Dr. Veera Talukdar
Principal, Shree Ram College
of Commerce and Science,
Datta Mandir Road, Bhandup,
West Mumbai, India

Dr. Amitabha Kar
Associate-Professor & Head,
Department of Human
Physiology, Women's College,
Govt. of Tripura, Agartala,
India

Dr. Partha Majumder
Biomedical Scientist &
Systems Biologist, Former
Principal
Scientist(Helixinfosystems),
Former Head of The
Department of Applied
Biotechnology &
Bioinformatics, Sikkim
Manipal University, Kolkata,
India

A glimpse on chronic lymphocytic leukemia and its insights of molecular heterogeneity

Dr. Veera Talukdar, Dr. Amitabha Kar, and Dr. Partha Majumder

Abstract
Chronic lymphocytic leukemia (CLL) has been consistently at the forefront of genetic research owing to its prevalence and the accessibility of sample material. Recently, genome-wide technologies have been intensively applied to CLL genetics, with remarkable progress. Single nucleotide polymorphism (SNP) arrays have identified recurring chromosomal aberrations, thereby focusing functional studies on discrete genomic lesions and leading to the first implication of somatic microRNA disruption in cancer. Next-generation sequencing (NGS) has further transformed our understanding of CLL by identifying novel recurrently mutated putative drivers, including the unexpected discovery of somatic mutations affecting spliceosome function. NGS has further enabled in-depth examination of the transcriptional and epigenetic changes in CLL that accompany genetic lesions, and has shed light on how different driver events appear at different stages of disease progression and clonally evolve with relapsed disease. In addition to providing important insights into disease biology, these inventionshave significant translational potential. They enhance prognosis by highlighting specific lesions associated with poor clinical outcomes (for example, driver events such as mutations in the splicing factor subunit gene *SF3B1*) or with increased clonal heterogeneity (for example, the presence of subclonal driver mutations). Here, we review new genomic discoveries in CLL and discuss their possible implications in the era of precision medicine.

Keywords: Chronic lymphocytic leukemia, monoclonal B-cell lymphocytosis, next-generation sequencing, somatic copy number alteration, tumor necrosis factor

Introduction
Treatment of chronic lymphocytic leukemia (CLL)
Chronic lymphocytic leukemia (CLL) is a low-grade B-cell malignancy, characterized by the accumulation of mature CD5$^+$/CD19$^+$/CD23$^+$ lymphocytes with weak surface expression of a monoclonal immunoglobulin (Ig) [1] in the peripheral blood, bone marrow, lymph nodes and spleen. It is diagnosed either incidentally (with an abnormally high white blood cell count) in asymptomatic patients, or due to symptoms that result from cytopenias, adenopathy or constitutional symptoms, as outlined by the 2008 International Workshop on CLL [2]. CLL is part of a spectrum of pathological conditions involving clonally proliferating B cells. It is thought to be preceded by monoclonal B-cell lymphocytosis (MBL), a state in which a smaller size B-cell clone is present, typically in the absence of symptoms [3]. At the other end of the spectrum, CLL may transform into a higher-grade malignancy, a process termed Richter's transformation, which is often associated with a dismal clinical outcome [4].

CLL possesses several features that place it at the forefront of cancer genetic research. First, it has high relevance as the most common leukemia in adults [5]. Second, the ability to easily procure primary tumor cells from the bloodstream facilitates the application of cutting-edge genetic methodologies. These technologies have been used to define the underlying biology of CLL (for instance, elucidating the cell of origin of this lymphoid malignancy [6]), as well as to explore clinical questions (such as how to predict clinical outcome in a highly variable disease on the basis of molecular indicators [7]). These investigations have yielded striking insights, including the first identification of a causative somatic microRNA alteration in cancer [8], as well as one of the first effective molecular prognostic schemes [9].

In parallel, there has been marked progress in the development of therapeutic options in CLL (extensively reviewed elsewhere [10-12]).

Correspondence
Dr. Partha Majumder
Biomedical Scientist &
Systems Biologist, Former
Principal Scientist
(Helixinfosystems), Former
Head of The Department of
Applied Biotechnology &
Bioinformatics, Sikkim
Manipal University, Kolkata,
India

While the general therapeutic paradigm in CLL remains based on the 'watch and wait' approach (that is, treatment is initiated only when symptoms occur) [13], clinicians now have an extensive array of effective options when treatment is required. For example, combination chemo-immunotherapy with fludarabine, cyclophosphamide and rituximab has yielded excellent long-term results [14]. Additionally, immunotherapy-based therapeutics such as alemtuzumab [15] and allogeneic stem-cell transplantation [16, 17] have been demonstrated to provide effective disease control in treatment-refractory or high-risk patients. Importantly, as CLL often affects elderly individuals, more tolerable therapeutic approaches have been successfully applied, such as lenalidomide [18] and bendamustine-based regimens [19]. Most recently, therapies targeting the B-cell-receptor signaling pathway, such as ibrutinib, have generated excitement as they have shown promising efficacy and tolerability in phase II clinical trials [20].

Despite the expansion of therapeutic options for CLL patients, which has improved patient survival, CLL remains largely incurable, and its course is difficult to predict. Furthermore, guidance about appropriate treatment selection on the basis of individual genetic and molecular abnormalities remains limited [21]. A full characterization of the CLL genomic landscape would enable several questions to be addressed. Can we accurately predict the course of the disease? Can we predict which patients will respond to which therapies? And can we use genomic information to target the therapy to the underlying genetic or other alterations? Over the past two years, genomic approaches have been intensively applied for studying this disease and have aided us in answering these important questions Here, we review the main findings of these investigations as well as their possible biological and clinical implications, focusing on key findings obtained by genomic technologies, such as the expanded compendium of somatic gene alterations and the characterization of clonal evolution and of the epigenetic landscape of CLL.

Somatic copy number alterations
The study of somatic copy number alterations (sCNAs), which are somatically acquired alterations of a genome that result in the cell having an abnormal number of copies of one or more sections of DNA, has revealed a high degree of molecular heterogeneity in CLL (reviewed extensively elsewhere [6, 7, 22]). Briefly, unlike other lymphoid tumors such as follicular lymphoma or diffuse large B-cell lymphoma, CLL is not characterized by a common translocation involving the Ig loci, but instead by specific recurrent sCNAs (such as chromosome 11q deletions (del(11q)), trisomy 12, del(13q) and del(17p)) that have been observed using comparative genome hybridization [23] and single nucleotide polymorphism (SNP)-array-based investigations [24]. Considering the near-diploid genome of CLL (only a small number of sCNAs are typically observed in CLL), these are probably causative events, as the finding of highly recurrent events against a backdrop of a low background sCNA rate testifies to significant selection and hence to a significant fitness advantage afforded to CLL cells by these lesions. Furthermore, they affect clinical outcome [9]: del(13q) is associated with a good prognosis whereas del(11q) and del(17p) are associated with a poor prognosis with present-day chemo-immunotherapy approaches. Lower frequency lesions have also been identified involving the *MYC* locus [25], the short arm of chromosome 8 [23], and lesions probably affecting *PIK3CA, NFKB2* and *MGA* [26, 27]. Allele-specific copy number quantification with SNP arrays has also enabled the discovery of frequent copy-neutral loss of heterozygosity in CLL, often resulting in biallelic hits (mutations or epigenetic alterations) in key CLL-related loci, and therefore potentially altering function [24]. For example, duplication of the allele containing the small del(13q) event may be concurrent with the loss of the sister normal allele.

Table 1: Recurrent sCNAs in peripheral blood primary CLL samples

sCNA	Frequency in CLL (%)	Likely target	Frequency in MBL (%)	More common IGHV status	Prognostic significance
del(13q14)	57 to 61 [26, 27]	*miR-15a/16*, encoded in an intron of *DLEU2*[8]. Its deletion leads to the release of BCL2 from microRNA-mediated down-regulation [28]	48	None	Good [9]
del(11q22.3)	6 to 27 [26, 27]	*ATM, BIRC3*	Rare	Unmutated	Poor [9]
Trisomy 12	11 to 12 [26, 27]	Unknown [7]	20	None	None [9]
del(17p)	6 to 8 [26, 27]	*TP53*	Rare	Unmutated	Poor [9]
amp(2p)	7 [27]	*XPO1, BCL11A*	Unknown	Unmutated	None [27]
amp(8q24.21)	5 [26, 27]	*MYC*	Unknown	Unmutated	Poor [26, 27]
del(15q15.1)	4 [27]	*MGA*	Unknown	Unknown	None [27]
del(10)(q24)	2 [27]	*NFKB2*	Unknown	Unknown	Unknown
del(18p)	3 [27]	Unknown	Unknown	Unmutated [27]	Unknown
del(6q)	7 [27]	Unknown	Unknown	Unknown	Unknown
amp(3q26.32)	6 [26]	*PIK3CA*	Unknown	Unknown	Poor [26]
del(8p)	5 [23, 26]	Unknown	Unknown	Unknown	Poor [26]

amp, amplification; sCNA, somatic copy number alterations; MBL, monoclonal B-cell lymphocytosis.

By measuring the affected portion of chromosomes across many CLL patient samples, and thus defining the size of minimally affected lesions, these methodologies have contributed to a mechanistic understanding of causative lesions in CLL. For instance, the minimal deleted region in del(13q14) focused functional investigation onto a small number of genetic elements, and ultimately led to the discovery that the microRNAs *miR-15a* and *miR-16-1*, encoded by an intron of *DLEU2*, have a causative role in CLL [8], perhaps through the release of the anti-apoptotic BCL2 protein from microRNA-mediated downregulation [28]. More recently, the case for *miR-15a/16-1* deletion having a causative role in CLL was strengthened with the generation of a CLL mouse model based on knockout of this

locus [29]. Significant variation in the size of the deleted region (from approximately 300 kilobases to more than 50 megabases) provides clues to additional contributing genetic components [30]. For example, adjacent hits within large monoallelic deletions (affecting, for example, the *RB1* gene) may have an important contributory role compared with a more isolated effect of the disruption of the microRNA cluster in the shorter biallelic deletions. While del(11q) and del(17p) impact the cellular network primarily due to the deletion of known tumor suppressor genes *ATM* and *TP53*, respectively, the mechanism by which trisomy 12 contributes to lympho-proliferation remains unknown [7]. This is due in part to the large size of the affected lesion (an entire chromosome), which limits the ability to focus investigations on a smaller number of genes; application of large RNA interference screen-based approaches, however, may reveal candidate genes.

Clinical application of this information yielded one of the earliest molecular classification schemata in cancer, predicting the course of disease based on the identity of the sCNA [9]. This is of particular importance in a disease like CLL where the clinical heterogeneity is enormous, with some patients remaining stable without treatment for years or even decades, while others follow a fulminant and treatment-refractory course. Higher genomic complexity - the presence of a high number of sCNAs - has also been associated with worse outcome, including shorter time to first therapy and lower overall survival rate [31, 32]. Nevertheless, in contrast to other tumors, CLL has a relative paucity of sCNAs [26]. This observation has led to the suggestion that somatic single nucleotide variants (sSNVs) and indels could play an important role in the pathogenesis of CLL, paving the way for the application of next-generation sequencing (NGS) technologies to this disease.

The genomic landscape of CLL probed with next-generation sequencing

NGS studies of the CLL genome [33, 34] have effectively elucidated the level of genomic complexity in CLL, and have revealed that the average number of non-silent mutations (that is, mutations that alter the protein sequence) is 10 to 20 per each sequenced CLL sample (out of approximately 1,000 somatic mutations per sample detected genome-wide). This is at least an order of magnitude lower than the number of lesions detected in the coding genomes of common epithelial cancers, such as lung cancer or melanoma [35]. Even among hematologic malignancies, the genomic complexity of CLL is relatively low, similar to that of acute leukemias [36]. The overwhelming majority of sSNVs involve C>T transitions at CpG sites, with some differences in mutation patterns between CLL with mutations in the Ig heavy variable region (*IGHV*-mutated) and CLL lacking *IGHV* mutations (*IGHV*-unmutated), suggestive of the involvement of aberrant somatic hypermutation with error-prone repair [33]. Importantly, the number of mutations in CLL samples from patients who received chemo-immunotherapy before sampling is not significantly increased [34]. These results suggest that, unlike several other cancers such as glioblastoma [37], CLL treatment does not substantially contribute to increased mutagenesis.

NGS has also uncovered an unusual form of genomic complexity in CLL, termed chromothripsis, which results from a massive genomic rearrangement event within a single region through an as yet unknown underlying mechanism [38]. Overall, chromothripsis was detected at a substantial frequency in CLL (approximately 2%) through inference from SNP-array data, and was seen almost exclusively in CLL with *IGHV*-unmutated status and with mutated *TP53*. This observation suggests that although genome integrity is largely preserved in CLL (as demonstrated by its typically near-diploid genome), catastrophic rearrangements can be tolerated and selected within a permissive genetic context. Perhaps unsurprisingly, chromothripsis has been associated with a worse prognosis [27].

Beyond the characterization of the mutational landscape in CLL, NGS has also been used to study, in an unbiased fashion, recurrent genetic alterations in CLL. Putative driver mutations, which are genetic lesions that are likely to confer a significant fitness advantage, have been identified (Tables 2 and 3). The first studies reported whole-genome [33] or whole-exome [39] sequencing of a handful of CLL samples, followed by targeted sequencing of coding mutations detected in these samples in larger validation cohorts. This approach uncovered several important putative drivers, including *MYD88* and *NOTCH1* mutations. An alternative approach using a larger initial cohort probed with whole-exome sequencing has enabled the discovery of a larger number of putative drivers [34, 40]. Collectively, these studies have demonstrated wide heterogeneity in the genetic lesions driving CLL transformation and progression, characterized by 'mountains' (that is, highly recurrent genes such as *TP53*) and 'hills' (infrequent but still statistically significant recurrent genes such as *XPO1*), as seen in other sequencing efforts [41].

Table 2: High-frequency recurrently mutated genes in CLL

Gene	Frequency in CLL (%)	Likely gene function	Frequency in MBL (%)	Richter's* or chemo-refractory cases	Gene mutation hotspots	More common IGHV status	Prognostic significance
TP53	7.5 to 13 [34, 39, 75]	Apoptosis, DNA repair	Rare [76]	Yes [46]	Inactivating mutations	Unmutated	Poor [46]
SF3B1	10 to 14 [34, 40, 75, 113]	Splicing factor	Rare [76]	Yes [46]	K700E	Unmutated	Poor [34, 40]
NOTCH1	10 to 17 [33, 34, 39, 40, 75, 113]	Developmental processes	Rare [76]	Yes [39]	P2515Rfs*4	Unmutated	Poor [46]
MYD88	3 to 8 [33, 34, 75]	TLR adaptor	Unknown	Unknown	L265P	Mutated	None
ATM	8 to 15 [34, 75,]	DNA repair	Unknown	Unknown	Inactivating mutations	Unmutated	Poor
BIRC3	4 [39, 59]	NF-κB pathway inhibitor	Absent [59]	Yes [59]	Inactivating mutations	Unmutated	Poor [46]

*Richter's transformation, in which CLL transforms to a higher-grade malignancy.

Table 3: Low-frequency recurrently mutated genes in CLL

Gene	Frequency (%)	Likely gene function
XPO1	2 to 4 [33, 75]	Nuclear export
CHD2	4 to 5 [40, 75]	Chromatin modification
POT1	3 to 5 [40, 75]	Telomere maintenance
HIST1H1E	3 [75]	Histone protein
NRAS	3 [75]	Cell growth
BCOR	3 [75]	Apoptosis regulation
ZMYM3	3 [34, 75]	Chromatin modification
RIPK1	3 [75]	Inflammatory pathway
SAMHD1	3 [75]	Innate immune response
KRAS	2 [75]	Cell growth
MED12	2 [75]	Gene transcription
ITPKB	2 [75]	B-cell signaling
DDX3X	2 [34, 75]	RNA helicase
EGR2	1 [75]	Transcription factor
FBXW7	3 [34, 75]	Ubiquitination
KLHL6	2 [33]	B-cell receptor signaling
MAPK1	3 [34]	MAP kinase
LRP1B	5 [40]	LDL receptor family

One of the earliest CLL drivers identified through NGS was *NOTCH1* [33, 34, 39]. *NOTCH1* encodes a ligand-activated transcription factor that regulates several downstream pathways important for the control of cell growth. One recurrent mutation (c.7544_7545fsdel) accounts for approximately 80% of all *NOTCH1* mutations and generates a premature stop codon in the PEST domain (a peptide rich in proline (P), glutamic acid (E), serine (S) and threonine (T), thought to act as a signal for protein degradation [42]), which normally limits the intensity and duration of NOTCH1 signaling [39]. Disruption of the PEST domain results in impaired NOTCH1 degradation, as it interferes with phosphorylation of the PEST domain of the receptor and its proteasomal degradation through the FBXW7-SCF ubiquitin ligase complex [43]. This in turn results in accumulation of an active NOTCH1 isoform, which is associated with a distinct transcriptional signature [33]. In CLL, the frequency of *NOTCH1* mutations is above 10%, and tends to occur in CLLs without *IGHV* mutation and with trisomy 12 [44], although it is important to note that the latter association was not found in another recent study [45]. In some studies, the presence of *NOTCH1* mutations provided independent prognostic information and identified a group of patients with intermediate-risk disease [46] and those in whom CLL was more likely to transform into high-grade lymphoma [47]. However, the effect size may not be as prominent as other CLL prognostic indicators, as further studies failed to show an independent prognostic value for the presence of these mutations [47, 48].

Another commonly mutated gene is *MYD88*, a critical adaptor molecule of the Toll-like receptor (TLR) complex [33, 34], seen in 3 to 8% of CLL cases. After TLR stimulation, MYD88 is recruited to the receptor as a homodimer and forms a complex with IRAK4, leading to activation of IRAK1 and IRAK2. This then leads to the downstream activation of TRAF6 and ultimately to phosphorylation of IκBα and activation of the central B-cell transcription factor, nuclear factor (NF)-κB [49, 50]. The recurrent *MYD88* mutation in CLL (L265P) imposes constitutive *MYD88-IRAK* signaling even in the absence of ligand-receptor binding, and thereby provides constitutive NF-κB activity. Of note, *MYD88* L265P mutations have been found exclusively in CLL with mutated *IGHV*. Exactly the same mutation has been identified in other malignancies of mature B cells such as diffuse large B-cell lymphoma [51], central nervous system lymphoma [52] and Waldenström's macroglobulinemia [53]. Furthermore, this aberration is potentially amenable to therapeutic targeting through direct inhibition of the MYD88-IRAK complex, through proteasomal inhibition [54] or even through the inhibition of Bruton's tyrosine kinase (BTK) [55].

Putative drivers can be further categorized based on the cellular pathways they involve. Recurrently mutated genes in CLL can be grouped into seven core cellular networks, in which the genes play well-established roles. As shown in Figure 1, these include DNA repair and cell-cycle control, Notch signaling, inflammatory pathways, Wnt signaling, RNA splicing and processing (found to be present in close to one-third of CLLs [56]), B-cell receptor signaling and chromatin modification. Pathway analysis may also be beneficial to detect commonly disrupted pathways that may be of high biological relevance but that do not contain a single highly recurrent gene, and may be missed by gene-centric analytic approaches. One such example is disruption of the Wnt pathway [34], a key player in CLL biology [57, 58].

Affected genes in CLL discovered through genomic sequencing studies can be grouped into seven core cellular pathways. Genes recurrently mutated in CLL samples are shown in red ovals, while genes found to be mutated in isolated samples but which did not reach statistical significance are shown as pink ovals. Affected cellular elements include four signaling pathways with a known role in B-cell biology: inflammatory pathways, B-cell receptor signaling, Notch signaling, and Wnt signaling. Notch and Wnt signaling both provide important pro-survival input for CLL cells, allowing them to evade apoptosis.In addition, they serve as an important bridge with the microenvironment, which is of particular importance in CLL, as manifested by relatively poor cell survival outside of the endogenous niche (for example, in *in vitro* or *in vivo* animal models) BCR signaling and inflammatory pathways may serve similar functions, and in addition may form optimal early targets for somatic mutations as they hijack physiologically active cellular pathways in relatively differentiated B cells [75]. In addition, three intranuclear processes are involved, including DNA repair, chromatic modification and RNA processing. Although the role of DNA repair disruptions has been extensively investigated, with multiple effects on pro-survival circuits, growth and genetic plasticitythe role of the other two intranuclear processes remains to be fully elucidated in CLL. IC, intracellular; C, cytoplasm.

Although the unbiased approach of whole-exome sequencing of large cohorts is highly effective at detecting putative drivers, it may still miss important drivers, either owing to lack of power to detect lower frequency events or to the patient characteristics of the investigated cohort. A striking example of such drivers is the case of *BIRC3*-inactivating mutations, which have not been detected in most of the large sequencing efforts. Targeted sequencing of the *BIRC3* coding sequence in CLL showed that *BIRC3* inactivation is particularly common in fludarabine-refractory patients (24%) [59]. BIRC3, along with TRAF2 and TRAF3, cooperates in negatively regulating MAP3K14, an activator of the non-canonical pathway of NF-κB signaling [60], and therefore *BIRC3* mutations result in constitutive NF-κB activation [59]. Thus, *BIRC3* mutations join *SF3B1* (described in the next section), *NOTCH1* and

TP53 as mutations that contribute to chemo-refractoriness [61]. This example highlights the need to include specific patient groups in sequencing efforts. Furthermore, it supports the idea that driver landscapes of similar types of malignancies can guide driver identification, as the study of *BIRC3* in CLL was prompted by its discovery in splenic marginal zone lymphoma [62].
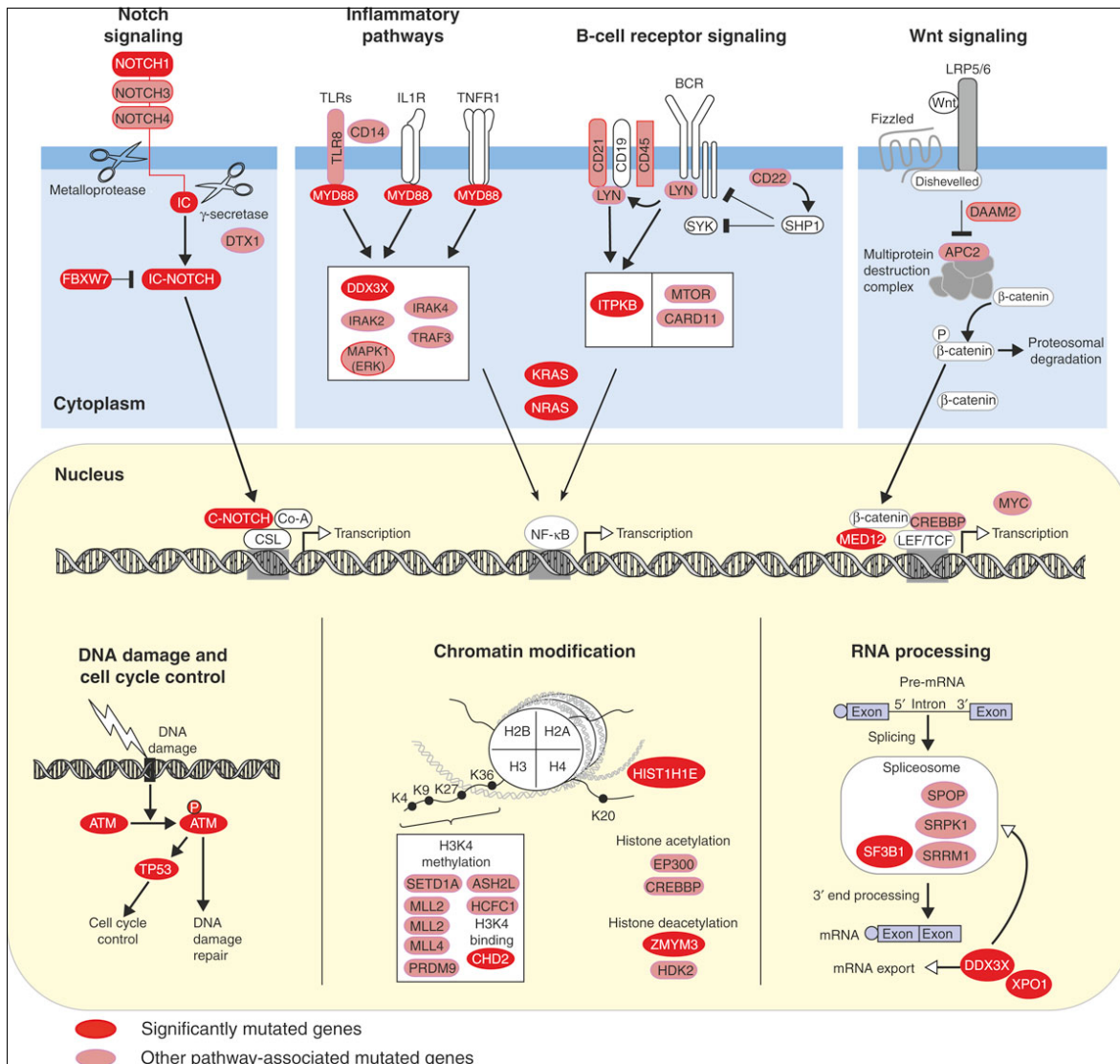

Fig 1

Spliceosome mutations are important driver events in CLL

One of the most unexpected and important findings arising from an unbiased NGS discovery approach was the identification of *SF3B1* as one of the most recurrently mutated genes in CLL [63]. *SF3B1* is a central component of the U2 spliceosome, which orchestrates the excision of introns from pre-mRNA to form mature mRNA [64]. Strikingly, *SF3B1* mutations are found in 10 to 14% of CLLs, particularly in CLL without *IGHV* mutation [34, 40]. This discovery coincided with the report of frequent somatic disruptions of the splicing machinery in myelodysplastic syndrome [65], thereby marking a new important path to oncogenesis in hematological malignancies [66-70], as well as in solid malignancies [71, 72]. The identification of a recurrently mutated gene in both unmutated *IGHV* CLL and myeloid malignancies may hint at a role of dysregulated hematopoietic stem/progenitor cells in some mature lymphoid malignancies [70].

The pathogenic role of *SF3B1* mutations is not only supported by its frequent occurrence in CLL, but also by the fact that mutations cluster in evolutionarily conserved hotspots within its carboxy-terminal repeat HEAT domains, whose function remains unknown [34]. *SF3B1* mutations potentially lead to a defective spliceosome complex that is incapable of performing the correct splicing steps. It has been reported that CLL cells with *SF3B1* mutations show defective splicing activity, with a high ratio of unspliced to spliced *BRD2* and *RIOK3* mRNA, transcripts that have previously been shown to require SF3b spliceosome activity [34, 73]. Elevated levels of truncated mRNA of the transcription factor *FOXP1* and additional proteins that are SF3b spliceosome targets have been reported in association with *SF3B1* mutation [40]. The precise mechanistic aspects of *SF3B1* mutation, however, are still under investigation. Of

note, in addition to *SF3B1* mutations, disruptions of other aspects of RNA processing have been observed in CLL, including recurrent mutations in *DDX3X* and *XPO1* [34], highlighting the importance of RNA processing in CLL.

Patients with *SF3B1*-mutated CLL have a shorter time to treatment, a shorter time to disease progression and lower overall survival rates [34, 40]. These mutations were also found in higher rates in patients with chemo-refractory CLL [69]. Other data indicate that the *SF3B1* mutation may be a later event in CLL, as it was observed to be acquired in patients with relapsed disease [74], or that it expands from a minor subclone to become the dominant subclone upon relapse [75]. Along the same lines, it has been suggested that it is rarely seen in MBL [76], a clonal condition that is thought to precede CLL, although the sample size, particularly of CLL samples with unmutated *IGHV*, may have been too small to adequately address this question. *SF3B1* mutations therefore may have a role in clonal evolution in CLL, emerging later in the disease course, and in relapsed or refractory disease.

Clonal evolution drives CLL progression
One of the main challenges for cancer therapeutics is the plasticity of cancer - its ability to adapt both to host defenses and to treatment. A central component of this plasticity is clonal evolution fueled by the coexistence of multiple subpopulations within the tumor [77]. These concepts were first demonstrated in CLL using cytogenetic technologies [78] and more recently using SNP arrays, which have also shown that relapsed disease is genetically altered compared with disease at diagnosis [23, 79, 80].

With the advent of NGS, clonal evolution has been characterized at unprecedented resolution using whole-genome sequencing of small cohorts of patients with a variety of cancers. In CLL, whole-genome sequencing was performed to track clonal heterogeneity in three CLL patients subjected to repeated cycles of therapy. Notably, three very different temporal patterns of repopulation of the leukemic cell mass emerged after therapy, varying from a stable equilibrium between five subpopulations over the course of years in one patient, to marked shifts, in which one minor subclone replaced the dominant clone entirely, in another. These findings suggest the existence in CLL of an intricate 'ecology' in which a complex interplay is present between intrinsic and extrinsic/environmental factors that control the balance between different subpopulations within the entire CLL population.

Recently, we investigated clonal evolution in CLL by using whole-exome sequencing [75]. The methodologies developed in this study enabled the analysis of a large cohort of samples involving 149 patients, including 18 cases that were followed longitudinally. By studying the allelic fraction of each mutation, the proportion of the subpopulation that harbored it among the entire cancer-cell mass was inferred, and each mutation event was classified as either clonal, meaning a mutation that affects all cancer cells (and corresponding to a founder mutation or an earlier mutation that underwent a complete selective sweep that eliminated all other cancer cells not bearing this mutation), or subclonal, which affects a subpopulation of cancer cells (representing events acquired later in the disease course).

This framework enabled the inference of the temporal order of genetic driver events in CLL, with the identification of earlier (for example, *MYD88* mutation) and later events (for example, *TP53* mutation) in disease progression. We also tracked clonal evolution longitudinally in 18 patients [75], observing that patients who received therapy had a higher rate of clonal evolution, suggesting that perhaps chemotherapy itself can hasten the evolutionary process. Finally, clonal heterogeneity was linked to adverse clinical outcome, adding a further dimension to current efforts to link discrete somatic mutations to outcome. These findings suggest that it is not only the presence or absence of a mutation that should be considered in analyses of the impact of mutations on clinical outcome, but also the size of the subpopulation a mutation affects. This finding has important clinical implications that can be tested in prospective clinical trials.

Beyond somatic genetic alterations: epigenetic changes in CLL
Cancer has traditionally been viewed as a disease driven by the accumulation of genetic mutations. This paradigm has been increasingly modified as cumulative evidence has suggested that the disruption of epigenetic regulatory mechanisms has a critical role in neoplastic transformation. In CLL, for example, epigenetic modifications have been implicated in the recurrent microRNA deregulation observed in *miR-15a/16* and the related *miR-29b*. Histone deacetylases were shown to be overexpressed in CLL, and mediate the epigenetic silencing of microRNAs through removal of the activating chromatin modification H3K4me2.

Perhaps the best-studied epigenetic modification in CLL has been direct DNA methylation, which occurs at the cytosine residue of the CpG dinucleotide in mammalian genomes. Patterns of DNA methylation can be inherited across generations of somatic cells as they are stably maintained through somatic cell division. This type of epigenetic alteration is at least as common as mutational events in the development of cancer. Published reports of epigenetic gene dysregulation in CLL include hypomethylation of *BCL2* and *TCL1* as well as silencing of *DAPK1* through promoter hypermethylation, which recapitulates a germline mutation found in a kindred of familial CLL

More recently, genome-wide platforms have been applied to the study of DNA methylation in CLL. DNA methylation arrays detect representative methylation sites across the entire genome and have been used to identify regions with differential methylation in CLL samples with mutated or unmutated *IGHV* status. Most of these differentially methylated regions have been reported to lie outside CpG islands, to remain stable over time and to involve multiple genes important in CLL biology, such as *ZAP70, NOTCH1* and *IBTK*, as well as epigenetic regulators (such as *DNMT3B*) and NF-κB/tumor necrosis factor (TNF) pathway genes. Similar investigations were performed comparing CLL samples with high and low CD38 expression, and found variable methylation in the *DLEU7* gene. Finally, pervasive methylation changes have been observed across numerous microRNA sites in CLL samples compared with normal B cells, which were associated with large changes in expression of these microRNAs.

Bisulfite conversion coupled with NGS has also been used to delineate DNA methylation across the entire genome at base-pair resolution.Using this method, methylation profiles have been shown to vary substantially between CLL with mutated versus unmutated *IGHV* status and to mirror

epigenetic differences seen between naive and memory B cells. The methylation patterns observed in the study allowed the authors to identify, in addition to the mutated and unmutated *IGHV* subsets, a third subset of CLL samples with distinct clinical behavior (an intermediate prognosis group, with a better prognosis than patients with *IGHV*-unmutated CLL and a worse prognosis than those with *IGHV* mutations), and an intermediate level of *IGHV* somatic hypermutation. Another method using bisulfite conversion focuses on a representative sample of CpG sites termed reduced representation bisulfite sequencing (RRBS). This method has been found to be highly informative, and is less costly than whole-genome bisulfite conversion. The application of RRBS to CLLhas shown that differentially methylated regions are enriched for transcription factors, including the homeobox family of proteins. Furthermore, DNA methylation serves to enhance particular critical pathways in CLL, such as Wnt signaling, by the simultaneous hypermethylation of pathway antagonists (for example, *DKK*) and hypomethylation of Wnt ligands and transcription factors (for example, *TCF7*), with the net result of decreased antagonist transcription and increased agonist transcription, respectively. Collectively, these studies have shown that DNA methylation probably plays a significant role in CLL biology.

Profiling the transcriptional landscape of CLL to understand the impact of genetic and epigenetic alterations on the cellular network
The various genetic and epigenetic alterations described earlier can affect the cellular network and lead to system-wide transcriptional changes. Studying the transcriptome enables an understanding of how mutations alter cellular behavior, and this should give a better idea of the ultimate phenotype. Expression arrays have been used to study CLL for many years in an effort to define subtypes related to clinical outcomes (reviewed extensively elsewhere. These methodologies have also been used to classify different subtypes (for example, *IGHV*-mutated versus *IGHV*-unmutated) as well as to try and identify the normal cellular counterpart of CLL (that is, the closest normal B-cell phenotype that may serve as a cell of origin for CLL).
A systems-level examination of the transcriptional landscape of CLL has the potential to reveal subsets of patients with disparate risks for CLL progression. By studying individual pathway disruptions, these pathways were shown to converge as patients progressed before treatment and to assume similar transcriptional profiles closer to the point at which they required treatment.Thus, the transcriptional profile of CLL can be reduced from a daunting number of individual genes to a handful of meaningful pathway annotations with important biological and clinical implications.
High-throughput RNA sequencing has enabled the harnessing of NGS technology for the study of transcriptional profiles. A pilot study compared RNA-sequencing data from a small number of samples with mutated versus unmutated *IGHV*, the most well-established prognostic factor in CLL. In addition to identifying 156 differentially expressed genes, the study identified a large number of differentially expressed non-coding RNAs as well as marked changes in splice variants between the two prognostic groups. Thus, this methodology is capable of providing a wealth of information in comparison with

microarray-based gene expression profiling, with the potential to demonstrate how genetic and epigenetic changes translate at the cellular network level.

Conclusions and future directions
The intensive application of NGS to the study of CLL has yielded remarkable insights over a short period of time, and it is likely that the exponential growth in our understanding of this disease will continue in the coming years. The use of these novel technologies has identified expected (for example, *TP53* and *ATM* mutations) and unexpected CLL drivers (for example, *SF3B1*), and has opened new avenues of research, such as the study of splicing abnormalities (Figure 2). NGS has also revealed the tremendous degree of genetic heterogeneity in CLL, both among patients and within individual leukemias over time.
Delineating the inter-patient genetic heterogeneity of CLL has high translational potential. First, novel genetic abnormalities such as *NOTCH1, SF3B1* and *BIRC3* mutations carry prognostic significance, and will probably be used in the future to predict the highly variable clinical course of CLL, beyond the established predictive factors such as *IGHV* mutation status and cytogenetic abnormalities [46]. Second, these lesions may also be informative regarding treatment stratification - similar to the use of *TP53* disruption today, which is known to be associated with chemo-refractory disease [2]. Finally, some of the genetic lesions identified by NGS represent attractive candidates for targeted therapy. NOTCH1, for example, is already being targeted by some drugs under development. The promising results obtained with inhibitors of BCR signaling (that is, the BTK inhibitor ibrutinib and the PI3K-δ inhibitor GS-1101 [20]) suggest that future research should also focus on how these drugs affect CLL cells with different driver lesions.
The emerging understanding of intra-tumoral genetic heterogeneity in CLL may also eventually have a clinical impact. Studying clonal evolution in relation to therapy could help us to refine our understanding of resistance mechanisms and repopulation kinetics. For example, studying the genomes of relapsed CLL compared with pre-treatment CLL patients could be informative with respect to specific lesions or mutations that are selected *in vivo* in the setting of therapeutic bottlenecks. Collecting multiple longitudinal samples throughout the disease and treatment process could highlight the comparative kinetics of different subpopulations, enhancing our understanding of the evolutionary process. It will also enable us to gain an understanding of the impact of targeting early clonal lesions compared with late aggressive subclonal drivers on therapeutic outcome. Finally, the suggestion that therapy itself can accelerate clonal evolution could influence the current paradigm of gene-specific discovery, by challenging us to conceive therapeutic strategies to directly address and anticipate clonal evolution, which has been demonstrated to affect clinical outcome [75].
Future directions for NGS-based studies will probably also include studying the entire continuum of CLL, from MBL to Richter's transformation [39, 61]. Studying MBL may be particularly informative regarding the nascent stages of CLL and the critical genetic steps required for transformation to CLL. In addition, focusing on distinct groups of patients, such as those with poor clinical outcome (rapid progression and poor treatment response), would assist in defining the

genetic elements that contribute to disease heterogeneity. Some of these have already been identified, such as the long-established role of mutations in *TP53* and *ATM*, as well as the more recent identification of the poor prognostic significance of *SF3B1* and *BIRC3* mutations. However, it is likely that other somatic events or specific mutation combinations can affect clinical phenotype, and a comprehensive mapping of these elements will improve prognostication. Pathway analysis, as portrayed in Figure 2, may also unravel how disruption of different parts of the cellular machinery can translate into altered clinical outcome.

Moreover, these technologies are likely to be applied to studying inherited predisposition for CLL, as this disease has a high incidence of familial cases. This area of investigation might provide important clues to the interaction between existing germline mutations and acquired somatic mutagenesis. Finally, probing the epigenetic profile of CLL is currently in its nascent stages and will likely lead to a better understanding of genome-wide levels of epigenetic modifications, as well as how different populations within the cancer-cell mass differ in their epigenetic profiles and how this affects functional diversity. For example, these epigenetic differences might lead to variations in proliferative capacity, pluripotent potential or ability to resist therapy.

Ultimately, a comprehensive understanding of the genetic basis of CLL will assist in stratifying patients and matching treatments with genetic lesions, with a goal of developing targeted therapies to improve CLL management. The wealth of emerging genetic data has great potential to provide new paths for improved treatment options for this disease, and will require focused translational efforts to enable the application of this knowledge into clinical care.

Acknowledgements

References
1. Parker TL, Strout MP. Chronic lymphocytic leukemia: prognostic factors and impact on treatment. Discov Med. 2011; 11:115-123.
2. Hallek M, Cheson B, Catovsky D, Caligaris-Cappio F, Dighiero G, Döhner 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; 111:5446-5456.
3. Shanafelt TD, Ghia P, Lanasa MC, Landgren O, Rawstron AC. Monoclonal B-cell lymphocytosis (MBL): biology, natural history and clinical management. Leukemia. 2010; 24:512-520.
4. Tsimberidou A, Keating M, Wierda W. Richter's transformation in chronic lymphocytic leukemia. Oncology. 2007; 2:265-271
5. Chiorazzi N, Rai KR, Ferrarini M. Chronic lymphocytic leukemia. N Engl J Med. 2005; 352:804-815.
6. Zenz T, Mertens D, Küppers R, Döhner H, Stilgenbauer S.From pathogenesis to treatment of chronic lymphocytic leukaemia. Nat Rev Cancer. 2010; 10:37-50.
7. Zenz T, Mertens D, Dohner H, Stilgenbauer S. Importance of genetics in chronic lymphocytic leukemia. Blood Rev. 2011; 25:131-137.
8. Calin G, Dumitru C, Shimizu M, Bichi R, Zupo S, Noch E, *et al.* Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc N atl AcadSci USA. 2002; 99:15524-15529.
9. Döhner H, Stilgenbauer S, Benner A, Leupolt E, Kröber A, Bullinger L *et al.* Genomic aberrations and survival in chronic lymphocytic leukemia. N Engl J Med. 2000; 343:1910-1916.
10. Hillmen P. Using the biology of chronic lymphocytic leukemia to choose treatment. Hematology Am Soc Hematol Educ Program. 2011; 2011:104-109.
11. Brown JR. The treatment of relapsed refractory chronic lymphocytic leukemia. Hematology Am Soc Hematol Educ Program. 2011, 110-118.
12. Lamanna N. Treatment of older patients with chronic lymphocytic leukemia. CurrHematolMalig Rep. 2012; 7:21-25.
13. CLL Trialists Collaborative Group: Chemotherapeutic options in chronic lymphocytic leukemia: a meta-analysis of the randomized trials. J Natl Cancer Inst. 1999; 91:861-868.
14. Hallek M, Fischer K, Fingerle-Rowson G, Fink AM, Busch R, Mayer J *et al.*Addition of rituximab to fludarabine and cyclophosphamide in patients with chronic lymphocytic leukaemia: a randomised, open-label, phase 3 trial. Lancet. 2010; 376:1164-1174.
15. Lozanski G, Heerema NA, Flinn IW, Smith L, Harbison J, Webb J, *et al.* Alemtuzumab is an effective therapy for chronic lymphocytic leukemia with mutations and deletions. Blood. 2004; 103(53):3278-3281.
16. Sorror M, Storer B, Sandmaier B, Maris M, Shizuru J, Maziarz R, *et al.* Five-year follow-up of patients with advanced chronic lymphocytic leukemia treated with allogeneic hematopoietic cell transplantation after nonmyeloablative conditioning. J ClinOncol. 2008; 26:4912-4920.
17. Sorror ML, Storer BE, Maloney DG, Sandmaier BM, Martin PJ, Storb R. Outcomes after allogeneic hematopoietic cell transplantation with nonmyeloablative or myeloablative conditioning regimens for treatment of lymphoma and chronic lymphocytic leukemia. Blood. 2008; 111:446-452.
18. Badoux XC, Keating MJ, Wen S, Lee BN, Sivina M, Reuben J *et al.* Lenalidomide as initial therapy of

elderly patients with chronic lymphocytic leukemia. Blood. 2011; 118:3489-3498.

19. Fischer K, Cramer P, Busch R, Bottcher S, Bahlo J, Schubert J *et al.* Bendamustine in combination with rituximab for previously untreated patients with chronic lymphocytic leukemia: a multicenter phase II trial of the German Chronic Lymphocytic Leukemia Study Group. J Clin Oncol. 2012; 30:3209-3216.

20. Woyach JA, Johnson AJ, Byrd JC. The B-cell receptor signaling pathway as a therapeutic target in CLL. Blood. 2012; 120:1175-1184.

21. Sellner L, Dietrich S, Dreger P, Glimm H, Zenz T. Can prognostic factors be used to direct therapy in chronic lymphocytic leukemia. Curr Hematol Malig Rep. 2012; 7:3-12.

22. Schnaiter A, Mertens D, Stilgenbauer S. Genetics of chronic lymphocytic leukemia. Clin Lab Med. 2011; 31:649-658.

23. Grubor V, Krasnitz A, Troge J, Meth J, Lakshmi B, Kendall J *et al.* Novel genomic alterations and clonal volution in chronic lymphocytic leukemia revealed by representational oligonucleotide microarray analysis (ROMA). Blood. 2009; 113:1294-1303.

24. Pfeifer D, Pantic M, Skatulla I, Rawluk J, Kreutz C, Martens U *et al.* Genome-wide analysis of DNA copy number changes and LOH in CLL using high-density SNP arrays. Blood. 2007; 109:1202-1210.

25. Schwaenen C, Nessling M, Wessendorf S, Salvi T, Wrobel G, Radlwimmer B *et al.* Automated array-based genomic profiling in chronic lymphocytic leukemia: development of a clinical tool and discovery of recurrent genomic alterations. Proc Natl Acad Sci USA. 2004; 101:1039-1044.

26. Brown JR, Hanna M, Tesar B, Werner L, Pochet N, Asara JM *et al.* Integrative genomic analysis implicates gain of PIK3CA at 3q26 and MYC at 8q24 in chronic lymphocytic leukemia. Clin Cancer Res. 2012; 18:3791-3802.

27. Edelmann J, Holzmann K, Miller F, Winkler D, Buhler A, Zenz T *et al.* High-resolution genomic profiling of chronic lymphocytic leukemia reveals new recurrent genomic alterations. Blood. 2012; 120:4783-4794.

28. Cimmino A, Calin G, Fabbri M, Iorio M, Ferracin M, Shimizu M *et al.* miR-15 and miR-16 induce apoptosis by targeting BCL2. Proc Natl Acad Sci USA. 2005; 102:13944-13949.

29. Klein U, Lia M, Crespo M, Siegel R, Shen Q, Mo T, *et al.* The DLEU2/miR-15a/16-1 cluster controls B cell proliferation and its deletion leads to chronic lymphocytic leukemia. Cancer Cell. 2010; 17:28-40.

30. Ouillette P, Erba H, Kujawski L, Kaminski M, Shedden K, Malek S. Integrated genomic profiling of chronic lymphocytic leukemia identifies subtypes of deletion 13-14. Cancer Res. 2008; 68:1012-1021.

31. Kujawski L, Ouillette P, Erba H, Saddler C, Jakubowiak A, Kaminski M *et al.* Genomic complexity identifies patients with aggressive chronic lymphocytic leukemia. Blood. 2008; 112:1993-2003.

32. Gunnarsson R, Isaksson A, Mansouri M, Goransson H, Jansson M, Cahill N *et al.* Large but not small copy-number alterations correlate to high-risk genomic aberrations and survival in chronic lymphocytic leukemia: a high-resolution genomic screening of newly diagnosed patients. Leukemia. 2010; 24:211-215.

33. 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; 475:101-105.

34. 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; 365:2497-2506.

35. Meyerson M, Gabriel S, Getz G. Advances in understandingcancergenomesthroughsecond-generation sequencing. Nat Rev Genet. 2010; 11:685-696.

36. Mardis E, Ding L, Dooling D, Larson D, McLellan M, Chen K *et al.* Recurring mutations found by sequencing an acute myeloid leukemia genome. N Engl J Med. 2009; 361:1058-1066.

37. Parsons D, Jones S, Zhang X, Lin J, Leary R, Angenendt P *et al.* An integrated genomic analysis of human glioblastomamultiforme. Science. 2008; 321:1807-1812.

38. Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, Mudie LJ *et al.* Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell. 2011; 144:27-40.

39. 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; 208:1389-1401.

40. 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. 2012; 44:47-52.

41. Wood L, Parsons D, Jones S, Lin J, Sjöblom T, Leary R *et al.* The genomic landscapes of human breast and colorectal cancers. Science. 2007; 318:1108-1113.

42. Oberg C, Li J, Pauley A, Wolf E, Gurney M, Lendahl U. The Notch intracellular domain is ubiquitinated and negatively regulated by the mammalian Sel-10 homolog. J Biol Chem. 2001; 276:35847-35853.

43. Aster JC, Pear WS, Blacklow SC. Notch signaling in leukemia. Annu Rev Pathol. 2008; 3:587-613.

44. Balatti V, Bottoni A, Palamarchuk A, Alder H, Rassenti LZ, Kipps TJ *et al.* NOTCH1 mutations in CLL associated with trisomy 12. Blood. 2012; 119:329-331.

45. Mansouri L, Cahill N, Gunnarsson R, Smedby KE, Tjonnfjord E, Hjalgrim H *et al.* TCH1 and SF3B1 mutations can be added to the hierarchical prognostic classification in chronic lymphocytic leukemia. Leukemia. 2013; 27:512-514.

46. Rossi D, Rasi S, Spina V, Bruscaggin A, Monti S, Ciardullo C *et al.* Integrated mutational and cytogenetic analysis identifies new prognostic subgroups in chronic lymphocytic leukemia. Blood. 2013; 121:1403-1412.

47. Villamor N, Conde L, Martinez-Trillos A, Cazorla M, Navarro A, Bea S *et al.* NOTCH1 mutations identify a genetic subgroup of chronic lymphocytic leukemia patients with high risk of transformation and poor outcome. Leukemia. 2012; 27:1100-1106.

48. Shedden K, Li Y, Ouillette P, Malek SN. Characteristics of chronic lymphocytic leukemia with somatically acquired mutations in NOTCH1 exon 34. Leukemia. 2012; 26:1108-1110.

49. Lin SC, Lo YC, Wu H. Helical assembly in the MyD88-IRAK4-IRAK2 complex in TLR/IL-1R signalling. Nature. 2010; 465:885-890.

50. Kawagoe T, Sato S, Matsushita K, Kato H, Matsui K, Kumagai Y *et al.* Sequential control of Toll-like receptor-dependent responses by IRAK1 and IRAK2. Nat Immunol. 2008; 9:684-691.

51. Ngo VN, Young RM, Schmitz R, Jhavar S, Xiao W, Lim KH, *et al.* Oncogenically active MYD88 mutations in human lymphoma. Nature. 2011; 470:115-119.

52. Gonzalez-Aguilar A, Idbaih A, Boisselier B, Habbita N, Rossetto M, Laurenge A *et al.* Recurrent mutations of MYD88 and TBL1XR1 in primary central nervous system lymphomas. Clin Cancer Res. 2012; 18:5203-5211.

53. Treon SP, Xu L, Yang G, Zhou Y, Liu X, Cao Y *et al.* MYD88 L265P somatic mutation in Waldenström's macroglobulinemia. N Engl J Med. 2012; 367:826-833.

54. Treon SP, Hunter ZR, Matous J, Joyce RM, Mannion B, Advani R, *et al.* Multicenter clinical trial of bortezomib in relapsed/refractory Waldenstrom's macroglobulinemia: results of WMCTG Trial 03-248. Clin Cancer Res. 2007; 13:3320-3325.

55. Qiu Y, Kung HJ. Signaling network of the Btk family kinases. Oncogene. 2000; 19:5651-5661.

56. Ramsay AJ, Rodriguez D, Villamor N, Kwarciak A, Tejedor JR, Valcarcel J *et al.* Frequent somatic mutations in components of the RNA processing machinery in chronic lymphocytic leukemia. Leukemia. 2012, doi: 10.1038/leu.2012.344

57. Shalek AK, Gaublomme JT, Wang L, Yosef N, Chevrier N, Andersen MS *et al.* Nanowire-mediated delivery enables functional interrogation of primary immune cells: application to the analysis of chronic lymphocytic leukemia. Nano Lett. 2012; 12:6498-6504.

58. Lu D, Zhao Y, Tawatao R, Cottam HB, Sen M, Leoni LM, *et al.* Activation of the Wnt signaling pathway in chronic lymphocytic leukemia. Proc Natl Acad Sci USA. 2004; 101:3118-3123.

59. Rossi D, Fangazio M, Rasi S, Vaisitti T, Monti S, Cresta S *et al.* Disruption of BIRC3 associates with fludarabine chemo refractoriness in TP53 wild-type chronic lymphocytic leukemia. Blood. 2012; 119:2854-2862.

60. Vallabhapurapu S, Karin M. Regulation and function of NF-κB transcription factors in the immune system. Annu Rev Immunol. 2009; 27:693-733.

61. Gaidano G, Foa R, Dalla-Favera R. Molecular pathogenesis of chronic lymphocytic leukemia. J Clin Invest. 2012; 122:3432-3438.

62. Rossi D, Deaglio S, Dominguez-Sola D, Rasi S, Vaisitti T, Agostinelli C *et al.* Alteration of BIRC3 and multiple other NF-κB pathway genes in splenic marginal zone lymphoma. Blood. 2011; 118:4930-4934.

63. Wan Y, Wu CJ. SF3B1 mutations in chronic lymphocytic leukemia. Blood. 2013.doi: 10.1182/blood-2013-02-427641,

64. Wahl MC, Will CL, Luhrmann R. The spliceosome: design principles of a dynamic RNP machine. Cell. 2009; 136:701-718.

65. Yoshida K, Sanada M, Shiraishi Y, Nowak D, Nagata Y, Yamamoto R *et al.* Frequent pathway mutations of splicing machinery in myelodysplasia. Nature. 2011; 478:64-69.

66. Makishima H, Visconte V, Sakaguchi H, Jankowska AM, Abu Kar S, Jerez A, *et al.* Mutations in the spliceosome machinery, a novel and ubiquitous pathway in leukemogenesis. Blood. 2012; 119:3203-3210.

67. Graubert TA, Shen D, Ding L, Okeyo-Owuor T, Lunn CL, Shao J *et al.* Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. Nat Genet. 2012; 44:53-57.

68. Papaemmanuil E, Cazzola M, Boultwood J, Malcovati L, Vyas P, Bowen D *et al.* Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. N Engl J Med. 2011; 365:1384-1395.

69. 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; 118:6904-6908.

70. Damm F, Thol F, Kosmider O, Kade S, Loffeld P, Dreyfus F *et al.* SF3B1 mutations in myelodysplastic syndromes: clinical associations and prognostic implications. Leukemia. 2012; 26:1137-1140.

71. Imielinski M, Berger AH, Hammerman PS, Hernandez B, Pugh TJ, Hodis E *et al.* Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. Cell. 2012; 150:1107-1120.

72. Harbour JW, Roberson ED, Anbunathan H, Onken MD, Worley LA, Bowcock AM. Recurrent mutations at codon 625 of the splicing factor SF3B1 in uveal melanoma. Nat Genet. 2013; 45:133-135.

73. Kotake Y, Sagane K, Owa T, Mimori-Kiyosue Y, Shimizu H, Uesugi M *et al.* Splicing factor SF3b as a target of the antitumor natural product pladienolide. Nat Chem Biol. 2007; 3:570-575.

74. Cazzola M, Rossi M, Malcovati L. Biologic and clinical significance of somatic mutations of SF3B1 in myeloid and lymphoid neoplasms. Blood. 2013; 121:260-269.

75. Landau D, Carter S, Stojanov P, McKenna A, Stevenson K, Lawrence M *et al.* Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. Cell. 2013; 152:714-726.

76. Greco M, Capello D, Bruscaggin A, Spina V, Rasi S, Monti S *et al.* Analysis of SF3B1 mutations in monoclonal B-cell lymphocytosis. HematolOncol. 2013; 31:54-55.

77. Nowell PC. The clonal evolution of tumor cell populations. Science. 1976; 194:23-28.

78. Shanafelt TD, Hanson C, Dewald GW, Witzig TE, LaPlant B, Abrahamzon J *et al.* Karyotype evolution on fluorescent in situ hybridization analysis is associated with short survival in patients with chronic lymphocytic leukemia and is related to CD49d expression. J ClinOncol. 2008; 26:e5-6.

79. Mullighan CG, Phillips LA, Su X, Ma J, Miller CB, Shurtleff SA *et al.* Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. Science. 2008; 322:1377-1380.

80. Ouillette P, Saiya-Cork K, Seymour EK, Li C, Shedden K, Malek SN. Clonal evolution, genomic drivers andeffects of therapy in chronic lymphocytic leukemia. Clin Cancer Res. 2013, doi:10.1158/1078-0432.CCR-13-0138.