TP53 mutations with low variant allele frequency predict short survival in chronic lymphocytic leukemia

Article (Accepted Version)


This version is available from Sussex Research Online: http://sro.sussex.ac.uk/id/eprint/102399/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

http://sro.sussex.ac.uk
Small *TP53* Mutated sub-clones predict short survival in Chronic Lymphocytic Leukemia

Riccardo Bomben¹, Francesca Maria Rossi¹, Filippo Vit¹², Tamara Bittolo¹, Tiziana D’Agaro¹, Antonella Zucchetto¹, Erika Tissino¹, Federico Pozzo¹, Elena Vendramini¹, Massimo Degan¹, Eva Zaina¹, Ilaria Cattarossi¹, Paola Varaschin¹, Paola Nanni¹, Michele Berton¹, Alessandra Braida¹, Jerry Polesel³, Jared A. Cohen¹, Enrico Santinelli⁴, Annalisa Biagi¹, Massimo Gentile⁵, Fortunato Morabito⁶, Gilberto Fronza⁸, Gabriele Pozzato⁹, Giovanni D’Arena¹⁰, Jacopo Olivieri¹¹, Pietro Bulian¹, Chris Pepper¹², Anna Hockaday¹³, Anna Schuh¹⁴, Peter Hillmen¹⁶, Davide Rossi¹⁷, Annalisa Chiarenza¹⁹, Francesco Zaja⁹, Francesco Di Raimondo¹⁹, Giovanni Del Poeta⁴, and Valter Gattei¹

¹Clinical and Experimental Onco-Haematology Unit, Centro di Riferimento Oncologico di Aviano (CRO), IRCCS, Aviano (PN), Italy;  
²Department of Life Science, University of Trieste, Trieste, Italy;  
³Unit of Cancer Epidemiology, Centro di Riferimento Oncologico di Aviano (CRO) IRCCS, Aviano, Italy  
⁴Division of Haematology, University of Tor Vergata, Rome, Italy;  
⁵Haematology Unit AO of Cosenza, Cosenza, Italy;  
⁶Biothecnology Research Unit, AO of Cosenza, Cosenza, Italy;  
⁷Haematology and Bone Marrow Transplant Unit, Haemato-Oncology Department, Augusta Victoria Hospital, East Jerusalem, Israel;  
⁸Mutagenesis and Cancer Prevention Unit, IRCCS Ospedale Policlinico San Martino, Genoa, Italy;  
⁹Department of Medical, Surgical and Health Sciences, University of Trieste, Italy;  
¹⁰Haematology Unit, Presidio Ospedaliero S. Luca, ASL Salerno, Italy;  
¹¹Clinica Ematologica, Centro Trapianti e Terapie Cellulari "Carlo Melzi" DISM, Azienda Ospedaliera Universitaria S. Maria Misericordia, Udine, Italy;  
¹²University of Sussex, Brighton and Sussex Medical School, Brighton, UK;  
¹³Clinical Trials Research Unit, Leeds Institute of Clinical Trials Research, University of Leeds, Leeds, UK;  
¹⁴Molecular Diagnostic Centre, Department of Oncology, University of Oxford, Oxford, UK;  
¹⁵Department of Oncology, University of Oxford, Oxford, UK;  
¹⁶Section of Experimental Haematology, Leeds Institute of Cancer and Pathology (LICAP), University of Leeds, Leeds, UK;  
¹⁷Haematology, Institute of Oncology Research, Bellinzona, Switzerland;  
¹⁸Oncology Institute of Southern Switzerland, Bellinzona, Switzerland;  
¹⁹Division of Haematology, Ferrarotto Hospital, Catania, Italy.
Correspondence: Riccardo Bomben, Clinical and Experimental Onco-Hematology Unit, Centro di Riferimento Oncologico, I.R.C.C.S., Via Franco Gallini 2, Aviano (PN), Italy; email: rbomben@cro.it; and Valter Gattei, Clinical and Experimental Onco-Hematology Unit, Centro di Riferimento Oncologico, I.R.C.C.S., Via Franco Gallini 2, Aviano (PN), Italy; email: vgatei@cro.it.

Abstract: 245 words

Text: 4,401 words

Running Title: Small TP53 Mutated sub-clones in CLL
Abstract

**Purpose.** In chronic lymphocytic leukemia (CLL), TP53 mutations associate with reduced survival and resistance to standard chemo-immunotherapy (CIT); nevertheless, the clinical impact of TP53 mutated subclones below the 10-15% variant allele frequency (VAF) remains unclear.

**Experimental Design.** By a training/validation approach, we retrospectively analyzed the clinical impact of TP53 mutations above (high-VAF) or below (low-VAF) the previously reported 10.0% VAF threshold, as determined by deep next-generation-sequencing (NGS). Results were validated in a cohort (n=251) of CLL treated with FCR or FCR-like regimens from two UK trials.

**Results.** In the training cohort 97/684 patients bore 152 TP53 mutations while in the validation cohort 71/536 patients had 109 TP53 mutations. In both cohorts, TP53 mutated patients experienced significantly shorter overall survival (OS) than TP53 wild-type (wt) patients irrespective of the TP53 mutation VAF. By combining TP53 mutation and 17p13.1 deletion (del17p) data in the total cohort (n=1,220), 113 cases were TP53 mutated only (73/113 with low-VAF mutations), 55 del17p/TP53 mutated (3/55 with low-VAF mutations), 20 del17p only, and 1,032 (84.6%) TP53 wt. A model including low-VAF cases outperformed the canonical model considering only high-VAF cases (c-indices 0.643 vs. 0.603, \( P<0.0001 \)), and improved the prognostic risk stratification of CLL-IPI. Similar results were obtained by circumscribing the analysis to CIT-treated cases from the retrospective cohort (n=552) or the UK trials cohort.

**Conclusion.** TP53 mutations impacted OS even when detected at very sub-clonal levels. This finding can be used to update the definition of TP53 mutated CLL for clinical purposes.
Translational Relevance

Evaluation of the TP53 mutational status is a central pillar of the prognostic algorithms used for the clinical management of CLL patients, predicting both disease progression and sub-optimal response to therapy. Next-generation sequencing allows detection of TP53 mutations at a limit far below the conventional sequencing threshold recommended by international guidelines. By a training/validation approach, here we demonstrate that, in the chemo-immuno-therapy setting, TP53 mutations associate with short overall survival irrespective of variant allele frequency (VAF) percentage, and that low-VAF mutations (<10.0% VAF) maintain the same deleterious impact as high-VAF mutations (≥10.0% VAF). Results are validated in an additional cohort of CLL patients treated with FCR-based regimens in the context of two UK CLL trials. These findings can be used to re-define the classification of TP53 mutated CLL patients, and imply that the prognostic scoring systems that include TP53 mutation evaluation, e.g. the CLL-IPI, should be updated accordingly.
**Introduction**

Disruption of the *TP53* gene, either by deletion at chromosome 17p13.1 (del17p) or mutations, represents the most important biomarker in chronic lymphocytic leukaemia (CLL) (1–4) given its inclusion into algorithms of proved prognostic relevance both in the context of chemo-immunotherapy (CIT; e.g. CLL International Prognostic Index, CLL-IPI) and novel target therapies (5,6). Its relevance as predictive factor is also underscored by the need of therapeutic strategies alternative to CIT for patients harbouring the lesion at the time of therapeutic choice (7–9).

Recent studies based on ultra-deep-next generation sequencing (NGS) have shown that *TP53* mutations can be present in tumor cell populations at very low variant allele frequency (VAF), far below the detection limit of Sanger sequencing (10–15), although the actual clinical impact of these subclonal *TP53* mutations is still a matter of debate (10–15). In fact, while some studies showed that patients bearing small *TP53* mutated subclones experience similar outcome of patients with more evident clonal mutations of *TP53* (10,12), other studies failed to confirm these findings (13,15,16). In addition, the TP53 Network of the European Research Initiative on CLL (ERIC) still recommends to avoid to call mutations below the threshold of 10%VAF given the justified concern for the possibility of false negative or false positive results below the 5-10% VAF range (12).

In this study, by applying a training/validation approach in a wide CLL cohort, analysed in a single institution for *TP53* mutations by ultra-deep NGS coupled with a rigorous bioinformatics pipeline, we provide evidence that *TP53* mutations impact on overall survival (OS) even if detected at very low subclonal levels. Results are further validated in an additional cohort of CLL patients treated with FCR or FCR-like regimens in the context of two UK CLL trials (17,18).
Material and Methods

CLL cohorts
The REMARK criteria were followed throughout this study (Table S1). The main body of this study is represented by a retrospective analysis of 1,220 CLL cases (552 treated with CIT as first-line therapy, 93 of them treated with new agents as second-line therapy) from a cohort of 1,736 CLL patients, all diagnosed and treated according to iwCLL guidelines (1,19), consecutively referred to a single institution (Clinical and Experimental Onco-Hematology Unit, Centro di Riferimento Oncologico, I.R.C.C.S., Aviano, Italy) for molecular and cytogenetic analyses from 2003 to 2019 (Figure 1; details of the referral policy in Supplemental Information). Clinical outcome data were updated as of December 2019. The median follow-up from CLL diagnosis was 77 months (interquartile range, IQR, 39-120 months), while the median follow-up from sampling was 46 months (IQR 22-84 months). In the case of patients undergoing first treatment, they were all analyzed before therapy (median time from sampling to first treatment, 4 month; IQR 1-18 months); all cases analysed in a sample collected after first-line treatment (n=174) were excluded from the final cohort (Figure 1). No difference was found in terms of OS by comparing patients whose sample was obtained within the first year of diagnosis (659 cases, 299 treated) and those from whom samples were obtained after the first year (561 cases, 253 treated; Figure S1A).

For the purposes of the present study, the whole cohort was split in two separate cohorts: a training cohort (684 patients, all referred from a single-center) that was utilized to set-up the NGS approach and bioinformatics pipeline for TP53 mutation detection, and a validation cohort (536 patients, referred from four different centers); Table S2 summarizes the demographics of the whole cohort, and of the separate training and validation cohorts. A similar median OS was observed in the training cohort compared to the validation cohort (both not reached; P=0.9048, Figure S1B).

As further validation, a random splitting of the whole cohort according to a 70:30 ratio was also performed (see Supplementary Information “Set-up of alternative training and validation cohorts”).

As an additional validation cohort, 251 CLL samples from two UK trials, ARCTIC and ADMIRE (Table S3) were included (17,18), in which patients were randomized to receive either FCR or FCR-like regimens without significant difference both in PFS and OS (P=0.5923 and P=0.6745, respectively; Figure S1CD). The median follow-up was 83.9 months with 150 progressions and 76 deaths, all deaths preceded by disease progression. In this cohort, 61/251 cases were treated with new agents as salvage therapy.

Further information is available in Supplemental Information.

Specimen characteristics
In accordance with the ERIC recommendations for TP53 disruption,(20) mutation analyses were always carried out from nearly pure (>80%) tumor cells, as obtained from peripheral blood (PB) specimens. When the percentage of leukemic cells, as determined by explorative flow cytometry (CD5+/CD19+) on PB, was below the threshold of 80%, a CD19-guided positive selection, using an autoMACS Pro Separator (Miltenyi Biotec, Bergisch Gladbach, Germany), was performed (Figure S3).
**TP53 mutations and bioinformatics pipeline of analysis**

A detailed description of the procedures for TP53 mutational status determination, including functional evaluation and the applied bioinformatics pipeline is available in the Supplemental Information, and summarized in Figure S2. In particular, the bioinformatics pipeline can be retrieved from the GitHub website (details in Supplemental Information). All the samples from the retrospective Italian cohort, and the prospective UK cohort have been sequenced and analyzed with the same pipeline at the Clinical and Experimental Onco-Hematology Unit (Aviano, Italy).

Briefly, analysis of TP53 mutations was performed by NGS with an amplicon-based strategy, covering exons 2-11, in keeping with the ERIC recommendations (20). Specific primers were designed with the Primer3 program, and modified according to the Illumina (San Diego, CA) protocol (Table S4). Amplicon libraries were generated using a modified Illumina protocol starting from 40 ng of DNA (~6,000 diploid genomes), a quantity capable to successfully detect mutations below the 1% VAF in the context of our procedures (Figure S3). Multiplex PCR products were generated using Phusion High-Fidelity DNA Polymerase (Thermo Scientific, Milan, Italy) and subsequently tagged with specific index according to modified procedures for NexteraXT (DNA Library Preparation kit, Illumina), as previously reported (21,22). Purified libraries were pooled, and paired-end sequenced in a MiSeq instrument (Illumina).

FASTQ files were aligned to the Hg19 reference with Burrows-Wheeler Aligner (BWA)-MEM algorithm, and allele variants were called by FreeBayes with non-stringent parameters (Figure S2) (23–26). A coverage ≥2,000X was obtained for each sample in 100% of the analyzed sequences (Figure S4). To calculate random/systematic errors, a database utilizing a subset of TP53 wild-type (wt) cases (n=362) was generated from the training cohort and utilized for comparisons. TP53 mutations were accepted if (both conditions fulfilled): i) with a VAF that outdistanced for at least 2.75 standard deviations the mean of the transformed VAF distribution related to any single nucleotide position of the TP53 sequence (Figure S5); ii) validated by Fisher’s exact test after Bonferroni correction (P<0.01). The minimal allelic fraction for TP53 mutation calling was 0.4% (Figure S6).

The IARC TP53 Database (http://p53.iarc.fr/) (27) was used to annotate TP53 mutations (Figure S3), and to functionally evaluate TP53 missense mutations for their capability to transcribe the CDKN1A gene (27,28). TP53 mutated cases with less than 2% VAF were all confirmed by a second independent NGS run starting from DNA; in selected cases, TP53 mutation VAF values below 2%, were validated by a different experimental approach, i.e. PCR with amplification-refractory mutation system (ARMS-PCR; Figure S7) (12).

**FISH analysis**

Interphase FISH was performed to detect del17p, 11q22.3 deletion (del11q), 13q14 deletion (del13p), and trisomy 12 (tris12) (29,30), as detailed in the Supplemental Information.

**Other CLL characterizations**
CLL patients were also characterized for age, sex, Rai/Binet staging, *IGHV* gene mutational status, *NOTCH1* mutations (retrospective cohort only) and CD49d expression (Table S2 and S3), as previously reported (21,31–34). Figure S8 summarizes the clinical impact (OS) for all these variables in the retrospective cohort (training/validation cohorts and composite cohort).

**Statistical analysis**

All statistical analyses were performed by using standard methods (35–37); details are reported in Supplemental Information. OS was computed from date of blood sampling to death (events) or last follow-up (censoring); analysis of OS from first treatment was measured from date of first treatment to date of death (event) or last follow-up (censoring). Progression-free survival (PFS) was calculated from the date of treatment initiation to progression (event) or last follow-up (censoring). Molecular studies were blinded to the study end points.
Results

Clinical impact of TP53 mutations
The clinical impact of TP53 mutations was investigated in a single-institution training cohort (n=684) and in a multicenter validation cohort (n=536), whose clinical and biological features are summarized in Table S2 and Figures S1AB and S8.

In the training cohort, the set-up of our NGS approach coupled with a robust bioinformatics pipeline of analysis led to the identification of a total of 152 TP53 mutations in 97 patients (range of mutations/patient: 1-11). By identifying TP53 mutated patients according to the VAF of the most prevalent TP53 mutation, the VAF range for TP53 mutated cases was 0.53-95.24% (Table S5 and Figure S9A). When the same approach was applied in the validation cohort, 109 TP53 mutations were identified in 71 cases (range mutation/patient: 1-6; VAF range of the prevalent mutation: 0.53-92.47%; Table S5 and Figure S9B), with no significant difference between the two cohorts (P=0.3824, Table S2).

Regarding OS, TP53 mutated patients experienced significantly shorter OS than TP53wt patients both in the training cohort (median OS: 80.0 months versus not reached; P<0.0001; Figure S9C), and in the validation cohort (median OS: 73.0 months versus not reached in TP53 mutated versus wt patients, respectively; P<0.0001; Figure S9D).

Given the impact of TP53 disruption in the prognostic stratification and therapeutic choice for CLL patients (3,5,19,20), ERIC guidelines suggest that only the variants with ≥10% VAF by NGS, should be reported (20). Here we verified whether TP53 mutations with low VAF, i.e. below the 10% VAF threshold, had a different clinical impact when compared with TP53 mutations detected at higher VAF.

In this regard, 56 out of 97 cases (57.7%), and 36 out of 71 cases (50.7%) were identified with “high-VAF” (i.e. with ≥10.0% VAF) for TP53 mutations in the training and the validation cohort, respectively. The remaining 41 and 35 cases were classified as “low-VAF” cases (i.e. with <10.0% VAF) (Table S2). Both in the context of the training and validation cohorts, cases with high-VAF or low-VAF TP53 mutations experienced shorter OS than TP53wt cases, with no difference between high-VAF and low-VAF TP53 mutated cases (Figure S9EF).

A similar result was obtained upon random split of the total cohort of 1,220 cases according to a 70:30 ratio.

When considering the combined cohort of 1,220 cases (Table S2), a total of 261 TP53 mutations (Table S5) in 168 cases were found (13.8%; mutations per patient: 1-11; Figure 2A), 92 classified as high-VAF cases (VAF range: 10.8-95.2%, median VAF=53.0%), and 76 as low-VAF cases (VAF range: 0.53-9.6%; median VAF=2.6%) (Figure 2A). Again, TP53wt CLL presented a significantly longer OS when compared with both of the TP53 mutated categories (P<0.0001), without difference between high-VAF and low-VAF TP53 mutated cases (P=0.0712; Figure 2B, left panel).
This observation was confirmed by limiting the analysis to newly diagnosed CLL, i.e. sampled within 6 months from diagnosis (n=539; Figure S10).

No major differences were found by comparing the clinical and biological features of cases with high-VAF versus low-VAF TP53 mutations (Table S6). Notably, the clinical consequence of TP53 mutations was similar when CLL patients with low-VAF TP53 mutations were stratified into three subclasses with different VAF, i.e. <1% VAF (22 cases), 1-5% VAF (42 cases), 5-10.0% VAF (12 cases) (Figure S11A). Moreover, the capacity of TP53 mutations to identify cases with shorter OS was also confirmed by separately considering cases that presented with a single or multiple TP53 mutations, either classified as high-VAF or low-VAF (Figure S11B).

These observations were further validated by Harrell’s c-indices comparison of TP53wt versus TP53 mutated CLL models. In particular, a model classifying as TP53 mutated both cases with high-VAF TP53 mutations and cases with low-VAF TP53 mutations significantly outperformed (c-index 0.643, range 0.606-0.686) a model where only cases with high-VAF TP53 mutations were considered as mutated (c-index 0.603, range 0.572-0.638; P<0.0001) according to current recommendations (20).

Finally, the presence of TP53 mutations, as evaluated by combining low-VAF and high-VAF mutations, remained independent OS predictor by multivariate analysis also after adjusting for possible confounders, including sex, age, Rai staging, and other biological factors (i.e. IGHV status, CD49d expression, NOTCH1 mutations, del11q and del17p. The same held true when the training and the validation cohorts were separately considered (Table S7).

Molecular profile of TP53 mutations

By considering all 261 mutations found in the combined cohort (VAF range: 0.4-95.2%; median VAF=3.2%), 100 (38.3%) were classified as high-VAF (VAF range: 10.8-95.2%; median VAF=48.9%), while 161 (61.7%) were low-VAF (VAF range: 0.4-9.8%; median VAF=1.3%) (Table S5 and Figure 2A). According to needle plot graphs (Figure 2C), no distribution differences along the TP53 coding sequence were observed between high-VAF and low-VAF TP53 mutations. When mutations were analyzed for their effect on the TP53 protein in terms of amino acid changes, according to previous reports (38-40), the most common mutations in TP53 coding region were missense mutations, followed by nonsense, frameshift, and splicing mutations, without any distribution difference between high-VAF and low-VAF mutations (P=0.65 Figure 2D). In this context, the residual capacity of missense mutations to transcribe the CDKN1A gene, as retrieved from the IARC TP53 database (27,28), was comparably low in both high-VAF (13.04%, 95% CI 4.72-16.92) and low-VAF (13.74%, 95% CI 6.93-20.10; P=0.20) mutations. Consistently, by splitting patients into different categories according to the effect on protein of the mutation with the highest VAF, a significantly shorter OS was observed in all instances compared to TP53wt cases, regardless of the different types of mutations (Figure 2B, left panel). This observation was confirmed also in the context of the training and validation cohorts, separately considered (Figure S12AB).
**TP53 mutations and del17p**

In the context of training (n=684), validation (n=536) and total (n=1,220) cohorts, cases bearing del17p accounted for 41, 34, and 75, respectively (Table S2), these cases experiencing significantly shorter OS according to the conventional hierarchical Dohner’s classification (41) (Figure S8AB).

Combining FISH data on del17p with TP53 mutation data (Table S2 and Table S6) in the total cohort, we obtained: 1,032 cases without any TP53 aberration, 20 cases with del17p only, 113 TP53 mutated only cases (73/113 with low-VAF TP53 mutations), and 55 cases bearing both del17p deletion and TP53 mutation (3/55 with low-VAF TP53 mutations; Figure 3A). Notably, low-VAF cases were underrepresented in the del17p/TP53-mutated category, probably for a relative VAF overestimation due to the loss of chromosome 17. Patients with TP53 mutations alone (median OS: 80.0 months) or concomitant TP53 mutations and del17p (median OS: 67.0 months) had significantly shorter OS than TP53wt cases (median OS: not reached; P<0.0001; Figure 3B). Conversely, CLL cases bearing del17p only (median OS: 107.0 months) were too few (n=20) and with too short follow-ups to draw any definitive conclusion (Figure 3B).

Finally, by splitting cases according to the presence of a single-hit (i.e. either TP53 mutations or del17p), or a double-hit (i.e. both TP53 mutations and del17p) TP53 disruption, both groups experienced significantly shorter OS intervals than TP53wt cases, without difference by subdividing these categories into low-VAF and high-VAF TP53 mutated cases (Figure S13AB); similar results were obtained by separately considering the training and the validation cohort (Figure S13C-F).

**Low-VAF TP53 mutations and CLL-IPI risk categories**

We also tested whether low-VAF TP53 mutations had an impact in the context of the risk categories identified by the CLL-IPI (17,18). Complete data to score CLL cases according to the CLL-IPI were available in 652 patients from the retrospective cohort. Again cases with low-VAF TP53 mutations had shorter OS than TP53wt cases, with no difference with cases bearing high-VAF TP53 mutations (Figure S14A). In this cohort, CLL-IPI was computed by considering as TP53 mutated either cases bearing high-VAF TP53 mutations only (standard CLL-IPI), or cases bearing both high-VAF and low-VAF TP53 mutations (CLL-IPI revisited). In particular, in standard CLL-IPI, 322, 183, 115, and 32 cases were assigned to the low, intermediate, high, and very high risk categories, respectively (Figure S14B). A CLL-IPI recalculated including low-VAF TP53 mutated patients yielded the overall switching towards the high/very high risk categories of 40 cases previously assigned to better categories (low risk category, 17 cases; intermediate risk category, 11 cases; high risk category, 12 cases), all these cases experiencing a lower 10-year OS than the cases remaining in the original category (Figure S14B). Although the CLL-IPI was highly performing in both configurations, the CLL-IPI revisited with the inclusion of low-VAF TP53 mutated cases (c-index 0.730) significantly outperformed the standard CLL-IPI (c-index 0.721; P<0.0001).

**TP53 mutations and del17p in treated CLL**

By focusing on treated patients from the retrospective cohort (n=552), and using as clinical readout the OS from the start of therapy, TP53 mutated CLL (64 cases; median OS: 54.0 months) or TP53 deleted/mutated CLL (39 cases; median OS: 57.0 months) again experienced significantly shorter OS when compared with TP53wt cases (median OS: not reached; P<0.0001 in both comparisons;
In this context, CLL cases bearing high-VAF and low-VAF TP53 mutations (61 and 42 cases, respectively) had similar OS intervals (median OS: 47.0 months, or 62.0 months, respectively; \( P=0.3170 \)), significantly shorter than TP53wt cases (median OS: not reached; \( P<0.0001 \); Figure 3D). Superimposable results were obtained by separately considering the training and the validation cohorts (Figure S15AB), as well as by limiting analyses to patients (n=368) whose samples were collected less than 12 months from treatment (Figure S15C).

In the context of the ARCTIC/ADMIRE cohort, a total of 65 TP53 mutations (VAF range: 0.57-86.8%; 38.5% high-VAF, 61.5% low-VAF; Table S8 and Figure 4A) were distributed in 40 out of 251 cases (15.9%; 1-8 mutations per patient), 18 of them (45.0%) with a VAF<10.0% (Table S8). No distribution differences along the TP53 coding sequence were observed between high-VAF and low-VAF TP53 mutations, and the most common mutations were missense mutations, both in the high-VAF and low-VAF TP53 mutation categories (\( P=0.3153 \); Figure 4BC) (38–40).

Again, CLL cases bearing high-VAF and low-VAF TP53 mutations had OS (Figure 4D) and PFS (Figure S16A) intervals, significantly shorter than TP53wt cases (median OS: 107.9 months; \( P<0.0001 \) and \( P=0.0058 \) versus high-VAF and low-VAF TP53 mutation cases, respectively; median PFS: 69.3 months; \( P<0.0001 \) and \( P=0.0045 \) versus high-VAF and low-VAF TP53 mutation cases, respectively).

Also in this setting, TP53 mutations (high-VAF and low-VAF combined) remained independent predictors of OS/PFS after adjustment for possible confounders (Table S9), and both single-hit and double-hit TP53 disruption marked patients subsets with worse OS compared to TP53wt cases, irrespective to low-VAF/high-VAF TP53 mutations (Figure 4E and Figure S16B).

**Evolution of TP53 mutated clones upon treatment**

The evolution of TP53 mutated clones was assessed by longitudinal NGS analysis of sequential PB samples collected from 13 patients before the first treatment (median time of sampling from first treatment, -1 month; range –68/0 months, Figure S17A) and at relapse (median time from treatment, 27 months, range 2.5-76.0 months, Figure 17B) corresponding to 14 total TP53 mutations (8 low-VAF and 6 high-VAF TP53 mutations; Table S10). In these cases, the TP53 mutated clone identified before treatment invariably increased at the time of second analysis (\( P=0.0001 \); Table S10), which was performed in the close proximity of 2nd- or 3rd-line therapy in 10/13 cases (Figure S17B). In particular, in the 3 out of 8 cases with low-VAF TP53 mutations, the small TP53 mutation identified at pre-treatment crossed the 10% threshold and became predominant at relapse (Table S10 and Figure S17A).
Discussion

The present study, by taking advantage of different retrospective and prospective CLL cohorts (17,18), all analyzed for TP53 mutations by a highly controlled and sensitive NGS approach, demonstrates that: i) TP53 mutations can be detected at a very subclonal level in a significant fraction of CLL; ii) these TP53 mutations detected at very subclonal level, even if they represent the sole TP53 aberration, have a clinical relevance by marking CLL cases with similar shorter OS as cases bearing a higher TP53 mutational burden. These findings were confirmed also by circumscribing analyses to CIT treated patients, including those from the prospective UK trials (17,18), and after adjustment with possible confounders both in the retrospective and ARCTIC/ADMIRE cohorts. Altogether, these results suggest that CLL patients, even if affected by a disease bearing low-VAF TP53 mutations, should not be given CIT as 1st-line therapy (10,12). In the case of patients treated with novel agents after 1st-line therapy (93/1,220 in the retrospective cohort, 61/251 in the ARCTIC/ADMIRE cohort), the lack of TP53 mutation re-testing at the time of starting of novel agent treatments does not allow to draw any conclusion regarding the impact of low-VAF TP53 mutations in this setting.

As a cutoff to discriminate between high-VAF and low-VAF TP53 mutations, we relied upon the 10% VAF cutoff, a cutoff consistent with ERIC suggestions, in line with the use of NGS rather than Sanger by most centers nowadays (12,20). The frequency of TP53 mutated cases observed here was 14%, including cases with high-VAF and low-VAF TP53 mutations, reduced to 7.5% if we considered as TP53 mutated cases only the cases bearing high-VAF mutations. So, about 6.5% of cases (76 cases in our series) would have been misclassified as TP53 wild-type, in keeping with other studies (10,12).

While in accordance with previous studies (10,12), our findings differ from others (17,18), allegedly due to different pre-analytical preparations and/or bioinformatics pipeline of analysis. For example, in Blakemore et al (15), the number of cases with low-VAF mutations is only 16 (3.2% of cases) with a minimum VAF of 1.43% (0.43% in the present study). Purification of samples with <80% purity, as routinely performed by us, by removing the possibility of under representation of mutations due to “contaminating” normal DNA, may have a role in the identification of cases with a very low mutational burden.

On the other hand, Brieghel et al (13), in a cohort of 290 cases, reported about 5% of cases with a TP53 mutational burden <1% VAF, as compared to the 1.8% of cases detected here; this was probably due to a declared sequencing depth (20,000X) far superior to the 2,000X of our study, which was consistent with routine use of TP53 mutation analyses in clinical practice (13). Being about 0.5% the lowest VAF value found as a single mutation in TP53 mutated CLL in our cohort, further studies are needed to evaluate the clinical impact of mutations with a VAF below 0.5%.

The ad-hoc bioinformatics pipeline generated in this study, available on GitHub for possible users, was based on the creation of a matrix dataset of TP53wt cases in order to eliminate the background noise due to random sequencing errors. This approach allowed to confidently detect TP53 mutations with a <1% VAF, which were validated by a high-sensitivity extra-assay method like allele-specific
PCR (12). Of note, in addition to our pipeline, each CLL sample entering our study was double-tested using the commercially available Miseq-Reporter software, which needs neither deep bioinformatics skills, nor any type of reference database. In keeping with the detection limit of the software, the Miseq-Reporter identified all the mutations called by our pipeline up to the lower limit of 1.3% VAF, leading to an overall misinterpretation of 30 cases out of 1,220 from the retrospective cohort (2.4%), all with VAF below the 1.3 threshold (R.B., unpublished observation).

Low-VAF TP53 mutations, as reported here, displayed a molecular distribution along the TP53 gene, and functional features of TP53 dysfunction, completely superimposable to those reported for high-VAF TP53 mutations (12,28). Moreover, in sequential samples from patients subjected to CIT treatments, the minor TP53 mutated clones were positively selected becoming the dominant leukemic population at relapse in some cases. This effect is in keeping with the notion that these mutations, although initially occurring in a minority of cells, are selected over time eventually resulting in the dominant population (42). The observation that low-VAF and high-VAF TP53 mutations had similar negative clinical impact is also in keeping with the known capability of TP53 mutated tumor cells to enhance the overall tumor cell fitness by influencing the tumor microenvironment and/or the TP53wt neoplastic component (43).

Whereas tumor suppressors are commonly inactivated by frameshift or nonsense mutations, the most frequent mutation type of TP53 in CLL is represented by missense mutations (10,12,13,40,44), an observation confirmed by our findings. Here, we demonstrated that in the context of CLL patients, all the identified TP53 mutants could be considered functionally equivalent from a clinical point of view. Grouping patients according to the type of TP53 mutation showed that each group experienced the same shorter OS when compared to TP53wt patients. Even if missense TP53 mutations were split in mutations occurring within the DNA-binding motifs (DBM, 59 cases in our retrospective composite cohort), and mutations occurring outside the DBM (66 cases in our cohort), as reported by Trbusek et al (44), no difference was found in term of median OS (73 months versus 78 months, R.B., unpublished observation). Overall, these data suggest that the effect of TP53 mutations in CLL seems to be neither related to the percentage of mutations nor to the different types of mutations.

When integrating mutational results for TP53 with del17p data, a small proportion of “17p deleted only cases” (20/1,220 cases in our total retrospective series, 1.6%) was observed, in keeping with literature data (3,45,46). Although the number of cases was low and with short follow-up intervals, the sole presence of del17p did not seem to associate with a significant negative impact on OS when compared with TP53wt, in line to what previously reported for other lympho-proliferative diseases (14,47,48), while the same was not observed in the so-called TP53 mutated only cases. In this regard, given the different sensitivity between the FISH versus the NGS methods, it cannot be excluded that small subclonal deletions of TP53, below the FISH detection limit, may actually have occurred in (some) TP53 mutated only cases.

In conclusion, we have demonstrated that in the CIT setting TP53 mutations confer a significantly shorter OS, irrespective of VAF percentage. These findings can be used for correct identification and clinical management of CLL patients bearing TP53 disruption, and, as shown here, may imply to reconsider those prognostic scoring systems, e.g. the CLL-IPI (5), that include TP53 disruption.
data in their evaluation. The clinical impact of this finding in the setting of CLL treatment with chemo-free regimens (49,50) remains to be established.
**Funding:** The present study is supported in part by: Progetto Ricerca Finalizzata PE-2016-02362756, and RF-2018-12365790, Italian Ministry of Health, Rome, Italy; Associazione Italiana Ricerca Cancro (AIRC), Investigator Grant IG-21687; Associazione Italiana contro le Leucemie, linfomi e mielomi (AIL), Venezia Section, Pramaggiore/Veneto Orientale Group, Italy; Linfo-check - Bando ricerca - contributo art. 15, comma 2, lett b) LR 17/2014; “5x1000 Intramural Program”, Centro di Riferimento Oncologico, Aviano, Italy; the Fondazione Umberto Veronesi, Post-doctoral Fellowships-year 2019; Italian Ministry of Health 5x1000 funds 2013, 2015, 2016; Current Research 2016; Compagnia S. Paolo Turin Italy project 2017.0526, and Blood Cancer UK (grant # 18005).

**Authors’ contribution:** R.B., designed the study, interpreted data, and wrote the manuscript F.M.R., F.V., T.D.A., T.B., A.Z., E.T., F.P., E.V., M.D., E.Z., I.C., P.V., P.N., M.B., A.B., J.A.C., G.F., D.R. performed and interpreted molecular studies, and contributed to data interpretation; F.V., J.P., and R.B. generated the bioinformatics pipeline of analysis, and performed statistical analyses; E.S., A.B., M.G., F.M., G.P., G.D.A., J.O., P.B, A.C., F.Z., F.D.R., G.D.P. collected clinical data and contributed to data interpretation; C.P., A.H., A.S., P.H., collected clinical data related to the U.K. trials and contributed to data interpretation; V.G. designed the study, interpreted data, and wrote the manuscript.

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


an analysis of 3307 cases. Leukemia 2017;31:705–11.


40. Sabapathy K, Lane DP. Therapeutic targeting of p53: All mutants are equal, but some mutants are more equal than others. Nat Rev Clin Oncol 2018;15:13–30.


Figure legends

Figure 1. Flow-chart representation of the study. Reported is a schematic representation of the flow-chart of the study with the number of patients analyzed.

Figure 2. Molecular profile and clinical impact of TP53 mutations in the total retrospective cohort. A) Distribution of TP53 mutations in the total retrospective cohort. Bar chart graph reports the percentage of VAF identified by NGS analysis for all 261 TP53 mutations sorted in descending order with regard to the percent of VAF. Black bars indicate TP53 mutations with the highest VAF in cases with multiple mutations. Gray bars indicate TP53 mutations with a lower VAF respect to the mutations with the highest VAF in the context of a single case with multiple mutations. According to a cut-off of 10.0% VAF (dotted lines) 100 TP53 mutations (92 cases) had more than 10.0% VAF (high-VAF TP53 mutations) and 161 TP53 mutations (76 cases) had less than 10.0% VAF (low-VAF TP53 mutations). B) Left-panel. Kaplan-Meier curves comparing OS probabilities of 1,052 TP53 wild-type cases (wt, green line), 92 cases with high-VAF TP53 mutations, i.e. more than 10.0% of VAF (red line), and 76 cases with low-VAF TP53 mutations, i.e. less than 10.0% of VAF (blue line). Right-panel. Kaplan-Meier curves comparing OS probabilities of 1,052 TP53 wt cases (wt, green line), 26 cases with TP53 frameshift mutations (blue line), 125 cases with missense mutations (red line), 9 cases with nonsense mutations (black line), and 8 cases with mutations affecting splice sites (splicing; purple line). For cases with more than one mutation the effect of the mutation with the highest VAF is reported. The number of patients in each group is reported; P value refers to log-rank test. C) Needle plot graph of high-VAF and low-VAF TP53 mutations along the TP53 coding sequence. Sequences referring to the transactivation domain, the DNA binding domain and the tetramerization domain of the TP53 protein are reported in green, red and blue, respectively. D) Pie-chart of mutations effect on the TP53 protein in terms of amino acid changes in the high-VAF and low-VAF TP53 mutation context.

Figure 3. TP53 mutations and 17p deletions in the total retrospective cohort. A) Pie-chart reporting the number of patients classified in four different categories according to the combination of TP53 mutations and 17p deletions. B) Kaplan-Meier curves comparing OS probabilities of 1,032 TP53 wild-type cases (wt, green line), 20 cases with del17p only (del17p_only, black line), 113 cases with TP53 mutations only (Mut_only, red line), and 55 cases with concomitant TP53 mutation and del17p (Mut&del17p, blue line). C) Kaplan-Meier curves comparing OS after treatment of 441 TP53 wt cases (wt, green line), 8 cases with del17p only (del17p_only, black line), 64 cases with TP53 mutations only (Mut_only, red line), and 39 cases with concomitant TP53 mutation and del17p (Mut&del17p, blue line). D) Kaplan-Meier curves comparing OS after treatment of 441 TP53 wt cases (wt, green line), 44 cases with TP53 mutations less than 12.5% of VAF (Mut<12.5%, blue line), and 59 cases with TP53 mutations more than 12.5% of VAF (Mut≥12.5%, red line). The number of patients in each group is reported; P value refers to log-rank test.

Figure 4. TP53 mutations and deletions in the ARCTIC/ADMIRE cohort. A) Distribution of TP53 mutations VAF in the ARCTIC/ADMIRE cohort (251 cases). Bar chart graph reports the percentage of VAF identified by NGS analysis for all 65 TP53 mutations sorted in descending order with regard to the percent of VAF. Black bars indicate TP53 mutations with the highest VAF in
cases with multiple mutations. Gray bars indicate TP53 mutations with a lower VAF respect to the mutations with the highest VAF in the context of a single case with multiple mutations. According to a cut-off of 10.0% VAF (dotted lines) 40 TP53 mutations (22 cases) had more than 10.0% VAF (high-VAF TP53 mutations) and 25 TP53 mutations (18 cases) had less than 10.0% VAF (low-VAF TP53 mutations). B) Needle plot graph of high-VAF and low-VAF TP53 mutations along the TP53 coding sequence. Sequences referring to the transactivation domain, the DNA binding domain and the tetramerization domain of the TP53 protein are reported in green, red and blue, respectively. C) Pie-chart of mutations effect on the TP53 protein in terms of amino acid changes in the high-VAF and low-VAF TP53 mutation context. D) Kaplan-Meier curves comparing OS of 211 TP53 wt cases (green line), 22 cases with high-VAF mutations, i.e. ≥10.0% VAF (red line), and 18 cases with low-VAF mutations, i.e. <10.0% VAF (blue line). E) Kaplan-Meier curves comparing OS of 201 TP53 wt cases (green line), 31 cases that present either TP53 mutations only, or del17p only (single, blue line), and 11 cases with a concomitant TP53 mutation and del17p (double, red line). The number of patients in each group is reported; P value refers to log-rank test.