

## Next-generation sequencing in chronic lymphocytic leukemia: recent findings and new horizons

Ana E. Rodríguez-Vicente<sup>1,2,\*</sup>, Vasilis Bikos<sup>3,\*</sup>, María Hernández-Sánchez<sup>2</sup>, Jitka Malcikova<sup>3,4</sup>, Jesús-María Hernández-Rivas<sup>2,5,6,#</sup> and Sarka Pospisilova<sup>3,4,#</sup>

<sup>1</sup>Department of Molecular and Clinical Pharmacology, University of Liverpool, Liverpool, United Kingdom

<sup>2</sup>IBSAL, IBMCC, Centro de Investigación del Cáncer, Universidad de Salamanca, CSIC, Hospital Universitario de Salamanca, Salamanca, Spain

<sup>3</sup>Central European Institute of Technology, Masaryk University, Brno, Czech Republic

<sup>4</sup>Department of Internal Medicine – Hematology and Oncology, Medical Faculty MU and University Hospital, Brno, Czech Republic

<sup>5</sup>Hematology Department, Hospital Universitario, Salamanca, Spain

<sup>6</sup>Department of Medicine, Universidad de Salamanca, Salamanca, Spain

\*These authors contributed equally to this work

#These authors share senior authorship

**Correspondence to:** Jesús-María Hernández-Rivas, **email:** jmhr@usal.es

**Keywords:** *chronic lymphocytic leukemia, next-generation sequencing, clonal evolution, immunogenetics, CLL prognosis*

**Received:** May 17, 2017

**Accepted:** July 12, 2017

**Published:** July 24, 2017

Copyright: Rodríguez-Vicente et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

### ABSTRACT

**The rapid progress in next-generation sequencing technologies has significantly contributed to our knowledge of the genetic events associated with the development, progression and treatment resistance of chronic lymphocytic leukemia patients. Together with the discovery of new driver mutations, next-generation sequencing has revealed an immense degree of both intra- and inter-tumor heterogeneity and enabled us to describe marked clonal evolution. Advances in immunogenetics may be implemented to detect minimal residual disease more sensitively and to track clonal B cell populations, their dynamics and molecular characteristics. The interpretation of these aspects is indispensable to thoroughly examine the genetic background of chronic lymphocytic leukemia. We review and discuss the recent results provided by the different next-generation sequencing techniques used in studying the chronic lymphocytic leukemia genome, as well as future perspectives in the methodologies and applications.**

### The emergence of next-generation sequencing in the field of the chronic lymphocytic leukemia

Chronic Lymphocytic Leukemia (CLL) is a disease that displays extreme clinical heterogeneity [1]. This heterogeneity reflexes CLL's marked molecular diversity [2], the analysis of which has led to the identification of a handful of biomarkers. Despite the significant progress in developing therapeutic options in CLL (extensively reviewed in [3, 4]), which has improved patient survival, CLL remains largely incurable and its course difficult to predict. Hence, the need to understand the disease's heterogeneous features is essential in order to characterize each CLL case and eventually proceed to more tailor-made therapeutic approaches. The technological advancements

emerging in the field of molecular analyses are providing the tools which will allow this to be done.

Next-generation sequencing technologies (NGS) have brought into play an unprecedented analytical depth to accommodate the characterization of the highly complicated genetic landscape of hematological cancers, with the case in point being CLL. Since the publication of the first complete cancer genome, obtained from a patient with acute myeloid leukemia [5], many groups have utilized NGS to further elucidate the biology of hematologic malignancies. These in turn have resulted in the assembly of large consortia such as the International Cancer Genome Consortium (ICGC) [6] and The Cancer Genome Atlas (TCGA) Research Network [7].

A glance back to the recent research on the elucidation of CLL genomics is enough to prove the usefulness of NGS technologies by virtue of the discovery of new drivers, including mutations in non-coding regions, and the elucidation of signaling pathways whose role was previously unknown or poorly understood [8–10].

Although CLL was originally considered a genetically stable disease, we now know that new genetic aberrations are acquired during clonal evolution and are characteristic of CLL development and relapse. NGS data have facilitated the detection of several tumoral subclones per sample and have illustrated different patterns of clonal evolution [2]. In addition, genetic loci can be sequenced in great depth and mutations existing at very low subclonal levels can be identified [11, 12]. The clones bearing these mutations can then be tracked during disease development and relapse after therapy.

Sequencing a specific genomic region in great depth also makes it possible to analyze the antibody repertoire more thoroughly, enabling a deeper understanding of the immune system with respect to the immunogenetic B cell features [13–15].

NGS allows us to explore the molecular pathways involved in disease pathogenesis and enables us to propose new genes which could be targeted for therapeutic purposes. Genes discovered using this approach may subsequently be included in prognostication gene panels, which in turn would ensure great coverage depth maximizing the results' clinical utility. Figure 1 summarizes the most prominent discoveries in CLL in this “NGS era” and describes how the different experimental set-ups can assist in finding answers to various scientific questions.

Here, we review the recent insights of the various NGS studies in CLL, paying particular attention to the use of this method to give us a profound understanding of the disease and its clinical applications.

## Applications of NGS to the study of the CLL genome

Whole-genome sequencing (WGS) and whole-exome sequencing (WES) data are now available for extended patient series with more than 800 published CLL exomes or genomes [2, 8–10, 16–21], Landau *et al.* and Puente *et al.* are amongst the studies providing the most comprehensive exploration of CLL's mutational landscape. These NGS approaches have revealed a vast genetic heterogeneity in CLL patients with a small number of genes mutated in approximately 10–15% of cases, and a large number of genes mutated at a lower frequency (< 10%) (intertumoral genetic heterogeneity). In parallel, as NGS enables the genome-wide detection of mutations between tumor cells, the genetic heterogeneity within malignant cells of the same patient (intratumoral heterogeneity) has also been accessed. Compared with solid tumors [22], CLL with a mutation rate of 0.60-0.87/Mb is a low genomic-complex

disease with an average of 15.3-26.9 somatic mutations per patient, according to Landau *et al.* [10] and Puente *et al.* [9], respectively. The early WGS/WES studies not only corroborated known CLL-associated alterations, such as somatic mutations in *TP53* and *ATM*, but also described a number of novel somatically mutated genes [8, 16, 17]. Amongst them, *NOTCH1* and *SF3B1* were identified as the most recurrently mutated genes with relatively higher frequencies than other candidates such as *MYD88*, *POT1*, *CHD2*, *XPO1*, *BIRC3*, *FBXW7* and *DDX3X*. Moreover, the most recent studies have increased the number of recurrently mutated genes, identifying previously unrecognized genes [9, 10]. It is noteworthy to mention some of them, such as *RPS15* [23], *EGR2* [24], *NFKBIE* [25] and *SETD2* [26]. The comparison between the two large cohort studies [9, 10] resulted in a significant overlap of 29 commonly mutated driver genes (Figure 2) albeit with existing discrepancies; the mutation frequencies were marked variable within common genes, and even some genes were exclusive of each study. These differences could be related to each cohort's clinical characteristics: Puente *et al.* [9] analyzed only untreated CLL samples whereas over 6% of the patients were treated before the sequencing in Landau *et al.* [10]. Another reason could be the different bioinformatics algorithms used in each study for variant calling as well as the definition of a mutated gene as a driver. It has been reported that the list of significantly mutated genes is different depending on the chosen computational method [27]. Incrementing the number of CLL cases analyzed by WGS or WES will identify even more significant mutated genes. Based on a saturation analysis and taking into account CLL background mutation rate, it has been estimated that an analysis of ~2000 samples would be enough to confidently identify recurrent mutated genes present in 1–2% of CLL patients [28].

Whereas WES provides information about coding DNA regions, WGS allows extensive detection of all abnormalities including non-coding regions. The application of WGS has led to the discovery of the most frequent recurrent non-coding mutation located in the 3' untranslated region of *NOTCH1* [9]. This splicing event is predicted to increase the stability of the NOTCH1 protein [9]. Furthermore, a small intergenic region of chromosome 9p13 was enriched for somatic mutations resulting in the reduced expression of the B-cell-specific transcription factor *PAX5* [9].

The large-scale comprehensive genetic characterization of CLL samples by WES and WGS studies has allowed us to better characterize the cell signaling pathways deregulated in CLL. Eight key cellular pathways (RNA metabolism, DNA damage, cell cycle control, apoptosis, NOTCH1 signaling, genome/chromatin structure, inflammatory pathway, B-cell signaling) have been described as recurrently mutated at different levels (Figure 2) [9, 10]. The emergence of NGS methodologies has also led to the discovery and extensive annotation

of CLL driver genes implicated in previously unknown CLL-related pathways such as RNA metabolism (*SF3B1*, *XPO1*, *RPS15*, *MGA*, *DDX3X*, *FUBP1*, *MED12*, *NXF1*). After an explosive growth in our understanding of CLL genetics, now it is time to assess the biological impact of the somatic mutations and how they affect cellular fitness. In this line, the biological relevance of some of them, such as mutations in *POT1*, *CHD2*, *MED12*, *SAMHD1*, has been further explored by functional analyses [29–32].

A particular pathway might be further analyzed in detail by targeted resequencing using a panel of chosen genes. Such a study involving the mutational screening of 18 core complex genes within the NF-κB pathway was performed in a cohort containing 315 cases and in concordance with the previous study, *NFKBIE* was found to be the most frequently mutated gene. It was further shown that the mutations were associated with inferior outcome [25]. On the same line, targeted NGS was employed on 9 genes related to the p53 pathway in a cohort of 180 CLL patients [33], and the identified mutations' impact on the transcriptional activation of p53 target genes was explored.

Genome-wide sequencing strategies have revealed novel candidate CLL driver genes within regions of copy number alterations. Apart from the known *TP53*/del(17p13) and *ATM-BIRC3*/del(11q22-23) associations, new regions have recently been identified, such as *SETD2*/del(3p21), *NFKB2*/del(10q24), *MGA*/del(15q15.1) [9, 10]. New structural variations have also been reported,

particularly structural rearrangements on chromosome 1, 2, 13, and 14 [9, 20]. Surprisingly, a small subset of CLL had complex rearrangements such as chromothripsis and chromoplexy [9, 20, 34].

An additional benefit to applying NGS approaches to the study of large cohorts is the enhanced power to explore the relationships between driver lesions based on the patterns of their co-occurrence. This analysis revealed significant relationships between several alterations, including not only the already known high co-occurrence of *NOTCH1* mutations/chromosome 12 trisomy, but also the previously undescribed high co-occurrence between mutations in *BIRC3* and trisomy 12, mutations in *SF3B1* and *POT1*, as well as mutations in *NOTCH1* and *MGA*. Other alterations, such as del(13q) with trisomy 12, co-occur less frequently [9, 10]. While interactions between genetic alterations could synergistically act to enhance tumor growth, lack of co-occurrence could indicate that alterations have highly similar downstream effects, and hence would lack further evolutionary advantage to the tumor cell.

### NGS to study clonal evolution in CLL

Genome-wide sequencing studies have also shed light on CLL pathogenesis. CLL is preceded by a pre-malignant state known as B-cell lymphocytosis (MBL). MBL was shown to carry mutations in some CLL drivers, and the existence of clonal evolution was associated with a shorter progression time to CLL [35, 36]. MBL cases

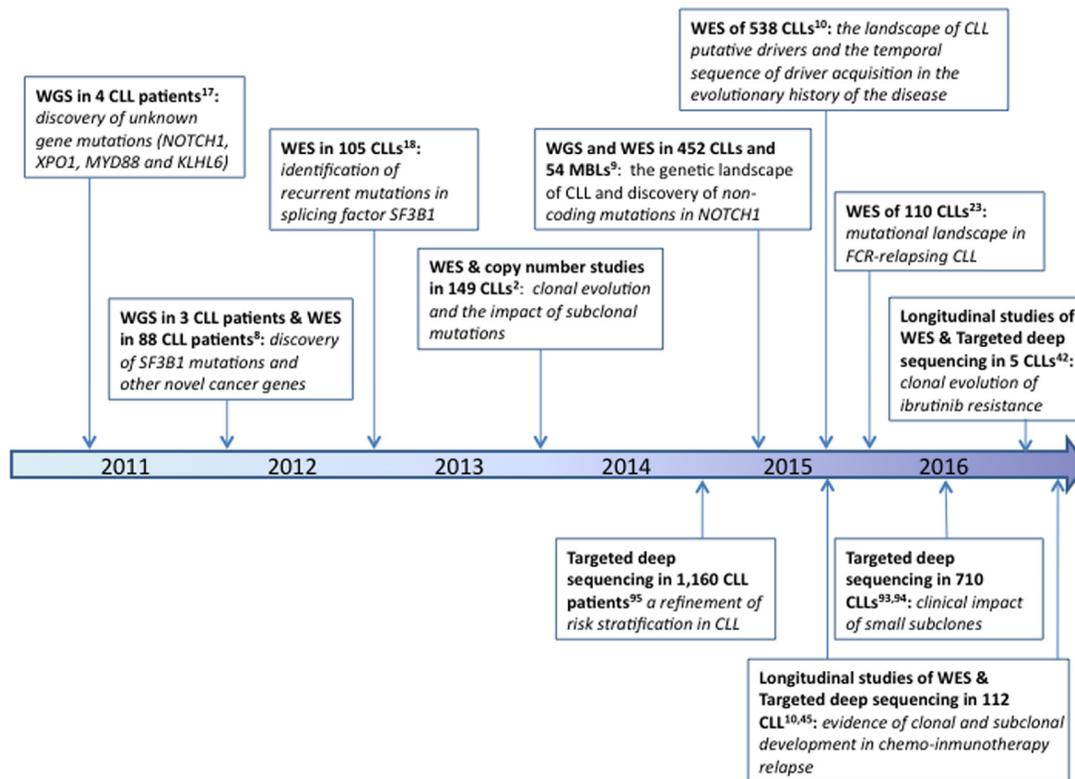
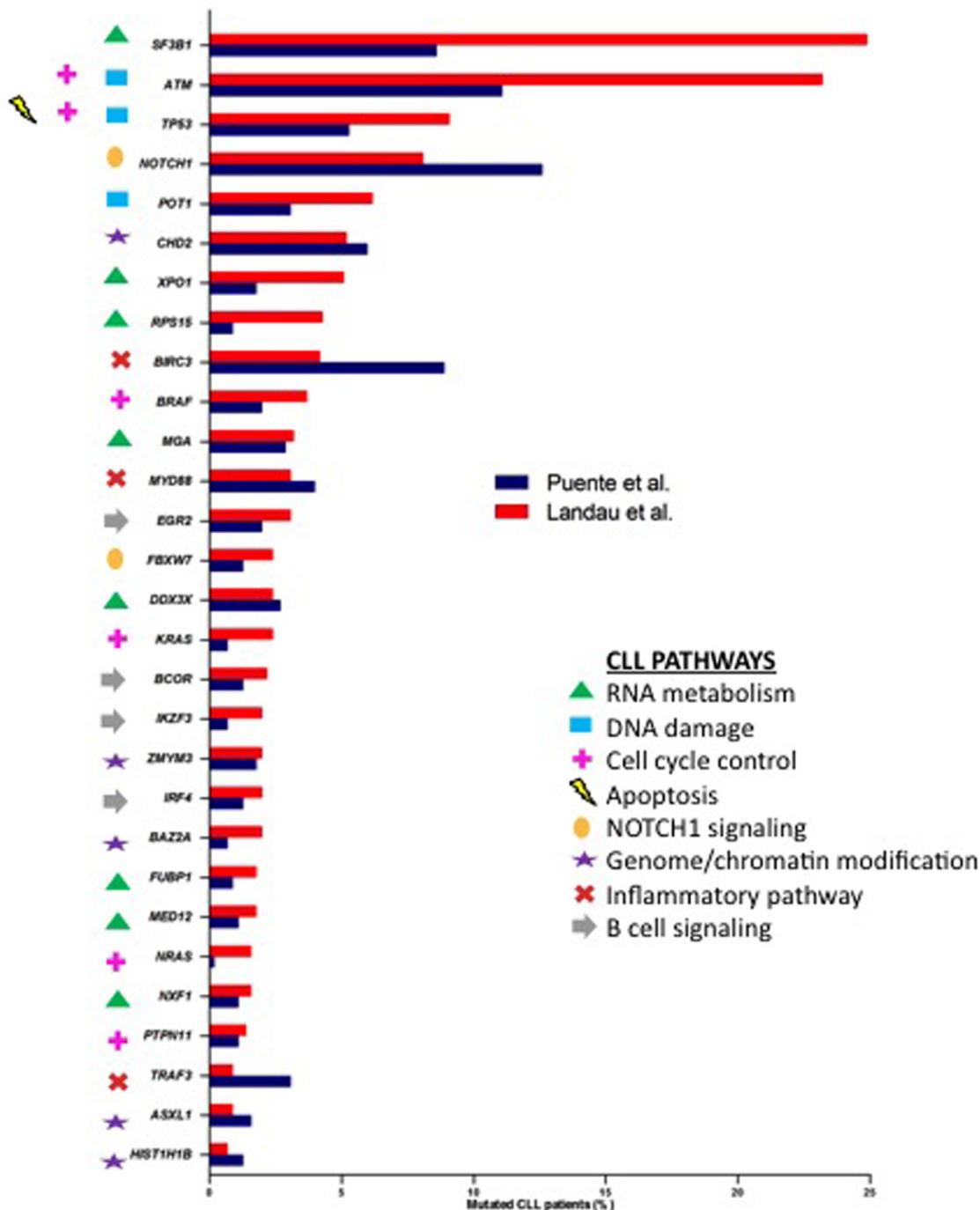


Figure 1: Timeline of the most relevant facts discovered in CLL by NGS studies.

have a lower driver alteration burden than CLL patients, consistent with a model in which MBL-to-CLL evolution is accomplished by the progressive accumulation of driver alterations [9]. Regarding the origin of this disease, further studies by NGS are still necessary to clarify the cell of origin since few recent studies have suggested that CLL genetic alterations are present not only in B lymphocytes but also in hematopoietic progenitors [37, 38].

NGS studies have revealed the conspicuous presence of multiple genetically defined subpopulations that fuel diverse evolution patterns. Primarily, two common clonal evolution patterns have been observed in CLL patients: linear evolution, in which one dominant clone acquires driver events over time, and branched evolution, in which several tumoral subclones coexist and evolve over time [2, 19]. In this line, WGS showed different temporal



**Figure 2: Percentages of samples affected by mutations in common CLL drivers from Puente *et al.* [9] (blue) and Landau *et al.* [10] (red) studies. Genes were marked with different symbols according to the biological pathways involved.**

repopulation patterns after therapy that deviate from a stable equilibrium among subpopulations, to marked shifts in which one minor subclone entirely replaced the dominant clone over time [19]. In a further investigation, WES data were analyzed using a computational approach to infer subclonal populations through integrating allelic fraction information about somatic mutations with local copy number and sample purity. This study demonstrated that genetic subclonal population features are linked to poorer clinical outcome, implying that an active evolutionary process is the underlying basis of an aggressive disease [2]. A temporal order of clonal and subclonal mutations corresponding to earlier and later events was suggested, in which the copy number variations such as 13q deletion and trisomy 12 were identified as consistently clonal early events while mutations in *ATM*, *BIRC3* or *TP53* were identified as subclonal later genetic alterations [2, 10]. Regarding biallelic inactivation of the *ATM*, a clonal del(11q) used to appear earlier than mutations in *ATM* which affects the remaining allele as a second hit [10].

The absence of an intervening therapy was largely associated with stable subclonal composition over time [39]. In contrast, chemotherapy exposure predominantly resulted in marked clonal evolution. WES data on matched samples collected at the time of first progression and relapse of fludarabine, cyclophosphamide, and rituximab (FCR) therapy have revealed large clonal shifts between pre-treatment and relapse samples in the majority of cases [10] and a recurrent mutation in *RPS15* in a large proportion of relapsed CLL cases [23]. Recently, WES was used to demonstrate that CLL patients with acquired resistance to ibrutinib therapy could not only have mutations in *BTK* at the binding site of ibrutinib or its immediate downstream partner PLCy2 [40] but also other potential alternative resistance mechanisms, such as del(8p) with additional driver mutations [41].

In order to monitor disease initiation, progression and relapse, it is important to define which clones are most biologically and clinically relevant. Targeted deep sequencing allows us to determine whether or not the driver mutations responsible for progression or relapse are already present in very small subclones at diagnosis. Longitudinal studies have already shown that the majority of driver mutations are already present in the initial stages, often as small subclones [10, 42–44]. Nevertheless, the clonal competition between the individual subclones carrying different mutations is a highly-complex process and mutations in the same gene can be selected in one patient and counter-selected in another [44]. The only consistent finding concerns *TP53*-mutated subclones, which were uniformly shown to increase upon relapse, suggesting a fitness advantage under therapeutic selective pressure [10, 43].

## Analysis of immunoglobulin genes using NGS

The importance of analyzing the molecular characteristics of the B cell receptor (BcR) in CLL patient management was shown by two seminal studies reporting that the immunoglobulin heavy chain (IGH) rearrangement's mutational load can be used as a prognosticator, distinguishing patients with a mutated BcR (M-CLL) and a more indolent disease course from CLL patients bearing unmutated clonotypic BcRs (U-CLL) [45, 46]. Subsequent studies shed light onto other BcR characteristics in CLL, such as VH-CDR3 stereotypy [47–51] as well as intraclonal diversification in a number of them [52, 53]. The SHM burden is one of the strongest prognosticators, but the disease heterogeneity is still present among the M-CLL and U-CLL major groups. The compartmentalization of CLL patients into subsets of stereotypy is a step towards the creation of more homogeneous prognostication groups. In fact, different subsets of stereotypy not only differ in terms of demographics, specific clinical observations and disease course [54–56] but also markedly differ in the presence or absence of certain genetic lesions (chromosomal or recurrent gene mutations). Cases in point are the high incidence of *SF3B1* mutations in the poor prognostic subset #2, or the asymmetric distribution of *NOTCH1* mutations in poor prognostic subsets in contrast to the high incidence of *MYD88* mutations in patients belonging to subsets characterized as prognostically good [57–61]. The advancements that massive parallel sequencing technologies bring to the field can assist in studying these crucial characteristics but also to more sensitively depict the immunogenetic background, study clonal diversity and dynamics and monitor the disease course [62, 63].

Introducing high-throughput sequencing technologies can improve disease monitoring by means of minimal residual disease (MRD) detection. Sensitivity in MRD detection is a critical issue, since half of the patients who have undergone allogeneic stem cell transplantation (allo-HCT) experience disease reoccurrence [64, 65], albeit associated with being long-term disease-free when the examination is negative for one year after allo-HCT [66, 67]. Logan *et al.* [68] reported achieving MRD sensitivity detection equal to 1:100,000 molecules. They concluded that disease relapses in cases with MRD positivity where the malignant clone was present in more than 1:100,000 sequences, less than one year after-HCT. Boyd *et al.* [69] successfully identified all clonal populations from several samples, to a level of 1:10,000 copies, and recapitulated the oligoclonal pattern of a CLL case, which was shown to exist by means of conventional sequencing and capillary electrophoresis analysis. Recently, a study was published by the European Initiative in CLL [70], where the ClonoSEQ assay for MRD detection was compared with flow cytometry

approaches and showed successful clone detection at the level of 1:10,000 copies.

NGS can also advance studies dealing with clonal diversity and intraclonal dynamics issues; Campbell *et al.* [71] attempted to track the evolution of the malignant clone by developing a bioinformatics algorithm to distinguish true somatic mutations from amplification and sequencing errors, based on the analysis of a non-polymorphic locus as a control, and reported unexpected intraclonal diversity not revealed by previous approaches. Gabriel *et al.* [72] attempted to recapitulate the immunogenetic profile in CLL patients. With a  $1.6 \times 10^{-4}$  detection sensitivity, they documented specific IG gene combinations with biased associations in CLL compared with healthy individuals and the intraclonal variation of the mutational burden.

A few studies dealt with the examination of multiple productive rearrangements in CLL. Kriangkum *et al.* [73] employed ImmunoSEQ, a multiplex PCR system to amplify CDR3 sequences using gDNA as a template in 26 CLL patients, to conclude that the allelic exclusion mechanism remained active in all cases examined, and that the presence of multiple clones is more frequent in M-CLL cases. Furthermore, they documented partner clones detected at a frequency greater than  $5 \times 10^9$  cells/L persisted over time despite treatment. More recently, Stamatopoulos *et al.* [74] employed the LymphoTrack IGH Somatic Hypermutation Assay Kit by Immunoscribe to show that the presence of multiple unrelated productive IGH rearrangements in CLL patients exceeds the so far estimated frequency, reaching a percentage of almost 25%. Moreover, based on the mutational status of these rearrangements, they categorized the patients into five different subgroups and claimed that this categorization improves disease stratification and patient management.

Finally, Blachly *et al.* [75] used next-generation RNA sequencing to compute the complete sequence of IG transcripts from unselected RNA-seq reads. With successful VH CDR3 region acquisition and the mutation status of the rearrangements along with a relative quantitation of oligoclonal samples they concluded that RNA-seq may simultaneously provide multidimensional data i.e. gene expression profile, mutation information and IGH Somatic hypermutation (SHM) status.

Sanger sequencing provided an essential image of the immunogenetic environment in CLL. The upgrade to the NGS era is still in its infancy; virtually all the studies so far are limited to analyzing only the heavy chain of the BcR. Although mutational status assessment and/or gene use in immunoglobulin light chain (*IGL*) rearrangement is not common practice, information about *IGH* and *IGL* associations could improve the research on CLL ontogeny, for example by strengthening the reliability of the functional analysis of antibody reactivity. Two individual groups attempted to fill that gap by combining single-cell PCR with next generation sequencing approaches to deep sequence pairs of IG heavy and light chains [76–79].

The 98% cut-off distinguishing M- and U-CLL was chosen arbitrarily, and borderline cases were always considered to be dealt with carefulness. NGS studies have documented the need for a more thorough classification [74], especially since there are suggestions for SHM analysis to be included in outline clinical tests [80]. This dictates the necessity to ensure a precise distinction between true somatic mutations and amplification/sequencing errors; the latter may generate artificial diversity and thus continuous efforts in building algorithms for error correction are taking place. However, such error corrections must be used with caution since they may inadvertently underestimate repertoire diversity by removing rare sequences [81]. A lab-bench advancement that may prove useful on this track, is the introduction of unique molecular identifiers (UMIs) in the sequenced fragments which will also assist a more detailed image of the clonal dynamics. To this end a methodology proposed by Khan *et al.* [82] and Turchaninova *et al.* [83] employed the introduction of UMIs at the reverse transcription level and applied asymmetrical sequencing on seven paired-end CLL libraries constructed from a specific number of cells/molecules in order for all molecules to be sequenced multiple times; they achieved full region sequencing with an estimated error rate of 1:10,000bp per UMI, since only molecules that were sequenced at least  $\times 5$  times were evaluated.

Other issues such as the low heterogeneity of IG sequences that can affect cluster formation and the biases introduced to PCR by the fact that some molecules are amplified more than others are currently under the spotlight. These issues and experimental set-ups to work around them are well-described and discussed by Georgiou *et al.* [84] At the same time the Euroclonality-NGS consortium has been established, with its primary objective being the development and standardization of NGS assays for clonality assessment, MRD analysis and repertoire analysis at the pre-analytical (wet-based) and post-analytical (bioinformatics) level.

## Clinical utilization of NGS in CLL

NGS techniques have provided a better knowledge of the genetic complexity underlying the clinical heterogeneity of CLL. Most findings indicate a prognostic role for some of the new gene mutations and emerging data suggest their role in predicting a therapeutic response and tracking disease relapse (reviewed in [85]) (Table 1). New genomic developments could improve patient outcome in two ways: (1) allowing for a more precise prognostic estimation and, (2) identifying novel therapeutic targets. For the use of prognostication in the clinical setting, targeted NGS is particularly suitable as it allows the concurrent evaluation of chosen genes, or parts of genes that are relevant to a given disease [86]. In addition, it is now possible to combine mutation analysis with copy-number analysis, thereby enabling the simultaneous

**Table 1: Clinical relevance of the most important mutated genes in CLL patients studied by NGS**

Gene mutation	Frequency	Association with	Prognosis	References
<i>TP53</i>	5.3–50%*	Del (17p13) UM- <i>IGHV</i>	Poor	2, 8, 17, 21, 23, 91, 93, 95
<i>ATM</i>	9–23.2%	Del (11q23) UM- <i>IGHV</i>	Poor	2, 8, 17, 23, 91, 93
<i>RPS15</i>	0.88–19.5%	UM- <i>IGHV</i>	Poor	9, 10, 21, 23, 45
<i>BIRC3</i>	4–15.5%	Del (11q22-q23) UM- <i>IGHV</i>	Poor	2, 16, 17, 21, 91, 93
<i>NOTCH1</i>	4–24.1%	Trisomy 12 UM- <i>IGHV</i>	Poor	11, 17, 18, 21, 23, 45, 91, 93
<i>SF3B1</i>	5–24.9%	Del (11q22-q23) ATM mutations UM- <i>IGHV</i> ZAP70 expression	Poor	2, 10, 17, 18, 21, 23, 45, 91, 93, 95
<i>MYD88</i>	1.5–10%	Del (13q14) M- <i>IGHV</i>	Good	2, 8, 10, 16–18, 45, 91, 95
<i>XPO1</i>	1.8–11.5%	UM- <i>IGHV</i>	Poor?	2, 10, 17, 18, 45, 91, 95
<i>FBXW7</i>	1.3–5%	Trisomy 12	Poor	2, 8, 10, 17, 45, 95
<i>SAMHD1</i>	2–9.8%	LOH on chromosome 20	Poor?	2, 10, 29, 45, 91

\*At the time of diagnosis, *TP53* mutations are reported in approximately 5% of patients. On the other hand, in patients with relapsed and refractory CLL, the prevalence can be up to 50% of patients.

detection of known copy-number aberrations. Several commercially available gene panels have been designed to detect either somatic or germline mutations in tumor-associated genes. As for CLL, there is currently only one commercially available CLL-specific gene panel. This amplicon-based panel allows single nucleotide as well as copy number variant analysis for the nine most prognostically relevant genes (Table 2). The efforts to design an optimal gene panel that could be widely used for CLL prognostication and, more importantly, treatment-effect prediction, are still on-going. So far, several groups have published results for studies employing a custom prognostic gene panel intended for clinical utilization (Table 2) [87–92]. Some of the recurrently mutated genes were common to all gene panels (*TP53*, *NOTCH1*, *SF3B1*), while others were found in one study only. The first study, reported by Jethwa *et al.* in 2013, used the designed NGS gene panel to study intratumor heterogeneity and clonal evolution, and documented the presence of convergent mutations [87]. Two other studies explored the suitability of NGS gene panels for diagnostic purposes [88, 89]. Guieze *et al.* [90] used their custom NGS gene panel to characterize relapsed/refractory CLL patients. They reported that concurrent gene mutations (“multiple-hit”) are frequent in patients with relapsed/refractory CLL and are associated with worse outcome. Rigolin *et al.* explored the co-occurrence of mutations and other clinical and biological characteristics in patients analyzed at diagnosis. They confirmed a negative impact of multi-hit profile and they reported association of mutation occurrence with adverse molecular and genetic findings including *IGVH* unmutated status, high-risk FISH results, and the presence of a complex karyotype [92]. Nadeu *et al.* used a gene panel to study the clinical impact and clonal evolution of subclonal mutations [91]. The role of small subclonal mutations is a timely question and the study by Nadeu *et al.* together with a report by Rasi *et al.* [93] consistently documented that only subclonal mutations in *TP53* and not *ATM*, *BIRC3*, *SF3B1* or *NOTCH1* genes

are of prognostic importance, thus further confirming the clinical significance of the small *TP53*-defective subclones suggested by previous studies focusing only on *TP53* gene [12, 94]. It is therefore now discussed whether patients with small *TP53*-defective subclones should also be appointed for non-chemotherapeutic treatment. This consideration is extremely important with respect to novel targeted therapies that appear to dramatically improve the outcome of CLL patients with *TP53* defects. To elucidate this issue, a prospective study conducted on a uniformly treated patient series is desirable.

Prospective multi-institution NGS studies of larger patient cohorts will allow us to integrate the newly discovered genetic lesions into a comprehensive prognostic model based on chromosomal abnormalities, gene mutations and also *IGHV* mutational status. This might improve patient prognostication and facilitate the introduction of testing into general practice. Efforts to devise an integrated prognostic model have already been made [95–97], although with partially discordant results, which may be related to differences in the design and analysis of the studies. Moreover, it is becoming evident that assessing the mutational complexity might be more informative than dissecting a specific prognosis for each driver as it was reported that cumulative number of driver mutations per tumor had a progressively worse effect [9, 90, 92].

Despite the still-widening list of prognostically significant genes, the only biomarker that influences the treatment strategy remains the presence of *TP53* defects [98], and to date, there has not been enough evidence to support the use of *NOTCH1*, *SF3B1*, *BIRC3* or *ATM* mutations, or CLL subclonal composition in making clinical decisions. Prospective trials with larger series of patients are needed to stratify CLL patients based on these mutations, in order to confirm their clinical impact and to address the possible use of these markers to adapt treatment.

Recent works have tried to recapitulate the biological consequences of some of the described mutations by

**Table 2: Summary of CLL-specific prognostic gene panel designed by different studies**

Study	Jethwa et al 2013 <sup>87</sup>	Sutton et al 2015 <sup>88</sup>	Vollbrecht et al 2015 <sup>89</sup>	Guize et al 2015 <sup>90</sup>	Nadeu et al 2016 <sup>91</sup>	Rigolin et al 2016 <sup>92</sup>	Commercially available panel		
Number of patients analyzed	178	188	136	114	406	200			
Library preparation method	Amplicon - multiplex PCR	HaloPlex (Agilent)	Amplicon - multiplex PCR	TruSeq Custom amplicon panel (Illumina)	Access-Array system, (Fluidigm)	Agilent HaloPlex Target Enrichment kit	Multiplex PCR (Multiplicom)		
Sequencing platform	GS Junior 454 (Roche)	HiSeq (Illumina)	MiSeq (Illumina)	MiSeq (Illumina)	MiSeq (Illumina)	Ion Torrent PGM (Life technologies)	MiSeq (Illumina), Ion PGM (Life Technologies)		
Limit of detection achieved	3%	10%	5%	8%, 30%*	0,3%	5%	Depending on samples per run analyzed		
Total number of genes	9	9	15	9	5	20	9		
Genes and regions covered	Genes included in more than one study	<i>ATM</i>	—	all coding exons	all coding exons	not specified	exons 2–63	exonic regions	all coding exons
		<i>BIRC3</i>	—	all coding exons	—	not specified	exons 2–9	exonic regions	all coding exons
		<i>MYD88</i>	exons 3, 5	all coding exons	all coding exons	not specified		exonic regions	all coding exons
		<i>NOTCH1</i>	exon 34	all coding exons	all coding exons	not specified	exons 26, 27, 34 and 3'UTR	exonic regions	all coding exons
		<i>PIK3CA</i>	exons 9, 20	—	exons 9–11, 20–21	—	—	exonic regions	—
		<i>SF3B1</i>	exons 14, 15	all coding exons	all coding exons	not specified	exons 14–16 and 18	exonic regions	all coding exons
		<i>TP53</i>	exons 4–10	all coding exons	all coding exons	not specified	exons 4–10	exonic regions	all coding exons
		<i>XPO1</i>	—	all coding exons	exons 12–13, 15,	not specified		exonic regions	all coding exons
	Other genes	<i>BRAF, EZH2, KRAS, NRAS</i> (selected gene regions)	<i>KLHL6, POT1</i> (all coding exons)	<i>BTK, CD79B, DDX3X, FBXW7, MAPK1, PIK3CD, PTEN, PTPN6</i> (selected gene regions)	<i>MED12, SAMDHI</i> (gene region not specified)	—	<i>BRAF, CDKN2A, PTEN, CDH2, DDX3X, FBXW7, KIT, KLHL6, KRAS, NRAS, POT1, ZMYM3</i> (exonic regions)	<i>FBXW7, POT1</i> (all coding exons)	

\*Depending on the sequencing depth

grouping together genes belonging to the same biological pathways, as the genetic activation or silencing of a pathway can occur on multiple levels, with a relatively consistent effect on the output of the overall pathway [23, 25, 33, 99]. The presence of recurrent mutations affecting the activity of relevant cellular signaling pathways could indicate a key role for these mutations in CLL, and targeting these genes and pathways could provide effective therapy. Future directions should go into understanding the pathway map, which could lead to successful targets, as it can be witnessed currently in the clinical use of BCR signaling inhibitors [100]. Whilst there are currently no agents in CLL clinical trials to target *NOTCH1* and *SF3B1*, they are clearly promising targets and a number of agents targeting these genes and pathways are under development [101, 102]. *NOTCH1* is a well-established therapeutic target in T-cell acute lymphoblastic leukemia (T-ALL), while it is possible that spliceosome inhibitors could have activity against *SF3B1* defects.

**Conclusions and future directions**

NGS approaches have expanded our knowledge of genetic changes in CLL and leukemic clonal architecture, and this could have important consequences for prognosis and the optimized management of patients. CLL’s mutational profile is characterized by a relatively small number of somatic mutations per patient, a few recurrent mutations at moderate frequency, and a long list of

recurrent low-frequency mutated genes. The abundance of these genomic aberrations in CLL has illustrated its huge biological heterogeneity. In fact, these genetic lesions are present not only in coding regions but also in non-coding regions. The advent of NGS has also allowed us to further refine the prognostic impact of *IGHV* rearrangement, enabling the identification of subclones with different *IGHV* rearrangements which would otherwise not be possible with Sanger sequencing, considerably less sensitive.

Technically, the NGS studies have proved the applicability, sensitivity and reproducibility of targeted sequencing approaches in CLL. However, several aspects need to be resolved before NGS can be implemented as part of routine diagnostics. Firstly, bioinformatics algorithms for data analysis represent an important obstacle. At present, multiple commercial as well as free software tools are available and many laboratories have developed their own in-house methodologies. Therefore, the inter-reproducibility of obtained results is an issue to be solved, particularly in the case of low-level subclonal mutations, and studies comparing the results obtained by different approaches are warranted and ongoing. Secondly, there is currently no agreement on the final list of genes to be included in the CLL-specific prognostication panel. The studies evaluating the prognostic impact of mutations in particular genes are to some extent contradictory, moreover, new potential driver mutations are continually being discovered and the gene selection may change with novel targeted therapies as the

new mechanisms of resistance to these therapies have already been emerging.

The exploitation of NGS in prognostication and treatment prediction seems likely to become widespread in the near future; however, despite the rapid progress, significant work on the validation of new marker optimization and harmonizing the process between laboratories is still required before it can be used routinely in a clinical setting.

## ACKNOWLEDGMENTS

We thank Matthew Smith and Francesco Muto for language editing.

## CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

## FUNDING

This work received funding from the European Union Seventh Framework Programme [FP7/2007-2013] under Grant Agreement n°306242-NGS-PTL and MSMT CR (2013–2015/no.7E13008). It was also supported by grants from the Spanish Fondo de Investigaciones Sanitarias FIS 09/01543, PI12/00281 and PI15/01471, Proyectos de Investigación del SACYL 355/A/09, COST Action EuGESMA (BM0801), Fundación Manuel Solórzano, Obra Social Banca Cívica (Caja Burgos), Fundación Española de Hematología y Hemoterapia (FEHH), and by grants (RD12/0036/0069 and RD12/0036/0044) from the Red Temática de Investigación Cooperativa en Cáncer (RTICC), Instituto de Salud Carlos III (ISCIII), Spanish Ministry of Economy and Competitiveness and European Regional Development Fund (ERDF) “Una manera de hacer Europa” (CEI 2010-1-0010).

AERV is supported by a grant from the Fundación Ramón Areces. MHS is fully supported by an Ayuda Predoctoral de la Junta de Castilla y León from the Fondo Social Europeo (JCYL-EDU/346/2013 PhD scholarship).

SP, VB and JM were supported by AZV grants from the Ministry of Health of the Czech Republic No. 15-31834A, 15-30015A and 16-29447A, MH CR project 65269705 and MSMT CR project CEITEC2020 LQ1601 under the National Sustainability Programme II.

## REFERENCES

1. Shanafelt TD, Geyer SM, Kay NE. Prognosis at diagnosis: integrating molecular biologic insights into clinical practice for patients with CLL. *Blood*. 2004; 103:1202–10. <https://doi.org/10.1182/blood-2003-07-2281>.
2. Landau DA, Carter SL, Stojanov P, McKenna A, Stevenson K, Lawrence MS, Sougnez C, Stewart C, Sivachenko A, Wang L, Wan Y, Zhang W, Shukla SA, et al. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell*. 2013; 152:714–26. <https://doi.org/10.1016/j.cell.2013.01.019>.
3. Barrientos JC. Sequencing of chronic lymphocytic leukemia therapies. *Hematology Am Soc Hematol Educ Program*. 2016; 2016:128–36. <https://doi.org/10.1182/asheducation-2016.1.128>.
4. Robak T, Stilgenbauer S, Tedeschi A. Front-line treatment of CLL in the era of novel agents. *Cancer Treat Rev*. 2016; 53:70–8. <https://doi.org/10.1016/j.ctrv.2016.12.007>.
5. Ley TJ, Mardis ER, Ding L, Fulton B, McLellan MD, Chen K, Dooling D, Dunford-Shore BH, McGrath S, Hickenbotham M, Cook L, Abbott R, Larson DE, et al. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature*. 2008; 456:66–72. <https://doi.org/10.1038/nature07485>.
6. International Cancer Genome C, Hudson TJ, Anderson W, Artez A, Barker AD, Bell C, Bernabe RR, Bhan MK, Calvo F, Eerola I, Gerhard DS, Guttmacher A, Guyer M, et al. International network of cancer genome projects. *Nature*. 2010; 464:993–8. <https://doi.org/10.1038/nature08987>.
7. Tomczak K, Czerwinska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol (Pozn)*. 2015; 19:A68–77. <https://doi.org/10.5114/wo.2014.47136>.
8. Wang L, Lawrence MS, Wan Y, Stojanov P, Sougnez C, Stevenson K, Werner L, Sivachenko A, DeLuca DS, Zhang L, Zhang W, Vartanov AR, Fernandes SM, et al. SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. *N Engl J Med*. 2011; 365:2497–506. <https://doi.org/10.1056/NEJMoa1109016>.
9. Puente XS, Bea S, Valdes-Mas R, Villamor N, Gutierrez-Abril J, Martin-Subero JI, Munar M, Rubio-Perez C, Jares P, Aymerich M, Baumann T, Beekman R, Belver L, et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature*. 2015; 526:519–524. <https://doi.org/10.1038/nature14666>.
10. Landau DA, Tausch E, Taylor-Weiner AN, Stewart C, Reiter JG, Bahlo J, Kluth S, Bozic I, Lawrence M, Bottcher S, Carter SL, Cibulskis K, Mertens D, et al. Mutations driving CLL and their evolution in progression and relapse. *Nature*. 2015; 526:525–530. <https://doi.org/10.1038/nature15395>.
11. Weissmann S, Roller A, Jeromin S, Hernandez M, Abaigar M, Hernandez-Rivas JM, Grossmann V, Haferlach C, Kern W, Haferlach T, Schnittger S, Kohlmann A. Prognostic impact and landscape of NOTCH1 mutations in chronic lymphocytic leukemia (CLL): a study on 852 patients. *Leukemia*. 2013; 27:2393–6. <https://doi.org/leu2013218>.
12. Malcikova J, Stano-Kozubik K, Tichy B, Kantorova B, Pavlova S, Tom N, Radova L, Smardova J, Pardy F, Doubek M, Brychtova Y, Mraz M, Plevova K, et al. Detailed analysis of therapy-driven clonal evolution of TP53 mutations in chronic lymphocytic leukemia. *Leukemia*. 2015; 29:877–85. <https://doi.org/10.1038/leu.2014.297>.

13. Bashford-Rogers RJ, Palser AL, Huntly BJ, Rance R, Vassiliou GS, Follows GA, Kellam P. Network properties derived from deep sequencing of human B-cell receptor repertoires delineate B-cell populations. *Genome Res.* 2013; 23:1874–84. <https://doi.org/10.1101/gr.154815.113>.
14. Glanville J, Zhai W, Berka J, Telman D, Huerta G, Mehta GR, Ni I, Mei L, Sundar PD, Day GM, Cox D, Rajpal A, Pons J. Precise determination of the diversity of a combinatorial antibody library gives insight into the human immunoglobulin repertoire. *Proc Natl Acad Sci USA.* 2009; 106: 20216–21. <https://doi.org/10.1073/pnas.0909775106>.
15. Weinstein JA, Jiang N, White RA 3rd, Fisher DS, Quake SR. High-throughput sequencing of the zebrafish antibody repertoire. *Science.* 2009; 324:807–10. <https://doi.org/10.1126/science.1170020>.
16. Fabbri G, Rasi S, Rossi D, Trifonov V, Khiabani H, Ma J, Grunn A, Fangazio M, Capello D, Monti S, Cresta S, Gargiulo E, Forconi F, et al. Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation. *J Exp Med.* 2011; 208:1389–401. <https://doi.org/10.1084/jem.20110921>.
17. Puente XS, Pinyol M, Quesada V, Conde L, Ordonez GR, Villamor N, Escaramis G, Jares P, Bea S, Gonzalez-Diaz M, Bassaganyas L, Baumann T, Juan M, et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature.* 2011; 475:101–5. <https://doi.org/10.1038/nature10113>.
18. Quesada V, Conde L, Villamor N, Ordonez GR, Jares P, Bassaganyas L, Ramsay AJ, Bea S, Pinyol M, Martinez-Trillos A, Lopez-Guerra M, Colomer D, Navarro A, et al. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *NatGenet.* 2011; 44:47–52.
19. Schuh A, Becq J, Humphray S, Alexa A, Burns A, Clifford R, Feller SM, Grocock R, Henderson S, Khrebtukova I, Kingsbury Z, Luo S, McBride D, et al. Monitoring chronic lymphocytic leukemia progression by whole genome sequencing reveals heterogeneous clonal evolution patterns. *Blood.* 2012; 120:4191–6. <https://doi.org/10.1182/blood-2012-05-433540>.
20. Kasar S, Kim J, Improgo R, Tiao G, Polak P, Haradhvala N, Lawrence MS, Kiezun A, Fernandes SM, Bahl S, Sougnez C, Gabriel S, Lander ES, et al. Whole-genome sequencing reveals activation-induced cytidine deaminase signatures during indolent chronic lymphocytic leukaemia evolution. *Nat Commun.* 2015; 6:8866. <https://doi.org/10.1038/ncomms9866>.
21. Messina M, Del Giudice I, Khiabani H, Rossi D, Chiaretti S, Rasi S, Spina V, Holmes AB, Marinelli M, Fabbri G, Piciocchi A, Mauro FR, Guarini A, et al. Genetic lesions associated with chronic lymphocytic leukemia chemo-refractoriness. *Blood.* 2014; 123:2378–88. <https://doi.org/10.1182/blood-2013-10-534271>.
22. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, Carter SL, Stewart C, Mermel CH, Roberts SA, Kiezun A, Hammerman PS, McKenna A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature.* 2013; 499:214–8. <https://doi.org/10.1038/nature12213>.
23. Ljungstrom V, Cortese D, Young E, Pandzic T, Mansouri L, Plevova K, Ntoufa S, Baliakas P, Clifford R, Sutton LA, Blakemore S, Stavroyianni N, Agathangelidis A, et al. Whole-exome sequencing in relapsing chronic lymphocytic leukemia: clinical impact of recurrent RPS15 mutations. *Blood.* 2015. <https://doi.org/10.1182/blood-2015-10-674572>.
24. Young E, Noerenberg D, Mansouri L, Ljungstrom V, Frick M, Sutton LA, Blakemore SJ, Galan-Sousa J, Plevova K, Baliakas P, Rossi D, Clifford R, Roos-Weil D, et al. EGR2 mutations define a new clinically aggressive subgroup of chronic lymphocytic leukemia. *Leukemia.* 2017. <https://doi.org/10.1038/leu.2016.359>.
25. Mansouri L, Sutton LA, Ljungstrom V, Bondza S, Arngarden L, Bhoi S, Larsson J, Cortese D, Kalushkova A, Plevova K, Young E, Gunnarsson R, Falk-Sorqvist E, et al. Functional loss of IkappaBepsilon leads to NF-kappaB deregulation in aggressive chronic lymphocytic leukemia. *J Exp Med.* 2015; 212: 833–43. <https://doi.org/10.1084/jem.20142009>.
26. Parker H, Rose-Zerilli MJ, Larrayoz M, Clifford R, Edelmann J, Blakemore S, Gibson J, Wang J, Ljungstrom V, Wojdacz TK, Chaplin T, Roghanian A, Davis Z, et al. Genomic disruption of the histone methyltransferase SETD2 in chronic lymphocytic leukaemia. *Leukemia.* 2016; 30:2179–86. <https://doi.org/10.1038/leu.2016.134>.
27. Tamborero D, Gonzalez-Perez A, Perez-Llamas C, Deu-Pons J, Kandoth C, Reimand J, Lawrence MS, Getz G, Bader GD, Ding L, Lopez-Bigas N. Comprehensive identification of mutational cancer driver genes across 12 tumor types. *Sci Rep.* 2013; 3:2650. <https://doi.org/10.1038/srep02650>.
28. Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, Meyerson M, Gabriel SB, Lander ES, Getz G. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature.* 2014; 505:495–501. <https://doi.org/10.1038/nature12912>.
29. Clifford R, Louis T, Robbe P, Ackroyd S, Burns A, Timbs AT, Wright Colopy G, Dreau H, Sigaux F, Judde JG, Rotger M, Telenti A, Lin YL, et al. SAMHD1 is mutated recurrently in chronic lymphocytic leukemia and is involved in response to DNA damage. *Blood.* 2014; 123:1021–31. <https://doi.org/blood-2013-04-490847>.
30. Kampjarvi K, Jarvinen TM, Heikkinen T, Ruppert AS, Senter L, Hoag KW, Dufva O, Kontro M, Rassenti L, Hertlein E, Kipps TJ, Porkka K, Byrd JC, et al. Somatic

- MED12 mutations are associated with poor prognosis markers in chronic lymphocytic leukemia. *Oncotarget*. 2015; 6:1884–8. <https://doi.org/10.18632/oncotarget.2753>.
31. Ramsay AJ, Quesada V, Foronda M, Conde L, Martinez-Trillos A, Villamor N, Rodriguez D, Kwarciak A, Garabaya C, Gallardo M, Lopez-Guerra M, Lopez-Guillermo A, Puente XS, et al. POT1 mutations cause telomere dysfunction in chronic lymphocytic leukemia. *Nat Genet*. 2013; 45: 526–30. <https://doi.org/10.1038/ng.2584>.
  32. Rodriguez D, Bretones G, Quesada V, Villamor N, Arango JR, Lopez-Guillermo A, Ramsay AJ, Baumann T, Quiros PM, Navarro A, Royo C, Martin-Subero JI, Campo E, et al. Mutations in CHD2 cause defective association with active chromatin in chronic lymphocytic leukemia. *Blood*. 2015; 126: 195–202. <https://doi.org/10.1182/blood-2014-10-604959>.
  33. Athanasakis E, Melloni E, Rigolin GM, Agnoletto C, Voltan R, Vozzi D, Piscianz E, Segat L, Dal Monego S, Cuneo A, Secchiero P, Zauli G. The p53 transcriptional pathway is preserved in ATM mutated and NOTCH1 mutated chronic lymphocytic leukemias. *Oncotarget*. 2014; 5: 12635–45. <https://doi.org/10.18632/oncotarget.2211>.
  34. Bassaganyas L, Bea S, Escaramis G, Tornador C, Salaverria I, Zapata L, Drechsel O, Ferreira PG, Rodriguez-Santiago B, Tubio JM, Navarro A, Martin-Garcia D, Lopez C, et al. Sporadic and reversible chromothripsis in chronic lymphocytic leukemia revealed by longitudinal genomic analysis. *Leukemia*. 2015; 29:758. <https://doi.org/10.1038/leu.2014.309>.
  35. Barrio S, Shanafelt TD, Ojha J, Chaffee KG, Secreto C, Kortum KM, Pathangey S, Van-Dyke DL, Slager SL, Fonseca R, Kay NE, Braggio E. Genomic characterization of high-count MBL cases indicates that early detection of driver mutations and subclonal expansion are predictors of adverse clinical outcome. *Leukemia*. 2017; 31:170–6. <https://doi.org/10.1038/leu.2016.172>.
  36. Ojha J, Secreto C, Rabe K, Ayres-Silva J, Tschumper R, Dyke DV, Slager S, Fonseca R, Shanafelt T, Kay N, Braggio E. Monoclonal B-cell lymphocytosis is characterized by mutations in CLL putative driver genes and clonal heterogeneity many years before disease progression. *Leukemia*. 2014; 28:2395–8. <https://doi.org/10.1038/leu.2014.226>.
  37. Damm F, Mylonas E, Cosson A, Yoshida K, Della Valle V, Mouly E, Diop M, Scourzic L, Shiraishi Y, Chiba K, Tanaka H, Miyano S, Kikushige Y, et al. Acquired initiating mutations in early hematopoietic cells of CLL patients. *Cancer Discov*. 2014; 4:1088–101. <https://doi.org/10.1158/2159-8290.CD-14-0104>.
  38. Quijada-Alamo M, Hernandez-Sanchez M, Robledo C, Hernandez-Sanchez JM, Benito R, Montano A, Rodriguez-Vicente AE, Quwaider D, Martin AA, Garcia-Alvarez M, Vidal-Mancenido MJ, Ferrer-Garrido G, Delgado-Beltran MP, et al. Next-generation sequencing and FISH studies reveal the appearance of gene mutations and chromosomal abnormalities in hematopoietic progenitors in chronic lymphocytic leukemia. *J Hematol Oncol*. 2017; 10:83. <https://doi.org/10.1186/s13045-017-0450-y>.
  39. Smith EN, Ghia EM, DeBoever CM, Rassenti LZ, Jepsen K, Yoon KA, Matsui H, Rozenzhak S, Alakus H, Shepard PJ, Dai Y, Khosroheidari M, Bina M, et al. Genetic and epigenetic profiling of CLL disease progression reveals limited somatic evolution and suggests a relationship to memory-cell development. *Blood Cancer J*. 2015; 5:e303. <https://doi.org/10.1038/bcj.2015.14>.
  40. Woyach JA, Furman RR, Liu TM, Ozer HG, Zapotka M, Ruppert AS, Xue L, Li DH, Steggerda SM, Versele M, Dave SS, Zhang J, Yilmaz AS, et al. Resistance mechanisms for the Bruton's tyrosine kinase inhibitor ibrutinib. *N Engl J Med*. 2014; 370:2286–94. <https://doi.org/10.1056/NEJMoa1400029>.
  41. Burger JA, Landau DA, Taylor-Weiner A, Bozic I, Zhang H, Sarosiek K, Wang L, Stewart C, Fan J, Hoellenriegel J, Sivina M, Dubuc AM, Fraser C, et al. Clonal evolution in patients with chronic lymphocytic leukaemia developing resistance to BTK inhibition. *Nat Commun*. 2016; 7:11589. <https://doi.org/10.1038/ncomms11589>.
  42. Rose-Zerilli MJ, Gibson J, Wang J, Tapper W, Davis Z, Parker H, Larrayoz M, McCarthy H, Walewska R, Forster J, Gardiner A, Steele AJ, Chelala C, 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. <https://doi.org/10.1038/leu.2016.10>.
  43. Amin NA, Seymour E, Saiya-Cork K, Parkin B, Shedden K, Malek SN. A Quantitative Analysis of Subclonal and Clonal Gene Mutations before and after Therapy in Chronic Lymphocytic Leukemia. *Clin Cancer Res*. 2016; 22:4525–35. <https://doi.org/10.1158/1078-0432.CCR-15-3103>.
  44. Ojha J, Ayres J, Secreto C, Tschumper R, Rabe K, Van Dyke D, Slager S, Shanafelt T, Fonseca R, Kay NE, Braggio E. Deep sequencing identifies genetic heterogeneity and recurrent convergent evolution in chronic lymphocytic leukemia. *Blood*. 2015; 125:492–8. <https://doi.org/10.1182/blood-2014-06-580563>.
  45. Damle RN, Wasil T, Fais F, Ghiotto F, Valetto A, Allen SL, Buchbinder A, Budman D, Dittmar K, Kolitz J, Lichtman SM, Schulman P, Vinciguerra VP, et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood*. 1999; 94:1840–7.
  46. Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood*. 1999; 94:1848–54.
  47. Messmer BT, Albesiano E, Efremov DG, Ghiotto F, Allen SL, Kolitz J, Foa R, Damle RN, Fais F, Messmer D, Rai KR, Ferrarini M, Chiorazzi N. Multiple distinct sets

- of stereotyped antigen receptors indicate a role for antigen in promoting chronic lymphocytic leukemia. *J Exp Med*. 2004; 200: 519–25. <https://doi.org/10.1084/jem.20040544>.
48. Agathangelidis A, Darzentas N, Hadzidimitriou A, Brochet X, Murray F, Yan XJ, Davis Z, van Gastel-Mol EJ, Tresoldi C, Chu CC, Cahill N, Giudicelli V, Tichy B, et al. Stereotyped B-cell receptors in one-third of chronic lymphocytic leukemia: a molecular classification with implications for targeted therapies. *Blood*. 2012; 119:4467–75. <https://doi.org/10.1182/blood-2011-11-393694>.
  49. Stamatopoulos K, Belessi C, Moreno C, Boudjograh M, Guida G, Smilevska T, Belhoul L, Stella S, Stavroyianni N, Crespo M, Hadzidimitriou A, Sutton L, Bosch F, et al. Over 20% of patients with chronic lymphocytic leukemia carry stereotyped receptors: Pathogenetic implications and clinical correlations. *Blood*. 2007; 109:259–70. <https://doi.org/10.1182/blood-2006-03-012948>.
  50. Murray F, Darzentas N, Hadzidimitriou A, Tobin G, Boudjogra M, Scielzo C, Laoutaris N, Karlsson K, Baran-Marzszak F, Tsaftaris A, Moreno C, Anagnostopoulos A, Caligaris-Cappio F, et al. Stereotyped patterns of somatic hypermutation in subsets of patients with chronic lymphocytic leukemia: implications for the role of antigen selection in leukemogenesis. *Blood*. 2008; 111:1524–33. <https://doi.org/10.1182/blood-2007-07-099564>.
  51. Tobin G, Thunberg U, Karlsson K, Murray F, Laurell A, Willander K, Enblad G, Merup M, Vilpo J, Juliusson G, Sundstrom C, Soderberg O, Roos G, et al. Subsets with restricted immunoglobulin gene rearrangement features indicate a role for antigen selection in the development of chronic lymphocytic leukemia. *Blood*. 2004; 104:2879–85. <https://doi.org/10.1182/blood-2004-01-0132>.
  52. Sutton LA, Kostareli E, Hadzidimitriou A, Darzentas N, Tsaftaris A, Anagnostopoulos A, Rosenquist R, Stamatopoulos K. Extensive intraclonal diversification in a subgroup of chronic lymphocytic leukemia patients with stereotyped IGHV4-34 receptors: implications for ongoing interactions with antigen. *Blood*. 2009; 114:4460–8. <https://doi.org/10.1182/blood-2009-05-221309>.
  53. Sutton LA, Papadopoulos G, Hadzidimitriou A, Papadopoulos S, Kostareli E, Rosenquist R, Tzovaras D, Stamatopoulos K. An entity evolving into a community: defining the common ancestor and evolutionary trajectory of chronic lymphocytic leukemia stereotyped subset #4. *Mol Med*. 2015; 20:720–8. <https://doi.org/10.2119/molmed.2014.00140>.
  54. Agathangelidis A, Vardi A, Baliakas P, Stamatopoulos K. Stereotyped B-cell receptors in chronic lymphocytic leukemia. *Leuk Lymphoma*. 2014; 55:2252–61. <https://doi.org/10.3109/10428194.2013.879715>.
  55. Baliakas P, Mattsson M, Stamatopoulos K, Rosenquist R. Prognostic indices in chronic lymphocytic leukaemia: where do we stand how do we proceed? *J Intern Med*. 2016; 279:347–57. <https://doi.org/10.1111/joim.12455>.
  56. Fabbri G, Khiabanian H, Holmes AB, Wang J, Messina M, Mullighan CG, Pasqualucci L, Rabadan R, Dalla-Favera R. Genetic lesions associated with chronic lymphocytic leukemia transformation to Richter syndrome. *J Exp Med*. 2013; 210:2273–88. <https://doi.org/10.1084/jem.20131448>.
  57. Baliakas P, Iskas M, Gardiner A, Davis Z, Plevova K, Nguyen-Khac F, Malcikova J, Anagnostopoulos A, Glide S, Mould S, Stepanovska K, Brejcha M, Belessi C, et al. Chromosomal translocations and karyotype complexity in chronic lymphocytic leukemia: a systematic reappraisal of classic cytogenetic data. *Am J Hematol*. 2014; 89:249–55. <https://doi.org/10.1002/ajh.23618>.
  58. Rossi D, Spina V, Bomben R, Rasi S, Dal-Bo M, Brusca A, Rossi FM, Monti S, Degan M, Ciardullo C, Serra R, Zucchetto A, Nomdedeu J, et al. Association between molecular lesions and specific B-cell receptor subsets in chronic lymphocytic leukemia. *Blood*. 2013; 121:4902–5. <https://doi.org/10.1182/blood-2013-02-486209>.
  59. Strefford JC, Sutton LA, Baliakas P, Agathangelidis A, Malcikova J, Plevova K, Scarfo L, Davis Z, Stalika E, Cortese D, Cahill N, Pedersen LB, di Celle PF, et al. Distinct patterns of novel gene mutations in poor-prognostic stereotyped subsets of chronic lymphocytic leukemia: the case of SF3B1 and subset #2. *Leukemia*. 2013; 27:2196–9. <https://doi.org/10.1038/leu.2013.98>.
  60. Sutton LA, Young E, Baliakas P, Hadzidimitriou A, Moysiadis T, Plevova K, Rossi D, Kminkova J, Stalika E, Pedersen LB, Malcikova J, Agathangelidis A, Davis Z, et al. Different spectra of recurrent gene mutations in subsets of chronic lymphocytic leukemia harboring stereotyped B-cell receptors. *Haematologica*. 2016; 101:959–67. <https://doi.org/10.3324/haematol.2016.141812>.
  61. Baliakas P, Hadzidimitriou A, Sutton LA, Rossi D, Minga E, Villamor N, Larrayoz M, Kminkova J, Agathangelidis A, Davis Z, Tausch E, Stalika E, Kantorova B, et al. Recurrent mutations refine prognosis in chronic lymphocytic leukemia. *Leukemia*. 2015; 29:329–36. <https://doi.org/10.1038/leu.2014.196>.
  62. Gawad C, Pepin F, Carlton VE, Klinger M, Logan AC, Miklos DB, Faham M, Dahl G, Lacayo N. Massive evolution of the immunoglobulin heavy chain locus in children with B precursor acute lymphoblastic leukemia. *Blood*. 2012; 120:4407–17. <https://doi.org/10.1182/blood-2012-05-429811>.
  63. Faham M, Zheng J, Moorhead M, Carlton VE, Stow P, Coustan-Smith E, Pui CH, Campana D. Deep-sequencing approach for minimal residual disease detection in acute lymphoblastic leukemia. *Blood*. 2012; 120:5173–80. <https://doi.org/10.1182/blood-2012-07-444042>.
  64. Ritgen M, Bottcher S, Stilgenbauer S, Bunjes D, Schubert J, Cohen S, Humpe A, Hallek M, Kneba M, Schmitz N, Dohner H, Dreger P, German CLLSG. Quantitative MRD monitoring identifies distinct GVL response patterns after allogeneic stem cell transplantation for chronic lymphocytic leukemia: results from the GCLLSG CLL3X

- trial. *Leukemia*. 2008; 22:1377–86. <https://doi.org/10.1038/leu.2008.96>.
65. Bottcher S, Ritgen M, Dreger P. Allogeneic stem cell transplantation for chronic lymphocytic leukemia: lessons to be learned from minimal residual disease studies. *Blood Rev*. 2011; 25:91–6. <https://doi.org/10.1016/j.blre.2011.01.001>.
  66. Rawstron AC, de Tute R, Jack AS, Hillmen P. Flow cytometric protein expression profiling as a systematic approach for developing disease-specific assays: identification of a chronic lymphocytic leukaemia-specific assay for use in rituximab-containing regimens. *Leukemia*. 2006; 20:2102–10. <https://doi.org/10.1038/sj.leu.2404416>.
  67. Farina L, Carniti C, Dodero A, Vendramin A, Raganato A, Spina F, Patriarca F, Narni F, Benedetti F, Olivieri A, Corradini P. Qualitative and quantitative polymerase chain reaction monitoring of minimal residual disease in relapsed chronic lymphocytic leukemia: early assessment can predict long-term outcome after reduced intensity allogeneic transplantation. *Haematologica*. 2009; 94:654–62. <https://doi.org/10.3324/haematol.2008.000273>.
  68. Logan AC, Gao H, Wang C, Sahaf B, Jones CD, Marshall EL, Buno I, Armstrong R, Fire AZ, Weinberg KI, Mindrinos M, Zehnder JL, Boyd SD, et al. High-throughput VDJ sequencing for quantification of minimal residual disease in chronic lymphocytic leukemia and immune reconstitution assessment. *Proc Natl Acad Sci USA*. 2011; 108: 21194–9. <https://doi.org/10.1073/pnas.1118357109>.
  69. Boyd SD, Marshall EL, Merker JD, Maniar JM, Zhang LN, Sahaf B, Jones CD, Simen BB, Hanczaruk B, Nguyen KD, Nadeau KC, Egholm M, Miklos DB, et al. Measurement and clinical monitoring of human lymphocyte clonality by massively parallel VDJ pyrosequencing. *Sci Transl Med*. 2009; 1:12ra23.
  70. Rawstron AC, Fazi C, Agathangelidis A, Villamor N, Letestu R, Nomdedeu J, Palacio C, Stehlikova O, Kreuzer KA, Liptrot S, O'Brien D, de Tute RM, Marinov I, et al. A complementary role of multiparameter flow cytometry and high-throughput sequencing for minimal residual disease detection in chronic lymphocytic leukemia: an European Research Initiative on CLL study. *Leukemia*. 2016; 30:929–36. <https://doi.org/10.1038/leu.2015.313>.
  71. Campbell PJ, Pleasance ED, Stephens PJ, Dicks E, Rance R, Goodhead I, Follows GA, Green AR, Futreal PA, Stratton MR. Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing. *Proc Natl Acad Sci USA*. 2008; 105:13081–6. <https://doi.org/10.1073/pnas.0801523105>.
  72. Niklas N, Proll J, Weinberger J, Zopf A, Wiesinger K, Krismer K, Bettelheim P, Gabriel C. Qualifying high-throughput immune repertoire sequencing. *Cell Immunol*. 2014; 288:31–8. <https://doi.org/10.1016/j.cellimm.2014.02.001>.
  73. Kriangkum J, Motz SN, Mack T, Beiggi S, Baigorri E, Kuppusamy H, Belch AR, Johnston JB, Pilarski LM. Single Cell Analysis and Next-Generation Immuno-Sequencing Show That Multiple Clones Persist in Patients with Chronic Lymphocytic Leukemia. *PLoS One*. 2015; 10:e0137232. <https://doi.org/10.1371/journal.pone.0137232>.
  74. Stamatopoulos B, Timbs A, Bruce D, Smith T, Clifford R, Robbe P, Burns A, Vavoulis DV, Lopez L, Antoniou P, Mason J, Dreau H, Schuh A. Targeted deep sequencing reveals clinically relevant subclonal IgHV rearrangements in chronic lymphocytic leukemia. *Leukemia*. 2016. <https://doi.org/10.1038/leu.2016.307>.
  75. Blachly JS, Ruppert AS, Zhao W, Long S, Flynn J, Flinn I, Jones J, Maddocks K, Andritsos L, Ghia EM, Rassenti LZ, Kipps TJ, de la Chapelle A, et al. Immunoglobulin transcript sequence and somatic hypermutation computation from unselected RNA-seq reads in chronic lymphocytic leukemia. *Proc Natl Acad Sci USA*. 2015; 112:4322–7. <https://doi.org/10.1073/pnas.1503587112>.
  76. Busse CE, Czogiel I, Braun P, Arndt PF, Wardemann H. Single-cell based high-throughput sequencing of full-length immunoglobulin heavy and light chain genes. *Eur J Immunol*. 2014; 44:597–603. <https://doi.org/10.1002/eji.201343917>.
  77. DeKosky BJ, Ippolito GC, Deschner RP, Lavinder JJ, Wine Y, Rawlings BM, Varadarajan N, Giesecke C, Dorner T, Andrews SF, Wilson PC, Hunicke-Smith SP, Willson CG, et al. High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. *Nat Biotechnol*. 2013; 31:166–9. <https://doi.org/10.1038/nbt.2492>.
  78. DeKosky BJ, Kojima T, Rodin A, Charab W, Ippolito GC, Ellington AD, Georgiou G. In-depth determination and analysis of the human paired heavy- and light-chain antibody repertoire. *Nat Med*. 2015; 21:86–91. <https://doi.org/10.1038/nm.3743>.
  79. Murugan R, Imkeller K, Busse CE, Wardemann H. Direct high-throughput amplification and sequencing of immunoglobulin genes from single human B cells. *Eur J Immunol*. 2015; 45:2698–700. <https://doi.org/10.1002/eji.201545526>.
  80. Parikh SA, Strati P, Tsang M, West CP, Shanafelt TD. Should IGHV status and FISH testing be performed in all CLL patients at diagnosis? A systematic review and meta-analysis. *Blood*. 2016; 127:1752–60. <https://doi.org/10.1182/blood-2015-10-620864>.
  81. Bolotin DA, Mamedov IZ, Britanova OV, Zvyagin IV, Shagin D, Ustyugova SV, Turchaninova MA, Lukyanov S, Lebedev YB, Chudakov DM. Next generation sequencing for TCR repertoire profiling: platform-specific features and correction algorithms. *Eur J Immunol*. 2012; 42:3073–83. <https://doi.org/10.1002/eji.201242517>.
  82. Khan TA, Friedensohn S, Gorter de Vries AR, Straszewski J, Ruscheweyh HJ, Reddy ST. Accurate and predictive antibody repertoire profiling by molecular amplification fingerprinting. *Sci Adv*. 2016; 2:e1501371. <https://doi.org/10.1126/sciadv.1501371>.

83. Turchaninova MA, Davydov A, Britanova OV, Shugay M, Bikos V, Egorov ES, Kirgizova VI, Merzlyak EM, Staroverov DB, Bolotin DA, Mamedov IZ, Izraelson M, Logacheva MD, et al. High-quality full-length immunoglobulin profiling with unique molecular barcoding. *Nat Protoc.* 2016; 11:1599–616. <https://doi.org/10.1038/nprot.2016.093>.
84. Georgiou G, Ippolito GC, Beausang J, Busse CE, Wardemann H, Quake SR. The promise and challenge of high-throughput sequencing of the antibody repertoire. *Nat Biotechnol.* 2014; 32:158–68. <https://doi.org/10.1038/nbt.2782>.
85. Sutton LA, Rosenquist R. Deciphering the molecular landscape in chronic lymphocytic leukemia: time frame of disease evolution. *Haematologica.* 2015; 100:7–16. <https://doi.org/10.3324/haematol.2014.115923>.
86. Hernandez JA, Hernandez-Sanchez M, Rodriguez-Vicente AE, Grossmann V, Collado R, Heras C, Puiggros A, Martin AA, Puig N, Benito R, Robledo C, Delgado J, Gonzalez T, et al. A Low Frequency of Losses in 11q Chromosome Is Associated with Better Outcome and Lower Rate of Genomic Mutations in Patients with Chronic Lymphocytic Leukemia. *PLoS One.* 2015; 10:e0143073. <https://doi.org/10.1371/journal.pone.0143073>.
87. Jethwa A, Hullein J, Stolz T, Blume C, Sellner L, Jauch A, Sill M, Kater AP, te Raa GD, Geisler C, van Oers M, Dietrich S, Dreger P, et al. Targeted resequencing for analysis of clonal composition of recurrent gene mutations in chronic lymphocytic leukaemia. *Br J Haematol.* 2013; 163: 496–500. <https://doi.org/10.1111/bjh.12539>.
88. Sutton LA, Ljungstrom V, Mansouri L, Young E, Cortese D, Navrkalova V, Malcikova J, Muggen AF, Trbusek M, Panagiotidis P, Davi F, Belessi C, Langerak AW, et al. Targeted next-generation sequencing in chronic lymphocytic leukemia: a high-throughput yet tailored approach will facilitate implementation in a clinical setting. *Haematologica.* 2015; 100:370–6. <https://doi.org/10.3324/haematol.2014.109777>.
89. Vollbrecht C, Mairinger FD, Koitzsch U, Peifer M, Koenig K, Heukamp LC, Crispatzu G, Wilden L, Kreuzer KA, Hallek M, Odenthal M, Herling CD, Buettner R. Comprehensive Analysis of Disease-Related Genes in Chronic Lymphocytic Leukemia by Multiplex PCR-Based Next Generation Sequencing. *PLoS One.* 2015; 10:e0129544. <https://doi.org/10.1371/journal.pone.0129544>.
90. Guieze R, Robbe P, Clifford R, de Guibert S, Pereira B, Timbs A, Dilhuydy MS, Cabes M, Ysebaert L, Burns A, Nguyen-Khac F, Davi F, Veronese L, et al. Presence of multiple recurrent mutations confers poor trial outcome of relapsed/refractory CLL. *Blood.* 2015; 126:2110–7. <https://doi.org/10.1182/blood-2015-05-647578>.
91. Nadeu F, Delgado J, Royo C, Baumann T, Stankovic T, Pinyol M, Jares P, Navarro A, Martin-Garcia D, Bea S, Salaverria I, Oldreive C, Aymerich 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. <https://doi.org/10.1182/blood-2015-07-659144>.
92. Rigolin GM, Saccenti E, Bassi C, Lupini L, Quaglia FM, Cavallari M, Martinelli S, Formigaro L, Lista E, Bardi MA, Volta E, Tammiso E, Melandri A, et al. Extensive next-generation sequencing analysis in chronic lymphocytic leukemia at diagnosis: clinical and biological correlations. *J Hematol Oncol.* 2016; 9:88. <https://doi.org/10.1186/s13045-016-0320-z>.
93. Rasi S, Khiabani H, Ciardullo C, Terzi-di-Bergamo L, Monti S, Spina V, Brusca A, Cerri M, Deambrogi C, Martuscelli L, Biasi A, Spaccarotella E, De Paoli L, et al. Clinical impact of small subclones harboring NOTCH1, SF3B1 or BIRC3 mutations in chronic lymphocytic leukemia. *Haematologica.* 2016; 101:e135-8. <https://doi.org/10.3324/haematol.2015.136051>.
94. Rossi D, Khiabani H, Spina V, Ciardullo C, Brusca A, Fama R, Rasi S, Monti S, Deambrogi C, De Paoli L, Wang J, Gattei V, Guarini A, et al. Clinical impact of small TP53 mutated subclones in chronic lymphocytic leukemia. *Blood.* 2014; 123:2139–47. <https://doi.org/10.1182/blood-2013-11-539726>.
95. Baliakas P, Hadzidimitriou A, Sutton LA, Rossi D, Minga E, Villamor N, Larrayoz M, Kminkova J, Agathangelidis A, Davis Z, Tausch E, Stalika E, Kantorova B, et al. Recurrent mutations refine prognosis in chronic lymphocytic leukemia. *Leukemia.* 2014; 29:329–36. <https://doi.org/10.1038/leu2014196>.
96. Jeromin S, Weissmann S, Haferlach C, Dicker F, Bayer K, Grossmann V, Alpermann T, Roller A, Kohlmann A, Haferlach T, Kern W, Schnittger S. SF3B1 mutations correlated to cytogenetics and mutations in NOTCH1, FBXW7, MYD88, XPO1 and TP53 in 1160 untreated CLL patients. *Leukemia.* 2014; 28:108–17. <https://doi.org/10.1038/leu2013263>.
97. Rossi D, Rasi S, Spina V, Brusca A, Monti S, Ciardullo C, Deambrogi C, Khiabani H, Serra R, Bertoni F, Forconi F, Laurenti L, Marasca R, et al. Integrated mutational and cytogenetic analysis identifies new prognostic subgroups in chronic lymphocytic leukemia. *Blood.* 2013; 121:1403–12. <https://doi.org/10.1182/blood-2012-09-458265>.
98. Pospisilova S, Gonzalez D, Malcikova J, Trbusek M, Rossi D, Kater AP, Cymbalista F, Eichhorst B, Hallek M, Dohner H, Hillmen P, van Oers M, Gribben J, et al. ERIC recommendations on TP53 mutation analysis in chronic lymphocytic leukemia. *Leukemia.* 2012; 26:1458–61. <https://doi.org/10.1038/leu.2012.25>.
99. Wang L, Shalek AK, Lawrence M, Ding R, Gaublot JM, Pochet N, Stojanov P, Sougnez C, Shukla SA, Stevenson KE, Zhang W, Wong J, Sievers QL, et al. Somatic mutation as a mechanism of Wnt/beta-catenin pathway activation in CLL. *Blood.* 2014; 124: 1089–98. <https://doi.org/10.1182/blood-2014-01-552067>.

100. Wiestner A. The role of B-cell receptor inhibitors in the treatment of patients with chronic lymphocytic leukemia. *Haematologica*. 2015; 100:1495–507. <https://doi.org/10.3324/haematol.2014.119123>.
101. Bonnal S, Vigevani L, Valcarcel J. The spliceosome as a target of novel antitumour drugs. *Nat Rev Drug Discov*. 2012; 11:847–59. <https://doi.org/10.1038/nrd3823>.
102. Groth C, Fortini ME. Therapeutic approaches to modulating Notch signaling: current challenges and future prospects. *Semin Cell Dev Biol*. 2012; 23:465–72. <https://doi.org/10.1016/j.semcdb.2012.01.016>.