Leukemia
https://doi.org/10.1038/s41375-018-0215-9

ARTICLE

Chronic lymphocytic leukemia

Clonal diversity predicts adverse outcome in chronic lymphocytic leukemia

Alexander C. Leeksma[1,2] · Justin Taylor[3] · Bian Wu[4,5] · Jeffrey R. Gardner[3] · Jie He[6] · Michelle Nahas[6] · Mithat Gonen[7] · Wendimagegn G. Alemayehu[8] · Doreen te Raa[1] · Tatjana Walther[5] · Jennifer Hüllein[5] · Sascha Dietrich[9] · Rainer Claus[10,11] · Fransien de Boer[12] · Koen de Heer[13] · Julie Dubois[1,2] · Maria Dampmann[14] · Jan Dürig[14] · Marinus H. J. van Oers[1] · Christian H. Geisler[15] · Eric Eldering[1,2] · Ross L. Levine[3] · Vincent Miller[6] · Tariq Mughal[6] · Nicole Lamanna[16] · Mark G. Frattini[16] · Mark L. Heaney[16] · Andrew Zelenetz[3] · Thorsten Zenz[17,5] · Omar Abdel-Wahab[3] · Arnon P. Kater[1,2]

Received: 7 March 2018 / Revised: 17 June 2018 / Accepted: 26 June 2018
© Springer Nature Limited 2018

Abstract

Genomic analyses of chronic lymphocytic leukemia (CLL) identified somatic mutations and associations of clonal diversity with adverse outcomes. Clonal evolution likely has therapeutic implications but its dynamic is less well studied. We studied clonal composition and prognostic value of seven recurrently mutated driver genes using targeted next-generation sequencing in 643 CLL patients and found higher frequencies of mutations in *TP53* (35 vs. 12%, $p < 0.001$) and *SF3B1* (20 vs. 11%, $p < 0.05$) and increased number of (sub)clonal ($p < 0.0001$) mutations in treated patients. We next performed an in-depth evaluation of clonal evolution on untreated CLL patients (50 "progressors" and 17 matched "non-progressors") using a 404 gene-sequencing panel and identified novel mutated genes such as *AXIN1*, *SDHA*, *SUZ12*, and *FOXO3*. Progressors carried more mutations at initial presentation (2.5 vs. 1, $p < 0.0001$). Mutations in specific genes were associated with increased (*SF3B1*, *ATM*, and *FBXW7*) or decreased progression risk (*AXIN1* and *MYD88*). Mutations affecting specific signaling pathways, such as Notch and MAP kinase pathway were enriched in progressive relative to non-progressive patients. These data extend earlier findings that specific genomic alterations and diversity of subclones are associated with disease progression and persistence of disease in CLL and identify novel recurrently mutated genes and associated outcomes.

Introduction

Within the past several years, whole exome sequencing [1–7] and whole genome sequencing [4, 8–12] studies have identified mutations in multiple genes including *TP53*, *NOTCH1*, *SF3B1*, *ATM*, *POT1*, *XPO1*, *BIRC3*, *BRAF*, *KRAS*, *MYD88, SAMHD1*, and *NRAS* as recurrently mutated drivers in chronic lymphocytic leukemia (CLL). Moreover, these studies demonstrated that a proportion of CLL patients harbor significant intra-tumoral genomic heterogeneity—defined as the presence of several subclones coexisting within an individual patient. The identification of subclonal mutations might have therapeutic implications related to targeting of specific genomic alterations. Moreover, detection of subclones in CLL itself predicts for adverse prognosis [3, 4, 13]. In addition, clonal evolution has been studied in patients receiving chemoimmunotherapy [1] or novel agents [14, 15], but remains less well-studied in untreated patients from diagnosis to initiation of treatment.

Given the diversity of clinical presentation of CLL in addition to the need for a clinically relevant method of testing clonal composition, we applied targeted next-generation sequencing (NGS) analysis to five cohorts of clinically annotated untreated and previously treated CLL

These authors contributed equally: Alexander C. Leeksma, Justin Taylor, Bian Wu, Thorsten Zenz, Omar Abdel-Wahab, Arnon P. Kater.

Electronic supplementary material The online version of this article (https://doi.org/10.1038/s41375-018-0215-9) contains supplementary material, which is available to authorized users.

✉ Omar Abdel-Wahab
abdelwao@mskcc.org

✉ Arnon P. Kater
a.p.kater@amc.uva.nl

Extended author information available on the last page of the article

Published online: 23 July 2018

patients. Comprehensive genomic profiling (CGP) was able to annotate somatic mutations, IGHV mutational status, and CN alterations in a single assay, when applied to progressive and non-progressive CLL patients.

Our data highlight the need for prospective analyses of serial genomic analysis of CLL, consideration of clonal composition in prognostication of CLL, and highlight several specific molecular pathways associated with progression of CLL for further study.

Methods

Study population

Peripheral blood (PB) and bone marrow (BM) samples as well as clinical data were obtained from five different patient cohorts: (1) 273 CLL patients from Heidelberg University Hospital; (2) 100 CLL patients from Essen University Hospital; (3) 84 CLL patients from the University Medical Center Freiburg, (4) 119 patients from the prospective HOVON68 (H68) clinical trial, and (5) 208 samples from 67 initially untreated CLL patients at Memorial Sloan Kettering Cancer Center (MSKCC) (see Supplemental excel file for patient characteristics). All patients matched standard diagnostic criteria for CLL and provided informed consent. Patients were evaluated for the presence of mutations based on the availability of DNA. Supplemental Methods list clinical variables, inclusion criteria and treatments administered to disease progression experiencing patients of whom DNA and RNA were used for longitudinal genetic analysis (Supplemental Table 1 shows the number of patients and samples in each category of disease presentation studied in this latter analysis).

Sample preparation and sequencing

Genomic DNA (gDNA) and RNA were extracted from CLL cells purified from peripheral blood mononuclear cells (PBMCs) by positive selection with anti-CD19+ beads for samples from Essen, Freiburg, and part of the samples from Heidelberg. In the other part of the samples from Heidelberg, white blood cell count (WBC) and absolute lymphocyte count (ALC) were used to approximate the percentage (%) of CLL cells in PBMCs. For samples from H68, CLL cells were purified from cryopreserved PBMNCs by negative selection using anti-CD3, CD14, and CD16. Genomic DNA and RNA from MSKCC samples were extracted from cryopreserved PB or BM MNCs. WBC and ALC were used to approximate the % of CLL cells in PB. Clinical histopathologic analysis of BM biopsies were used to approximate the % of CLL cells in BM MNCs (Supplemental Table 2 lists these characteristics).

Excluding the MSKCC samples, targeted sequencing and data analysis were performed as previously described [16]. A targeted sequencing panel of seven recurrently mutated genes covering exons with mutation hotspots including splice sites of *BRAF* (exons 11 and 15), *KRAS* (exons 2 and 3), *MYD88* (exons 3 and 5), *NOTCH1* (exon 34), *NRAS* (exons 2 and 3), *SF3B1* (exons 14 and 15), as well as exons 4–10 of *TP53* was designed. Sequencing libraries were prepared and subjected to NGS on the GS Junior 454 platform (Roche, Penzberg, Germany) according to manufacturer's protocols (Roche). Sequence alignment, variant calling, and annotation were performed using GS Amplicon Variant Analyzer software version 2.5 (Roche). Filtering process was conducted based on bidirectional sequencing, variant depth, variant allele frequency (VAF), dbSNP, and COSMIC database (see Supplemental Fig. 1 for Bioinformatics pipeline plus workflow). The variants after filtering were manually reviewed in GS Run Browser software version 2.5 (Roche).

For MSKCC samples, gDNA was sequenced for all classes of genomic alterations in 404 genes recurrently mutated in cancer, including CLL (FoundationOne Heme; Supplemental Table 3), as well as selected introns of 32 commonly rearranged genes to capture fusions which may be missed with targeted mRNA sequencing (RNA-seq; Supplemental Table 4). RNA-seq was performed to capture a broad range of fusions for 265 genes (Supplemental Table 5) and to sequence the rearranged heavy variable segment. This sequence was analyzed using IgBlast to determine the IGHV clone and to quantitate somatic hypermutation. Sequencing was performed on the HiSeq2000 instrument (Illumina) with 40-bp paired reads to a median depth of 466×. Resultant sequences were analyzed for base substitutions, insertions, deletions, copy-number (CN) alterations, and selected gene fusions as previously described [17]. Validation of the FoundationOne Heme panel was done by profiling of 3696 hematological tumors [18]. Known germline variants from the 1000 Genomes Project (dbSNP135) were removed, and previously confirmed somatic alterations deposited in the Catalog of Somatic Mutations in Cancer (COSMIC); v62) were included for study [19]. All inactivating events (i.e., truncations and deletions) in known tumor-suppressor genes were also included.

Clonality analysis

Based on the use of purified tumor cells and/or % of total lymphocytes (to correct for the purity of CLL cells in PBMCs), we inferred the likely clonal/subclonal structure of mutations using variant allele frequency (VAF). VAF observed from sequencing was normalized by actual CLL cell content for samples with low tumor burden (<80%) and

no local CN change. Considering *TP53* variants with del (17)(p13.1) or *KRAS* variants with trisomy 12, VAFs were corrected using the copy number information from FISH results (VAF corrected = VAF observed * CNT/CNN). CNT is local CN in the tumor cell which was estimated from FISH results and CNN is local CN in normal cells which is always equal to 2. We classified variants into clonal and subclonal mutations and conservatively defined subclonal mutations to either have VAF of 1–40% or 60–90% and clonal to either have VAF of 41–59% or >90% respectively.

Statistical analysis

Pearson's Chi-square test was used to compare categorical factors and unpaired *t*-test was used to compare continuous variables between groups of patients. VAFs were analyzed using rank-based methods. Multiple comparisons were taken into account using single-step max-T procedure. For each of the genes that were found to be significantly different between initial samples progressors and non-progressors, multivariable logistic regression models were used to assess whether they were independently associated with progression adjusting for IGHV mutational status as well as del(11q) and del(17p).

Overall survival (OS) was calculated using the Kaplan–Meier estimator and Kaplan–Meier curves were compared by logrank testing. Independent risk factors for OS were identified by multivariable analysis using Cox's proportional hazards regression with backward selection including the following variables: genomic aberrations (del (17p) and trisomy12), *TP53*, *NOTCH1*, and *BRAF* mutation. All factors which were found to be associated with survival in univariate analysis were included in the model. A p-value < 0.05 was considered as statistically significant.

Results

Distinct clonal and molecular characteristics of untreated vs. treated CLL patients

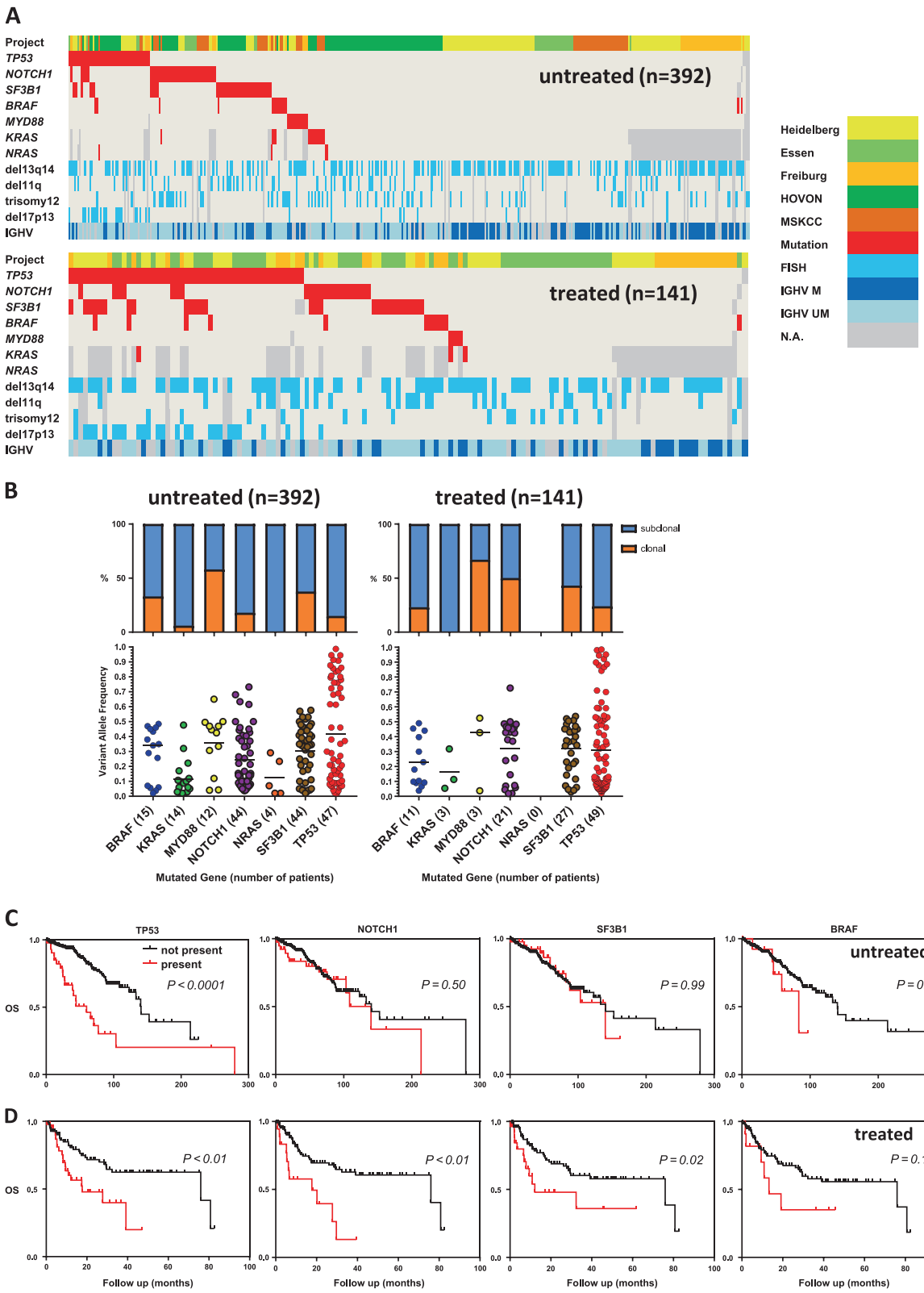
We sequenced seven CLL driver genes across 643 well-annotated CLL patients. The frequency of mutations in these seven genes in this cohort were as follows: *TP53* 111/639 (17.4%), *NOTCH1* 81/639 (12.7%), *SF3B1* 80/635 (12.6%), *BRAF* 29/642 (4.5%), *KRAS* 28/470 (3.8%), *MYD88* 16/637 (2.5%), and *NRAS* 4/475 (0.8%) (see Supplemental Fig. 2 for mutation hotspots). Frequencies and distribution of mutations, chromosomal aberrations, and IGHV mutational status are depicted in Supplemental Fig. 3A.

In total, 207/643 (32%) CLL patients harbored at least one subclonal mutation, 86/643 (13%) patients only carried clonal mutations and, in 349/643 (54%) patients, no mutation was detected in any of the 7 sequenced genes. VAF of mutations identified in *TP53*, *SF3B1*, *NOTCH1*, *BRAF*, *MYD88*, *KRAS*, *NRAS* are depicted in Supplemental Fig. 3B. We next compared untreated ($n = 392$) and treated ($n = 141$) CLL patients, excluding those where treatment status at date of sampling could not be confirmed. The mean number of mutations was significantly higher in treated patients compared to untreated patients ($p < 0.0001$; Fig. 1a). Moreover, the mean number of subclonal mutations was higher in treated patients compared to untreated patients ($p < 0.0001$; Fig. 1b). Mutations in specific genes were enriched in previously treated patients as compared to untreated patients: *TP53* (35 vs. 12%, $p < 0.001$) and *SF3B1* (20 vs. 11%, $p < 0.05$) as was already identified previously [1, 20, 21].

As expected, significant correlations were detected between specific mutations and CN alterations [1, 21, 22]. Patients harboring del(17p) frequently had a *TP53* mutation (83 vs. 11%, $p < 0.001$) and patients with del(11q) often had an *SF3B1* mutation (24 vs. 12%, $p = 0.002$) whereas trisomy 12 was seen less frequently in the setting on an *SF3B1* mutation (4 vs 19%, $p = 0.002$). Additionally, there was enrichment of specific mutations based on IGHV mutational status. For example, *MYD88* mutations were enriched in CLL with mutated IGHV ($p = 0.002$), whereas mutations in *BRAF* ($p < 0.001$), *NOTCH1* ($p = 0.002$), *SF3B1* ($p = 0.009$), and *TP53* ($p = 0.001$) were associated with unmutated IGHV (Supplemental Fig. 3C). Thus, sequencing a focused panel of putative driver genes across a large cohort of clinically annotated CLL patients confirmed a higher mutation load in previously treated patients with increased frequency of *TP53* and *SF3B1* mutations following treatment.

Differential prognostic value of mutations on survival in untreated and treated patients

In the total cohort, *TP53* ($p < 0.001$), *NOTCH1* ($p < 0.01$) and *BRAF* mutations ($p = 0.01$) were associated with adverse OS in univariate analysis. The presence of *SF3B1* mutations did not predict for poorer survival (Supplemental Fig. 4). Also, mutations in *MYD88*, *KRAS*, and *NRAS* were not associated with significant differences in survival while *MYD88* mutations showed a trend towards better prognosis ($p = 0.07$; data not shown). Using a multivariable analysis (MVA), mutations in *TP53* and *BRAF* retained statistical significance taking into account IGHV mutational status, trisomy12, del(17p) as well as mutations in *TP53*, *NOTCH1*, and *BRAF* (Supplemental Table 6). When only data of untreated patients are included, a similar trend was



observed (Fig. 1c), but this only revealed a significant effect for *TP53* mutations ($p < 0.01$). Also, *TP53* subclonal mutations were associated with adverse outcome ($p < 0.001$) (Supplemental Fig. 4). In contrast, for treated patients a worse survival was associated with *TP53* ($p < 0.01$), *NOTCH1* ($p < 0.001$) and *SF3B1* ($p = 0.02$) mutations

◀ **Fig. 1** Frequency of recurrent genetic alterations and their correlations with clinical characteristics in untreated and treated CLL patients. **a** Mutational analysis of seven putative driver genes in CLL was performed on a total of $n = 392$ untreated (upper panel) and $n = 141$ treated (lower panel) CLL patients from five different cohorts. Rows represent project, mutational status of putative CLL driver genes (*TP53*, *SF3B1*, *NOTCH1*, *BRAF*, *MYD88*, *KRAS*, *NRAS*), del13q14, del11q, trisomy 12, del17p13 and IGHV status. Columns represent individual patients. Upper row: color-coding is based on type of cohort (yellow = "Heidelberg cohort", light green = "Essen cohort", light orange = "Freiburg cohort", dark green = "HOVON68 cohort", brown = "MSKCC cohort"). Second-eighth row: gene mutation status for *TP53*, *SF3B1*, *NOTCH1*, *BRAF*, *MYD88*, *KRAS* and *NRAS* (white absence of mutation; red presence of mutation. Ninth-twelfth row: FISH analysis for deletions in 13q and 11q, trisomy 12, and deletions in 17p (white, absence of aberration; light blue presence of aberration). 13th row: IGHV mutation status (light blue IGHV unmutated, dark blue IGHV mutated). Gray boxes indicate analysis not performed. **b** Subclonal vs. clonal % (upper panel) and variant allele frequency (VAF) (lower panel) of mutations identified in the 7 analyzed putative CLL driver genes (*TP53*, *SF3B1*, *NOTCH1*, *BRAF*, *MYD88*, *KRAS*, *NRAS*) of mutated genes in untreated ($n = 392$) vs. treated ($n = 141$) CLL patients. VAF was corrected for percentage of CLL cells in sample (as described in Methods). Bar indicates the median VAF for each mutated gene. Kaplan–Meier plots of effect *TP53*, *SF3B1*, *NOTCH1*, *BRAF* on overall survival in **c** untreated and **d** treated CLL patients at date of sampling for mutational analysis

(Fig. 1d) and both *TP53* ($p < 0.05$) and *NOTCH1* ($p < 0.001$) subclonal mutations were associated with adverse outcome (Supplemental Fig. 4).

Overall, these data indicate that the mutational landscape in CLL is dynamic and suggest that chemoimmunotherapy affects clonal composition likely through clonal evolution as shown by others [1]. In order to study the extent of clonal evolution prior to treatment and the impact of this evolution on progression, we next aimed to longitudinally compare the mutational landscape of progressive vs. non-progressive CLL patients prior to any therapy.

Deep targeted sequencing of CLL identifies novel recurrent genetic alterations

We studied 67 CLL patients who initially presented with CLL without clinical indication for therapy more comprehensively using DNA and RNA sequencing. Fifty patients ("progressors") eventually developed progressive disease requiring the initiation of chemoimmunotherapy [23]. Seventeen patients with similar baseline characteristics who did not progress over a similar time frame ("non-progressors") were selected as comparators. Patients were identified in a retrospective fashion and non-progressors were specifically selected such that prognostic parameters, including Rai Stage, FISH/cytogenetic characteristics, and IGHV mutational status, were matched for progressors and non-progressors at initial presentation (Supplemental Table 1).

A median of two samples were sequenced per patient with a range of 1–6 samples per patient (Supplemental Fig. 5). The length of time between initial presentation and first serial analysis was similar between progressors and non-progressors (448 days vs. 469 days, respectively ($p = 0.87$). Moreover, the period of observation between most recent sample used for sequencing and study cut-off date was similar between progressors and non-progressors (2.7 years for progressors and 2.6 years for non-progressors, $p = 0.84$).

A custom targeted DNA- and RNA-sequencing platform (FoundationOne Heme panel) was applied to the 208 CLL samples from the 67 patients. In addition to mutational analysis on 404 genes, the assay included characterization of IGHV mutations as well as CN alterations including detection of trisomy 12, del(11q), del(13q), and del(17p). Both mutation status and CN alterations could be verified by PCR and FISH respectively, showing 100 and 97% concordance (Supplemental Figure 6).

In total, we identified 199 non-synonymous mutations across 82 protein-coding regions (Fig. 2a and Supplemental Table 7). A histogram of all genes affected by non-synonymous changes is shown in Fig. 2a. The most commonly mutated genes were *NOTCH1*, *TP53*, *KRAS*, *SF3B1*, *ATM*, *BCOR*, *BRAF*, and *FBXW7*. A number of genomic alterations not previously reported in >5% [3, 6, 9] and mutations not previously reported in CLL were identified, including *CREBBP*, *SPEN*, *AXIN1*, *SDHA*, *SUZ12*, and *FOXO3*. The majority of genes (57%) were mutated in only 1 sample, providing further evidence of the molecular heterogeneity of CLL noted in prior studies [1, 3, 6, 13, 22].

We found the median VAF of all mutations, following correction for the proportion of CLL cells in the samples utilized for sequencing, to be 22% (range 1–100%), (Fig. 2b). In fact, the distribution of VAFs did not follow a normal distribution with nearly half of the mutations being present at a VAF ≤ 20% (Supplemental Figure 7). These data suggest that the majority of somatic mutations in CLL are not present in the dominant CLL clone. Similarly, 54.5% of the *TP53* mutations were present at a VAF of ≤2%. Similar findings could be identified for other common mutations such as *NOTCH1*, where 31.1% of mutations were present at a VAF of <10% and would not have been identified by Sanger sequencing.

In this subset of patients used for comprehensive genomic profiling, mutations in *ATM* ($p = 0.003$), *BCOR* ($p = 0.02$), *SPEN* ($p = 0.008$) and total number of genomic alterations (CN alterations plus mutations; $p < 0.0001$), were enriched in CLL patients with unmutated IGHV (Supplemental Figure 8). In addition to correlations between gene mutations and IGH*V* mutational status, enrichment of specific mutations and CN alterations (Supplemental Figure 9 and Supplemental Table 8) was seen. As has been

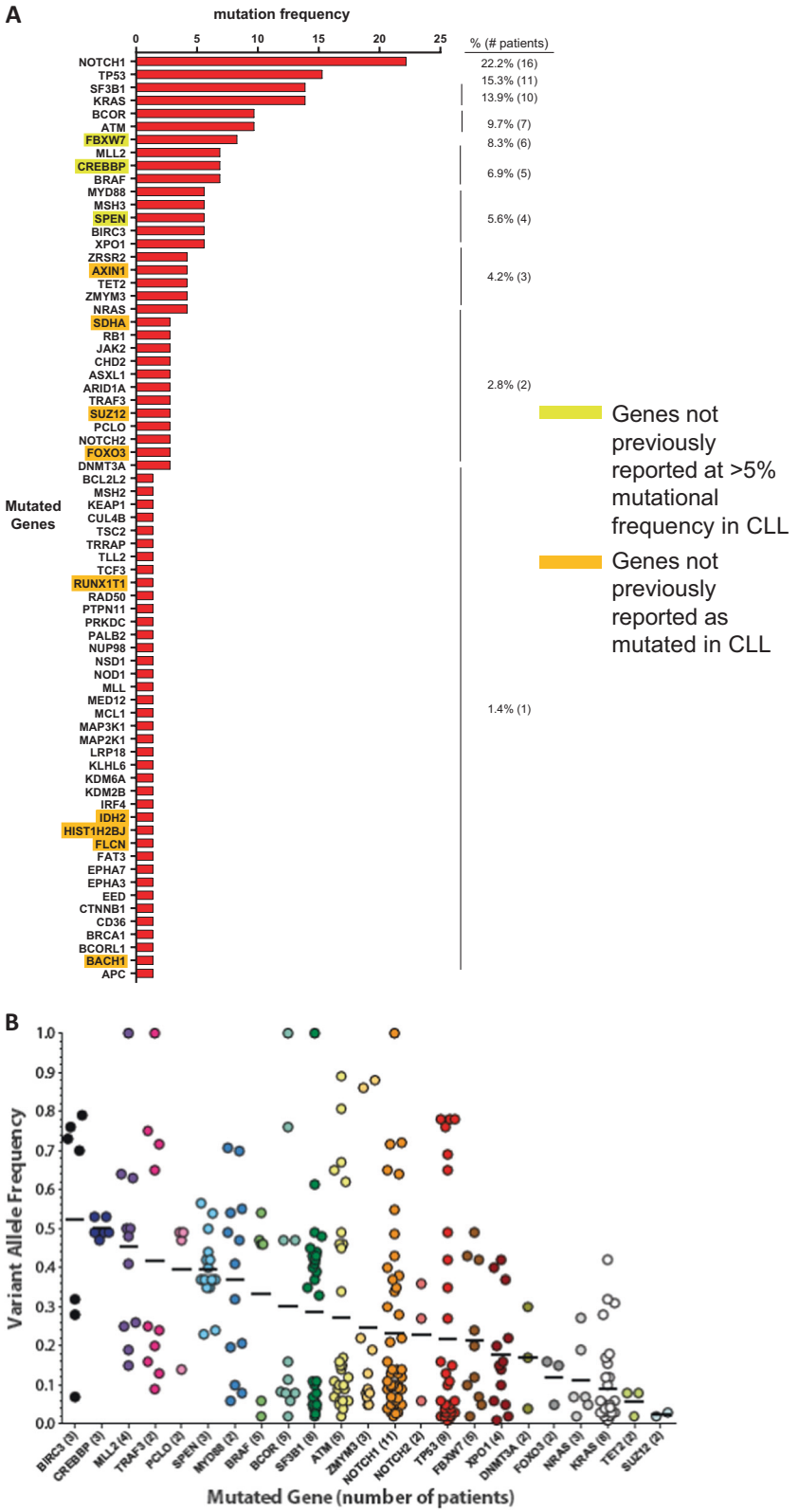


Fig. 2 Extensive, high-depth targeted sequencing analysis of CLL identifies novel recurrent genomic alterations. **a** Histogram of mutated genes identified at any time point in the 208 CLL samples sequenced. Indicated are the genes not previously identified to be mutated at >5% mutational frequency in CLL (yellow) and genes not previously identified to be mutated in CLL (orange). **b** Variant allele frequency (VAF) of genes mutated in ≥2 samples in CLL patients from **a**. VAF was corrected for percentage of CLL cells in sample (see Methods). Bar indicates the median VAF for each mutated gene

previously reported, there was a significant correlation between specific CN alterations and somatic mutations in key tumor suppressor genes targeted by these CN alterations, such as between *ATM* mutations, as well as *BIRC3* mutations, and del(11q) ($p < 0.001$ and $p = 0.001$, repectively; Supplemental Table 8). A previously described

[24] significant correlation between *MYD88* mutations and del(13q) ($p < 0.01$) was observed and in addition to identifying the previously described association between *NOTCH1* and trisomy 12 ($p < 0.01$) [25], we also detected an association between trisomy 12 and mutations in *SPEN* ($p = 0.04$) (Supplemental Table 8), a negative regulator of Notch signaling.

In addition to DNA-based analysis of somatic mutations and CN alterations, targeted combined DNA and RNA-seq was also performed to assess for rearrangements. This detected two patients with *IGH-BCL2* rearrangements (with similar breakpoints as previously described in CLL) [26, 27] (Supplemental Table 7) and no other rearrangements, consistent with prior data on the rarity of rearrangements in CLL [28].

Somatic mutations enriched at disease progression

We next longitudinally compared the evolution of somatic mutations between progressors and non-progressors. There was a significantly higher number of mutations at initial presentation in progressors as compared to non-progressors (median of 2.5 in progressors compared with 1 in non-progressors ($p < 0.0001$); Fig. 3a). Number of mutations was higher in progressor samples at any timepoint in disease, including timepoints following chemotherapy treatment (Fig. 3a). Given that progressors had significantly more mutations than non-progressors, we next evaluated whether specific genes were mutated more frequently in progressors. In univariate analysis, mutations in *SF3B1*, *ATM*, and *FBXW7* were significantly enriched in the initial sample of progressors relative to non-progressors (Fig. 3b and Supplemental Figure 10). Conversely, mutations in *AXIN1* and *MYD88* were significantly enriched in non-progressors, suggesting that mutations in these genes may be biomarkers of indolent CLL. We next repeated this analysis using MVA taking into account IGHV mutational status and CN status of 17p and 11q with mutations in *AXIN1* ($p = 0.006$), *SF3B1* ($p = 0.008$) and *ATM* ($p = 0.009$) retaining statistical significance. We also compared VAF of recurrently mutated genes between progressors and non-progressors (Fig. 3c). A number of genetic events were enriched in allelic frequency in progressors relative to non-progressors but none reached statistical significance.

Mutations affecting specific pathways are associated with disease course

Prior genomic analyses of CLL have grouped recurrently mutated genes into core biologic pathways [1, 3]. We therefore next compared mutations affecting these core pathways amongst different courses of disease (non-progressors, initial sample progressors, pre-treatment progression samples, and post-treatment progression samples). Mutations affecting Notch signaling, DNA repair, and RNA processing were significantly enriched in the initial sample from progressive relative to non-progressive patients (Fig. 4). Conversely, mutations affecting Wnt/β-catenin were enriched in non-progressors. Mutations affecting the DNA repair pathway were increasingly enriched with each phase of disease among progression patients, suggesting that chemotherapy treatment was associated with selection of CLL clones with mutations impacting DNA repair (Fig. 4). In contrast, mutations affecting MAP kinase signaling were enriched at disease progression but decreased following chemotherapy treatment.

Given that we identified frequent mutations in the MAP kinase pathway and Notch signaling we next examined specific genes mutated in each of these pathways. We identified *NOTCH1* to be mutated at a frequency of 22.2% of CLL patients which is higher than the 12.5% mutational frequency observed in our total dataset and the 10–17% noted in several recent studies [1, 3, 21, 29, 30] (Fig. 5a). The subclonal nature of mutations affecting the Notch pathway was evident in the fact that several CLL patients here harbored co-occuring mutations in *NOTCH1*, *FBXW7*, and/or *NOTCH2* (Fig. 5a).

In addition to the Notch pathway mutations, mutations affecting MAP kinase signaling were most commonly subclonal with a median overall VAF of 5% (range 1–47%; Fig. 5d). We identified numerous patients with simultaneous activating mutations in *NRAS*, *KRAS*, *BRAF*, and/or *MAP2K1* (Fig. 5e). We also observed that *NRAS*, *KRAS* and *BRAF* mutation frequency was higher in this cohort compared to what was seen in the total analyzed dataset. As mentioned earlier, recent work has highlighted the importance of subclonal *TP53* mutations in CLL [31], which may co-exist with CN loss of *TP53*. We therefore examined the frequency of CN changes and somatic mutations affecting DNA damage and repair pathway members (Fig. 5g). This analysis revealed frequent overlap of somatic mutations involving *TP53* as well as concomitant CN loss of *TP53* (Supplemental Table 8) as previously noted in CLL [31]. Similar overlap of somatic mutations as well as CN loss of *ATM* was also seen (Fig. 5g). These data further highlight the need for genomic analysis of both somatic and cytogenetic loss in clinically important genes in CLL patients such as *TP53* and *ATM*.

Targeted sequencing traces clonal evolution of CLL

Genomic alterations present with a VAF ≤ 10% were identified in 61% of CLL patients, suggesting that the vast majority of CLL patients may harbor subclonal genomic

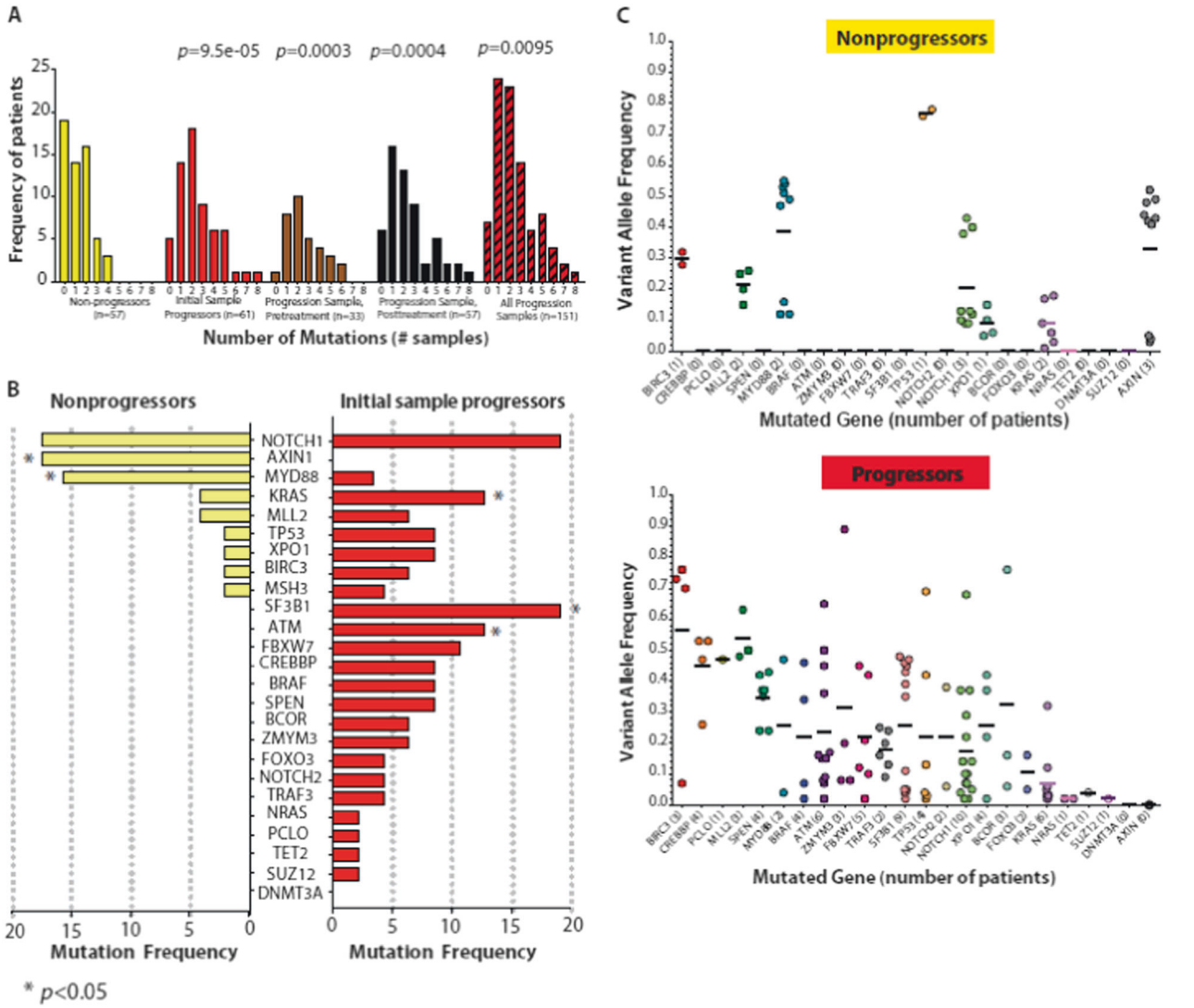


Fig. 3 Total number of somatic mutations at initial presentation as well as mutations in specific genes is associated with disease progression or lack of disease progression in CLL. **a** Histogram of the number of mutations at each stage of disease (non-progression ($n = 57$ samples from 17 patients), initial sample from progressors ($n = 61$ samples from 51 patients) progression samples pretreatment ($n = 33$ samples from 27 patients, progression samples post-treatment ($n = 56$ samples from 25 patients), progression samples from any time point ($n = 89$ samples from 51 patients)). p-value (Wilcoxon Rank Sum test, with continuity correction) indicates comparison of progression sample to non-progressors. **b** Mutational frequency of individual recurrently mutated genes in non-progressors relative to initial sample from progressors. Asterisks indicate p-value < 0.05 comparing the two groups (univariate analysis; Pearson X-squared test). **c** Variant allele frequency (VAF) of mutations identified in non-progressors (on top) and progressors (bottom). The mutated gene, with number of patients bearing mutations, is noted in the x-axis. VAF is corrected for CLL cell burden as described in the Methods section

alterations. Although some subclonal genetic events present at initial presentation became undetectable following chemotherapy this was concomitant with emergence of additional events at the time of progression. For example, an initially untreated CLL patient with del(11q) plus 2 independent *ATM* mutations underwent therapy with the PI-3Kδ inhibitor idelalisib (GS-1101) plus rituximab. Treatment resulted in normalization of blood counts and clinical improvement in adenopathy. Serial genomic analysis indicated the presence of at least 2 CLL clones pretreatment. While idelalisib/rituximab eliminated 1 clone, treatment was associated with persistence of the other initial clone and emergene of 1–2 additional clones (Figure S11A).

The identification of multiple subclonal mutations and/or CN alteration affecting the same gene was seen within many individual patients and provides further evidence for convergent evolution in CLL [1, 13, 16, 32]. For example, one patient presented with 2 separate *NOTCH1* mutations at initial presentation as well as with a subclonal *TP53* mutation (Figure S11B). Chemoimmunotherapy (6× pentostatin, cyclophosphamide and rituximab) failed to induce remission and was associated with development

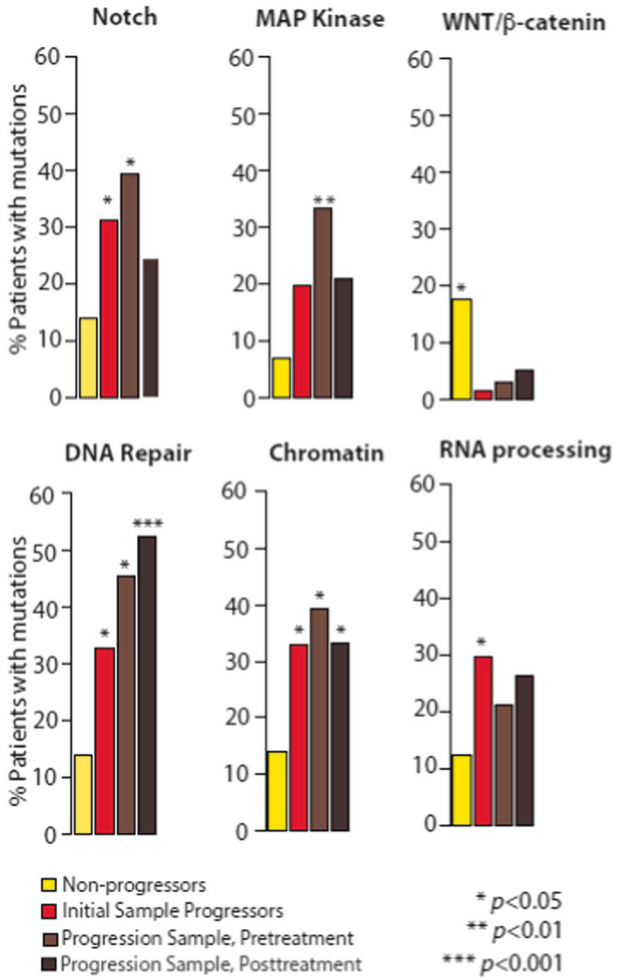


Fig. 4 Mutations affecting specific genetic pathways are associated with disease progression or lack of disease progression in CLL. Each colored bar indicates a phase of disease (non-progressors, initial sample progressors, progression samples pretreatment, progression samples post-treatment). Asterisk represents p-value < 0.05 comparing the frequency of mutations per patient in each pathway (Pearson X-squared test). Genes included in each pathway are listed in Supplemental Table 9

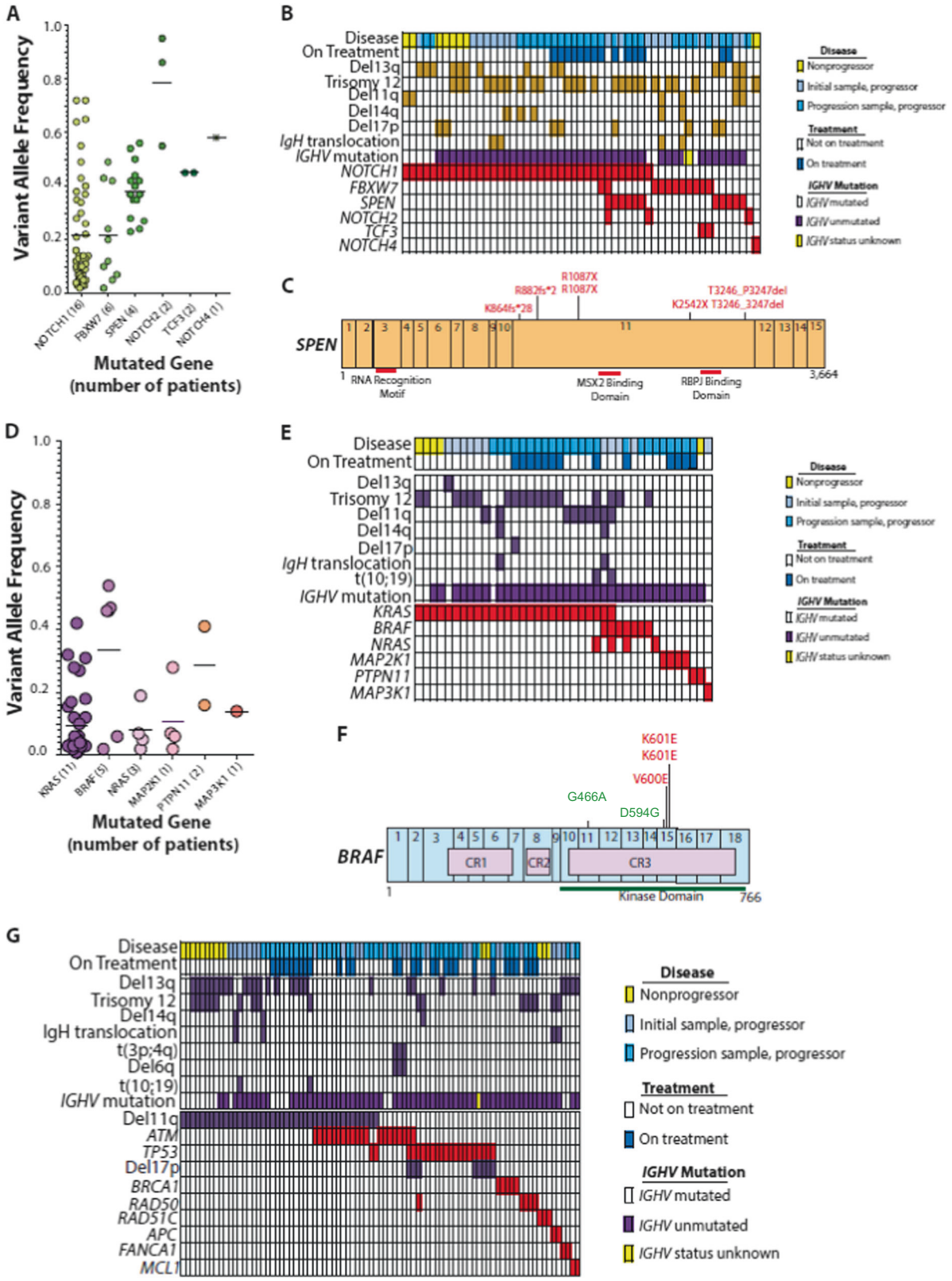
of del(17p) in addition to the pre-existing single *TP53* mutation. Further treatment (6× pentostatin, cyclophosphamide, rituximab, and mitoxantrone plus lenalidomide, ofatumumab and R-CVP) was then associated with development of 2 additional subclonal *TP53* mutations resulting in the presence of 4 genomic alterations affecting *TP53* in this patient. These data demonstrate the ability of targeted NGS encompassing mutations, CN alterations, and IGHV mutational status to identify clonal events associated with disease progression and relapse in CLL.

Discussion

In the largest cohort of CLL patients analyzed by targeted NGS to date, we found that somatic genetic alterations in untreated CLL patients provide prognostic information independent of currently used clinical parameters. In addition, using a large targeted NGS panel in a cohort of stable vs. progressive CLL patients, we observed mutations in genes not previously identified to be frequently mutated in CLL as well as previously identified putative driver genes that segregated by disease progression (*SF3B1* and *ATM*) or lack of progression (*AXIN1*) in multivariate analysis.

The depth of sequencing combined with the large number of genes sequenced facilitated identification of subclonal genomic alterations in nearly every CLL patient studied. This study further supports the recent finding that subclonal genomic alterations are much more common than previously appreciated in CLL and that sequencing only a small panel of genes may miss many relevant mutations [13]. Next-generation sequencing in CLL has been helpful in identifying genetic subgroups of patients with particular features and clinical outcomes. Here we identified mutations enriched in untreated CLL patients that ultimately progressed including mutations in *SF3B1*, *KRAS* and *ATM*. In addition, we also identified mutations associated with matched patients that did not progress during the same interval including mutations in *MYD88* and *AXIN1*. Mutations in *MYD88* were previously noted to be associated with del (13q), mutated IGHV, and a clinically indolent form of CLL [24]. *AXIN1* is a negative regulator of the Wnt/β-catenin signaling pathway, a pathway known to be recurrently affected by mutations in CLL [33, 34]. Further studies focused on sequencing of *AXIN1*, *KRAS*, and other genes noted here in larger cohorts of untreated CLL patients will be critical in validating these potential biomarkers of disease progression.

Although the clinical impact of subclonal *TP53* mutations in both untreated and treated CLL has been recently demonstrated [31], further efforts to understand the prognostic relevance of subclonal mutations in *NOTCH1*, *SF3B1*, *ATM* and other genes with clinical importance in CLL will be needed. Several of these mutated genes may be key therapeutic targets. For instance *BRAF* mutations were associated with shorter survival. It should be noted that most of the *BRAF* mutations noted in this study and others affect codons outside of V600, but are expected to result in high B-Raf kinase activity [35]. They therefore may confer sensitivity to MEK inhibition [36, 37] or to mutant-specific BRAF inhibitors if located at V600 [38]. Recurrent mutations affecting MAP kinase signaling in genes including *NRAS*, *KRAS*, *BRAF*, *PTPN11*, and *MAP2K1* were seen in 25% of patients, highlighting a need to understand the therapeutic impact of MAP kinase targeted therapies more broadly in CLL as well. Although the entire coding region of *BTK* and *PLCG2* was sequenced, no *BTK* and *PLCG2* mutations were identified in this cohort, consistent with a



recent report that *BTK* mutations rarely exist preceding BTK inhibitor therapy or may be below the level of detection of most assays [39].

We also demonstrate that common CN alterations often co-exist with mutations in genes targeted by CN loss in CLL (e.g., the occurrence of *TP53* mutations and del(17p) and

◀ **Fig. 5** Distribution of mutations in Notch, MAP kinase, and DNA damage/repair pathways as determined by targeted, high-depth mutational analysis. **a** VAF of mutations in genes encoding members of the Notch pathway (VAF corrected for CLL burden in sample at time of sequencing as noted in Methods). **b** Distribution of genes identified as mutated in **a**. **c** Mutations in *SPEN* identified in CLL patients from this study. **d** VAF of mutations in genes encoding members of the MAP kinase pathway. **e** Distribution of genes identified as mutated in **d**. **f** Mutations in *BRAF* identified in CLL patients from this study (mutations known to result in increased kinase activity in red, and those known to result in a non-active kinase in green). **g** Distribution of mutations in select genes encoding members of DNA damage/repair pathways

ATM mutations and del(11q)). Currently, detection of CN alterations is performed by interphase FISH while IGHV mutational status is performed by PCR in most clinical settings. Thus, the addition of molecular analysis by NGS would necessitate the use of at least three modalities for genetic characterization of CLL. We show that CGP can simultaneously detect CN alterations, IGHV mutational status, and the mutational status of recurrently mutated genes in CLL. Application of this technique to serial samples revealed the persistence of somatic mutations in some CLL patients who clinically respond to therapy as well as the emergence of novel mutations and CN alterations with relapse.

The limitations of this study include the retrospective design and the limited amount of non-progressors followed longitudinally; however, our results support the recent observation of evolutionary capacity as indicator of adverse outcome [15]. Despite our findings, it remains incompletely understood whether treatment influences clonal selection as a result of evolutionary pressure, or whether more mutations at baseline only reflect more unstable genomes. Our results here would suggest that both potentially play a role given that untreated patients that ultimately progressed had higher mutation frequency and more complex clonal composition at diagnosis.

In conclusion, we demonstrate that recurrent gene mutations are enriched in chemotherapy refractory patients and we observe a more complex mutational landscape in CLL progressors vs. non-progressors. CGP may be a valuable alternative tool for identifying actionable genetic alterations that predict for disease relapse and in elucidating the impact of therapy on clonal composition.

Acknowledgements ACL is supported by the van der Laan foundation. JT is supported by an American Society of Hematology (ASH) Research Training Award for Fellows, an ASCO Young Investigator Award, and the ASH and Harold Amos Medical Faculty Development Program of the Robert Wood Johnson Foundation. JG was supported by the Vogelstein Fund at MSKCC. OAW is supported by NIH/NCI R01 (1R01HL128239-01 and 1R01 CA201247-01) and support from the Cycle for Survival. This work was also supported by the Peter Solomon Fund at MSKCC. TZ was supported by a Mildred Scheel Professorship, the "Monique Dornonville de la Cour Stiftung" and the Transcan project. Many thanks go to the participating patients and we would like to acknowledge all participating institutions in the HOVON68 CLL trial, the data managers and statisticians of the HOVON data center, as well as the HOVON68 CLL study funding sources, the participating German centers and MSKCC.

Author contributions ACL, JT, BW, JRG, MN, JH, MG, OAW, TZ, and APK were responsible for conception and design of the study. ACL, BW, JT, JRG, RLL, VM, and TM were involved in analysis and interpretation of data. MN, JH, MH, DR, WGA, TW, JH, FB, KH, JD, NL, MGF, SD, RC, JD, CHG, MHJO, MLH, and AZ, TZ, OAW, and APK were involved in acquisition of patient materials. ACL, BW, JRG, and MG perfomed statistical analyses. ACL, BW, and JRG made the figures. ACL, JT, TZ, OAW, and APK wrote the first draft; and all authors were involved in revising the manuscript and approved the final version.

Compliance with ethical standards

Conflict of interest MN, JH, VM, and TM are employees of Foundation Medicine. RLL and OAW are consultants for Foundation Medicine. The remaining authors declare that they have no conflict of interest.

References

1. Landau DA, Tausch E, Taylor-Weiner AN, Stewart C, Reiter JG, Bahlo J, et al. Mutations driving CLL and their evolution in progression and relapse. Nature. 2015;526:525–30.
2. Ljungstrom V, Cortese D, Young E, Pandzic T, Mansouri L, Plevova K, et al. Whole-exome sequencing in relapsing chronic lymphocytic leukemia: clinical impact of recurrent RPS15 mutations. Blood. 2016;127:1007–16.
3. Landau DA, Carter SL, Stojanov P, McKenna A, Stevenson K, Lawrence MS, et al. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. Cell. 2013;152:714–26.
4. Nadeu F, Delgado J, Royo C, Baumann T, Stankovic T, Pinyol M, et al. Clinical impact of clonal and subclonal TP53, SF3B1, BIRC3, NOTCH1, and ATM mutations in chronic lymphocytic leukemia. Blood. 2016;127:2122–30.
5. Rose-Zerilli MJ, Gibson J, Wang J, Tapper W, Davis Z, Parker H, et al. Longitudinal copy number, whole exome and targeted deep sequencing of 'good risk' IGHV-mutated CLL patients with progressive disease. Leukemia. 2016;30:1301–10.
6. Quesada V, Conde L, Villamor N, Ordonez GR, Jares P, Bassaganyas L, et al. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. Nat Genet. 2011;44:47–52.
7. Fabbri G, Rasi S, Rossi D, Trifonov V, Khiabanian H, Ma J, et al. Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation. J Exp Med. 2011;208:1389–401.
8. Schuh A, Becq J, Humphray S, Alexa A, Burns A, Clifford R, et al. Monitoring chronic lymphocytic leukemia progression by whole genome sequencing reveals heterogeneous clonal evolution patterns. Blood. 2012;120:4191–6.
9. Puente XS, Pinyol M, Quesada V, Conde L, Ordonez GR, Villamor N, et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. Nature. 2011;475:101–5.
10. Burns A, Alsolami R, Becq J, Timbs A, Bruce D, Robbe P, et al. Whole-genome sequencing of chronic lymphocytic leukaemia reveals distinct differences in the mutational landscape between IgHV(mut) and IgHV(unmut) subgroups. Leukemia. 2017;3:332–42.

11. Kasar S, Kim J, Improgo R, Tiao G, Polak P, Haradhvala N, et al. Whole-genome sequencing reveals activation-induced cytidine deaminase signatures during indolent chronic lymphocytic leukaemia evolution. Nat Commun. 2015;6:8866.
12. Zhao Z, Goldin L, Liu S, Wu L, Zhou W, Lou H, et al. Evolution of multiple cell clones over a 29-year period of a CLL patient. Nat Commun. 2016;7:13765.
13. Nadeu F, Clot G, Delgado J, Martin-Garcia D, Baumann T, Salaverria I, et al. Clinical impact of the subclonal architecture and mutational complexity in chronic lymphocytic leukemia. Leukemia. 2017;32:645–53.
14. Burger JA, Landau DA, Taylor-Weiner A, Bozic I, Zhang H, Sarosiek K, et al. Clonal evolution in patients with chronic lymphocytic leukaemia developing resistance to BTK inhibition. Nat Commun. 2016;7:11589.
15. Landau DA, Sun C, Rosebrock D, Herman SEM, Fein J, Sivina M, et al. The evolutionary landscape of chronic lymphocytic leukemia treated with ibrutinib targeted therapy. Nat Commun. 2017;8:2185.
16. Jethwa A, Hullein J, Stolz T, Blume C, Sellner L, Jauch A, et al. Targeted resequencing for analysis of clonal composition of recurrent gene mutations in chronic lymphocytic leukaemia. Br J Haematol. 2013;163:496–500.
17. Lipson D, Capelletti M, Yelensky R, Otto G, Parker A, Jarosz M, et al. Identification of new ALK and RET gene fusions from colorectal and lung cancer biopsies. Nat Med. 2012;18:382–4.
18. He J, Abdel-Wahab O, Nahas MK, Wang K, Rampal RK, Intlekofer AM, et al. Integrated genomic DNA/RNA profiling of hematologic malignancies in the clinical setting. Blood. 2016;127:3004–14.
19. Forbes SA, Bindal N, Bamford S, Cole C, Kok CY, Beare D, et al. COSMIC: mining complete cancer genomes in the catalogue of somatic mutations in cancer. Nucleic Acids Res. 2011;39:D945–50.
20. Rossi D, Bruscaggin A, Spina V, Rasi S, Khiabanian H, Messina M, et al. Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. Blood. 2011;118:6904–8.
21. Baliakas P, Hadzidimitriou A, Sutton LA, Rossi D, Minga E, Villamor N, et al. Recurrent mutations refine prognosis in chronic lymphocytic leukemia. Leukemia. 2015;29:329–36.
22. Wang L, Lawrence MS, Wan Y, Stojanov P, Sougnez C, Stevenson K, et al. SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. N Engl J Med. 2011;365:2497–506.
23. Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, Dohner H, et al. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. Blood. 2008;111:5446–56.
24. Martinez-Trillos A, Pinyol M, Navarro A, Aymerich M, Jares P, Juan M, et al. Mutations in TLR/MYD88 pathway identify a subset of young chronic lymphocytic leukemia patients with favorable outcome. Blood. 2014;123:3790–6.
25. Balatti V, Bottoni A, Palamarchuk A, Alder H, Rassenti LZ, Kipps TJ, et al. NOTCH1 mutations in CLL associated with trisomy 12. Blood. 2012;119:329–31.
26. Adachi M, Cossman J, Longo D, Croce CM, Tsujimoto Y. Variant translocation of the bcl-2 gene to immunoglobulin lambda light chain gene in chronic lymphocytic leukemia. Proc Natl Acad Sci USA. 1989;86:2771–4.
27. Dyer MJ, Zani VJ, Lu WZ, O'Byrne A, Mould S, Chapman R, et al. BCL2 translocations in leukemias of mature B cells. Blood. 1994;83:3682–8.
28. Ferreira PG, Jares P, Rico D, Gomez-Lopez G, Martinez-Trillos A, Villamor N, et al. Transcriptome characterization by RNA sequencing identifies a major molecular and clinical subdivision in chronic lymphocytic leukemia. Genome Res. 2014;24:212–26.
29. Stilgenbauer S, Schnaiter A, Paschka P, Zenz T, Rossi M, Dohner K, et al. Gene mutations and treatment outcome in chronic lymphocytic leukemia: results from the CLL8 trial. Blood. 2014;123:3247–54.
30. Del Giudice I, Rossi D, Chiaretti S, Marinelli M, Tavolaro S, Gabrielli S, et al. NOTCH1 mutations in +12 chronic lymphocytic leukemia (CLL) confer an unfavorable prognosis, induce a distinctive transcriptional profiling and refine the intermediate prognosis of +12 CLL. Haematologica. 2012;97:437–41.
31. Rossi D, Khiabanian H, Spina V, Ciardullo C, Bruscaggin A, Fama R, et al. Clinical impact of small TP53 mutated subclones in chronic lymphocytic leukemia. Blood. 2014;123:2139–47.
32. Ojha J, Ayres J, Secreto C, Tschumper R, Rabe K, Van Dyke D, et al. Deep sequencing identifies genetic heterogeneity and recurrent convergent evolution in chronic lymphocytic leukemia. Blood. 2015;125:492–8.
33. Wang L, Shalek AK, Lawrence M, Ding R, Gaublomme JT, Pochet N, et al. Somatic mutation as a mechanism of Wnt/beta-catenin pathway activation in CLL. Blood. 2014;124:1089–98.
34. Messina M, Del Giudice I, Khiabanian H, Rossi D, Chiaretti S, Rasi S, et al. Genetic lesions associated with chronic lymphocytic leukemia chemo-refractoriness. Blood. 2014;123:2378–88.
35. Yao Z, Yaeger R, Rodrik-Outmezguine VS, Tao A, Torres NM, Chang MT, et al. Tumours with class 3 BRAF mutants are sensitive to the inhibition of activated RAS. Nature. 2017;548:234–8.
36. Heidorn SJ, Milagre C, Whittaker S, Nourry A, Niculescu-Duvas I, Dhomen N, et al. Kinase-dead BRAF and oncogenic RAS cooperate to drive tumor progression through CRAF. Cell. 2010;140:209–21.
37. Dahlman KB, Xia J, Hutchinson K, Ng C, Hucks D, Jia P, et al. BRAF(L597) mutations in melanoma are associated with sensitivity to MEK inhibitors. Cancer Discov. 2012;2:791–7.
38. Zhang C, Spevak W, Zhang Y, Burton EA, Ma Y, Habets G, et al. RAF inhibitors that evade paradoxical MAPK pathway activation. Nature. 2015;526:583–6.
39. Fama R, Bomben R, Rasi S, Dal Bo M, Ciardullo C, Monti S, et al. Ibrutinib-naive chronic lymphocytic leukemia lacks Bruton tyrosine kinase mutations associated with treatment resistance. Blood. 2014;124:3831–3.

Affiliations

Alexander C. Leeksma[1,2] · Justin Taylor[3] · Bian Wu[4,5] · Jeffrey R. Gardner[3] · Jie He[6] · Michelle Nahas[6] · Mithat Gonen[7] · Wendimagegn G. Alemayehu[8] · Doreen te Raa[1] · Tatjana Walther[5] · Jennifer Hüllein[5] · Sascha Dietrich[9] · Rainer Claus[10,11] · Fransien de Boer[12] · Koen de Heer[13] · Julie Dubois[1,2] · Maria Dampmann[14] · Jan Dürig[14] · Marinus H. J. van Oers[1] · Christian H. Geisler[15] · Eric Eldering[1,2] · Ross L. Levine[3] · Vincent Miller[6] ·

Tariq Mughal[6] · Nicole Lamanna[16] · Mark G. Frattini[16] · Mark L. Heaney[16] · Andrew Zelenetz[3] · Thorsten Zenz[17,5] · Omar Abdel-Wahab[3] · Arnon P. Kater[1,2]

[1] Department of Hematology and Lymphoma and Myeloma Center Amsterdam, Academic Medical Center Amsterdam, University of Amsterdam, Amsterdam, The Netherlands

[2] Department of Experimental Immunology, Cancer Center Amsterdam and Infection and Immunity Institute of Amsterdam, Amsterdam, The Netherlands

[3] Human Oncology and Pathogenesis Program and Leukemia Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA

[4] Cancer Center, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[5] Department of Molecular Therapy in Haematology & Oncology, German Cancer Research Center (dkfz) and National Center for Tumor Diseases (NCT), Heidelberg, Germany

[6] Foundation Medicine, Inc., Cambridge, MA, USA

[7] Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA

[8] HOVON Data Center, Erasmus MC, Daniel Den Hoedkliniek, Rotterdam, The Netherlands

[9] Department of Hematology, University Hospital Heidelberg, Heidelberg, Germany

[10] Department of Hematology, Oncology and Stem Cell Transplantation, University Medical Center, Faculty of Medicine, University of Freiburg, Freiburg, Germany

[11] Department of Haematology/Oncology, Augsburg Hospital, Augsburg, Germany

[12] Department of Internal medicine, Ikazia Hospital, Rotterdam, The Netherlands

[13] Department of Internal medicine, Flevo Hospital, Almere, The Netherlands

[14] Department of Haematology, University Hospital Essen, Essen, Germany

[15] Department of Hematology, Rigshospitalet, Copenhagen, Denmark

[16] Department of Medicine, Columbia University Medical Center, New York, USA

[17] Department of Haematology, University Hospital Zürich and University of Zürich, Zürich, Switzerland