented in most diagnostic laboratories, despite evidence demonstrating its clinical relevance. Our results indicate that routine diagnostic workflows do not necessarily require advanced sequencing technologies for accurate determination of *TP53* allelic status. Standard NGS approaches, combining VAF assessment with cytogenetic profiling, appear sufficient to identify patients with biallelic *TP53* inactivation, supporting classification under the 5th WHO edition and informing allo-HSCT decision-making. Nevertheless, as demonstrated here, two *TP53* mutations may, in rare cases, reside on the same allele, justifying the use of LRS in selected patients.

References

1. Khoury JD, Solary E, Abla O, et al. The 5th edition of the World Health Organization classification of haematolymphoid tumours: myeloid and histiocytic/dendritic neoplasms. Leukemia. 2022;36(7):1703-1719.

2. Germing U, Adea L, Fontenay M, Haase D. Myelodysplastic neoplasm with biallelic TP53 inactivation. In: WHO classification of tumours: haematolymphoid tumours. Vol 2024. 5th Edition. Volume 11:80-82.

3. Arber DA, Orazi A, Hasserjian RP, et al. International consensus classification of myeloid neoplasms and acute leukemias: integrating morphologic, clinical, and genomic data. Blood. 2022;140(11):1200-1228.

4. Bernard E, Nannya Y, Hasserjian RP, et al. Implications of TP53 allelic state for genome stability, clinical presentation and outcomes in myelodysplastic syndromes. Nat Med. 2020;26(10):1549-1556.

5. Zhang L, Abro B, Campbell A, Ding Y. TP53 mutations in myeloid neoplasms: implications for accurate laboratory detection, diagnosis, and treatment. Lab Med. 2024;55(6):686-699.

6. Lontos K, Saliba RM, Kanagal-Shamanna R, et al. TP53-mutant variant allele frequency and cytogenetics determine prognostic groups in MDS/AML for transplantation. Blood Adv. 2025;9(11):2845-2854.

7. Urrutia S, Wong TN, Link DC. A clinical guide to TP53 mutations in myeloid neoplasms. Blood. 2025;146(18):2157-2167.

8. Lodé L, Ameur A, Coste T, et al. Single-molecule DNA sequencing of acute myeloid leukemia and myelodysplastic syndromes with multiple TP53 alterations. Haematologica. 2018;103(1):e13-e16.

9. Hu T, Chitnis N, Monos D, Dinh A. Next-generation sequencing technologies: An overview. Hum Immunol. 2021;82(11):801-811.

10. Bahaj W, Kewan T, Gurnari C, et al. Novel scheme for defining the clinical implications of TP53 mutations in myeloid neoplasia. J Hematol Oncol. 2023;16(1):91.

11. Nagata Y, Makishima H, Kerr CM, et al. Invariant patterns of clonal succession determine specific clinical features of myelodysplastic syndromes. Nat Commun. 2019;10(1):5386.

12. da Silva-Coelho P, Kroeze LI, Yoshida K, et al. Clonal evolution in myelodysplastic syndromes. Nat Commun. 2017;8(1):15099.

13. Marks JA, Wang X, Fenu EM, Bagg A, Lai C. TP53 in AML and MDS: the new (old) kid on the block. Blood Rev. 2023:60:101055.

14. Donehower LA, Soussi T, Korkut A, et al. Integrated analysis of TP53 gene and pathway alterations in the cancer genome atlas. Cell Rep. 2019;28(5):1370-1384.e5.

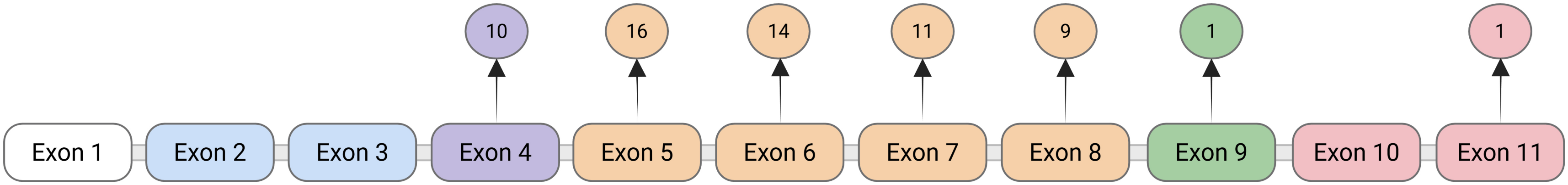
Figure Legends

Figure 1. *TP53* mutation landscape in the study cohort. Data from 29 patients with myelodysplastic neoplasms (MDS) harboring 62 *TP53* mutations identified by short-read sequencing (SRS) are summarized. The **upper panel** shows the distribution and number of mutations across the *TP53* gene. Exons are depicted as colored blocks, with colors indicating functional domains. The **lower panel** displays the cumulative variant allele frequencies (VAFs) per patient, previously determined using SRS. Created in BioRender. https://BioRender.com/2ulsm0y

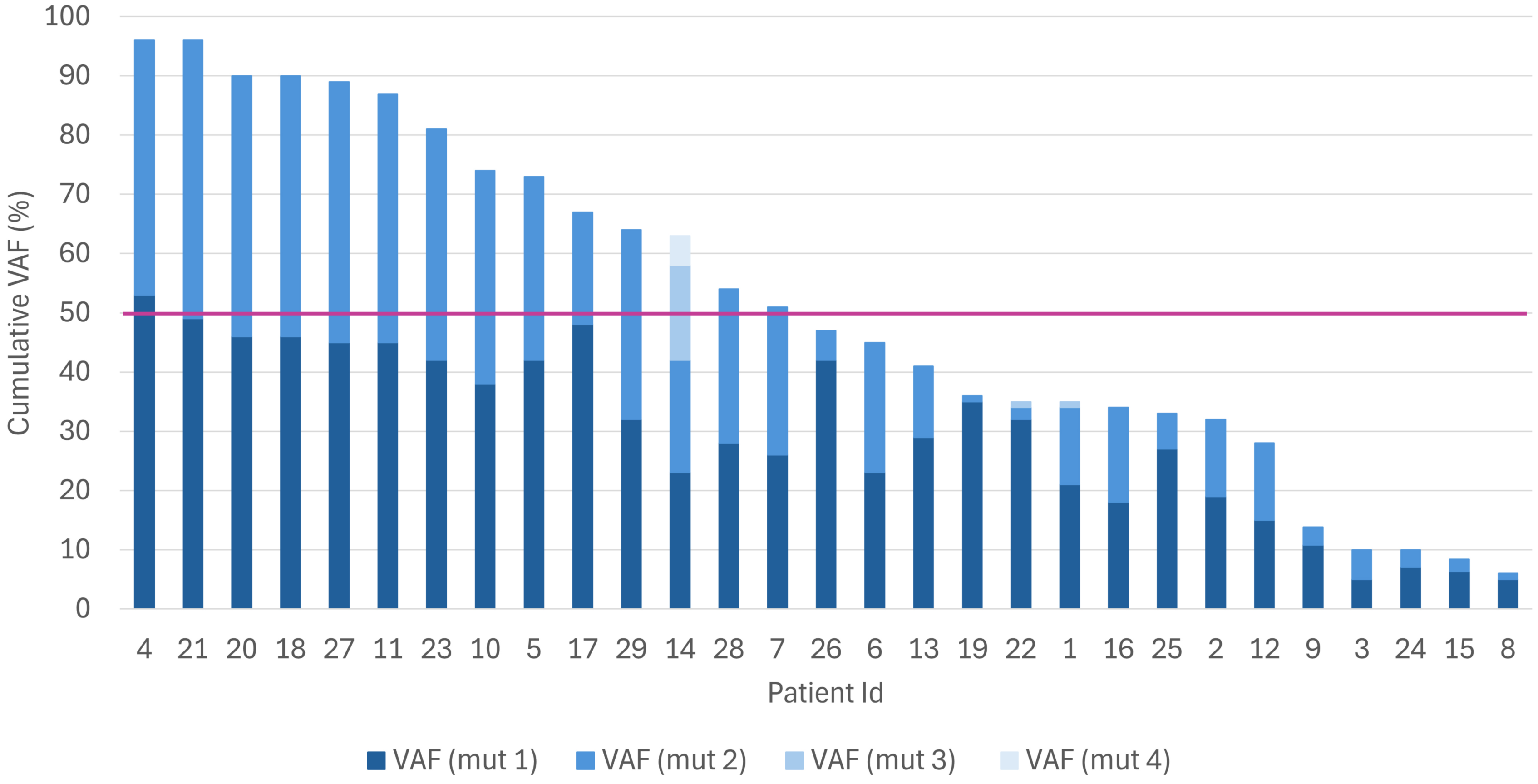
Figure 2. Study workflow. The figure outlines the experimental design. Twenty-nine patients with myelodysplastic neoplasms (MDS) and multiple *TP53* mutations (n = 62) were analyzed using long-read sequencing (Oxford Nanopore platforms). The workflow included DNA extraction, *TP53* amplification, library preparation, sequencing, and alignment to the hg38 reference genome. Allelic status was determined by phasing full-length reads to distinguish monoallelic from multi-hit configurations using Integrative Genomics Viewer (IGV). Created in BioRender. https://BioRender.com/0gt5jxl

Figure 3. *TP53* allelic status in myelodysplastic neoplasms. The **upper panel** shows the distribution of *TP53* allelic configurations among 29 patients with myelodysplastic neoplasms (MDS). One patient (3.4%) exhibited monoallelic *TP53* inactivation, whereas the remaining 28 patients (96.6%) were classified as multi-hit, harboring *TP53* mutations on different alleles. According to Bernard et al[4] and the 5th WHO classification,[1] biallelic *TP53* inactivation was defined as cumulative VAF above 50% based on SRS. For the remaining 14 multi-hit samples, biallelic inactivation could not be excluded; these cases may represent either biallelic events or multiple independent clones carrying distinct *TP53* mutations. The **lower panels** show representative IGV visualizations of two *TP53* inactivation patterns: **monoallelic** (Id 3: NM_000546.6: c.542G>C, p.(Arg181Pro), VAF: 5%, NM_000546.6:c.565G>C, p.(Ala189Pro), VAF: 5%) and **multi-hit** (Id 6: NM_000546.6: c.380C>T, p.(Ser127Phe), VAF: 23%, NM_000546.6: c.614A>C, p.(His205Pro VAF: 22%). Allelic status was determined by phasing full-length reads to distinguish mutations on the same allele from those on different alleles. Monoallelic inactivation (Id 3) involved two low-VAF mutations in the bulk sample, whereas multi-hit inactivation (Id 6) comprised two distinct mutations on separate alleles. NM_000546.6, hg38. **A**: Total number of reads sorted by the first mutation. **B**: Allele-specific analysis including BAM files filtered to include only reads harboring the second mutation. **C**: Allele-specific analysis including BAM files filtered to include only reads harboring the first mutation. Created in BioRender. https://BioRender.com/mue1iac

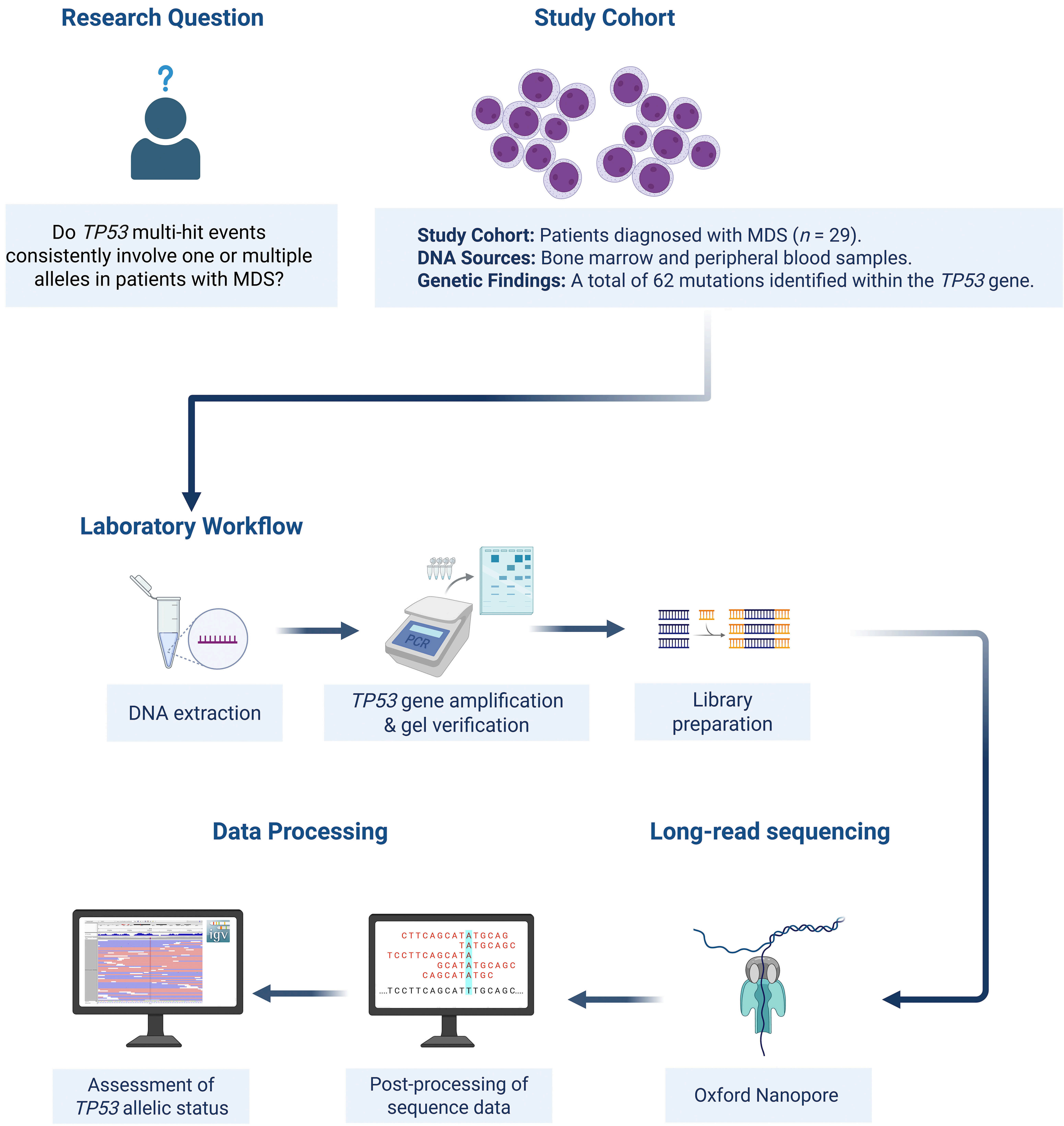
TP53 Mutation Count, Distribution and Localization



Distribution and number of mutations (n=62) within the *TP53* gene



Cumulative VAF per Patient in MDS Cohort (n=29)



Research Question

Do *TP53* multi-hit events consistently involve one or multiple alleles in patients with MDS?

Study Cohort

Study Cohort: Patients diagnosed with MDS (n = 29).
DNA Sources: Bone marrow and peripheral blood samples.
Genetic Findings: A total of 62 mutations identified within the *TP53* gene.

Laboratory Workflow

DNA extraction

TP53 gene amplification & gel verification

Library preparation

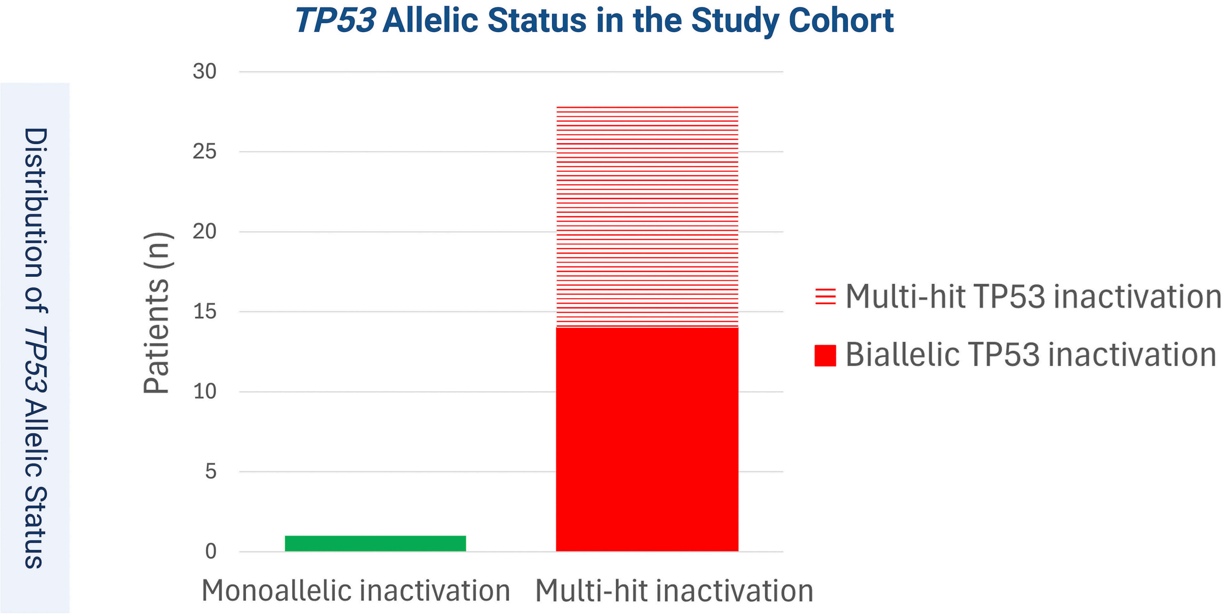
Long-read sequencing

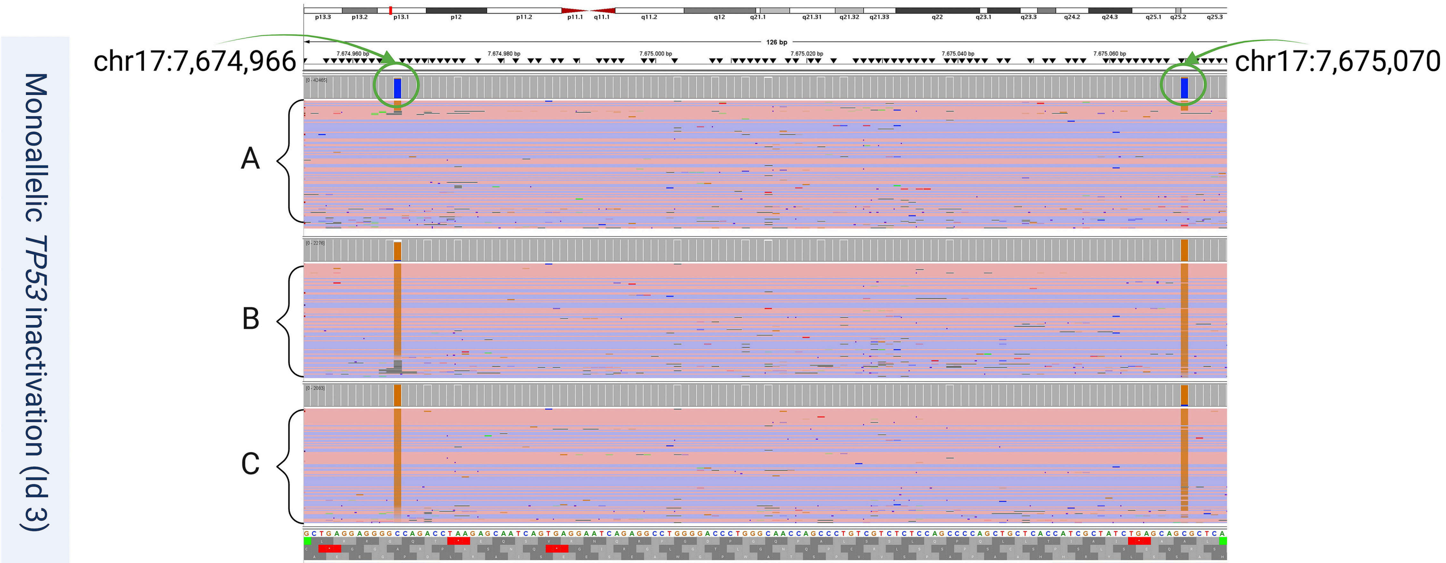
Oxford Nanopore

Data Processing

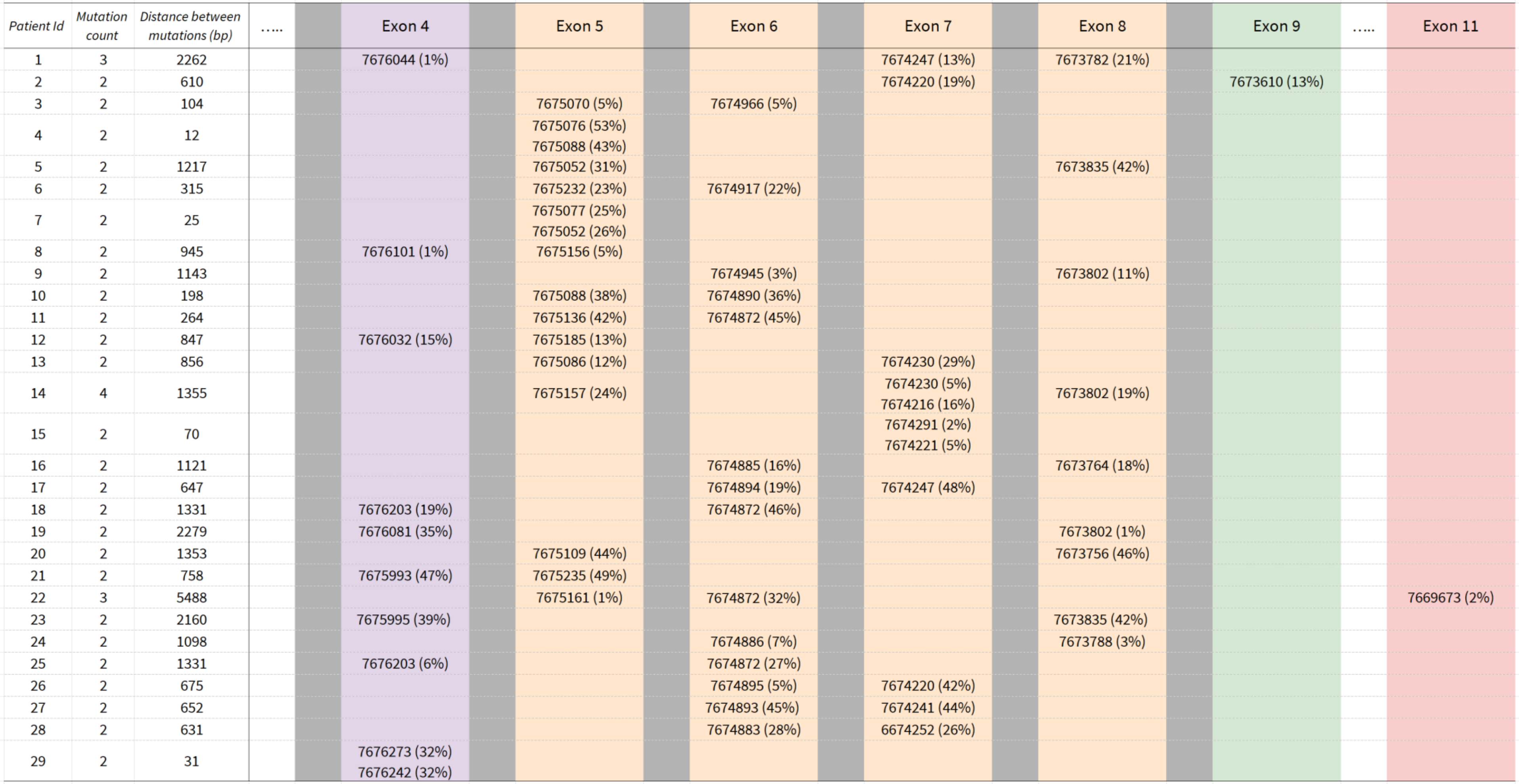
Post-processing of sequence data

Assessment of *TP53* allelic status


TP53 Allelic Status in the Study Cohort
Distribution of *TP53* Allelic Status
Patients (n)
30
25
20
15
10
5
0
Monoallelic inactivation
Multi-hit inactivation
Multi-hit TP53 inactivation
Biallelic TP53 inactivation


Monoallelic *TP53* inactivation (Id 3)
chr17:7,674,966
chr17:7,675,070
126 bp
A
B
C
TP53


Multi-hit *TP53* inactivation (Id 6)
chr17:7,674,917
chr17:7,675,232
359 bp
A
B
C
TP53



Patient Id	Mutation count	Distance between mutations (bp)	…..		Exon 4		Exon 5		Exon 6		Exon 7		Exon 8		Exon 9	…..	Exon 11
1	3	2262			7676044 (1%)						7674247 (13%)		7673782 (21%)				
2	2	610									7674220 (19%)				7673610 (13%)		
3	2	104					7675070 (5%)		7674966 (5%)								
4	2	12					7675076 (53%) 7675088 (43%)										
5	2	1217					7675052 (31%)						7673835 (42%)				
6	2	315					7675232 (23%)		7674917 (22%)								
7	2	25					7675077 (25%) 7675052 (26%)										
8	2	945			7676101 (1%)		7675156 (5%)										
9	2	1143							7674945 (3%)				7673802 (11%)				
10	2	198					7675088 (38%)		7674890 (36%)								
11	2	264					7675136 (42%)		7674872 (45%)								
12	2	847			7676032 (15%)		7675185 (13%)										
13	2	856					7675086 (12%)				7674230 (29%)						
14	4	1355					7675157 (24%)				7674230 (5%) 7674216 (16%)		7673802 (19%)				
15	2	70									7674291 (2%) 7674221 (5%)						
16	2	1121							7674885 (16%)				7673764 (18%)				
17	2	647							7674894 (19%)		7674247 (48%)						
18	2	1331			7676203 (19%)				7674872 (46%)								
19	2	2279			7676081 (35%)								7673802 (1%)				
20	2	1353					7675109 (44%)						7673756 (46%)				
21	2	758			7675993 (47%)		7675235 (49%)										
22	3	5488					7675161 (1%)		7674872 (32%)								7669673 (2%)
23	2	2160			7675995 (39%)								7673835 (42%)				
24	2	1098							7674886 (7%)				7673788 (3%)				
25	2	1331			7676203 (6%)				7674872 (27%)								
26	2	675							7674895 (5%)		7674220 (42%)						
27	2	652							7674893 (45%)		7674241 (44%)						
28	2	631							7674883 (28%)		6674252 (26%)						
29	2	31			7676273 (32%) 7676242 (32%)												

Figure S1. ***TP53 mutation landscape in the study cohort.*** *Data from 29 patients with myelodysplastic neoplasms (MDS) harboring 62 TP53 mutations identified by short-read sequencing (SRS) are summarized. The figure shows the TP53 gene structure with mutation sites and corresponding positions in the coding sequence (NM_000546.6, hg38); variant allele frequencies (VAFs) are shown in parentheses. Exons are depicted as colored blocks, with colors indicating functional domains.*

Table S1. ***DNA Extraction, NGS Panels, Sequencing Platforms, and TP53 Primer Specifications***

DNA extraction protocols, NGS-panels and sequencing platforms used

DNA extraction protocol	**NGS panel and sequencing platform (short-read sequencing)**	
Maxwell RSC Cell DNA (AS1370)	Oncomine™ Myeloid Research Assay - Chef Ready (A36941)	Ion Proton
QIAsymphony DSP DNA Midi Kit (96) (ID. 937255)	Sophia DDM™ Myeloid Solution (BS0103ILLCSML01-032)	Illumina MiSeq
	AmpliSeq for Illumina Myeloid Panel Customized with 9 extra genes (reference)	Illumina NextSeq 500

PCR primers for TP53 exons 2–11: sequences, binding positions, and sizes of primers and amplicons

Name	**Sequence**	**Start**	**End**	**Exon**	**Span (bp)**
*TP53*_US_For	5'-TGGGCCAGCAGAGACTTGACAAC	7.669.004	7.669.026		23
*TP53*_DS_Rev	5'-TCCCCACTTTTCCTCTTGCAGCAG	7.676.620	7.676.643		24
*TP53*_US_For + *TP53*_DS_Rev		7.669.004	7.676.643	2-11	7640