

## Single-molecule DNA sequencing of acute myeloid leukemia and myelodysplastic syndromes with multiple TP53 alterations

Laurence Lodé, Adam Ameur, Thibault Coste, Audrey Ménard, Steven Richebourg, Jean-Baptiste Gaillard, Yannick Le Bris, Marie-Christine Béné, Thierry Lavabre-Bertrand, Thierry Soussi

### ▶ To cite this version:

Laurence Lodé, Adam Ameur, Thibault Coste, Audrey Ménard, Steven Richebourg, et al.. Single-molecule DNA sequencing of acute myeloid leukemia and myelodysplastic syndromes with multiple TP53 alterations. Haematologica, 2017, 103 (1), pp.e13-e16. 10.3324/haematol.2017.176719 . hal-02318127

### HAL Id: hal-02318127 https://hal.sorbonne-universite.fr/hal-02318127

Submitted on 16 Oct 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés. Published Ahead of Print on October 27, 2017, as doi:10.3324/haematol.2017.176719. Copyright 2017 Ferrata Storti Foundation.



# Single-molecule DNA sequencing of acute myeloid leukemia and myelodysplastic syndromes with multiple TP53 alterations

by Laurence Lodé, Adam Ameur, Thibault Coste, Audrey Ménard, Steven Richebourg, Jean Baptiste Gaillard, Yannick Le Bris, Marie Christine Béné, Thierry Lavabre-Bertrand, and Thierry Soussi

Haematologica 2017 [Epub ahead of print]

Citation: Lodé L, Ameur A, Coste T, Ménard A, Richebourg S, Gaillard JB, Le Bris Y, Béné MC, Lavabre-Bertrand T, and Soussi T. Single-molecule DNA sequencing of acute myeloid leukemia and myelodysplastic syndromes with multiple TP53 alterations. Haematologica. 2017; 102:xxx doi:10.3324/haematol.2017.176719

Publisher's Disclaimer.

*E-publishing ahead of print is increasingly important for the rapid dissemination of science. Haematologica is, therefore, E-publishing PDF files of an early version of manuscripts that have completed a regular peer review and have been accepted for publication. E-publishing of this PDF file has been approved by the authors. After having E-published Ahead of Print, manuscripts will then undergo technical and English editing, typesetting, proof correction and be presented for the authors' final approval; the final version of the manuscript will then appear in print on a regular issue of the journal. All legal disclaimers that apply to the journal also pertain to this production process.*  Letter to the editor:

Single-molecule DNA sequencing of acute myeloid leukemia and myelodysplastic syndromes with multiple *TP53* alterations

Laurence Lodé <sup>1,2</sup>, Adam Ameur <sup>3</sup>, Thibault Coste<sup>1,5</sup>, Audrey Ménard <sup>2</sup>, Steven Richebourg <sup>2,4</sup>, Jean-Baptiste Gaillard <sup>5</sup>, Yannick Le Bris <sup>2</sup>, Marie-Christine Béné <sup>2</sup>, Thierry Lavabre-Bertrand <sup>5</sup> and Thierry Soussi <sup>6-9</sup>

Affiliations:

1 - Hematology Laboratory, University Hospital, Montpellier, F-34000, France

2 - Hematology Laboratory, University Hospital, Nantes, F-44093, France

3- Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, Uppsala, Sweden.

4 - Cytogenetics, Hôpital du Saint Sacrement, CHU de Québec, Université Laval, Department of molecular medicine, Quebec, Canada

5 - Department of Clinical Cytology and Cytogenetics, Nîmes University Hospital, Nîmes, France.

6 - Sorbonne Université, UPMC Univ Paris 06, F- 75005 Paris, France

7 - INSERM, U1138, Centre de Recherche des Cordeliers, Paris, France

8 - Department of Oncology-Pathology, Karolinska Institutet, Cancer Center Karolinska (CCK) R8:04, Stockholm SE-171 76, Sweden

9: Correspondence: thierry.soussi@ki.se

**Correspondence:** Thierry Soussi, Sorbonne Université, UPMC Univ Paris 06, F- 75005 Paris, France or Department of Oncology-Pathology, Karolinska Institutet, Cancer Center Karolinska (CCK) R8:04, Stockholm SE-171 76, Sweden thierry.soussi@ki.se

Keywords: TP53 mutation; AML; MDS; Next Generation Sequencing; MDS

L.L. and A.A. contributed equally to this study.

Although the frequency of *TP53* mutations in hematologic malignancies is low, these mutations have a high clinical relevance and are usually associated with poor prognosis. Somatic *TP53* mutations have been detected in up to 73.3% of cases of acute myeloid leukemia (AML) with complex karyotype and 18.9% of AML with other unfavorable cytogenetic risk factors. <sup>1</sup> AML with *TP53* mutations, and/or chromosomal aneuploidy has been defined as a distinct AML subtype. In low-risk myelodysplastic syndromes (MDS), TP53 mutations occur at an early disease stage and predict disease progression. <sup>2</sup> *TP53* mutation diagnosis is now part of the revised European LeukemiaNet (ELN) guidelines.<sup>3,4</sup>

The use of next generation sequencing (NGS), particularly ultra-deep sequencing, has led to the discovery that patients with either MDS or AML (either *de novo*, secondary or therapy-related) present multiple *TP53* mutations, suggesting that several *TP53* independent clones may co-exist. <sup>5</sup> Patient follow-up also reveals a highly dynamic evolution of these mutations during disease progression in treated and untreated patients. <sup>6,7</sup> This observation is in line with the recent recognition that human tumors harbor an extensive genetic intratumoral heterogeneity. <sup>8</sup>. These findings will likely have implications for therapy and biomarker discovery and determination of genetic complexity is becoming part of clinical decision-making processes in the age of precision medicine.

In this report, *in silico* analysis of the UMD *TP53* database showed that *TP53* variants detected in patients with multiple *TP53* alterations are fully oncogenic. Furthermore, using long-range single-molecule real-time (SMRT) sequencing on AML and MDS patients harboring multiple *TP53* mutations, we showed that all of these variants are localized on different subclones, emphasizing the considerable tumor heterogeneity in these patients.

The 2017 release of the UMD\_TP53 database contains the mutation status of 75,448 patients, including 922 cases of AML and 899 cases of MDS (Online Supplemental Table S1). 9.10 Among these patients, 158 MDS cases (22.3%) and 99 AML cases (13%) harbor more than one TP53 variant in their tumors, higher than the rate observed in solid tumors (Online Supplemental Table S1). Chronic Lymphocytic Leukemia (CLL) patients also harbor a high frequency of tumors with multiple TP53 mutations. This feature has been observed increasingly frequently over recent years with the advent of deep sequencing techniques. Whether or not all of the multiple variants identified in these patients are truly deleterious or comprise a mix of driving and passenger mutations has never been addressed. The UMD TP53 database includes quantitative functional data for all TP53 missense variants and can therefore be used to determine whether patients with multiple TP53 mutations frequently harbor non-deleterious TP53 variants. Analysis of the 257 AML and MDS cases with more than one TP53 variant showed that the majority (98%) of these variants are true deleterious TP53 mutations with complete loss of function and not simply random passenger mutations co-selected during tumor progression (Online Supplemental Table S2 and Online Supplemental Figure **S1a to e).** A few non-deleterious variants have been identified, but they are likely very rare non-somatic polymorphisms.

To further demonstrate the presence and the dynamics of multiple independent tumor clones in AML and MDS, we have developed a novel, third-generation single-molecule real-time (SMRT) sequencing assay using the Pacific Biosciences platform with long-read lengths that span the most frequently mutated region of the *TP53* gene. Sanger sequencing cannot be used to define the allelic distribution of multiple *TP53* variants. This is also true for standard NGS if the two variants are more than 200 base pairs apart. On

the other hand, SMRT analysis can be used to phase mutations located multiple kilobases apart directly from sequencing reads.

Eleven patients harboring multiple *TP53* mutations in their tumors were enrolled. For 3 patients, sequential samples were available to assess the evolution of the various variants. The *TP53* status of these patients was already defined according to stringent clinical criteria using either Sanger sequencing or standard NGS (**Online Supplemental Material and Online Supplemental Table S3**). In silico analysis of all these variants using the UMD *TP53* database showed that they were true deleterious *TP53* mutations that have already been described in various types of cancer (**Online Supplemental Table S1**).

The majority of mutations detected for clinical evaluation were readily identified by SMRT, except for 2 variants that were not included in the amplicon used for analysis. SMRT identified 5 mutations that were not identified by clinical analysis (**Online Supplemental Table S3 and Online Supplemental Figure S2**). Manual examination of the sequencing data performed for clinical analysis confirmed that 2 of these mutations were detected at a frequency below the cut-off used for the analysis (**Table 1**). Most of the remaining mutations detected by SMRT were present at a very low frequency (**Online Supplemental Figure S2**). The Variant allele frequency (VAF) observed for each variant detected by the two analyses was remarkably similar, despite being performed in different centers according to very different methodologies (**Online Supplemental Figure S3**).

Our analysis shows that all oncogenic *TP53* variants were located in different alleles **(Table 1 and Online Supplemental Figures S4a to S4k).** For two samples, patient Fr7, sample 7b and patient Fr2, the close proximity of two *TP53* variants allowed analysis of the alignment obtained after standard NGS and confirmed that these mutations were carried by different alleles **(Online Supplemental Figures S5 and S6).** For 2 samples, the allelic distribution was also confirmed by the observation that the different *TP53* variants were associated with different *TP53* haplotypes **(Online Supplemental Figure S4j and k, patients Fr10 and Fr11).** 

Figure 1 shows a typical result observed for two samples collected 5 years apart from a patient with multiple *TP53* mutations. Of note, the diagnostic sample was negative with standard NGS and was therefore not used for SMRT analysis. In the first sample analysed by SMRT, clinical analysis identified two pathogenic *TP53* mutations confirmed by SMRT. SMRT analysis also identified two novel *TP53* mutations at very low frequency and showed that the 4 variants were distributed in different *TP53* molecules. The two novel variants were readily identified by both methodologies (2 variants) and 2 additional variants were found at low frequency by SMRT (**Figure 1**). All these variants, carried by different *TP53* molecules, were true driver mutations already identified in multiple tumor types, as shown by their high frequency in the *TP53* database. This dynamic evolution of the various subclones can also lead to the elimination of certain subclones, as shown for patient Fr10 with the disappearance of *TP53* variants (**Online Supplemental Figure S4J**).

Using both *in silico* analysis and SMRT sequencing, we demonstrate that the presence of multiple subclones with different *TP53* variants is a common feature in AML and MDS. All *TP53* variants detected in MDS and AML patients by SMRT sequencing are true, physically independent *TP53* variants, confirming the results of indirect computational studies currently used to infer cancer heterogeneity. It is highly likely that each TP53 variant belongs to independent subclone arising from a wild-type TP53 founder clone. The observation of multiple subclones with different *TP53* variants in these patients suggests the occurrence of a specific genetic background in the founder clones that requires *TP53* inactivation for further progression. All of these subclones present a highly dynamic evolution, but it remains to be determined whether this evolution is driven by treatment, a

natural characteristic of the tumor or both. A recent study on 1,514 MDS patients after stem-cell transplantation showed that 283 patients (19%) had at least one oncogenic *TP53* mutation and a poor overall survival. <sup>11</sup> One hundred two (36%)of these patients had more than one *TP53* variant (range 2-6). It is likely that the use of a sensitive methodology for DNA sequencing will reveal that tumors with multiple *TP53* variants constitute a general feature raising potential problems for treatment options. Finally, in this report, we demonstrate the efficiency of SMRT sequencing for the analysis of complex samples. The rapid progress in NGS, combining longer reads, increased sensitivity and decreased costs, will allow investigation of the whole sequence of clinically relevant genes in a single analysis. Long-read RNA-seq analysis could also be used to address the issue of *TP53* alternative spliced transcripts that have already been described to be of clinical interest in AML.<sup>12</sup>.

**Acknowledgments:** This work was supported by Radiumhemmets Forskningsfonder and the Swedish Cancer Society (Cancerfonden) to TS. SMRT sequencing was performed by the National Genomics Infrastructure (NGI) hosted by SciLifeLab Uppsala. The authors are most grateful to the Molecular Hematology team and to the IRCNA Tumor Bank (CHU de Nantes, Institut de Cancérologie de l'Ouest, Saint-Herblain F44800, France) for their assistance.

**Contribution:** T.S. and L.L. designed the research; T.C., A.M., S.R., J.B.G., Y.L., M.C.B. and T.L.B. were actively involved in patient care or routine diagnostic procedures; A.A. performed the SMRT sequence and bioinformatics analysis; all authors reviewed and approved the manuscript

Conflict-of-interest disclosure: The authors declare no competing financial interests.

**Correspondence:** Thierry Soussi, Sorbonne Université, UPMC Univ Paris 06, F- 75005 Paris, France or Department of Oncology-Pathology, Karolinska Institutet, Cancer Center Karolinska (CCK) R8:04, Stockholm SE-171 76, Sweden thierry.soussi@ki.se

**ORCID profiles:** orcid.org/0000-0001-8184-3293

### References

- 1. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic Classification and Prognosis in Acute Myeloid Leukemia. N Engl J Med. 2016;374(23):2209-2221.
- 2. Jadersten M, Saft L, Smith A, et al. TP53 Mutations in Low-Risk Myelodysplastic Syndromes With del(5q) Predict Disease Progression. J Clin Oncol. 2011;29(15):1971-1979.
- 3. Döhner H, Estey E, Grimwade D, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood. 2017;129(4):424-447.
- 4. Malcovati L, Hellström-Lindberg E, Bowen D, et al. Diagnosis and treatment of primary myelodysplastic syndromes in adults: recommendations from the European LeukemiaNet. Blood. 2013;122(17):2943-2964.
- 5. Stengel A, Kern W, Haferlach T, Meggendorfer M, Fasan A, Haferlach C. The impact of TP53 mutations and TP53 deletions on survival varies between AML, ALL, MDS and CLL: an analysis of 3307 cases. Leukemia. 2017;31(3):705-711.
- 6. Mossner M, Jann JC, Nowak D, et al. Prevalence, clonal dynamics and clinical impact of TP53 mutations in patients with myelodysplastic syndrome with isolated deletion (5q) treated with lenalidomide: results from a prospective multicenter study of the german MDS study group (GMDS). Leukemia. 2016;30(9):1956-1959.
- 7. da Silva-Coelho P, Kroeze LI, Yoshida K, et al. Clonal evolution in myelodysplastic syndromes. Nat Commun. 2017;815099.
- 8. Marusyk A, Almendro V, Polyak K. Intra-tumour heterogeneity: a looking glass for cancer? Nat Rev Cancer. 2012;12(5):323-334.
- Leroy B, Ballinger ML, Baran-Marszak F, et al. Recommended Guidelines for Validation, Quality Control, and Reporting of TP53 Variants in Clinical Practice. Cancer Res. 2017;77(6):1250-1260.
- 10. Leroy B, Anderson M, Soussi T. TP53 mutations in human cancer: database reassessment and prospects for the next decade. Hum Mutat. 2014;35(6):672-688.
- 11. Lindsley RC, Saber W, Mar BG, et al. Prognostic Mutations in Myelodysplastic Syndrome after Stem-Cell Transplantation. N Engl J Med. 2017;376(6):536-547.
- 12. Anensen N, Hjelle SM, Van Belle W, et al. Correlation analysis of p53 protein isoforms with NPM1/FLT3 mutations and therapy response in acute myeloid leukemia. Oncogene. 2012;31(12):1533-1545.

Patient	Sample	Sex/Age (diagnosis)	Hematologic malignancy <sup>1</sup>	Treatment	<i>TP</i> 53 alteration (clinical) <sup>2</sup>	<i>TP53</i> alteration (SMRT) <sup>3</sup>	<i>TP53</i> Mutation <sup>4</sup>	Allelic distribution <sup>5</sup>
Er1	Er1	M/77	De novo MK-	No (at diagnosis,	Yes	Yes	c.673-2A>G	NR
111	111		AML	Treatment         K-       No (at diagnosis, before treatment)         R-       Yes (Lenalidomide)         Yes (Lenalidomide)       Image: Comparison of the second	Yes	No <sup>6</sup>	c.897_912del15	NR
Er2	Er2	M/62	De novo MK-	No (at diagnosis,	Yes	Yes	c.673-2A>T	Yes
112	112	101/03	AML	before treatment)	Yes	Yes	c.743G>A	Yes
Er2	Er2	E/72	De novo MK-	No (at diagnosis,	Yes	Yes	c.413C>T	Yes
FIS	гіз	F// 3	AML	before treatment)	Yes	Yes	c.794T>C	Yes
	Er4	E/70	De novo MK-	No (at diagnosis,	Yes	Yes	c.395A>G	Yes
F14	F14	F//O	AML	<ul> <li>K- No (at diagnosis, before treatment)</li> <li>R- Yes (Lenalidomide)</li> </ul>	Yes	Yes	c.824G>T	Yes
Er6	Cr5	E/72	De novo MK-	No (at diagnosis,	Yes	Yes	c.637C>T	Yes
FID	FID	F//3	AML	<ul> <li>IK- No (at diagnosis, before treatment)</li> <li>IK- No (at diagnosis, before treatment)</li> <li>.R- Yes</li> </ul>	Yes	Yes	c.455C>T	Yes
Er6	Ere	E/75	s-AML (post LR-	Yes	Yes	No	c.524G>A	NR
FIO	FIO	F// J	MDS del5q)	(Lenalidomide)	Yes	Yes	c.844C>T	NR
					Yes	Yes	c.314G>T	Yes
	Er7o				Yes	Yes	c.743G>A	Yes
	гиа				No	Yes	c.818G>A	Yes
<b>F-7</b>		E/70		Yes	No	Yes	c.742C>T	Yes
		F//3	LK-IMDS delog	(Lenalidomide)	Yes	Yes	c.314G>T	Yes
					Yes	Yes	c.743G>A	Yes
	FI/D				Yes	Yes	c.584T>A	Yes
					Yes	Yes	c.614A>G	Yes

<sup>&</sup>lt;sup>1</sup> MK-AML: Monosomal karyotype AML; LR-MDS: Low-risk-MDS

<sup>2</sup> Identification of *TP53* mutations for clinical analysis using conventional Sanger sequencing (VAF cut-off: 10-15%) or standard NGS (VAF cut-off: 1%).

<sup>3</sup> Identification of *TP53* mutations using SMRT sequencing

<sup>4</sup> *TP 53* variant description using the NM\_000546.5 reference. A full description of the mutations and their consequences is presented in supplementary Table S1.

<sup>5</sup> Yes: mutations are located on different alleles; NR: not relevant.

<sup>6</sup> Mutation outside the amplicon used for SMRT analysis

<sup>&</sup>lt;sup>7</sup> Samples Fr7a and Fr7b were taken at an interval of 59 months

					Yes	Yes	c.818G>A	Yes
					Yes	Yes	c.833C>G	Yes
					No	Yes	c.524G>A	Yes
					No	Yes	c.659A>G	Yes
Er0	Er0	E/72	s-AML (post LR-	Yes	Yes	Yes	c.725G>T	NR
FIO	FIO	F/12	MDS del5q)	(Lenalidomide)	Yes	No	c.920-1G>A	NR
	Er0o <sup>8</sup>		s-AML (post LR-		Yes	Yes	c.421T>G	Yes
Er0	FI9a	M/76	MDS del5q)	Yes	Yes <sup>9</sup>	Yes	c.711G>T	Yes
FIS	ErOb		s-AML (post LR-	Yes	Yes	Yes	c.421T>G	Yes
	FIBD		MDS del5q)		Yes	Yes	c.711G>T	Yes
					Yes	Yes	c.743G>A	Yes
	Fr10a		LR-MDS del5q	Yes (Lenalidomide)	Yes	Yes	c.844C>T	Yes
Fr10 <sup>10</sup>		M/69			Yes	Yes	c.817C>T	Yes
	Er10b		s-AML (post LR-	Yes	Yes	Yes	c.743G>A	Yes
	FIIUD		MDS del5q)	Yes     Yes       Image: West (Lenalidomide)     Yes       Image: West (Lenalidomide)     Yes       Image: West (Lenalidomide)     Yes	Yes	Yes	c.844C>T	Yes
				N	Yes <sup>7</sup>	Yes	c.659A>G	Yes
	Fr11a	F/85	LR-MDS del5q	Yes (Lenalidomide)	Yes <sup>12</sup>	Yes	c.840A>T	NR
Er11 <sup>11</sup>					No	Yes	c.701A>G	Yes
				X	Yes	Yes	c.701A>G	Yes
	Fr11b	F/85	LR-MDS del5q	Yes (Lenalidomide)	Yes	Yes	c.659A>G	Yes
				(Lenalidomide)	Yes	Yes	c.840A>T	Yes

Table 1: Patient characteristics. More information is available in **Supplemental Table S1**.

<sup>&</sup>lt;sup>8</sup> Sample 9a and 9: DNA extracted from a frozen pellet from whole blood leukocytes or a cytogenetic pellet from bone marrow, respectively (same sampling date)
<sup>9</sup>Mutation detected by SMRT and identified at very low frequency by reviewing the standard NGS data

<sup>&</sup>lt;sup>10</sup> Sample Fr10a and Fr10b were taken at an interval of 7 months

<sup>&</sup>lt;sup>11</sup> Sample Fr11a and Fr11b were taken at an interval of 9 months

<sup>&</sup>lt;sup>12</sup> Mutation detected at high frequency in the second sample and identified at low frequency by reviewing the data of the first sample.

#### Figure 1: Clinical course and TP53 mutation analysis on patient Fr7

A: Patient Fr7 was diagnosed in February 2007 with lower risk MDS with del5q with no *TP53* mutations detected at a cut-off of 1%. One year after initiation of red blood cell (RBC) transfusion for anemia, patient Fr7 was treated with lenalidomide at a dosage of 5 mg/day, but rapidly experienced major adverse effects, leading to reduction of the dosage to 5 mg every other day (or less in a context of poor adherence). Five months after onset of therapy, 2 *TP53* mutations were identified together with improvement of anemia. Five years later, in July 2013 (and until the present time), patient Fr7 was still alive and with no disease progression to secondary AML despite growth of at least 6 TP53-mutant clones, suggesting possible clonal equilibrium due to competition between the numerous mutant clones.

B: Sanger sequencing and/or standard NGS analysis is shown in the upper part with 2 and 6 mutations in sample 7a and 7b, respectively. No allelic distribution can be inferred from this type of analysis. SMRT sequencing (lower part) provides an accurate picture of the allelic distribution of each *TP53* variant, as well as the remaining wt allele. The frequencies of the 9 different alleles are shown in brackets.

Red triangle: *TP53* variants identified by both types of analysis. White triangle: *TP53* variants detected only by SMRT sequencing.







wt (35.2 %)

## Single-molecule DNA sequencing of acute myeloid leukemia and myelodysplastic syndromes with multiple *TP53* alterations

Laurence Lodé, Adam Ameur, Thibault Coste, Audrey Ménard, Steven Richebourg, Jean-Baptiste Gaillard, Yannick Le Bris, Marie-Christine Béné, Thierry Lavabre-Bertrand and Thierry Soussi

### SUPPLEMENTAL MATERIAL

### Table of contents

### A <u>Supplemental Methods</u>

Patients and Samples	Page 3
SMRT sequencing of TP53 amplicons	Page 3
Detection of SNPs and mutations in SMRT sequencing data	Page 3
Analysis of the phasing of SNPs and mutations	Page 3
In silico analysis of TP53 variants in AML and MDS	Page 4

### B Supplemental Tables

Supplemental Table S1: frequency of patients with multiple *TP53* variants in the UMD TP53 database.

Page 5 **Supplemental Table S2:** frequency of tumors with multiple *TP53* variants according to the type of alterations in AML and MDS.

Page 6

Supplemental Table S3: TP53 mutations in AML and MDS patients. Supplied as an Excel file.

### C <u>Supplemental Figures</u>

**Supplemental Figure S1a to f**: analysis of *TP53* variant loss of function in AML and MDS tumors with multiple *TP53* mutations

Page 7-19

**Supplemental Figure S2:** Venn diagram showing the mutational concordance of validated somatic variants based on the sequencing strategy.

Page 20

**Supplemental Figure S3:** variant allele frequency (VAF) observed for all *TP53* variants identified by both standard NGS and SMRT methodologies

Page 21

Supplemental Figure S4a to k: detailed analysis of the 11 patients included in this study.

Page 22-37

**Supplemental Figure S5:** visualization of NGS alignment, confirming that the two variants are located on different alleles for patient Fr2.

Page 38

**Supplemental Figure S6:** visualization of NGS alignment, confirming that the two variants are located on different alleles for patient Fr7

Page 39

Supplemental Figure S7: strategy used for the analysis of TP53 mutations

Page 40

### D References to Supplemental Material

Page 41

### A Supplemental Methods

### Patients and Samples

This retrospective study was conducted on 11 patients shown to harbor multiple TP53 mutations in their tumors (Table 1 and Supplemental Table S3). Six patients were diagnosed with refractory anemia between 1996 and 2010 and a diagnosis of lower risk MDS with del(5g) was confirmed at Nantes University Hospital (5 patients) or Nîmes University Hospital (1 patient) by conventional cytogenetics and/or FISH analysis. Four patients experienced disease progression to secondary acute myeloid leukemia (s-AML). The remaining five patients were diagnosed with poor-risk de novo monosomal karyotype acute myeloid leukemia (MK-AML) at Nîmes University Hospital between 1998 and 2011 (Supplemental Table S3). Multiple samples obtained during the course of the disease were available for 4 of the 11 patients and were sequential for 3 patients. TP53 status was established by two certified TP53 centers either by Sanger sequencing, standard NGS or both, using stringent criteria specific for clinical analysis (Table 1 and Supplemental Table S2). <sup>1,2</sup> Peripheral blood or bone marrow samples were stored as frozen cell pellets or cytogenetic pellets before therapy for MK-AML and before and/or after initiation of lenalidomide therapy for Lower Risk MDS del5q. TP53 short-read libraries were prepared using either primer plates from the IRON-II study network, sequenced on a GS-Junior (Roche, Basel, Switzerland) with data processed as described by Kohlmann et al.<sup>1</sup>, or a custom TSCA design (Illumina, San Diego, CA, USA) sequenced on a MiSeq (Illumina, San Diego, CA, USA), in which case, reads were processed with a homemade analysis pipeline including VarScan v2.3.6 or an analysis pipeline provided by Illumina including the MiSeqReporter Suite and Variant Studio annotation tool. With a depth of coverage between 800X and 9,500X (after resequencing of selected samples), a minimum of 10 mutated bidirectional reads were taken into account allowing a VAF detection threshold of 1% to 2%. The VAF detection threshold was set to a lower limit of >1% for bidirectional reads, according to a recent study investigating the assay's lower limit of detection<sup>2</sup>. Mutations with VAF>10% allowed cross-validation with Sanger Sequencing using a VAF detection threshold of 10-15%.

### SMRT sequencing of TP53 amplicons

A 2.8 Kb amplicon that encompasses exons 4 to 8 was used for SMRT analysis (**Supplemental Figure S7**). This region includes the majority of the mutations detected in the patients, as well as some common *TP53* SNP useful for phasing the various mutations on the two alleles.

The *TP53* amplicons (11 patients, 15 samples) underwent DNA damage repair and end-repair before ligation of hairpin adaptors to generate SMRTbell<sup>™</sup> libraries for circular consensus sequencing. Libraries were then subjected to exo treatment and PB AMPure bead wash procedures for clean-up. Each library was loaded onto one SMRTcell<sup>™</sup> and sequenced on the PacBio RS II instrument using C4 chemistry, P6 polymerase and a 240-min movie time. <sup>3</sup>

### Detection of SNPs and mutations in SMRT sequencing data

SNPs and mutations were identified by a two-step procedure. First, the 'Minor and Compound Variants' plugin (v2.3.0 of SMRT Analysis) was executed on each sample. This resulted in a total of 84 positive variants in all 25 samples. We then performed a more stringent analysis of each of these mutations by counting the number of reference/alternative alleles occurring in the CCS read using a 20 bp window surrounding each mutation. This counting-based method is a sensitive approach to determine exact mutation frequencies, as previously demonstrated <sup>3</sup>. Forty-three of the original 84 variants were detected at a frequency of at least 0.5% in at least one sample.

### Analysis of the phasing of SNPs and mutations

Custom R scripts were used to determine the clonal composition of mutations and SNPs. We counted the number of CCS reads comprising all possible combinations of reference/alternative variants and obtained a read count for all different *TP53* molecules present in each sample. To remove any chimeric molecules introduced by aberrations in the PCR step, we first determined the phasing of homozygous and heterozygous SNPs from the information provided by the molecule with the highest read count, and then filtered out all molecules discordant with this SNP phasing pattern. Next, we removed any remaining molecules that could be explained by a single jump between different molecules during PCR, i.e. molecules with a phasing pattern that can be created by concatenation of two other molecules with higher read counts.

### In silico analysis of TP53 variants in AML and MDS

The 2017 release of the *TP53* mutation database contains 82,134 *TP53* mutations, from 75,448 patients including those from 1,821 cases with AML or MDS.<sup>4,5</sup>. The database includes records for each tumor, indicating the number and description of each variant.

The database also includes functional data for most missense mutations. Residual transactivating activity for WAF, MDM2, BAX, 14-3-3-s, AIP, GADD45, NOXA and P53R2 promoters was originally published by Kato et al <sup>6</sup>. The residual transcriptional activity of mutant p53 was always compared to wild-type p53 for the same promoter (%).

For nonsense and splice variants as well as indels (*TP53* null), this value was set to 0 as no *TP53* protein is generally expressed.

Supplemental Table S1: Frequency of patients with multiple TP53 variants in the UMD TP53 database.

Cancer type	MM	SM	Total	MM Frequency
Acute myeloid leukemia	99	664	763	13.00
Myelodysplastic syndrome	158	550	708	22.30
Chronic lymphocytic leukemia	181	1742	1923	9.41
Head and Neck SCC	454	4385	4839	9.40
Lung (NSCLC)	400	7142	7542	5.30
Colorectal carcinoma	420	7794	8214	5.10
Gastric carcinoma	85	1542	1627	5.20
Ovarian carcinoma	120	4223	4343	2.70
Pancreatic carcinoma	35	1686	1721	2.03

Analysis was performed using the 2017 release of the UMD *TP53* database (82,134 *TP53* mutations and 75,448 patients ). For each cancer type, the number of patients with either one (SM) or more than 1 (MM) *TP53* variants is reported.

**Supplemental Table S2**: Frequency of patients with multiple *TP53* variants according to the type of alterations in AML and MDS .

Tumors with multiple TP53 variants	Number
2 single nucleotide substitutions	161
1 single nucleotide substitution and 1 null variant	70
3 different variants (all types)	19
4 different variants (all types)	5
5 different varints (all types)	1
6 differents variants (all types)	1

Two hundred fifty-seven patients with MDS and AML carry multiple *TP53* mutations (**Supplemental Table S1**). Two different nucleotide substitutions were detected in 161 tumors (62%), while, in 70 patients, the second event was a *TP53* null event (either splice or indel variants).

**Supplemental Figure S1 a to f**: analysis of *TP53* variant loss of function in AML and MDS tumors with multiple *TP53* mutations

**Figure S1a:** analysis of AML and MDS tumors with two different *TP53* single nucleotide substitutions

**Figure S1b:** analysis of AML and MDS tumors with one *TP53* single nucleotide substitution and one frameshift TP53 variant

Figure S1c: analysis of AML and MDS tumors with 3 different TP53 alterations

Figure S1d: analysis of AML and MDS tumors with 4 different TP53 alterations

Figure S1e: analysis of AML and MDS tumors with 5 different TP53 alterations

Figure S1f: analysis of AML and MDS tumors with 6 different TP53 alterations

Residual transactivating activity for WAF (W), MDM2 (M), BAX (B), 14-3-3- $\sigma$  (14), AIP (A), GADD45 (G), NOXA (N) and P53R2 (P) ranges from 0 (red) to 100% (green). The frequency of the variant in the database is shown as both a bar and a number in the right part of the Figure.



### **Residual TP53 activity**

Sample_ID	Disease	Variant	w	м	В	14	Α	G	N	Р	Frequency_graph	i i
OHN-GM	AML	p.V172F										75
OHN-GM	AML	p.C238Y										289
	A N 41	n 1/172E										75
OHN-GM-Tum	AMI	p.v1/2F n C238Y										75 289
	,E	p.02501										200
3	AML	p.H178P										30
3	AML	p.R290H										75
											-	
6	AML	p.Q167*									-	113
0	AIVIL	μ.κ24ου										.500
92-886	AML	p.R273C									2	168
92-886	AML	p.R306*										612
14	MDS	p.H193R										341
14	MDS	p.1195T										359
5	АМІ	n V143M										115
5	AML	p.V274A										63
		•										
DA4	AML	p.H178P										30
DA4	AML	p.R290H										75
D£	MDC	n C220V										200
DS	MDS	p.C2381 n R2481										289
23	11125	p.112-102										257
КВ	MDS	p.Q192*										322
КВ	MDS	p.Y220C									1	230
43	AML	p.R156H										45
43	AIVIL	p.cz//ł									•	54
47	AML	p.V143M										115
47	AML	p.V216M										273
											L	
CMK	AML	p.D49H										20
СМК	AML	p.M133K										44
MOLM-16	АМІ	n V173M										241
MOLM-16	AML	p.C238S										41
P31-FUJ	AML	p.R196*										837
P31-FUJ	AML	p.Y236C										256
DK1	A N 41	n N220D										122
RK4	AML	p.N239D p.S261T										152
		P.0-0-1										-
RK8	AML	p.C135S										13
RK8	AML	p.M246K										26
24	MDC	- 02720									-	4.00
34 34	MDS	р.кz/зс n R273H									2	200
54	NUD5	p.nz7511										.500
49	MDS	p.R158H										310
49	MDS	p.R273H									2	300
											h	
394	AML	p.C135S										24
394	AIVIL	p.1VI246K										26
403	AML	p.N239D										132
403	AML	p.S261T										8
AML047T	AML	p.V143M										115
AML047T	AML	p.V216M										273
	ΔΝЛΙ	n ∩52*										<u>م</u>
AML096T	AML	p.Q32 p.W91*										56
											F	
MOLM-16	AML	p.V173M										241
MOLM-16	AML	p.C238S									I	41

### Supplemental Figure S1a (part 1)

Sample_ID	Disease	Variant	w	м	В	14	Α	G	Ν	Р	Frequency_grap	h
P31-FUJ	AML	p.R196*										837
131-103	AWL	p.1250C										250
768	AML	p.G245D										470
768	AML	p.R248Q										2500
867	AML	p.G105D										10
867	AML	p.V157F										577
886	ΔΜΙ	n H179R										565
886	AML	p.R248Q										2500
888 888	AML	p.V1/3G n I195T										42 359
000	,	p										
892	AML	p.Y205N										18
892	AML	p.D281Y										49
894	AML	p.Y163C										499
894	AML	p.M237I										362
898	AML	p.Y220C										1230
898	AML	p.Y327*										13
019	0.041	n C121D									h	,
918 918	AML	p.R306*										612
		·										
930 920	AML	p.R273H										2300
930	AIVIL	h.v200										012
947	AML	p.F134L									<u> </u>	61
947	AML	p.I195T										359
964	AML	p.V173M										241
964	AML	p.Y220C										1230
GSM630070	AN/I	n R156H										45
GSM630979	AML	p.C277Y										54
<b>CCL</b> 4C20005												
GSM630985 GSM630985	AML	p.V143M p.V216M										115 273
		P									—	
GSM631041	AML	p.Q52*										40
GSIVI031041	AIVIL	p.war.									<b>I</b>	50
MDS38	MDS	p.C238Y										289
MDS38	MDS	p.R248L										297
MDS39	MDS	p.Q192*										322
MDS39	MDS	p.Y220C										1230
UPN-1	MDS	n.R158H										310
UPN-1	MDS	p.R273H										2300
UDN 10	MDC	n V162C										400
UPN-10	MDS	p.r103C p.C275Y										243
UPN-16	MDS MDS	p.R248W										1949
OFIN-17	WD3	p.11170F									l .	30
UPN-24	MDS	p.Y163C										499
UPN-24	MDS	p.Y234C										396
UPN-25	MDS	p.Y220C										1230
UPN-25	MDS	p.Q331H									l	19
UPN-26	MDS	p.K132O									li de la companya de	49
UPN-26	MDS	p.G262V										80
	MDC	n \/17214										2/11
UPN-27	MDS	p. V 17 5 1VI p. H214R										219

### Supplemental Figure S1a (part 2)

Sample_ID	Disease	Variant	w	м	В	14	Α	G	Ν	Р	Frequency_grap	h
UPN-9	MDS	p.L43*										9
UPN-9	MDS	p.C238Y										289
27125	A N 41	m T110A									I.	2
37135		p.1118A										۲ 612
57155		p.11500										012
VMK9	MDS	p.V143M										115
VMK9	MDS	p.H179R										565
TCGA-AB-2908	AML	p.C141W									<u> </u>	41
TCGA-AB-2908	AML	p.Q317*										157
176267	MDC											2240
176267	MDS	p.K1/5H										3319
1/0207	IVID3	μ.κ207Q									•	45
369682	MDS	p.O144P										21
369682	MDS	p.V216M										273
558896	MDS	p.V143M									l i i i i i i i i i i i i i i i i i i i	115
558896	MDS	p.Y163C										499
45		14055										~
15	MDS	p.1195F										124
15	ND3	p.v2/2L										134
2	MDS	p.C238Y										289
2	MDS	p.R282W										1568
4	MDS	p.A161D									<u> </u>	48
4	MDS	p.S241C										129
c		62205									-	
6	MDS	p.C238F										2200
0	IVID3	p.ĸz75⊓										2300
BLE1	MDS	p.A161T										185
BLE1	MDS	p.C176Y										279
BLE2	MDS	p.Y220C										1230
BLE2	MDS	p.S241Y									1	89
	MDC	n D1E1T										
BLE7 BLE7	MDS	p.F1311 n C275Y										243
DEL,	WID'S	p.ez/51										240
BLE9	MDS	p.C238Y										289
BLE9	MDS	p.R273H										2300
35	AML	p.H179R										565
35	AML	p.V274F										139
7	<b>ΔΝ/Ι</b>	n R110P										61
7	AML	p.N1107 p.N2471										39
		P										
AML-46	AML	p.R248Q										2500
AML-46	AML	p.E258A										19
											-	
10	MDS	p.C238F										147
10	IVIDS	р.кz/зп										2300
11	MDS	p.C238Y										289
11	MDS	p.R282W										1568
12	MDS	p.S127F										173
12	MDS	p.R249G									I	112
40		<b>D2</b> + 0 <i>C</i>										~-
13	MDS	p.R248G										83 125
13	COINI	p.n2000										132
14	MDS	p.A161D										48
14	MDS	p.S241C										129
4	MDS	p.P222S										10
4	MDS	p.R248Q										2500

### Supplemental Figure S1a (part 3)

Sample_ID	Disease	Variant	w	м	В	14	Α	G	N	Р	Frequency_graph
7 7	MDS MDS	p.C176R p.G245S									54 1156
8 8	MDS MDS	p.P152L p.V216M									258 273
11 11	AML AML	p.E171* p.R248Q									85 2500
17 17	AML AML	p.E171G p.P177L									17 78
1 1	AML AML	p.L194R p.Y205S									224 40
14 14	AML AML	p.Y236C p.R249G									256 112
21 21	AML AML	p.R175H p.H193Y									3319 148
23 23	AML AML	p.H179R p.R280*									565 27
13 13	MDS MDS	p.H179R p.A189V									565 28
MOLM-16 MOLM-16	AML AML	p.V173M p.C238S									241 41
013-18 013-18	AML AML	p.K132E p.R273H									81 2300
037-02 037-02	AML AML	p.N131Y p.V173M									23 241
391-02 391-02	AML AML	p.Y236C p.C238Y									256 289
449-03 449-03	AML AML	p.R110L p.C275R									150 60
560-03 560-03	AML AML	p.C238S p.G266R									41 192
664-03 664-03	AML AML	p.V173L p.R248Q									197 2500
AML27 AML27	AML AML	p.C176Y p.C238Y									279
AML45 AML45	AML AML	p.R175H p.R282W									3319 1568
AML80 AML80	AML AML	p.P152L p.R248W									258
MI-AML-047 MI-AML-047	AML AML	p.V143M p.V216M									115 273
MI-AML-096 MI-AML-096	AML AML	p.Q52* p.W91*									40 87
MI-AML-250 MI-AML-250	AML AML	p.R196* p.C275W									837
377512 377512	t-MDS / t-AML t-MDS / t-AML	p.L130V p.R273C									60 2168
530447 530447	t-MDS / t-AML t-MDS / t-AML	p.K139N p.R248Q									26 2500

### Supplemental Figure S1a (part 4)

Sample_ID	Disease	Variant	w	м	в	14	Α	G	Ν	Р	Frequency_g	raph
1_39	AML	p.C242Y										162
1_39	AML	p.R110P										61
1 73	AMI	n R248W										1949
1_73	AML	p.A138D										9
1_74	AML	p.R273C										2168
1_/4	AIVIL	p.ĸ175G										91
229	AML	p.V216M										273
229	AML	p.F134L									l	20
220	A.N.41	- F270C									h	52
330	AMI	p.F270C n M246I										53
000	,	p										
BM12	MDS	p.R175H										3319
BM12	MDS	p.W146*										131
PD11154a	AMI	n Y220C										1230
PD11154a	AML	p.W146*										131
PD11168a	AML	p.M237I										362
PD11168a	AML	p.K164E										66
21	AML	p.R248W										1949
22	AML	p.R248G										83
DFCI-006382-3581	MDS	p.C176Y										279
51 CI-000582-558.	IVIDS	p.1220C										1250
5273	AML	p.P152L										258
5273	AML	p.M237I										362
7071	A.N.41	n 1/10 <b>0</b> *									h	12
7071	AML	p.R196P										62
		P									-	
3022518	MDS	p.C176Y										279
3022518	MDS	p.R175G										91
3054448	MDS	n H179D										71
3054448	MDS	p.K132R										202
3116963	MDS	p.R248Q										2500
3116963	MDS	p.¥220C										1230
3139981	MDS	p.Y220C										1230
3139981	MDS	p.L194R										224
			_								-	
3195456	MDS	p.R267W										119
5195450	IVID5	p.AIUIT									-	185
3250152	MDS	p.R248W										1949
3250152	MDS	p.R2801										68
											h	
3280694	MDS	p.I162N										30
5280094	IVIDS	μ.κ2/5Π										2500
3310889	MDS	p.Y220H										65
3310889	MDS	p.R282W										1568
2202656	MDC	- D10111										75
3382656	MDS	p.K181H p.H179R										565
5562656	WD5	p.111/51										505
3429739	MDS	p.P151T									I	56
3429739	MDS	p.\$90P									l	7
3/137575	MDS	n V220C										1720
3437575	MDS	p.1220C p.D281N										113
		·										
3474059	MDS	p.R175H										3319
3474059	MDS	p.Y126C										95

### Supplemental Figure S1a (part 5)

Sample_ID	Disease	Variant	w	м	В	14	Α	G	Ν	Р	Frequency_gra	ph
3490378	MDS	p.I195T										359
3490378	MDS	p.R273C										2168
3550478	MDS	n T253A									l	11
3550478	MDS	p.1255A p.G105V										27
5556176		p.01001									1	_,
3594211	MDS	p.H168P									l	31
3594211	MDS	p.A276P										37
			_								-	
3642291	MDS	p.S241C										129
3042291	IVID3	p.v210ivi										2/3
3658115	MDS	p.Y220C										1230
3658115	MDS	p.R273Н										2300
3700016	MDS	p.R248Q										2500
3700016	MDS	p.R273H										2300
3876517	MDS	n C238F										147
3876517	MDS	p.V173M										241
8010773	MDS	p.R175G										91
8010773	MDS	p.Y163C										499
9011227	MDS	n C2449										170
8011237	MDS	p.02443 p.F113I										10
		P									ļ	
8063897	MDS	p.H214D									l	15
8063897	MDS	p.R175H										3319
		¥22.44										
8088920	MDS	p.Y234H										2200
8088920	IVID3	p.nz/3fi										2300
8137149	MDS	p.E258*									1	70
8137149	MDS	p.V143M										115
8142412	MDS	p.Y220N										52
8142412	MDS	p.v1/3M										241
8189833	MDS	p.Y220S									1	70
8189833	MDS	p.P152L										258
8248094	MDS	p.A159P										136
8248094	MDS	p.P152L										258
8251//29	MDS	n R273C										2168
8251429	MDS	p.Q136E										41
		1										
8290369	MDS	p.K164N										15
8290369	MDS	p.G266R										39
8202262	MDS	n P2490										2500
8293363	MDS	μ.κ246Q n C277Y										2500 54
0293903	WID'S	p.ez//1									•	
8540417	MDS	p.M246V										154
8540417	MDS	p.M237I										362
0563004	MDC										I.	
8563904	MDS	p.P177H p.V143M										11
8303904	WID5	p. v 1451vi										115
8602801	MDS	p.H193R										341
8602801	MDS	p.R273H										2300
8693966	MDS	p.H179Q										40
8693966	MDS	p.P151S									-	236
8707610	MDS	p.R175H										3319.
8707610	MDS	p.V143M										115
8817369	MDS	p.G199E										35
8817369	MDS	p.H179N									I	70

### Supplemental Figure S1a (part 6)

Sample_ID	Disease	Variant	w	м	В	14	Α	G	Ν	Ρ	Frequency_gra	ph
8882668	MDS	p.R174W										38
8882668	MDS	p.R273H										2300
0000427	MDC	~ V220C										1220
8898427	MDS	p.Y220C										1230
8898427	IVIDS	р.н179к										505
8970135	MDS	n C242F										219
8970135	MDS	p.R175H									-	3319
		P										
8976124	MDS	p.Y234C										396
8976124	MDS	p.V216M										273
8989984	MDS	p.Q317*										157
8989984	MDS	p.H179Q										40
											_	
9889514	MDS	p.D281H										98
9889514	MDS	p.P152Q										16
0052661	MDC	- 110FT										250
9952661	MDS	p.11951										359
9952661	IVIDS	р.кт/зн										3319
0073131	MDS	n P106G										10
9973131	MDS	n C141Y										221
5575151	ND5	p.01411									-	221
9980522	MDS	p.R306*										612
9980522	MDS	p.R282W										1568
CLL090	MDS	p.E286K										303
CLL090	MDS	p.H179R										565
AML_113	AML	p.R196*										837
AML_113	AML	p.D281N										113
AML_135	AML	p.R248G										83
AML_135	AML	p.G245V										239
ANAL 145	0.041	n D772U										2200
ANI 145		p.K273H										2300
ANIL_145	ANIE	p.1205D										45
AMI 547	AMI	p.Y234C										396
AML 547	AML	p.Y234H										113
		P										
AML_70	AML	p.R273H										2300
AML_70	AML	p.M133T										24
1037	AML / MDS	p.H179P										20
1037	AML / MDS	p.R273H										2300
											-	
1062	AML / MDS	p.H193L										187
1062	AML / MDS	p.L257P										44
1068		n C1/11										221
1068		p.01411										440
1000	ANIE / WIDS	p.11275E										440
MDS143	MDS	p.L265P										60
MDS143	MDS	p.R248Q									-	2500
MDS555	MDS	p.P278A									l i i i i i i i i i i i i i i i i i i i	84
MDS555	MDS	p.V143A										44
											_	
LB 382	AML	p.R213Q										165
LB 382	AML	p.Y220C										1230
5												~ ~
D	AIVIL	p.C1355										154
U	AIVIL	p.1v1240V										104

### Supplemental Figure S1a (part 7)

Sample_ID	Disease	Variant	W	М	В	14	Α	G	Ν	Р	Frequency_graph	
3709100	MDS	p.E171G										17
3709100	MDS	Splice_site										107
3751801	MDS	p.E258D										20
3751801	MDS	Splice_site									j.	28
7803089	MDS	Splice site										118
7803089	MDS	Splice_site									F	20
7884287	MDS	n 0104*										85
7884287	MDS	Splice_site									ī	54
7901520	MDS	n \/272E									1	24
7891530	MDS	Splice_site									i i	55
7020626	MDC	Splica cita										110
7939626	MDS	Frameshift_del										8
8040244	MDS	p.V272L									I	36
8040244	MDS	Splice_site									I	38
8223840	MDS	p.V173M										241
8223840	MDS	Splice_site									l l	80
8738788	MDS	n R282W									1	568
8238788	MDS	Splice_site										18
8500527	MDS	n D010*										214
8590527	MDS	Splice_site										18
		B2001/									-	
8592327	MDS	p.R280K										206
8592327	WIDS	spice_site									1	44
8610838	MDS	p.V173M										241
8610838	MDS	Splice_site										58
8677482	MDS	Splice site										78
8677482	MDS	Splice_site									1	80
8706141	MDS	p.R248W									1	949
8706141	MDS	Splice_site										69
0000026	MDS	n M246V										154
8809836	MDS	Splice_site									F	14
8021622	MOG	- 0200*										<b>C12</b>
8921623	MDS	p.R306* Splice_site										612 118
				_								
9889870 9889870	MDS	Splice_site Frameshift_del										56 4
9978666 9978666	MDS MDS	p.I232F Splice_site										27 55
											_	
CLL125 CLL125	MDS MDS	Frameshift_del Frameshift_del										173 73
AML_500 AML_500	AML AML	p.R282W Splice_site									1	1568 7
1072	AML / MDS	p.C124R Splice site										13 11
10/2	AIVIL / IVIDS	Spice_site									ł	
1085	AML / MDS	p.P250L										150
1085	AML / MDS	Splice_site									)	17
800684	AML / MDS	Splice site									1	56
800684	AML / MDS	p.A161T										185

### Supplemental Figure S1b (part 1)

Sample_ID	Disease	Variant	w	м	В	14	Α	G	Ν	Р	Frequency_graph
40	MDS	p.N239D									132
40	MDS	Frameshift_ins									1
12/		n C1255									24
134	t-MDS / t-AMI	Frameshift indel									4
10.	<i>cbby c ic</i>	indirectioninder									•
37	t-MDS / t-AML	p.R248Q									2500
37	t-MDS / t-AML	Frameshift_ins									2
_			_								
3	MDS	p.R282P									62
3	IVIDS	Frameshint_der									/
OCI-M1	AML	p.L145R									25
OCI-M1	AML	Splice_site									20
			_								-
RK14	AML	Splice_site									80
RK14	AML	p.Y220H									65
400	ΔΜΙ	Solice site									80
400	AML	p.Y220H									65
		P									
13	AML	Splice_site									78
13	MDS	p.R283P									98
10		- 142271									202
19		p.IVI2371 Frameshift ins									362
15	AIVIE	Traniesinit_ins									
OCI-M1	AML	p.L145R									25
OCI-M1	AML	Splice_site									20
11	AML	p.R158H									310
11	AIVIL	Frameshift_Ins									1
899	AML	p.A138V									121
899	AML	Frameshift ins									1
900	AML	p.Y220C									1230
900	AML	Frameshift_del									48
907	AN/I	Solica sita									15
907	AML	p.P278S									301
		P									
933	AML	p.H179R									565
933	AML	Frameshift_del									5
	MDC										
UPN-11	MDS	P.G200E Frameshift ins									221
01111	WID5	Tunicshint_ins									-
TCGA-AB-2829	AML	Splice_site									42
TCGA-AB-2829	AML	p.R280G									135
TCGA-AB-2878	AML	p.S215G									71
TCGA-AB-2676	AIVIL	Frameshint_der									0
TCGA-AB-2938	AML	p.H179R									565
TCGA-AB-2938	AML	Frameshift_del									48
137404	MDS	p.V272L									134
137404	MDS	Frameshift_del									1
693881	MDS	Solice site									64
693881	MDS	p.M237I									362
20	MDS	p.G245S									1156
20	MDS	Frameshift_indel									2
Λ	AN41	n \/216M									נדר
4	AML	Frameshift del									2/3

### Supplemental Figure S1b (part 2)

Sample_ID	Disease	Variant	w	м	В	14	Α	G	Ν	Р	Frequency_graph
002-27	AML	p.V172G									9
002-27	AML	Frameshift_del									7
262.01	A.M.I	Splica sita									٥
363-01	AMI	n C275Y									243
505 01	,	p:02/01									
198041	t-MDS / t-AML	p.Y163H									76
198041	t-MDS / t-AML	Frameshift_del									2
400992	t-MDS / t-AML	p.R273C									2168
400992	t-MDS / t-AML	Frameshift_del									3
122697		n P206*									612
433687	t-MDS / t-AMI	Frameshift del									32
837334	t-MDS / t-AML	p.G112R									1
837334	t-MDS / t-AML	Frameshift_ins									1
											-
889867	t-MDS / t-AML	p.L265P									60
889867	t-MDS / t-AML	Frameshift_del									34
1 42	A.M.I	Splico sito									110
1_42		Erameshift del									118
1_42	7.0012	Tranconnt_uci									-
1 54	AML	p.C135F									188
1_54	AML	Splice_site									80
5_110	AML	p.V216M									273
5_110	AML	Splice_site									55
0011170-		= C2C2V									
PD11178a	AMI	p.0202v Frameshift del									
10111/00	AWL	Trainesinit_der									-
PD11213a	AML	p.S215R									56
PD11213a	AML	Splice_site									80
PD11215a	AML	Frameshift_del									21
PD11215a	AML	Frameshift_ins									1
PD93173	AMI	Solica sita									42
PD9312a	AML	p.R175H									3319
		P									
3115973	MDS	Splice_site									80
3115973	MDS	Splice_site									28
			_	_	_				_	_	-
3157019	MDS	p.G244C									144
3157019	MDS	Frameshift_ins									3
33707/0	MDS	n R273H									2300
3329749	MDS	Frameshift ins									1
											•
3333824	MDS	p.Y103*									14
3333824	MDS	Splice_site									107
				_							
3431933	MDS	Splice_site									60
3431933	MDS	Frameshift_del									1
3468465	MDS	n G245D									470
3468465	MDS	Splice site									25
3490509	MDS	Frameshift_del									7
3490509	MDS	Frameshift_ins									3
											<b>L</b> _
3556079	MDS	p.A159P									136
3556079	MDS	Frameshift_ins									1
3586309	MDC	n 0165*									144
3586308	MDS	Snlice site									144 52
5555500		oplice_site									r 55
3668658	MDS	Splice_site									14
3668658	MDS	Splice_site									44

### Supplemental Figure S1b (part 3)

Sample_ID	Disease	Variant	w	М	В	14	Α	G	Ν	Р	Frequency_graph	
CG2	AML	p.Y236H										44
CG2	AML	p.M237V			1.1	1.1	1.1	1.1	1.1	1.1		34
CG2	AML	p.R248Q										2500
AML250T	AML	p.Y163C		1.1								499
AML250T	AML	p.R196*										837
AML250T	AML	p.C275W										30
	MDC	5 D17EU										2210
	IVIDS	p.R175H								-		5519
	MDS	p.c176K		-	-							54
VIVIK4	IVID3	p.C176W	1.11	1.1	1.1						•	00
6	MDS	n R175H										2210
6	MDS	n C238Y	-									289
6	MDS	n G245D		-		-	-		-	-		470
Ū		p.02.100										
3058001	MDS	p.C238Y										289
3058001	MDS	p.M237l										362
3058001	MDS	Frameshift ins										2
3083841	MDS	p.I195T		-								359
3083841	MDS	p.L194R		-								224
3083841	MDS	p.C135Y										220
3140470	MDS	p.S241P										32
3140470	MDS	p.Y220C										1230
3140470	MDS	p.R249K										58
3278690	MDS	p.R283C			1.1	-	1.1	-	-	1		113
3278690	MDS	p.G244S										170
3278690	MDS	Frameshift_del										24
3598212	MDS	p.G245V										239
3598212	MDS	p.R273H								1.0		2300
3598212	MDS	Frameshift_ins										1
			_									
3771132	MDS	p.R306*										612
3771132	MDS	p.R213*										1214
3771132	MDS	p.R280T										268
2000 420	MDC											
3809429	MDS	p.R248W										1949
3809429	IVIDS	p.C238R		1.1								59
3809429	IVIDS	p.K132N										89
8372075	MDS	n P212*										121/
8372975	MDS	p.N213										1214
8372975	MDS	Splice site										52
8372373	IVID 5	spilce_site										52
8527776	MDS	n 0167*										113
8527776	MDS	p.V274A									ī	63
8527776	MDS	Splice site									-	9
8536126	MDS	p.M237I										362
8536126	MDS	Frameshift_del										6
8536126	MDS	Frameshift_del										1
8742054	MDS	p.R248W										1949
8742054	MDS	p.R175H								1		3319
8742054	MDS	p.R273H										2300
8866992	MDS	p.G244V										56
8866992	MDS	p.R213Q										165
8866992	MDS	p.C275Y										243
			_								_	
8919150	MDS	p.H193R		1.1								341
8919150	MDS	p.R175H		1.1								3319
8919150	MDS	p.R273H								1.1		2300
011447	1400	- 114000										
CLL117	MDS	p.H193D		1.1	1.1							32
CLL117	MDS	Splice_site										1
CLL11/	NDS	р.к248Q										2500
	MDC	n   10/P										224
MDS155	MDS	p.L194N	-									565
MDS155	MDS	n C176S										53
		p.01/00										

### Supplemental Figure S1c

Supplemental material

Sample_ID	Disease	Variant	W	М	В	14	Α	G	Ν	Р	Frequency_graph
PD9260a	AML	p.A276P									37
PD9260a	AML	p.R248Q									2500
PD9260a	AML	p.R196Q			_	-	_	-	I	I	24
PD9260a	AML	p.K132R		-	_	-		-	_	_	202
	0										
3024120	MDS	p.R248Q									2500
3024120	MDS	p.R248W									1949
3024120	MDS	p.Y220H		-							65
3024120	MDS	p.R273H								_	2300
	0										
3143975	MDS	Splice_site									2
3143975	MDS	p.M237I		-	_						121
3143975	MDS	Frameshift_del									3
3143975	MDS	Framshift_ins									1
	0										
3282470	MDS	p.R248Q									2500
3282470	MDS	p.H179R	_	I	-	I	-	I.		_	565
3282470	MDS	p.C275Y									243
3282470	MDS	p.R273H								_	2300
	0										
8001517	MDS	p.K319N		-	-	-		-			2
8001517	MDS	Frameshift_del									2
8001517	MDS	p.K291N	_	_							2
8001517	MDS	Frameshift_del									6
	0										
8753539	MDS	p.G245D		-	I		_		_	_	470
8753539	MDS	p.I195T		_	_	_	_		_	_	359
8753539	MDS	p.R273C									2168
8753539	MDS	p.F134L									14
											•

### **Supplemental Figure S1d**

Sample_ID	Disease	Variant	w	м	В	14	Α	G	Ν	Р	Frequency_graph
8205391	MDS	p.E285K			1		1	1		1	498
8205391	MDS	p.R248Q									2500
8205391	MDS	p.R175G						1		1	91
8205391	MDS	p.V272M			-	_	-	1		1	321
8205391	MDS	Splice_site									58

### **Supplemental Figure S1e**



### **Supplemental Figure S1f**

**Supplemental Figure S2:** Venn diagram showing the mutational concordance of validated somatic variants based on the sequencing strategy



- \* Mutation outside the amplicon used for SMRT analysis
- \*\* Mutation detected at very low frequency by SMRT (1-2%)
- \*\*\* Mutation detected at high frequency in the recurrent sample and identified at low frequency after manual examination of the primary sample.
- \*\*\*\* Mutation detected by SMRT and identified at very low frequency by manual examination of the short-read NGS data.



**Supplemental Figure S3**: Variant allele frequency (VAF) observed for all *TP53* variants identified by both classical NGS and SMRT methodologies

Supplemental Figure S4a to k: Detailed analysis of the 11 patients included in this study.

For each patient, 3 sections are available i.e. clinical information, sequencing and haplotype

**Clinical information:** this section includes age, disease information, treatment and 17p status

**Sequencing**: Sanger sequencing and/or standard NGS analysis is shown in the left part. No allelic distribution can be inferred from this type of analysis. SMRT sequencing (right part) provides an accurate picture of the allelic distribution of each *TP53* variant, as well as the remaining wt allele. The frequencies of the different alleles are shown in brackets.

Red triangle:	TP53 variants identified by both types of analysis.
White triangle:	TP53 variants detected only by SMRT sequencing.
Yellow triangle:	TP53 Variants detected after manual examination but
	below the cut-off used for clinical validation
Black triangle : sequencing	TP53 variants outside the amplicon used for the long range
Blue triangle:	TP53 variants not detected by long range sequencing.

**Haplotype:** allelic distribution of all *TP53* variants (germline and somatic) according to the SMRT analysis

Somatic *TP53* variants are shown in red. Biallelic germline variants (SNP) are shown in white (allele 1) and green (allele 2) to make a distinction for heterozygote cases (see cases Fr10).

Disease:de novo MK-AMLAge:77Treatment:none (diagnosis)17p status:no deletion

### Sequencing



▶ Variant outside the amplicon used for the long range sequencing



rs1800370	rs1042522	rs1794287	rs2909430	rs1625895	c.673-2A>G	rs12947788	rs12951053	Frequency 56.6 %	Reads 5504
rs1800370	rs1042522	rs1794287	rs2909430	rs1625895		rs12947788	rs12951053	43.3 %	4206
	Sor	natic varian	t	-	Germli	ne SNP			

Disease:de novo MK-AMLAge:63Treatment:none17p status:deletion (CGH array)

### Sequencing

٥

# Sanger/NGS (short reads)

c.743G>A (47 %)

c.673-2A>T (41 %)

Variant detected by both analyses

## Haplotype

									Frequency	Reads
rs1800370	rs1042522	rs1794287	rs2909430	rs1625895		c.743G>A	rs12947788	rs12951053	43.8 %	4776
rs1800370	rs1042522	rs1794287	rs2909430	rs1625895	c.673-2A>T		rs12947788	rs12951053	40.8 %	лллл
					_				40.0 %	
rs1800370	rs1042522	rs1794287	rs2909430	rs1625895			rs12947788	rs12951053	15.0.4	
									15.3 %	1666
	Sor	natic varia	nt		Ge Ge	ermline SI	NP			
		natio varia					••			



**SMRT** sequencing (long reads)

c.743G>A (43.8 %) c.673-2A>T (40.8 %) wt (15.3 %)



Disease:de novo MK-AMLAge:73Treatment:none17p status:deletion (FISH)

## Sequencing



### Variant detected by both analyses

## Haplotype

									Frequency	Reads
rs1800370	rs1042522	rs1794287	rs2909430	c.413C>T	rs1625895	rs12947788	rs12951053		68.6 %	5672
rs1800370	rs1042522	rs1794287	rs2909430		rs1625895	rs12947788	rs12951053		27.5 %	2278
rs1800370	rs1042522	rs1794287	rs2909430		rs1625895	rs12947788	rs12951053	c.794T>C	3.76 %	311
•	Soma	atic variant		ľ	Ger	mline SNF	5			

Disease:de novo MK-AMLAge:78Treatment:none17p status:no deletion (FISH)

### Sequencing





									Frequency	Reads
rs1800370	rs1042522	rs1794287	rs2909430	c.395A>G	rs1625895	rs12947788	rs12951053		36.6 %	1134
rs1800370	rs1042522	rs1794287	rs2909430		rs1625895	rs12947788	rs12951053	c.824G>T	36.0 %	1117
rs1800370	rs1042522	rs1794287	rs2909430		rs1625895	rs12947788	rs12951053		27.3 %	847
	Sor	natio varia	nt			ormling SN	D			
	- 30	nalic vana	i it				1			

Disease:de novo MK-AMLAge:73Treatment:none17p status:deletion (FISH)

### Sequencing



▶ Variant detected by both analyses

									Frequency	Reads
rs1800370	rs1042522	rs1794287	rs2909430		c.637C>T	rs1625895	rs12947788	rs12951053	72.3 %	3908
rs1800370	rs1042522	rs1794287	rs2909430			rs1625895	rs12947788	rs12951053	22.1 %	1198
	_								22.1 )0	
rs1800370	rs1042522	rs1794287	rs2909430	c.455C>T		rs1625895	rs12947788	rs12951053	E 40 9/	202
									5.42 %	293
	S	omatic var	iant			Germline S	NP			
					<u> </u>					

Disease:s-AML (post LR-MDS del5q)Age:75Treatment:Lenalidomide17p status:no deletion (karyotype)

### Sequencing

b



▶ Variant detected by both analyses

Variant not detected by long range sequencing

rs1800370	rs1042522	rs1794287	rs2909430	rs1625895		rs12947788	rs12951053	Frequency 50.9 %	Reads 2965
rs1800370	rs1042522	rs1794287	rs2909430	rs1625895	c.844C>T	rs12947788	rs12951053	49.0 %	2853
	Som	atic variant			Germlir	ne SNP			

Patient Fr7 Sample 7a and 7b

Disease:	LR-MDS del5q
Age:	73
<u>Treatment</u>	sample 7a: Lenalidomide
	sample 7b: Lenalidomide
<u>17p status:</u>	no deletion

### Sequencing



Variant detected only by the long range sequencing

Variant detected by both analyses

## Haplotype

•

7a												
2 01											Frequency	Reads
rs1800370	rs1042522		rs1794287	rs2909430	rs1625895				rs12947788	rs12951053	69.7 %	4646
rs1800370	rs1042522	c.314G>T	rs1794287	rs2909430	rs1625895				rs12947788	rs12951053	22.5 %	1501
rs1800370	rs1042522		rs1794287	rs2909430	rs1625895		c.743G>A		rs12947788	rs12951053	6.17 %	411
rs1800370	rs1042522		rs1794287	rs2909430	rs1625895	c.742C>T			rs12947788	rs12951053	<1%	51
rs1800370	rs1042522		rs1794287	rs2909430	rs1625895			c.818G>A	rs12947788	rs12951053	<1%	50

Aug 2008 (7a)

.

7b

∨ Jul 2013 (7b)

rs1800370	rs1042522		rs1794287	rs2909430					rs1625895		rs12947788	rs12951053			Frequency 35.2 %	Reads 2280
rs1800370	rs1042522		rs1794287	rs2909430					rs1625895	c.743G>A	rs12947788	rs12951053			28.3 %	1835
rs1800370	rs1042522	c.314G>T	rs1794287	rs2909430					rs1625895		rs12947788	rs12951053			22.9 %	1487
rs1800370	rs1042522		rs1794287	rs2909430		c.584T>A			rs1625895		rs12947788	rs12951053			6.99 %	453
rs1800370	rs1042522		rs1794287	rs2909430					rs1625895		rs12947788	rs12951053	c.818G>A		2.91 %	189
rs1800370	rs1042522		rs1794287	rs2909430			c.614A>G		rs1625895		rs12947788	rs12951053			1.29 %	84
rs1800370	rs1042522		rs1794287	rs2909430					rs1625895		rs12947788	rs12951053		c.833C>G	< 1 %	64
rs1800370	rs1042522		rs1794287	rs2909430				c.659A>G	rs1625895		rs12947788	rs12951053			<1%	53
rs1800370	rs1042522		rs1794287	rs2909430	c.524G>A				rs1625895		rs12947788	rs12951053			< 1 %	30



Somatic variant



Germline SNP

Figure S4g

Disease:s-AML (post LR-MDS del5q)Age:72Treatment:Lenalidomide17p status:deletion

### Sequencing

.

.



Variant detected by both analyses

rs1800370	rs1042522	rs1794287	rs2909430	rs1625895		rs12947788	rs12951053	Frequency 50.6 %	Reads 3518
rs1800370	rs1042522	rs1794287	rs2909430	rs1625895	c.725G>T	rs12947788	rs12951053	49.4 %	3439
-	Somati	c variant			Germli	ine SNP			

Disease:s-AML (post LR-MDS del5q)Age:76Treatment:Lenalidomide17p status:deletion (partial)

### Sequencing



Variant detected after manual examination but below the cut-off used for clinical validation

Variant detected by both analyses

Sample 9a: frozen pellet from whole blood leukocytes Sample 9b: cytogenetic pellet from bone marrow (same timepoints)

## Haplotype

									Frequency	Reads
rs1800370	rs1042522	rs1794287	rs2909430		rs1625895		rs12947788	rs12951053	57.8 %	5358
rs1800370	rs1042522	rs1/94287	rs2909430	c.4211>G	rs1625895		rs12947788	rs12951053	33.4 %	3094
rs1800370	rs1042522	rs1794287	rs2909430		rs1625895	c.711G>T	rs12947788	rs12951053	8.73 %	809
									Frequency	Reads
rs1800370	rs1042522	rs1794287	rs2909430		rs1625895		rs12947788	rs12951053	Frequency 56.4 %	Reads 4168
rs1800370	rs1042522	rs1794287	rs2909430		rs1625895		rs12947788	rs12951053	Frequency 56.4 %	Reads 4168
rs1800370	rs1042522 rs1042522	rs1794287 rs1794287	rs2909430 rs2909430	c.421T>G	rs1625895		rs12947788	rs12951053	Frequency 56.4 %	Reads 4168 2296
rs1800370	rs1042522 rs1042522	rs1794287 rs1794287	rs2909430 rs2909430	c.421T>G	rs1625895 rs1625895		rs12947788	rs12951053	Frequency 56.4 % 31.1 %	Reads 4168 2296
rs1800370	rs1042522 rs1042522	rs1794287 rs1794287	rs2909430 rs2909430	c.421T>G	rs1625895	c.711G>T	rs12947788	rs12951053	Frequency 56.4 % 31.1 %	Reads 4168 2296





Germline SNP

Figure S4i Supplemental material

Disease:	sample 10a: LR-MDS del5q
	sample 10b: s-AML (post LR-MDS del5q)
<u>Age:</u>	69
<u>Treatment</u>	sample 10a: Lenalidomide
	sample 10b: Lenalidomide
<u>17p status:</u>	no deletion (karyotype)

### Sequencing



Variant detected only by the long range sequencing

Variant detected by both analyses

### Patient Fr10 Sample 10a and 10b

## Haplotype

•

rs1800370	rs1042522	rs1794287	rs2909430	rs1625895			rs12947788	rs12951053			Frequency 59.6 %	Reads 1394
rs1800370	rs1042522	rs1794287	rs2909430	rs1625895			rs12947788	rs12951053			22.0 %	516
rs1800370	rs1042522	rs1794287	rs2909430	rs1625895			rs12947788	rs12951053		c.844C>T	10.8 %	253
rs1800370	rs1042522	rs1794287	rs2909430	rs1625895		c.743G>A	rs12947788	rs12951053			3.68 %	86
rs1800370	rs1042522	rs1794287	rs2909430	rs1625895			rs12947788	rs12951053	c.817C>T		2.61 %	61
rs1800370	rs1042522	rs1794287	rs2909430	rs1625895		c.743G>A	rs12947788	rs12951053			< 1 %	13
rs1800370	rs1042522	rs1794287	rs2909430	rs1625895	c.742C>T		rs12947788	rs12951053			< 1 %	12
rs1800370	rs1042522	rs1794287	rs2909430	rs1625895	c.742C>T		rs12947788	rs12951053			< 1 %	1

### August 2008 (10a)

V

.

### March 2009 (10b)

rs1800370 rs1042522 rs1794287 rs2909430 rs1625895 c.743G>A rs12947788 rs12951053	<1%	5
rs1800370 rs1042522 rs1794287 rs2909430 rs1625895 c.743G>A rs12947788 rs12951053	3.60 %	77
rs1800370 rs1042522 rs1794287 rs2909430 rs1625895 rs12947788 rs12951053	9.60 %	205
rs1800370 rs1042522 rs1794287 rs2909430 rs1625895 rs12947788 rs12951053 c.844C>T	26.8 %	574
rs1800370 rs1042522 rs1794287 rs2909430 rs1625895 rs12947788 rs12951053	Frequency 59.6 %	Reads

Figure S4j

Disease:	sample 1: LR-MDS del5q
	sample 2: LR-MDS del5q )
<u>Age:</u>	85
<u>Treatment</u>	sample 11a: Lenalidomide
	sample 11b: Lenalidomide
<u>17p status:</u>	no deletion (karyotype)

### Sequencing

Sanger/NGS (Sh	ort reads)	SMI	RT sequei	ncing (long reads)
<ul> <li>▼ c.659A&gt;G (?)</li> <li>▼ c.840A&gt;T (?)</li> <li>(Sanger only)</li> </ul>	2014 (11a)	•	4	c.659A>G (18.9%) c.701A>G (8.43%) wt (a) (45.7) %) wt (b) (26.3) %)
<ul> <li>▼ c.701A&gt;G (29%)</li> <li>▼ c.659A&gt;G (27%)</li> <li>▼ c.840A&gt;T (5%)</li> </ul>	∨ 2015 (11b)	•	₽ <b>₩</b>	c.659A>G (33.6.9%) c.701A>G (20.4%) c.840A>T (2.14%) wt (a) 24.4%) wt (b) (18 %)

Variant detected only by the long range sequencing

- Variant detected by both analyses
- Variant not detected by long range sequencing

### Patient Fr11 Sample 11a and 11b

## Haplotype

•

										Frequency	Reads
rs18	00370	rs1042522	rs1794287	rs2909430		rs1625895		rs12947788	rs12951053	45.7 %	1948
rs18	00370	rs1042522	rs1794287	rs2909430		rs1625895		rs12947788	rs12951053	26.3 %	1121
rs18	00370	rs1042522	rs1794287	rs2909430	c.659A>G	rs1625895		rs12947788	rs12951053	18.9 %	808
		_			_						
rs18	00370	rs1042522	rs1794287	rs2909430		rs1625895	c.701A>G	rs12947788	rs12951053	8 43 %	359
										0.10 %	000
rs18	00370	rs1042522	rs1794287	rs2909430	c.659A>G	rs1625895		rs12947788	rs12951053	~ 1 %	20
										< 1 %	20

Oct 2014 (11a)

•

### V Juil 2015 (11b)

										Frequency	Reads
rs1800370	rs1042522	rs1794287	rs2909430	c.659A>G	rs1625895		rs12947788	rs12951053		33.6 %	974
rs1800370	rs1042522	rs1794287	rs2909430		rs1625895		rs12947788	rs12951053		24.4 %	707
rs1800370	rs1042522	rs1794287	rs2909430		rs1625895	c.701A>G	rs12947788	rs12951053		20.4.9/	502
										20.4 /8	393
rs1800370	rs1042522	rs1794287	rs2909430		rs1625895		rs12947788	rs12951053		19.0 %	501
										10.0 /8	521
rs1800370	rs1042522	rs1794287	rs2909430		rs1625895		rs12947788	rs12951053	c.840A>T	0149/	60
										2.14 %	02
rs1800370	rs1042522	rs1794287	rs2909430	c.659A>G	rs1625895		rs12947788	rs12951053		1.27 %	37

Germline SNP

Somatic variant

**Supplemental Figure S5:** Visualization of NGS alignment, confirming that the two variants are located on different alleles for patient Fr2.



Patient Fr2

**Supplemental Figure S6:** Visualization of NGS alignment, confirming that the two variants are located on different alleles for patient Fr7



Patient Fr7 Sample 7b

#### Α rs1800370 rs1042522 rs1625895 rs12947788 rs12951053 rs2909430 v , 5 6 8 7 4 rs1794287 Rev For

Supplemental Figure S7: strategy used for the analysis of TP53 mutations

В

SNP	cDNA_variant	Genomic_variant (HG19)	Protein variant
rs1800370	c.108G>A	chr17:g.7579579G>A	p.P36=
rs1042522	c.215C>G	chr17:g.7579472C>G	p.P72R
rs1794287	c.376-283T>C	chr17:g.7578837T>C	p.(=)
rs2909430	c.376-91G>A	chr17:g.7578645G>A	p.(=)
rs1625895	c.672+62A>C	chr17:g.7578115A>C	p.(=)
rs12947788	c.782+72C>T	chr17:g.7577427C>T	p.(=)
rs12951053	c.782+92T>G	chr17:g.7577407T>G	p.(=)

С

	Sequence	Coordinates (HG19)
Forward primer	5' cctggtcctctgactgctct 3'	7579626-7579607
Reverse primer	5' tacctcgcttagtgctccct 3	7577035-7577016

Supplemental material