



HAL
open science

The Broad Spectrum of TP53 Mutations in CLL: Evidence of Multiclinality and Novel Mutation Hotspots

Grégory Lazarian, Bernard Leroy, Floriane Theves, Myriam Hormi, Rémi Letestu, Virginie Eclache, Giulia Tueur, Adam Ameer, Audrey Bidet, Pascale Cornillet-Lefebvre, et al.

► To cite this version:

Grégory Lazarian, Bernard Leroy, Floriane Theves, Myriam Hormi, Rémi Letestu, et al.. The Broad Spectrum of TP53 Mutations in CLL: Evidence of Multiclinality and Novel Mutation Hotspots. Human Mutation, 2023, 2023, pp.1-17. 10.1155/2023/4880113 . hal-04147818

HAL Id: hal-04147818

<https://hal.science/hal-04147818>

Submitted on 1 Jul 2023

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.

Title

The broad spectrum of *TP53* mutations in CLL: evidence of multiclonality and novel mutation hotspots.

Running title (60 max)

Spectrum of *TP53* mutations in CLL

Grégory Lazarian¹, Bernard Leroy², Floriane Theves³, Myriam Hormi¹, Rémi Letestu¹, Virginie Eclache¹, Giulia Tueur¹, Adam Ameur⁴, Audrey Bidet⁵, Pascale Cornillet-Lefebvre⁶, Frédéric Davi³, Eric Delabesse⁷, Marie-Hélène Estienne⁸, Pascaline Etancelin⁹, Olivier Kosmider¹⁰, Sophy Laibe¹¹, Marc Muller¹², Nathalie Nadal¹³, Dina Naguib¹⁴, Cédric Pastoret¹⁵, Stéphanie Poulain¹⁶, Pierre Sujobert¹⁷, Lauren Veronese¹⁸, Samia Imache¹, Valérie Lefebvre¹, Florence Cymbalista¹, Fanny Baran-Marszak¹, Thierry Soussi^{2,19,20}. On behalf of the French Innovative Leukemia Organization (FILO)

1 Service d'hématologie biologique, CHU Hôpital Avicenne, Bobigny, 93000, France

2 Sorbonne Université, place Jussieu, 75005 Paris, France

3 Service d'hématologie Clinique, CHU Pitié Salpêtrière, Paris 75013

4 Science for Life Laboratory, Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden

5 Service d'hématologie biologique, CHU Bordeaux-Haut Lévêque, Bordeaux, 33000, France

6 Service d'hématologie biologique, CHU Hôpital Robert Debré, Reims, 51100, France

7 Laboratoire d'Hématologie, CHU Toulouse; Inserm 1037, CNRS, Université Toulouse III-Paul Sabatier, Centre de Recherches en Cancérologie de Toulouse, Toulouse, France

8 Service d'hématologie biologique, CHRU Tours, Tours, 37044, France

9 Laboratoire de Génétique Oncologique, Département de Biopathologie, Centre Henri Becquerel, 76038 ROUEN

10 Service d'hématologie biologique, CHU Hôpital Cochin, Paris, 75014, France

11 XPath Cytogénétique et biologie moléculaire, Marseilles, 13010, France

12 Laboratoire de génétique médicale, CHRU Nancy Hôpitaux de Brabois, Vandoeuvre-Lès-Nancy, 54511, France

13 Service de génétique chromosomique et moléculaire, CHU Dijon, Dijon, 21000, France

14 Service d'hématologie biologique, CHU Caen, Caen, 14033, France

15 Service d'hématologie biologique, CHU Rennes, Rennes, 35000, France

16 Service d'Hématologie Cellulaire, CHRU Lille, Lille, 59000, France

17 Service d'Hématologie Cellulaire, Hospices civils de Lyon, Lyon, 69002, France

18 Service de cytogénétique médicale, CHU Clermont-Ferrand, Clermont-Ferrand, 63000, France

19 Centre de Recherche Saint-Antoine , UMRS_938, « Hematopoietic and Leukemic Development » , Sorbonne Université, 75012 Paris

20 Dept of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, SE-751 85 Uppsala, Sweden

Correspondence:

Corresponding Authors: Fanny Baran-Marszak Service d'hématologie biologique, CHU Hôpital Avicenne, Bobigny, 93000, France; email fanny.baran-marszak@aphp.fr or Thierry Soussi, Equipe « Hematopoietic and Leukemic Development » UMRS_938 Sorbonne Université, Centre de Recherche Saint-Antoine 75012 Paris Paris France France: email: thierry.soussi@sorbonne-universite.fr or thierry.soussi@igp.uu.se

Members of the French Innovative Leukemia Organization (FILO):

Thérese Aurrant, Charles Herbaux, Guillaume Cartron, Caroline Dartigeas, Alain Delmer, Jehan Dupuis, Marie Sarah Dilhuydy, Emmanuelle Ferrant, Pierre Feugier, Franck Genevieve, Sophie De Guibert, Romain Guieze, Véronique Leblond, Vincent Levy, Stephane Leprêtre, Fatiha Merabet, Anne-Sophie Michallet, Florence Nguyen-Khac, Anne Quinquenel, Sophie Raynaud, Daniel Re, Catherine Thieblemont, Olivier Tournilhac, Xavier Troussard, Malgorzata Truchan-Graczyk, Lise Willems, Loic Ysebaert, Jean-Marc Zini.

Author contributions:

Conception and design:

Florence Cymbalista, Fanny Baran-Marszak and Thierry Soussi.

Acquisition of data (acquired and managed patients, provided facilities, etc.):

Grégory Lazarian, Floriane Theves, Myriam Hormi, Rémi Letestu, Virginie Eclache, Audrey Bidet, Pascale Cornillet-Lefebvre, Frédéric Davi, Eric Delabesse, Marie-Hélène Estienne, Pascaline Etancelin, Olivier Kosmider, Sophie Laibe, Laurence Lode, Marc Muller, Nathalie Nadal, Dina Naguib, Cédric Pastoret, Stephanie Poulain, Pierre Sujobert, Lauren Veronese, Samia Imache, Valérie Lefebvre, Florence Cymbalista, Fanny Baran-Marszak and the members of the French Innovative Leukemia Organization

Development of methodology:

Adam Ameur

Analysis and interpretation of data (e.g. statistical analysis, biostatistics, computational analysis):

Grégory Lazarian, Bernard Leroy, Floriane Theves, Florence Cymbalista, Fanny Baran-Marszak and Thierry Soussi.

Writing, review, and/or revision of the manuscript:

Grégory Lazarian Florence Cymbalista, Fanny Baran-Marszak and Thierry Soussi.

Study supervision:

Florence Cymbalista, Fanny Baran-Marszak and Thierry Soussi.

Acknowledgments

This study was supported by Cancéropole Île-de-France (convention n°2019-1-EMERG-22-INSERM 6-1) for TS and FBM. We graciously thank Gisela Barbany and Anh-nhi Tran (Karolinska Institute) for providing sample and clinical data. Samples from CLL patients from university hospital of Toulouse have been obtained after informed consent and stored at the HIMIP collection (BB-0033-00060). According to the French law, HIMIP collections has been declared to the Ministry of Higher Education and Research (DC 2008-307 collection 1) and obtained a transfer agreement (AC 2008-129) after approbation by the "Comité de Protection des Personnes Sud-Ouest et Outremer II" (ethical committee).

Abstract

TP53 aberrations are a major predictive factor of resistance to chemoimmunotherapy in chronic lymphocytic leukemia (CLL) and an assessment of them before each line of treatment is required for theranostic stratification. Acquisition of subclonal *TP53* abnormalities underlies the evolution of CLL. To better characterize the distribution, combination and impact of *TP53* variants in CLL, 1,056 *TP53* variants collected from 683 patients included in a multicenter collaborative study in France were analyzed and compared to UMD_CLL, a dataset built from published articles collectively providing 5,173 *TP53* variants detected in 3,808 patients.

Our analysis confirmed the presence of several CLL-specific hotspot mutations, including a two-base-pair deletion in codon 209 and a missense variant at codon 234, the latter being associated with alkylating treatment. Our analysis also uncovered a novel unidentified CLL-specific variant in the splice acceptor signal of intron 6 leading to the use of a cryptic splice site, similarly utilized by *TP53* to generate p53psi, a naturally truncated p53 isoform localized in the mitochondria. Examination of both UMD_CLL and several recently released large-scale genomic analyses of CLL patients confirmed that this splice variant is highly enriched in this disease when compared to other cancer types. Using a *TP53*-specific single-nucleotide polymorphism, we also confirmed that copy-neutral loss of heterozygosity is frequent in CLL. This event can lead to misinterpretation of *TP53* status. Unlike other cancers, CLL displayed a high proportion of patients harboring multiple *TP53* variants. Using both *in silico* analysis and single molecule smart sequencing, we demonstrated the coexistence of distinct subclones harboring mutations on distinct alleles. In summary, our study provides a detailed *TP53* mutational architecture in CLL and gives insights into how treatments may shape the genetic landscape of CLL patients.

Introduction

Chronic lymphocytic leukemia (CLL) is a highly heterogeneous disease in terms of clinical outcomes and chemoimmunotherapy responses ^[1]. Despite improvements in care, CLL remains incurable. Even after prolonged responses to therapy, patients will relapse and thus need multiple lines of treatment ^[2]. Since the first publication of *TP53* alterations in CLL in the early 1990s, numerous teams have confirmed the high variety and prevalence of *TP53* alterations in the pathology. Predominantly, these alterations manifest as a deletion of the gene on the short arm of chromosome 17 (17p13.1) and a missense mutation in the second allele ^[3, 4]. *TP53* alterations are relatively infrequent in treatment-naïve CLL patients (10%), but their incidence may reach 50% to 60% in those with fludarabine-refractory disease. In a seminal paper published in 2000, Döhner *et al.* showed that 17p deletion (del(17p)) was associated with markedly decreased survival and that it predicted impaired response to chemoimmunotherapy ^[5]. Their findings were confirmed in many subsequent studies, making *TP53* status the main predictive marker in CLL for the selection of appropriate treatments ^[6–8]. The detection of del(17p) and *TP53* mutations has become an integral part of routine diagnostics and should be performed before any administration of treatment. The advent of next generation sequencing (NGS) has changed the CLL paradigm ^[9]. First, it has led to a better understanding of the heterogeneous nature of the disease, with the discovery of multiple driver genes associated with its development. Second, concerning *TP53*, it has shed light on new features such as the occurrence of minor clones, which had remained undetectable by conventional Sanger sequencing ^[10–15]. The clinical value of these minor clones is still under investigation ^[16]. In a previous study, using a cohort of 336 *TP53* mutated CLL patients, we uncovered a novel *TP53* mutation hotspot in codon 234 associated with chlorambucil treatment ^[17].

In the present study, we collected retrospective data on *TP53* mutated patients from centers affiliated with the French Innovative Leukemia Organization-CLL (FILO) in GBMHH (French Molecular Biology Group in Hematology) laboratories. All centers contributing to the present work had GBMHH or ERIC (European Research Initiative on CLL) quality control certification ^[18].

Results

The FILO cohort and UMD_CLL dataset

To our knowledge, the FILO cohort is, as of this writing, the largest aggregation of *TP53* mutated CLL patients. For the present study, it provided 1,056 *TP53* variants collected from 683 patients analyzed either by conventional Sanger analysis (172 patients, 196 *TP53* variants) or by NGS (511 patients, 860 *TP53* variants) (**Figure 1, Supplementary Table S1 to S3**). Only variants located in the major transcript (NM_000546) and targeting the main *TP53* isoform (NP_000537) will be discussed here. To extend our comparison and analysis, we used the UMD_CLL dataset. This latter includes all CLL patients from UMD_TP53 (excluding the FILO cohort) and provides 5,173 *TP53* variants (3,808 patients) that have been manually curated to remove duplicate entries (**Material and methods**). The FILO cohort includes a mix of variants, extending from rare variants to mutation hotspots, a classical schema seen also in UMD_CLL and other types of cancer (**Figure 2A**). As expected, most of the infrequent or unique variants were associated with insertions or deletions (Indel) (**Figure 2A**). *TP53* variant pathogenicity was defined using the ACMG criteria included in the UMD_TP53 database^[19]. Pathogenic (P) or likely pathogenic (LP) variants were identified in 82% of the patients in the FILO cohort. The remaining ones in that cohort were labeled as variants of uncertain significance (VUS); no benign or likely benign variants were observed therein (**Figure 2B**). In contrast, 45 patients included in UMD_CLL carried a variant defined initially as benign. However, when verified against the recent release of the new infrequent *TP53* single nucleotide polymorphism (SNP) datasets, 43 (95.5%) of those variants were redefined as benign polymorphisms (**Supplementary Table S4**)^[20]. These variants were removed from all datasets used for the subsequent analyses. Using a *TP53*-specific grading system based on the recurrence of *TP53* variants in multiple independent genomic repositories (cancer shared datasets: CSD), we have previously defined a set of 480 variants as certified oncogenic variants (**see Material and methods**)^[19]. In the FILO cohort and UMD_CLL, 52% and 55% of variants respectively were certified as oncogenic (**Figure 2C**).

A 1% variant allele frequency (VAF) cut off was used for NGS data included in the FILO cohort, but it did not lead to the selection of spurious mutations as the VAF of uncommon *TP53* variants was similar to that of frequent variants (**Figure 2D**), and furthermore similar among the ACMG classes (**Figure 2E**). We note that 315 patients displayed a VAF between 1% and 5% and thus would not have been identified via conventional sequencing. This included 71 patients with single and 244 patients with multiple mutations. The most common *TP53* mutations in the FILO cohort were missense mutations, accounting for 73% of the variants. Frameshift, inframe, nonsense and splice variants accounted respectively for 11%, 2%, 6% and 8% for the CLL cohorts. Those percentages were highly similar to those found in UMD_CLL or other cancer types (**Supplementary Figure S1**). Analysis of the mutational events targeting the *TP53* gene in patients from both the UMD database and the CLL cohorts showed that they displayed a high frequency of GC>AT and AT>GC transitions (**Supplementary Figure S2A**). GC>AT transitions were predominantly associated with hotspot variants localized in CpG dinucleotides and common to all types of cancer. In contrast, the high frequency of AT>GC transitions found in both the FILO cohort and the CLL patients included in UMD_TP53 was not observed in other cancer types (**Supplementary Figure S2A**). The same pattern was observed when analyzing data from

whole genome sequencing in four CLL samples, indicating that this pattern of mutations is indeed characteristic of CLL (**Supplementary Figure S2B**).

FISH analysis of del(17p) was performed in 208 (61%) of the next-generation sequenced cases. VAF for *TP53* variants was expectedly higher than 50% in some patients with del(17p). Strikingly however, VAF was also greater than 50% (range: 52%–98%) in 11 cases showing no del(17p), suggesting a partial or total replacement of the wildtype *TP53* locus by the mutant allele. To infer the haplotype of these tumors, we studied 11 SNPs of the *TP53* locus covered by NGS (**Supplementary Figure S3A**). An analysis of eight cases with no del(17p) and a VAF >50%, showed tumor homozygosity for all SNPs. As expected, heterozygous SNPs were identified in the eight informative cases with VAF <50% and no del(17p) (**Supplementary Figure S3B**).

TP53 hotspot variants in CLL

TP53 variants in the whole FILO cohort were predominantly distributed in the DNA-binding domain, with classical CpG-related hotspots at codons 175, 248 and 273, a distribution similar to that observed in other cancers, including hematological malignancies (**Supplementary Figure S4A**). Nevertheless, three CLL-specific mutation hotspots were observed in the FILO cohort and validated in UMD_CLL. The first was located at codon 234 (NP_000537_p.Tyr234His, NM_000546_c.700T>C), which has a very low frequency of mutation in the majority of cancers (**Supplementary Figure S4B**). Our previous study on 336 patients showed that this non-functional *TP53* variant is found mainly in patients treated with chlorambucil (CLB), an alkylating drug that had been widely used to treat CLL patients before the development of individualized therapy^[17]. A survey of the literature and data from UMD_CLL also showed an excess of mutations at codon 234 in CLL compared to other cancer types (**Supplementary Figure S4B**). In the present cohort, we also noticed that two CLB-treated patients carried two different substitutions at codon 234 (NP_000537_p.Tyr234Cys and NP_000537_p.Tyr234Ser) on different alleles (**Supplementary Figure S5A**). Remarkably, this type of event appeared to be particularly rare in any type of cancer in UMD_TP53 (**Supplementary Figure S5B**): only six of the 225,000 patients included in the whole database carried multiple missense variants at codon 234. Among them are one AML and five CLL patients, with four of these latter having been treated with CLB^[21–23] (**Supplementary Figure 5B**). The absence of this variant in recent cohorts of CLL patients in which none received CLB also supports the association between this codon and that treatment. Our finding is reminiscent of the association between exposure to certain carcinogens (aflatoxin B1 in hepatocellular carcinoma and benzo(a)pyrene in lung cancer) and *TP53* hotspot variants in codon 249 or 157^[24].

The frequency of frameshift mutations arising from deletions, insertions or duplications ranges from 5% to 8% among the various types of cancer (**Supplementary Figure S1**). The majority of frameshift variants in *TP53* are scattered along the p53 protein at very low frequency without any obvious hotspots, but a variant at codon 209 (c.626_627del) leading to premature termination (NP_000537_p.Arg209LysfsTer6) was found to be enriched in CLL (**Figure 3A and Supplementary Figure S6**). This variant was observed previously in a cohort of 254 CLL patients^[25] and the present analysis showed that it is highly specific to CLL: its frequency ranged between 1% and 3% for most cancer types but reached 17% in both the FILO cohort and UMD_CLL (**Figure 3B**). This difference was highly significant compared to all other cancer types and defined as a bona fide hotspot for CLL ($p < 0.0001$, chi test). The concerned region includes an inverted repeat that could explain the high mutability of the sequence *in vivo* (**Figure 3C**). As this variant is observed in multiple reports

using different methodologies such as conventional Sanger sequencing or NGS, an association with a methodological bias is unlikely. We also noticed that frameshift mutations at codon 210 in the FILO cohort were also elevated compared to other cancer types (4% versus 0.4% or 0.23% in colorectal and breast cancer respectively). Whether this high frequency in codons 209 and 210 is due to a paucity of frameshift mutations in other regions of *TP53* in CLL, or to an increase of this event in CLL, is currently unknown but as discussed below, a CLL-specific selection cannot be excluded.

The third hotspot is located in the splice acceptor signal of intron 6 (**Figure 4A and B**). Splice mutations (alterations of the canonical ± 1 or ± 2 splice sites) were underestimated when Sanger sequencing was used to assess *TP53* status. However, more recent exome or whole genome analysis showed that they account for 3% to 6% of *TP53* inactivation in UMD_*TP53* with no significant differences among the various histological groups (**Supplementary Figure S7A and B**). During the analysis of the FILO cohort, we noticed that variants at position NM_000546_c.673-2, the acceptor signal for intron 6, were highly enriched in CLL compared to other cancer types and represented 22% of CLL splice variants (**Figure 4A and B**). This was observed in the FILO cohort and in UMD_CLL. Catherwood *et al.* have recently described a cohort of 303 *TP53* mutated patients (429 *TP53* variants) [26]. Unreported by those authors, we noticed that splice variants at position NM_000546_c.673-2A were also the major splice variants. In both UMD_*TP53* and UMD_CLL, no other splice site hotspots were identified when analyzing all other cancer types (**Supplementary Figure S8**). Any one cancer type is characterized by specific patterns of mutational signatures arising from the various mutational processes that have occurred in the tumor (**Supplementary figure S2**). Analysis of the mutational events at position NM_000546_c.673-2 in various types of cancer showed that the transition A>G is always the major mutational event. In contrast, the three potential substitutions (the transition NM_000546_c.673-2A>T, and the two transversions NM_000546_c.673-2A>G and NM_000546_c.673-2A>C) were equally frequent in CLL (**Figure 4C**). It is therefore unlikely that they arise from a specific mutational process; rather, these results more likely suggest specific selection for these three splice variants. RNA-based studies on tumor samples or cell lines bearing these splice variants have shown that a cryptic acceptor splice site located 49 base pairs upstream of the canonical splice site is preferentially used (**Figure 5**) [27, 28]. In normal cells bearing no *TP53* mutation, this cryptic site is used upon specific stress to generate an alternative transcript expressing p53psi, a truncated isoform that localizes in the mitochondria and displays pro-proliferative activities despite being unable to bind to DNA and transactivate canonical *TP53* target genes (**Figure 5**) [29]. Although the expression of this alternative splice variant is inducible upon specific signals, the splice mutation leads to constitutive expression with a potential oncogenic effect. This particular activity of p53psi has also been observed for two other *TP53* variants, i.e., NP_000537_p.Arg196Ter and NP_000537_p.Arg213Ter, the most frequent nonsense variants observed in CLL. That observation suggests that variants truncated in this region are associated with this gain of function [30]. We note that the putative protein NP_000537_p.Arg209LysfsTer6 expressed by hotspot variant NM_000546_c.626_627del may display the same property (**Figure 5**). Therefore, it is tempting to explain the second and third CLL hotspot mutations described in this section by the specific selection of truncated *TP53* variants with a gain of function specific to CLL.

CLL patients harbor multiple subclones with different *TP53* mutations

The presence of multiple *TP53* mutations in tumors is, generally speaking, uncommon. However, an analysis of the most recent issue of UMD_TP53 showed that multiple mutations were far more frequent in CLL, with a wide range of mutations per patient (**Figure 6 and Supplementary figure S9A-D**). Previous studies using FASAY, a *TP53*, specific functional assay, had already identified patients carrying different *TP53* variants and the use of NGS has further expanded this observation. In the NGS subset of the FILO cohort, 160 patients (31%) were shown to express multiple *TP53* variants with 81, 41 and 38 patients showing two, three or more than three variants per tumor (range 2 to 14) (**Figure 6B**). The same trend was observed in independent reports included in UMD_CLL with *TP53* mutations ranging from 2 to 37 per patient (**Supplementary figure S9B-D**). It is generally assumed that this high burden of *TP53* mutations is associated with the important clonal heterogeneity of CLL. However, that assumption has never been fully investigated.

For patients from the FILO cohort analyzed via NGS and expressing multiple *TP53* variants, the cumulated VAF of those variants never exceeded 100%, suggesting that most of them were on different alleles or in independent subclones (**Figure 6C**). This was confirmed by the analysis of individual DNA sequencing reads harboring mutations within the same exon and no more than 50 nucleotides apart, as all *TP53* variants were always in a trans configuration confirming their status as driver mutations (passenger mutations would have been distributed randomly in cis and trans configurations) (**Figure 6E and Supplementary Figure S10A to S10P**). Patient AVC_62, with del(17p) and 10 different *TP53* variants was particularly informative as several of the mutations were in close proximity with no cis configuration (**Figure 6E**). Five CLL samples from another institution (Karolinska Hospital) were also analyzed using a third-generation, single-molecule real-time (SMRT) sequencing platform (RS II instrument, PacBio, Menlo Park, California) offering long read lengths able to span the most-frequently mutated region of the *TP53* gene^[31] These samples, which had been previously tested by NGS and shown to harbor multiple *TP53* variants, also showed only trans configuration (**Figure 6F and supplementary Figure S11A to S11E**). Tumors carrying two *TP53* variants may result from biallelic mutations. However, because we observed a similar frequency of these multimutated tumors for cases with or without del(17p) (**Supplementary Fig. S12**), it appears more likely that such mutations are associated with multiple subclones. Although non-functional *TP53* variants are associated with CLL, it is well established that there is also an important heterogeneity among the various *TP53* variants with either a simple loss of function, a dominant negative activity or, for some variants, a gain of function that can vary among cancer types. Multiple classical hotspot variants at codons 175, 248 or 273 can be observed in the same patient, indicating that these alterations are unrelated and ruling out the possibility of any associated bystander effect that could drive the selection/ expansion of weaker variants (**Supplementary Figure S13**). The frequency of these classical hotspot variants as single alterations ranged from 50% to 70%, which accords with the frequency of tumors expressing a unique *TP53* variant (**Figure 7B-C**). Furthermore, their VAF distribution ranges widely from 1% to values greater than 95% (**Figure 7E**). In contrast, hotspot splice variants at position 673-2A are found predominantly in *TP53* polymutated patients with VAFs never exceeding 50% (**Figure 7B to F**). This observation holds true for FILO, UMD_CLL, and data from Catherwood *et al.*

Discussion

TP53 mutation in CLL is a paradigm for translational research. First, it is among the few cancers where *TP53* alterations have been incorporated into routine clinical diagnostics to improve patient stratification and, second, the landscape of *TP53* mutations in CLL finds no equivalent in other cancer types and inspires research to improve our understanding of this still incurable disease [8, 32–34]. We present here an analysis of the largest set of *TP53* mutations in CLL patients to date with 1,056 *TP53* variants from 683 patients recruited within the centers of the French Innovative Leukemia Organization-CLL (FILO), and 5,173 variants from UMD_CLL, a carefully curated database of *TP53* mutations.

As compared to all other cancer types, our analysis highlights the highly specific landscape of *TP53* alterations in CLL, one comprising three distinct features: i) a high prevalence of *TP53* mutated minor clones (VAF below 5%); ii) an important intratumoral heterogeneity with multiple subclones expressing different *TP53* variants; and iii) multiple CLL-specific mutation hotspots.

Yeast functional assays and thereafter NGS have shown that CLL patients often harbor minor clones expressing pathogenic *TP53* variants (VAF 1%–5%) that progress during the course of disease [10, 21, 35, 36]. There are still some controversial issues regarding the limit of detection (LOD) used for reporting *TP53* variants. The latest recommendations from the *TP53* Network of the European Research Initiative on Chronic Lymphocytic Leukemia (ERIC) advocate for the use of 10% VAF [8]. However, multiple studies have reported variants at lower VAFs [10, 13, 37]. In the present study, 14% of the patients analyzed via NGS harbored a single *TP53* variant with a VAF between 1% and 5% that would have been missed by conventional Sanger sequencing (LOD = 10%–15%). In the era of targeted therapies, chemoimmunotherapy is not recommended for CLL cases with *TP53* mutation or 17p deletion [38]. Thus, the determination of a cut-off able to identify patients with mutated or unmutated *TP53* has become important for therapeutic choices. In their recent laboratory practice recommendations, the GBMHM (*Groupe des biologistes moléculaires des hémopathies malignes*) established rules and quality control standards for the validation of a clinically-applicable cut-off value (between 1% and 2%) [18]. The present work shows that this value did not lead to the inclusion of spurious variants with no impact on *TP53* activity, as variants expressed by the above-mentioned minor clones are similar to those found in larger clones with analogous hotspot variant distribution. Moreover, considering the deleterious effect of chemoimmunotherapy in *TP53*-mutated cases and the availability of Btk inhibitors, it seems adequate to consider patients harboring small *TP53*-mutated variants as candidates for targeted therapies [16].

In the setting of our study, most of the *TP53* mutations targeted the DNA-binding domain of the protein and led to *TP53* loss of function in a fashion similar to other cancer types. However, we did uncover some CLL-specific particularities observed both in the FILO cohort and the CLL data in the UMD_database. For example, a high, 20% frequency of AT>GC transition in *TP53* was observed in both datasets. That specific transition is seen at a frequency of only 5% to 10% in solid tumors. Similar observations of frequencies of AT>GC transition between 20% and 25% distributed over the entire genome have been made using whole genome sequencing in four untreated CLL patients as well as in 11 patients with monoclonal B-cell lymphocytosis and five patients with ultra-stable CLL (>10 years without progression from initial diagnosis) [39, 40]. The etiology of this particular signature is currently unknown but its detection in untreated patients indicates that it may be a specific feature of CLL.

In this study, using specific *TP53* SNPs, we identified eight patients without del(17p) but with copy-neutral (CN) LOH (also known as uniparental disomy, UDP). However, our numbers likely underestimate the frequency of CN-LOH in CLL. Because of the presence of multiple clones with different *TP53* variants at low frequency, we cannot exclude CN-LOH in minor clones and only single cell sequencing can provide a more accurate characterization of the landscape of this particular genetic event. The question as to whether clones with *TP53* mutations and CN-LOH have the same clinical value as those with a single mutation with or without del(17p) remains unanswered. Also of importance is that *TP53* variants associated with CN-LOH can easily be mistakenly considered as clones with a single *TP53* variant without LOH. This may blur several types of analyses such as the definition of dominant-negative activity of mutant *TP53* toward the wild-type protein.

Three CLL-specific variants were highlighted by the present analysis. First, we discovered that variants in the splice acceptor signal position NM_000546_c.673-2 are specifically enriched in CLL compared to all other cancer types and found predominantly in patients carrying multiple *TP53* variants, a specific feature observed both in the FILO cohort and in UMD_CLL. Although alterations in splice donor sites are predicted to be deleterious, RNA sequencing data analysis from different types of tumors or cell lines have shown that these particular *TP53* mutations lead to the use of an intronic cryptic splice site, with a partial intron retention and the synthesis of a specific protein isoform, p53psi^[27, 29]. Upon specific stress, p53psi, lacking a nuclear localization signal, is translocated to the mitochondria and interacts with cyclophilin-D, which leads to an increase in mitochondrial pore permeability and reactive oxygen production^[29]. Whether or not TP53psi is expressed in CLL patients and leads to a specific phenotype is currently unknown. However, the high specificity of this variant in CLL indicates that it may merit investigation. Aberrant splicing is a common feature in CLL and recurrent mutations in SF3B1 can be found in 5% of patients at presentation and climb to 20% during disease progression. These mutations are predominantly subclonal events and associated with more aggressive disease and shorter survival. All these features are similar to those observed for *TP53* but alterations in these two genes have been shown to be exclusive. Deep analysis of splicing events associated with SF3B1 mutations shows that they induce the use of cryptic 3'AG signal splices leading to aberrant splicing and partial retention of the 3' intronic region, which is similar to the outcome of the *TP53* splice hotspot variant at position c.673-2A^[41]. *TP53* was not shown to be among the aberrantly spliced mRNA in SF3B1 mutated patients, but it could be mirrored by these specific hotspot mutations.

TP53 frameshift mutations are, generally speaking, rare. With a frequency ranging between 1% and 5% of all frameshift mutations, variant NM_000546_c.626_627del (NP_000537_p.Arg209LysfsTer6) is nonetheless one of the most frequent of them. Here, in the setting of CLL, we found that the frequency of this variant reached 16% (both in the FILO and the UMD datasets) and that it was the main hotspot for frameshift mutation. The frequency of NM_000546_c.626_627del in CLL had already been brought to light in two independent studies^[25, 42]. Our results thus strengthen that observation and clearly define this variant as a CLL-specific hotspot. The coding region residues 625–630 contain an inverted repeat sequence of four nucleotides separated by a four-base spacer, a structure known to cause insertions and/or deletions during replication^[43]. Whether or not this particular event is linked to a specific genetic defect in CLL remains to be determined. It should be noted that the putative protein expressed by this frameshift variant is highly similar to p53psi or variant NP_000537_p.Arg213Ter, suggesting that all these variants resulting

from different mutational events could be selected for a yet to be discovered gain of function in CLL. The third CLL-specific variant is located at codon 234 with a high prevalence of tyrosine-to-cysteine substitution. This variant, found at very low frequency in all tumor types, has been shown to be specifically associated to CLB treatment [17]. It remains to be determined whether this variant results directly from a mutational event provoked by CLB or is specifically selected during the treatment.

The present study also shed light on the important clonal heterogeneity of CLL, with 113 cases (33%) presenting multiple pathogenic *TP53* variants (up to 11 in a single patient). In a previous study, we showed that tumors with multiple *TP53* mutations were more frequent in lymphoma and leukemia than they were in solid tumors [31]. In that work, we also used SMRT sequencing to demonstrate that *TP53* variants were always distributed in different alleles in acute lymphocytic leukemia (ALL) and myelodysplastic syndrome (MDS), confirming the clonal heterogeneity of those tumors. Using the same methodology here, we confirmed that these mutations were also located in different alleles in all of the analyzed CLL samples. This supplementary observation strongly supports the multiclonality of these tumors.

The intratumoral genetic heterogeneity in CLL raises the question of the impact of treatment on the selection or acquisition of *TP53* mutations. Longitudinal studies have described the acquisition of *TP53* abnormalities and complex karyotypes in treated relapsed or refractory patients [36]. CLL patients who undergo therapy will ultimately relapse, resulting in the need for multiple lines of therapy with various combinations of drugs. We observed a correlation between the number of *TP53* variants per patient and previous treatments. Any line of treatment significantly increased the number of mutations per patient but some treatment types did so more than others. Any regimen including the continuous administration of the alkylating agent CLB, whether alone or with other subsequent treatments, dramatically increased the number of mutations per patient.

The analysis of *TP53* gene alteration in CLL acts as a magnifying glass on two important features: first, heterogeneity in the setting, with a high number of minor clones; and second, *TP53* addiction with multiple variants selected during disease evolution. Recent studies using single-cell analyses of CLL tumors have identified a high level of genetic and epigenetic heterogeneity beyond *TP53* alteration. Those observations confirm that CLL is not a stable entity, but rather a dynamic disease characterized by a heterogeneous subclonal architecture with a complex course over time shaped by endogenous and exogenous selection pressures [2, 44].

In conclusion, our data emphasize the importance of an adequate limit of detection when using NGS for patient stratification. They also highlight the important effects of treatment on clonal heterogeneity and the specific deleterious impact of continuous chlorambucil on both the type and the number of mutations (Box 1). Considering our results, we feel it is necessary to recommend an assessment of the clonal architecture of *TP53* mutations at each line of treatment, in order to limit the use of therapies promoting clonal evolution. In the future, the accumulation of data following long-term targeted therapies is warranted to optimize treatment sequences.

Material and methods

Patient cohort

For the FILO cohort, we collected the retrospective data of *TP53* mutated patients from centers affiliated with the French Innovative Leukemia Organization-CLL (FILO) in GBMHM (French Molecular Biology Group in Hematology) laboratories. All centers contributing to the present work had GBMHM or ERIC (European Research Initiative on CLL) quality control certification [18]. Compared to the NGS cohort used for the previous analysis of codon 234 (336 patients, 568 *TP53* variants), the present cohort included 175 new *TP53* mutated patients, enabling the identification of a total of 860 *TP53* variants. We also included 172 patients (196 *TP53* variants) analyzed via Sanger sequencing (exons 4 to 8). For NGS analysis, either Illumina or Ion Torrent technologies were used with a VAF of 1% as recommended by the GBMHM.

Polymorphisms were carefully excluded using the new *TP53* SNP data included in the most recent version of the UMD_*TP53* database [20, 45].

The UMD_*TP53* database and *in silico* analysis of *TP53* variants

The latest issue of the database (2022_R1) includes 207,168 *TP53* variants and data from multiple large-scale tumor analyses such as TCGA GENIE and MSKSCC. This version of UMD_*TP53* now includes the OncoTree cancer classification developed by Kundra *et al.* [46]. It includes nearly 900 tumor types classified into 32 organ sites. Having both tumor types and OncoTree classification in UMD_*TP53* increased the specificity of the various analyses. For an accurate analysis, data from all CLL studies were checked manually to remove duplicate entries resulting from the use of the same patients in independent studies. Thus, in all, 4,084 *TP53* somatic variants were extracted and included in UMD_*CLL*. No specific filtering was applied to remove any variant during this process. FILO data were not included in UMD_*CLL* for the analyses performed in the present study. For nine studies using NGS (excluding the present analysis), VAF for *TP53* variants were included in the database. The classification of *TP53* variant pathogenicity, in accordance with ACMG criteria and based on population data and *TP53*-specific functional information, has been previously described [19].

SMRT sequencing of *TP53* amplicons

A 2.8 Kb amplicon encompassing exons 4 to 8 was used for SMRT analysis. This region includes the majority of the mutations detected in patients and several common *TP53* SNPs useful for phasing the various mutations on the two alleles.

The *TP53* amplicons underwent DNA damage repair and end repair before ligation of hairpin adaptors to generate SMRTbell 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 and sequenced on the PacBio RS II instrument using C4 chemistry, P6 polymerase and a 240-minute movie time. The detection and phasing analyses of *TP53* have been previously described [31].

Statistical analysis

Statistical analyses were performed with GraphPad Prism 9.0 (GraphPad Software, Inc., San Diego, CA).

References

1. Lazarian G, Guièze R, Wu CJ Clinical Implications of Novel Genomic Discoveries in Chronic Lymphocytic Leukemia. *J Clin Oncol.* 2017;35:984-993
2. Bosch F, Dalla-Favera R Chronic lymphocytic leukaemia: from genetics to treatment. *Nat Rev Clin Oncol.* 2019;16:684-701
3. Fenaux P, Preudhomme C, Lai JL, Quiquandon I, Jonveaux P, Vanrumbeke M, Sartiaux C, Morel P, Loucheux-Lefebvre MH, Bauters F Mutations of the p53 gene in B-cell chronic lymphocytic leukemia: a report on 39 cases with cytogenetic analysis. *Leukemia.* 1992;6:246-250
4. Gaidano G, Ballerini P, Gong JZ, Inghirami G, Neri A, Newcomb EW, Magrath IT, Knowles DM, Dalla-Favera R p53 mutations in human lymphoid malignancies: association with Burkitt lymphoma and chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A.* 1991;88:5413-5417
5. Döhner H, Stilgenbauer S, Benner A, Leupolt E, Kröber A, Bullinger L, Döhner K, Bentz M, Lichter P Genomic aberrations and survival in chronic lymphocytic leukemia. *N Engl J Med.* 2000;343:1910-1916
6. Stilgenbauer S, Schnaiter A, Paschka P, Zenz T, Rossi M, Döhner K, Bühler A, Böttcher S, Ritgen M, Kneba M, Winkler D, Tausch E, Hoth P, Edelmann J, Mertens D, Bullinger L, Bergmann M, Kless S, Mack S, Jäger U, Patten N, Wu L, Wenger MK, Fingerle-Rowson G, Lichter P, Cazzola M, Wendtner CM, Fink AM, Fischer K, Busch R, Hallek M, Döhner H Gene mutations and treatment outcome in chronic lymphocytic leukemia: results from the CLL8 trial. *Blood.* 2014;123:3247-3254
7. Eichhorst B, Fink AM, Bahlo J, Busch R, Kovacs G, Maurer C, Lange E, Köppler H, Kiehl M, Sökler M, Schlag R, Vehling-Kaiser U, Köchling G, Plöger C, Gregor M, Plesner T, Trneny M, Fischer K, Döhner H, Kneba M, Wendtner CM, Klapper W, Kreuzer KA, Stilgenbauer S, Böttcher S, Hallek M, international GOI, German CLL Study Group GCLLSG First-line chemoimmunotherapy with bendamustine and rituximab versus fludarabine, cyclophosphamide, and rituximab in patients with advanced chronic lymphocytic leukaemia (CLL10): an international, open-label, randomised, phase 3, non-inferiority trial. *Lancet Oncol.* 2016;17:928-942
8. Malcikova J, Tausch E, Rossi D, Sutton LA, Soussi T, Zenz T, Kater AP, Niemann CU, Gonzalez D, Davi F, Gonzalez Diaz M, Moreno C, Gaidano G, Stamatopoulos K, Rosenquist R, Stilgenbauer S, Ghia P, Pospisilova S, European RIOCLLERICN ERIC recommendations for TP53 mutation analysis in chronic lymphocytic leukemia-update on methodological approaches and results interpretation. *Leukemia.* 2018;32:1070-1080
9. Guièze R, Wu CJ Genomic and epigenomic heterogeneity in chronic lymphocytic leukemia. *Blood.* 2015;126:445-453
10. Rossi D, Khiabani H, Spina V, Ciardullo C, Brusca A, Famà R, Rasi S, Monti S, Deambrogi C, De Paoli L, Wang J, Gattei V, Guarini A, Foà R, Rabadan R, Gaidano G Clinical impact of small TP53 mutated subclones in chronic lymphocytic leukemia. *Blood.* 2014;123:2139-2147
11. Blakemore SJ, Clifford R, Parker H, Antoniou P, Stec-Dziedzic E, Larrayoz M, Davis Z, Kadalyayil L, Colins A, Robbe P, Vavoulis D, Forster J, Carr L, Morilla R, Else M, Bryant D, McCarthy H, Walewska RJ, Steele AJ, Chan J, Speight G, Stankovic T, Cragg MS, Catovsky D, Oscier DG, Rose-Zerilli MJ, Schuh A, Strefford JC Clinical significance of

TP53, BIRC3, ATM and MAPK-ERK genes in chronic lymphocytic leukaemia: data from the randomised UK LRF CLL4 trial. *Leukemia*. 2020;34:1760-1774

12. Malcikova J, Pavlova S, Kunt Vonkova B, Radova L, Plevova K, Kotaskova J, Pal K, Dvorackova B, Zenatova M, Hynst J, Ondrouskova E, Panovska A, Brychtova Y, Zavacka K, Tichy B, Tom N, Mayer J, Doubek M, Pospisilova S Low-burden TP53 mutations in CLL: Clinical impact and clonal evolution within the context of different treatment options. *Blood*. 2021

13. Bomben R, Rossi FM, Vit F, Bittolo T, D'Agaro T, Zucchetto A, Tissino E, Pozzo F, Vendramini E, Degan M, Zaina E, Cattarossi I, Varaschin P, Nanni P, Berton M, Braida A, Polesel J, Cohen JA, Santinelli E, Biagi A, Gentile M, Morabito F, Fronza G, Pozzato G, D'Arena G, Olivieri J, Bulian P, Pepper C, Hockaday A, Schuh A, Hillmen P, Rossi D, Chiarenza A, Zaja F, Di Raimondo F, Del Poeta G, Gattei V TP53 Mutations with Low Variant Allele Frequency Predict Short Survival in Chronic Lymphocytic Leukemia. *Clin Cancer Res*. 2021;27:5566-5575

14. Nadeu F, Clot G, Delgado J, Martín-García D, Baumann T, Salaverria I, Beà S, Pinyol M, Jares P, Navarro A, Suárez-Cisneros H, Aymerich M, Rozman M, Villamor N, Colomer D, González M, Alcoceba M, Terol MJ, Navarro B, Colado E, Payer ÁR, Puente XS, López-Otín C, López-Guillermo A, Enjuanes A, Campo E Clinical impact of the subclonal architecture and mutational complexity in chronic lymphocytic leukemia. *Leukemia*. 2018;32:645-653

15. Brieghel C, Kinalis S, Yde CW, Schmidt AY, Jønson L, Andersen MA, da Cunha-Bang C, Pedersen LB, Geisler CH, Nielsen FC, Niemann CU Deep targeted sequencing of TP53 in chronic lymphocytic leukemia: clinical impact at diagnosis and at time of treatment. *Haematologica*. 2019;104:789-796

16. Lazarian G, Cymbalista F, Baran-Marszak F Impact of Low-Burden TP53 Mutations in the Management of CLL. *Front Oncol*. 2022;12:841630

17. Lazarian G, Theves F, Hormi M, Letestu R, Eclache V, Bidet A, Cornillet-Lefebvre P, Davi F, Delabesse E, Estienne MH, Etancelin P, Kosmider O, Laibe S, Lode L, Muller M, Nadal N, Naguib D, Pastoret C, Poulain S, Sujobert P, Veronese L, Imache S, Lefebvre V, Cymbalista F, Baran-Marszak F, Soussi T, French Innovative Leukemia Organisation FILO TP53 mutations at codon 234 are associated with chlorambucil treatment in chronic lymphocytic leukemia. *Am J Hematol*. 2022;97:E159-E162

18. Sujobert P, Le Bris Y, de Leval L, Gros A, Merlio JP, Pastoret C, Huet S, Sarkozy C, Davi F, Callanan M, Thieblemont C, Sibon D, Asnafi V, Preudhomme C, Gaulard P, Jardin F, Salles G, Macintyre E The Need for a Consensus Next-generation Sequencing Panel for Mature Lymphoid Malignancies. *Hemasphere*. 2019;3:e169

19. Soussi T, Leroy B, Devir M, Rosenberg S High prevalence of cancer-associated TP53 variants in the gnomAD database: A word of caution concerning the use of variant filtering. *Hum Mutat*. 2019;40:516-524

20. Soussi T Benign SNPs in the Coding Region of TP53: Finding the Needles in a Haystack of Pathogenic Variants. *Cancer Res*. 2022;82:3420-3431

21. Malcikova J, Stano-Kozubik K, Tichy B, Kantorova B, Pavlova S, Tom N, Radova L, Smardova J, Pardy F, Doubek M, Brychtova Y, Mraz M, Plevova K, Diviskova E, Oltova A, Mayer J, Pospisilova S, Trbusek M Detailed analysis of therapy-driven clonal evolution of TP53 mutations in chronic lymphocytic leukemia. *Leukemia*. 2015;29:877-885

22. Landau DA, Carter SL, Stojanov P, McKenna A, Stevenson K, Lawrence MS, Sougnez C, Stewart C, Sivachenko A, Wang L, Wan Y, Zhang W, Shukla SA, Vartanov A,

Fernandes SM, Saksena G, Cibulskis K, Tesar B, Gabriel S, Hacohen N, Meyerson M, Lander ES, Neuberg D, Brown JR, Getz G, Wu CJ Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell*. 2013;152:714-726

23. Metzeler KH, Herold T, Rothenberg-Thurley M, Amler S, Sauerland MC, Görlich D, Schneider S, Konstandin NP, Dufour A, Bräundl K, Ksienzyk B, Zellmeier E, Hartmann L, Greif PA, Fiegl M, Subklewe M, Bohlander SK, Krug U, Faldum A, Berdel WE, Wörmann B, Büchner T, Hiddemann W, Braess J, Spiekermann K, AMLCG SG Spectrum and prognostic relevance of driver gene mutations in acute myeloid leukemia. *Blood*. 2016;128:686-698

24. Loeb LA, Harris CC Advances in chemical carcinogenesis: a historical review and prospective. *Cancer Res*. 2008;68:6863-6872

25. Zenz T, Vollmer D, Trbusek M, Smardova J, Benner A, Soussi T, Helfrich H, Heuberger M, Hoth P, Fuge M, Denzel T, Habe S, Malcikova J, Kuglik P, Truong S, Patten N, Wu L, Oscier D, Ibbotson R, Gardiner A, Tracy I, Lin K, Pettitt A, Pospisilova S, Mayer J, Hallek M, Dohner H, Stilgenbauer S TP53 mutation profile in chronic lymphocytic leukemia: evidence for a disease specific profile from a comprehensive analysis of 268 mutations. *Leukemia*. 2010;24:2072-2079

26. Catherwood MA, Wren D, Chiecchio L, Cavalieri D, Donaldson D, Lawless S, ElHassadi E, Hayat A, Cahill MR, O'Shea D, Sargent J, Stewart P, Maurya M, Quinn J, Murphy P, de Castro DG, Mills K, Cross NCP, Forconi F, Iyengar S, Schuh A, Thornton P TP53 Mutations Identified Using NGS Comprise the Overwhelming Majority of TP53 Disruptions in CLL: Results From a Multicentre Study. *Front Oncol*. 2022;12:909615

27. Smeby J, Sveen A, Eilertsen IA, Danielsen SA, Hoff AM, Eide PW, Johannessen B, Hektoen M, Skotheim RI, Guren MG, Nesbakken A, Lothe RA Transcriptional and functional consequences of TP53 splice mutations in colorectal cancer. *Oncogenesis*. 2019;8:35

28. Palmisano A, Vural S, Zhao Y, Sonkin D MutSpliceDB: A database of splice sites variants with RNA-seq based evidence on effects on splicing. *Hum Mutat*. 2021;42:342-345

29. Senturk S, Yao Z, Camiolo M, Stiles B, Rathod T, Walsh AM, Nemajerova A, Lazzara MJ, Altorki NK, Krainer A, Moll UM, Lowe SW, Cartegni L, Sordella R p53 Ψ is a transcriptionally inactive p53 isoform able to reprogram cells toward a metastatic-like state. *Proc Natl Acad Sci U S A*. 2014;111:E3287-96

30. Shirole NH, Pal D, Kasthuber ER, Senturk S, Boroda J, Pisterzi P, Miller M, Munoz G, Anderluh M, Ladanyi M, Lowe SW, Sordella R TP53 exon-6 truncating mutations produce separation of function isoforms with pro-tumorigenic functions. *Elife*. 2017;6

31. Lodé L, Ameur A, Coste T, Ménard A, Richebourg S, Gaillard JB, Le Bris Y, Béné MC, Lavabre-Bertrand T, Soussi T Single-molecule DNA sequencing of acute myeloid leukemia and myelodysplastic syndromes with multiple TP53 alterations. *Haematologica*. 2018;103:e13-e16

32. Albi E, Capasso A, Schiattone L, Ghia P, Scarfò L Are we finally getting personal? Moving towards a personalized approach in chronic lymphocytic leukemia. *Semin Cancer Biol*. 2022;84:329-338

33. Soussi T, Baliakas P Landscape of TP53 Alterations in Chronic Lymphocytic Leukemia via Data Mining Mutation Databases. *Front Oncol*. 2022;12:808886

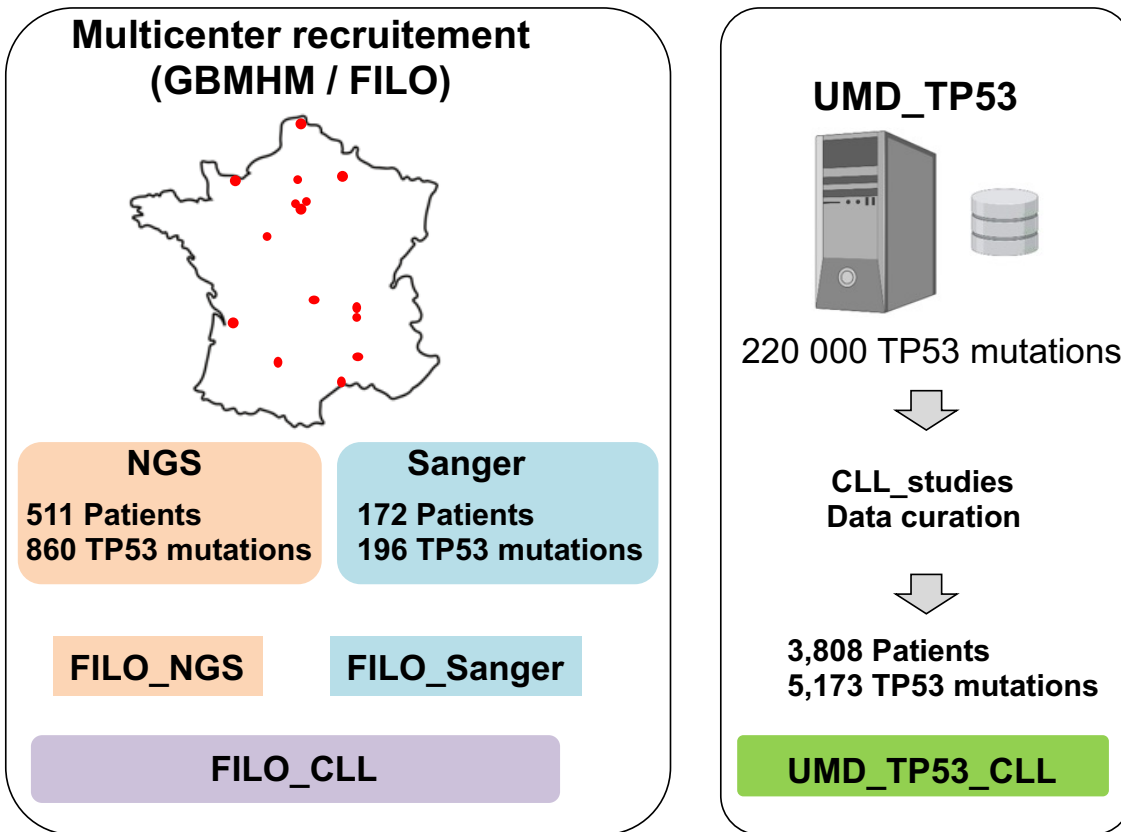
34. Ten Hacken E, Wu CJ Understanding CLL biology through mouse models of human genetics. *Blood*. 2021;138:2621-2631

35. Gruber M, Bozic I, Leshchiner I, Livitz D, Stevenson K, Rassenti L, Rosebrock D, Taylor-Weiner A, Olive O, Goyette R, Fernandes SM, Sun J, Stewart C, Wong A, Cibulskis C, Zhang W, Reiter JG, Gerold JM, Gribben JG, Rai KR, Keating MJ, Brown JR, Neuberg

- D, Kipps TJ, Nowak MA, Getz G, Wu CJ Growth dynamics in naturally progressing chronic lymphocytic leukaemia. *Nature*. 2019;570:474-479
36. Lazarian G, Tausch E, Eclache V, Sebaa A, Bianchi V, Letestu R, Collon JF, Lefebvre V, Gardano L, Varin-Blank N, Soussi T, Stilgenbauer S, Cymbalista F, Baran-Marszak F TP53 mutations are early events in chronic lymphocytic leukemia disease progression and precede evolution to complex karyotypes. *Int J Cancer*. 2016;139:1759-1763
37. Pandzic T, Ladenvall C, Engvall M, Mattsson M, Hermanson M, Cavelier L, Ljungström V, Baliakas P Five Percent Variant Allele Frequency Is a Reliable Reporting Threshold for TP53 Variants Detected by Next Generation Sequencing in Chronic Lymphocytic Leukemia in the Clinical Setting. *Hemasphere*. 2022;6:e761
38. Hallek M Chronic lymphocytic leukemia: 2020 update on diagnosis, risk stratification and treatment. *Am J Hematol*. 2019;94:1266-1287
39. Puente XS, Pinyol M, Quesada V, Conde L, Ordonez GR, Villamor N, Escaramis G, Jares P, Bea S, Gonzalez-Diaz M, Bassaganyas L, Baumann T, Juan M, Lopez-Guerra M, Colomer D, Tubio JM, Lopez C, Navarro A, Tornador C, Aymerich M, Rozman M, Hernandez JM, Puente DA, Freije JM, Velasco G, Gutierrez-Fernandez A, Costa D, Carrio A, Guijarro S, Enjuanes A, Hernandez L, Yague J, Nicolas P, Romeo-Casabona CM, Himmelbauer H, Castillo E, Dohm JC, de Sanjose S, Piris MA, de Alava E, San Miguel J, Royo R, Gelpi JL, Torrents D, Orozco M, Pisano DG, Valencia A, Guigo R, Bayes M, Heath S, Gut M, Klatt P, Marshall J, Raine K, Stebbings LA, Futreal PA, Stratton MR, Campbell PJ, Gut I, Lopez-Guillermo A, Estivill X, Montserrat E, Lopez-Otin C, Campo E Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature*. 2011;475:101-105
40. Agathangelidis A, Ljungström V, Scarfò L, Fazi C, Gounari M, Pandzic T, Sutton LA, Stamatopoulos K, Tonon G, Rosenquist R, Ghia P Highly similar genomic landscapes in monoclonal B-cell lymphocytosis and ultra-stable chronic lymphocytic leukemia with low frequency of driver mutations. *Haematologica*. 2018;103:865-873
41. Darman RB, Seiler M, Agrawal AA, Lim KH, Peng S, Aird D, Bailey SL, Bhavsar EB, Chan B, Colla S, Corson L, Feala J, Fekkes P, Ichikawa K, Keaney GF, Lee L, Kumar P, Kunii K, MacKenzie C, Matijevic M, Mizui Y, Myint K, Park ES, Puyang X, Selvaraj A, Thomas MP, Tsai J, Wang JY, Warmuth M, Yang H, Zhu P, Garcia-Manero G, Furman RR, Yu L, Smith PG, Buonamici S Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3' Splice Site Selection through Use of a Different Branch Point. *Cell Rep*. 2015;13:1033-1045
42. Newcomb EW, el Roubi S, Thomas A A unique spectrum of p53 mutations in B-cell chronic lymphocytic leukemia distinct from that of other lymphoid malignancies. *Mol Carcinog*. 1995;14:227-232
43. Zou X, Morganella S, Glodzik D, Davies H, Li Y, Stratton MR, Nik-Zainal S Short inverted repeats contribute to localized mutability in human somatic cells. *Nucleic Acids Research*. 2017;45:11213-11221
44. Gaiti F, Chaligne R, Gu H, Brand RM, Kothen-Hill S, Schulman RC, Grigorev K, Risso D, Kim KT, Pastore A, Huang KY, Alonso A, Sheridan C, Omans ND, Biederstedt E, Clement K, Wang L, Felsenfeld JA, Bhavsar EB, Aryee MJ, Allan JN, Furman R, Gnirke A, Wu CJ, Meissner A, Landau DA Epigenetic evolution and lineage histories of chronic lymphocytic leukaemia. *Nature*. 2019;569:576-580
45. Doffe F, Carbonnier V, Tissier M, Leroy B, Martins I, Mattsson JSM, Micke P, Pavlova S, Pospisilova S, Smardova J, Joerger AC, Wiman KG, Kroemer G, Soussi T Identification

and functional characterization of new missense SNPs in the coding region of the TP53 gene. *Cell Death Differ.* 2021;28:1477-1492

46. Kundra R, Zhang H, Sheridan R, Sirintrapun SJ, Wang A, Ochoa A, Wilson M, Gross B, Sun Y, Madupuri R, Satravada BA, Reales D, Vakiani E, Al-Ahmadie HA, Dogan A, Arcila M, Zehir A, Maron S, Berger MF, Viaplana C, Janeway K, Ducar M, Sholl L, Dogan S, Bedard P, Surrey LF, Sanchez IH, Syed A, Rema AB, Chakravarty D, Suehnholz S, Nissan M, Iyer GV, Murali R, Bouvier N, Soslow RA, Hyman D, Younes A, Intlekofer A, Harding JJ, Carvajal RD, Sabbatini PJ, Abou-Alfa GK, Morris L, Janjigian YY, Gallagher MM, Soumerai TA, Mellinghoff IK, Hakimi AA, Fury M, Huse JT, Bagrodia A, Hameed M, Thomas S, Gardos S, Cerami E, Mazor T, Kumari P, Raman P, Shivdasani P, MacFarland S, Newman S, Waanders A, Gao J, Solit D, Schultz N OncoTree: A Cancer Classification System for Precision Oncology. *JCO Clin Cancer Inform.* 2021;5:221-230



Landscape of TP53 mutations in CLL

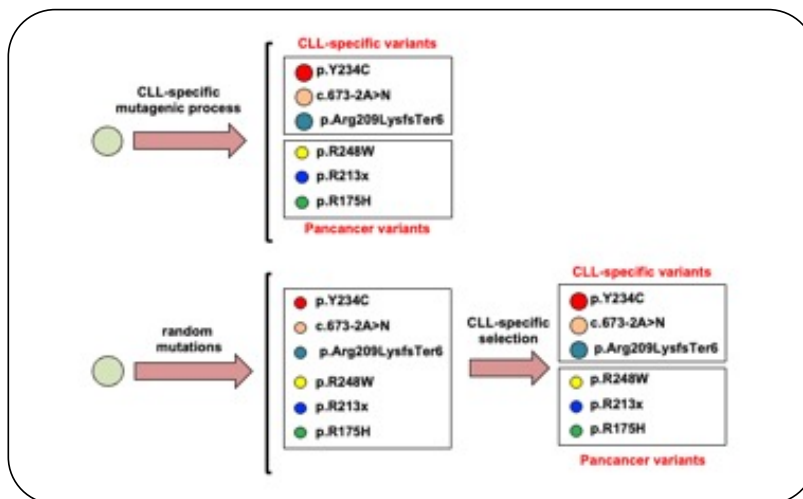
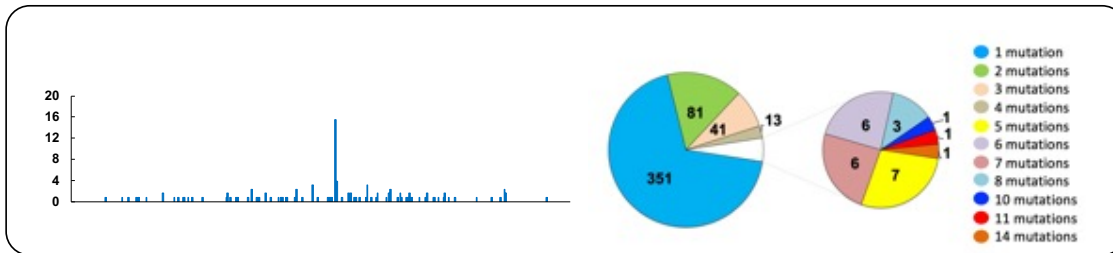


Figure 1: Workflow diagram of the study.

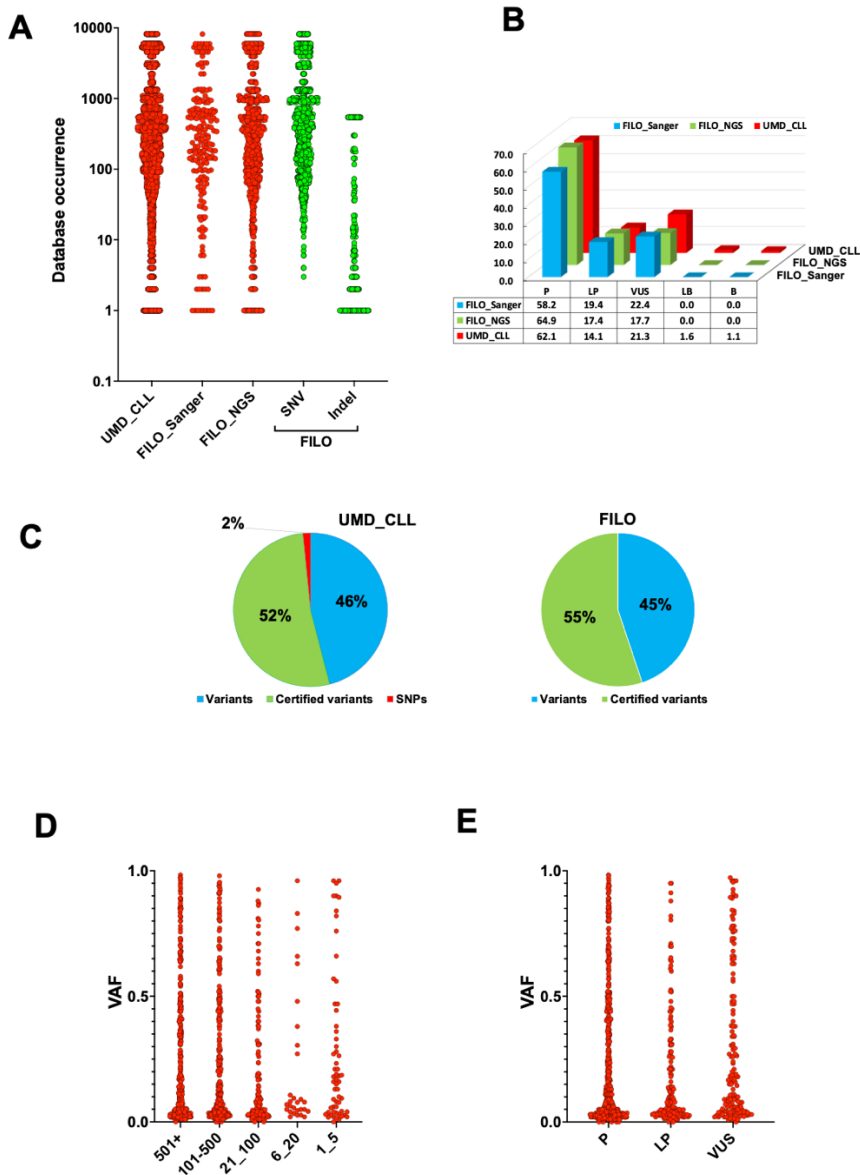
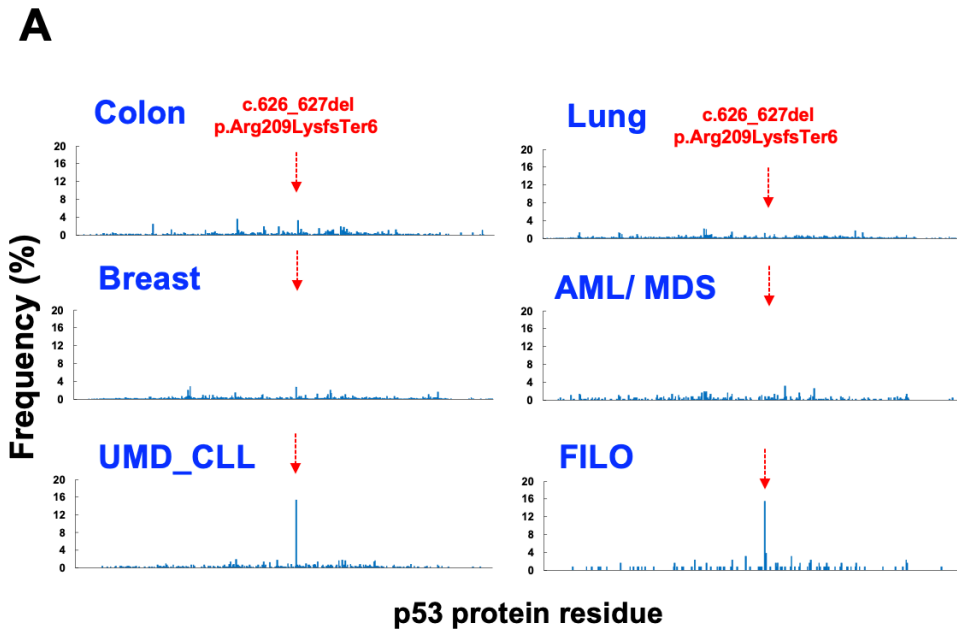
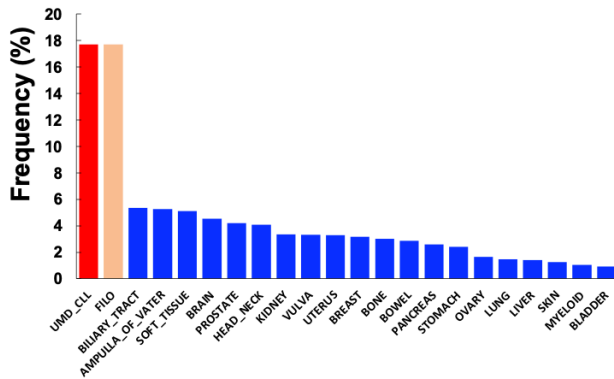


Figure 2: Validation of the FILO cohort.

A: The frequency of *TP53* variants in the FILO cohort is similar to those included in UMD_CELL. The occurrence of each *TP53* variant extracted from UMD_TP53 is shown on the Y-axis of the graph for three datasets: UMD_CELL: all *TP53* variants from CLL patients included in the UMD_TP53 database, excluding FILO data (5173 variants); FILO_NGS / FILO_Sanger: data from the subsets of FILO patients analyzed via NGS or Sanger sequencing respectively. Single nucleotide variants (SNV) are more frequent than frameshift mutations (Indel) **B:** The majority of *TP53* variants in the FILO cohort are either pathogenic (P) or likely pathogenic (LP). *TP53* variants in both cohorts were classified according to the pathogenicity data included in the UMD_TP53 database (B: benign; LB: likely benign; VUS: variant of uncertain significance). **C:** More than 50% of *TP53* found in CLL, in both UMD_CELL and the FILO data set, are certified deleterious *TP53* variants. UMD_CELL includes several benign polymorphisms (SNP) that were removed for all subsequent analyses (Supplementary table S4). **D:** VAF is similar depending on the frequency of the *TP53* variants. The number of reported cases for each *TP53* variant is indicated in the x-axis. The y-axis corresponds to the VAF of each *TP53* variant. **E:** VAF is similar among the 3 classes (P, LP and VUS) of *TP53* variants. The y-axis corresponds to the VAF of each *TP53* variant.



B



C

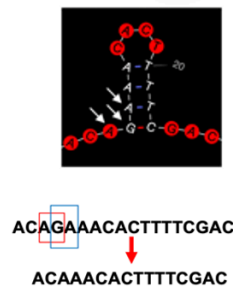


Figure 3: Variant NM_000546_c.626_627del (NP_000537_p.Arg209LysfsTer6) is a hotspot mutation in CLL.

A: Distribution of frameshift mutations at each codon of the *TP53* protein in various types of cancer. Only insertions and deletions are analyzed. The same scale is used for all the analyses to emphasize the CLL hotspot (see **Supplementary Figure S5** for more information). AML/MDS: acute myeloid Leukemia and myelodysplastic syndrome. Lung: NSCLC and SCLC. **B:** Frequency of NM_000546_c.626_627del (NP_000537_p.Arg209LysfsTer6) in various cancer types included in the UMD_TP53. **C:** Potential hairpin structure associated with an inverted repeat in regions 626–628. Two potential mutational events can lead to the same mutation with deletion of either AG or GA depicted by the three arrows. In both the UMD_T53 database and FILO cohort, variant NM_000546_c.626_627del was detected using different methodologies (conventional Sanger sequencing or NGS) precluding any methodological bias.

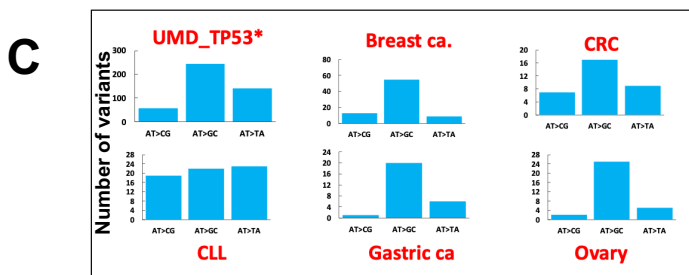
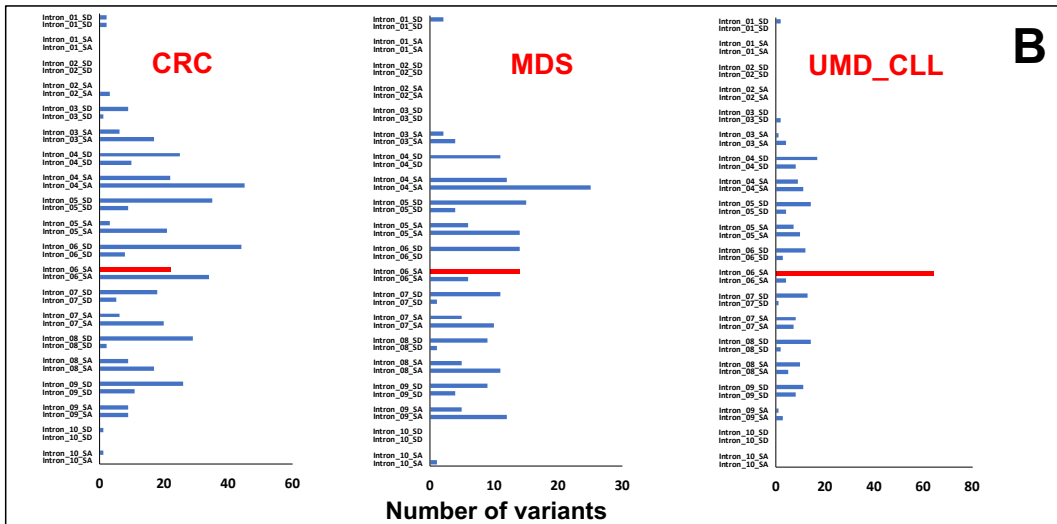
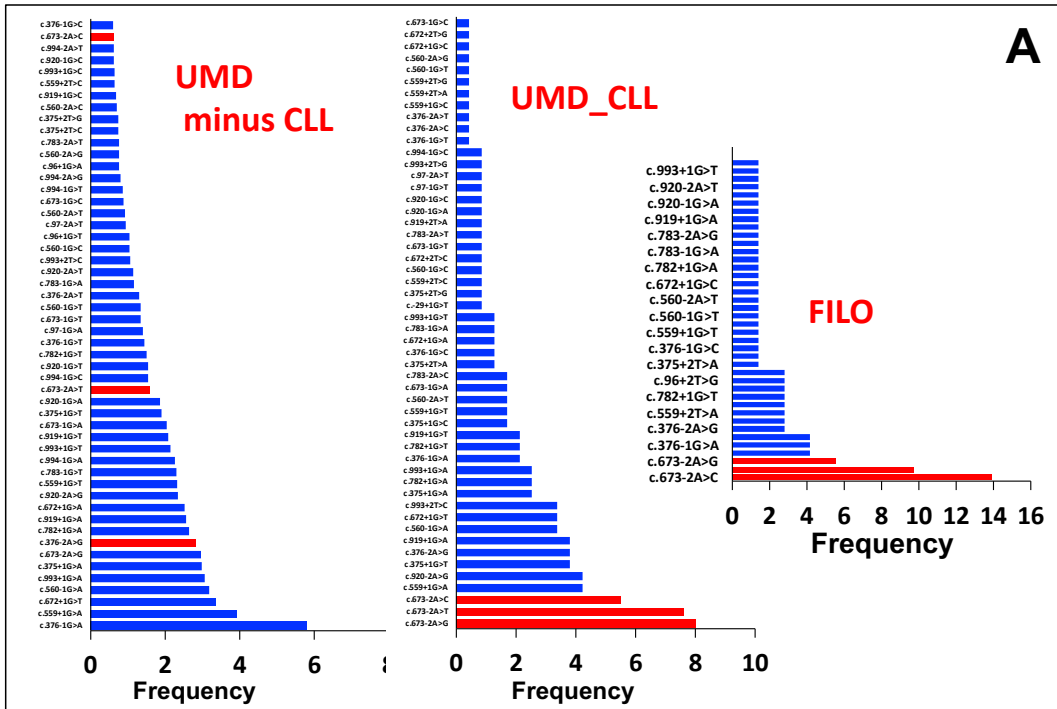


Figure 4: CLL hotspot mutations in intron 6 splice acceptor signal.

A: Frequency of *TP53* mutation in various splice signals included in UMD_TP53 (left panel), UMD_CLL (middle panel) and FILO (right panel). Variants at position NM_000546_c.673-2A are shown in red. **B:** Distribution of mutations in the 20 splice signals of the 10 introns of the *TP53* gene. **C:** Mutational events at position NM_000546_c.673-2 in various types of cancer. UMD_TP53*: whole *TP53* database without CLL. CRC: colorectal carcinoma. MDS: myelodysplastic syndrome.

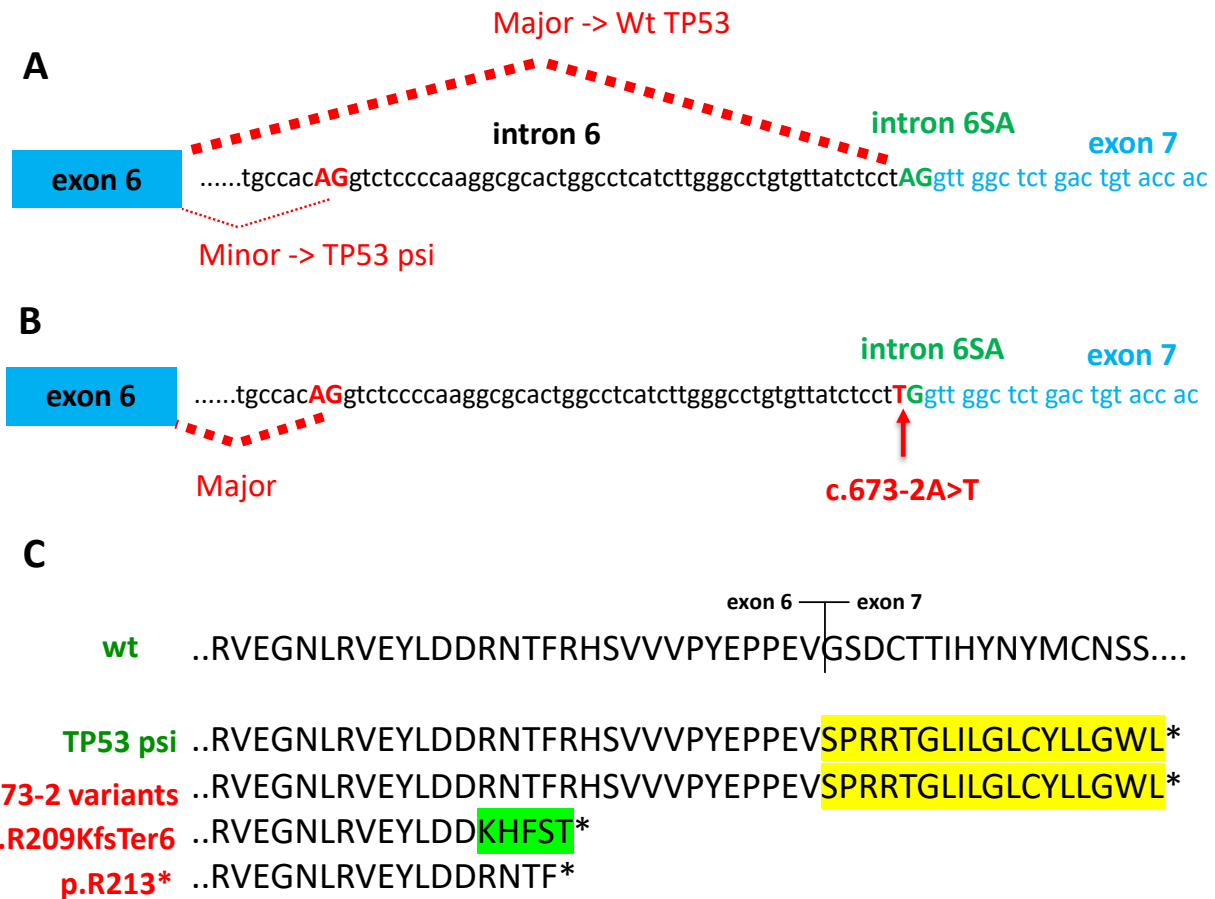


Figure 5: Alternative splicing and mutation consequences in *TP53* intron 6.

A: In unstressed normal cells, full-length wild-type p53 (NP_000537.3, 393 residues) derives from a splice event occurring between exons 6 and 7 (in blue) of the major RNA transcript (NM_000546) using the major splice acceptor site (in green). Upon specific stress, an alternative splice occurs between a cryptic acceptor splice site (in red) localized in the 3' region of intron 6 leading to the synthesis of a shorter *TP53* isoform (TP53psi). **B:** Mutations at position NM_000546_c.673-2 lead to the inactivation of the original acceptor site and the utilization of the cryptic splice acceptor used to generate TP53psi. This consequence has been observed with RNA sequencing analysis in multiple tumors or cell lines bearing variants at position NM_000546_c.672-2A [27, 28]. **C:** Putative *TP53* protein variants resulting from various events leading to a truncation of TP53. P53psi or putative variants resulting from a mutation at position NM_000546_c.673-2 bear a new carboxy terminus rising from the translation of intron 6 and finishing with the stop codon in the beginning of exon 7, which is translated in a different reading frame compared to wt p53 (highlighted in yellow). The putative protein, NP_000537_p.Arg209LysfsTer6, expressed by the hotspot variant NM_000546_c.626_627del ends in exon 6 with 5 extra amino acids (highlighted in green). NM_000546_c.637C>T, a hotspot variant found in every type of cancer, truncates *TP53* at codon 213 (NP_000537_p.Arg213Ter).

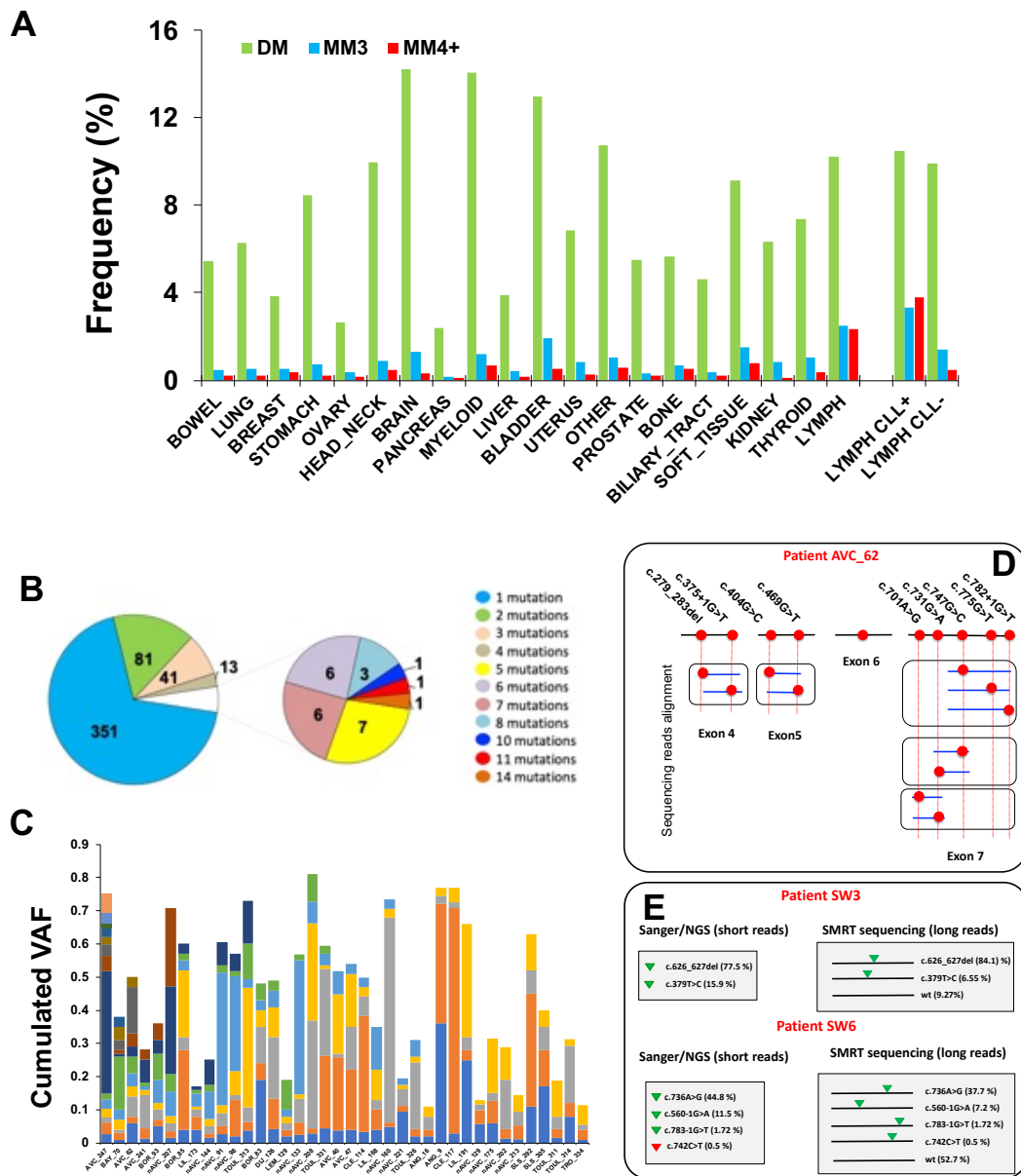


Figure 6: CLL patients are polymutated.

A: *TP53* cancer types classified according to OncoTree were analyzed for tumors carrying two (DM), three (MM3) or more than three (MM4+) *TP53* variants. Lymphoid tumors were split into two subgroups including (CLL+) or excluding (CLL-) CLL patients. **B:** Distribution of the number of mutations in tumors from the NGS subset of the FILO cohort. **C:** Cumulated VAF from polymutated patients from the NGS subset of the FILO cohort. **D:** *TP53* variants are distributed on different chromosomes in the tumor of patient AVC-62. Manual examination of the sequence alignment of NGS data was performed for each exon. For 3 pairs of variants and 1 triplet, *TP53* variants are in a trans configuration. **E:** SMRT sequencing shows that *TP53* mutations are on different alleles for patients SW3 and SW6. Standard NGS analysis is shown on the left. No allelic distribution can be inferred from this type of analysis. SMRT sequencing (right) 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. Green triangle: *TP53* variants identified by both types of analyses. Red triangle: *TP53* variants not detected by long-range sequencing (see also **Supplementary Figure S10a to S10e** for more patients analyzed by SMRT sequencing).

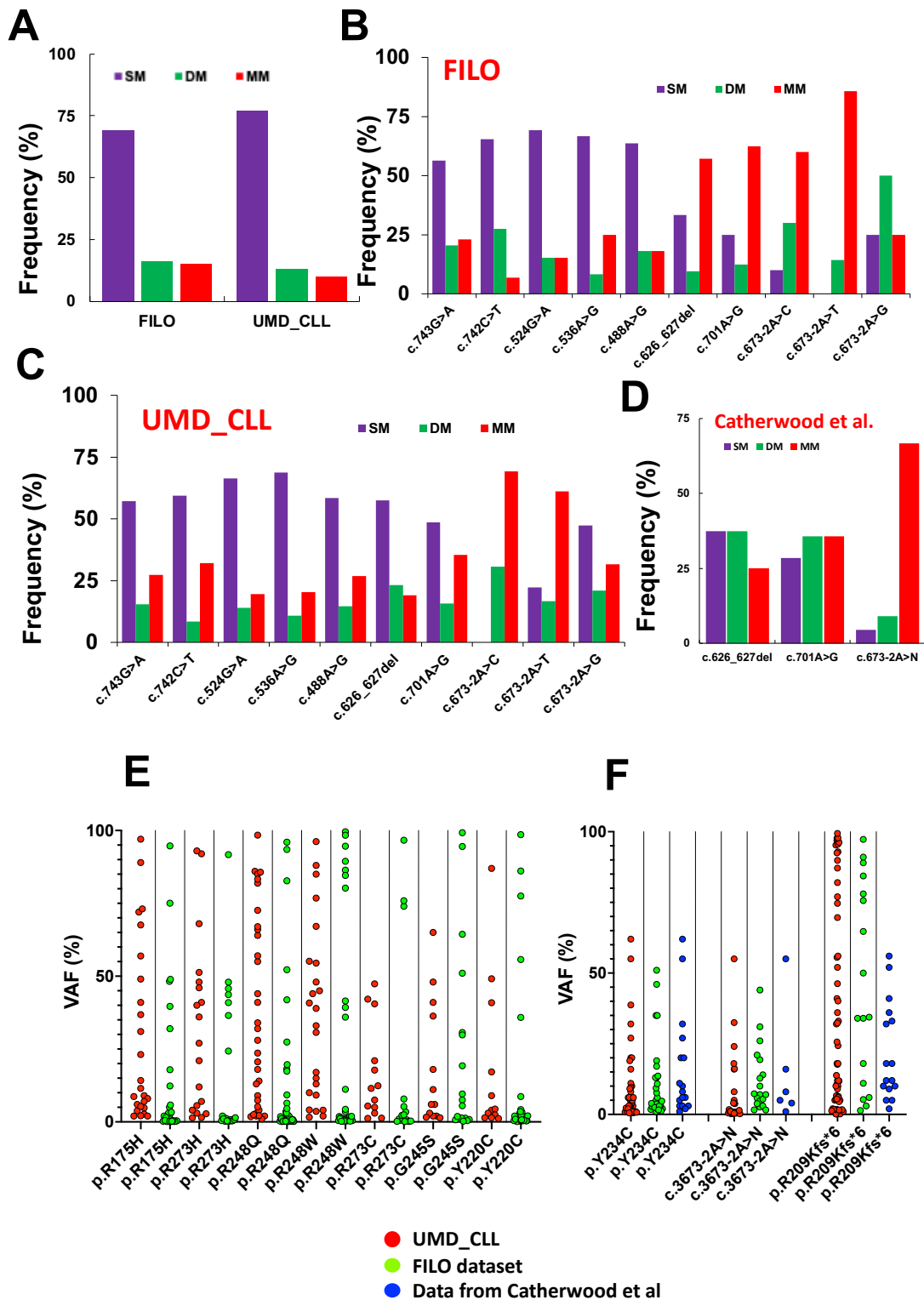


Figure 7: CLL-specific *TP53* variants are observed predominantly in low VAF polymutated patients.

A: Frequency of patients with one (SM), two (DM) or more than two (MM) *TP53* mutations per tumor. **B to D:** Frequency of individual *TP53* variants in tumors bearing one (SM), two (DM) or more than two (MM) *TP53* variants in the FILO cohort (B), UMD_CLL (C) or the dataset of Catherwood *et al.* **E and F:** VAF distribution for classical *TP53* hotspot variants (E) or CLL specific variants (F) in different datasets.

Box 1

TP53 variants in CLL: Unresolved questions and potential research studies

What is the origin of *TP53* mutation intratumoral heterogeneity in CLL?

How do subclones with different *TP53* mutations compete with each other during natural CLL progression or during treatment?

How are CLL- specific *TP53* variants generated and/or selected during CLL progression?

What is the functional and the potential clinical relevance of the various hotspot variants such as NM_000546_c.626_627del or the splice variants in intron 6 detected in CLL?