

Published in final edited form as:

*Semin Hematol.* 2014 July ; 51(3): 177–187. doi:10.1053/j.seminhematol.2014.05.004.

## Evolving Understanding of the CLL Genome

Michaela Gruber, MD, PhD<sup>1,2,3</sup> and Catherine J. Wu, MD<sup>1,2,4</sup>

<sup>1</sup>Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA

<sup>2</sup>Broad Institute, Cambridge, MA, USA

<sup>3</sup>Department of Internal Medicine I, Division of Hematology and Haemostaseology, Medical University of Vienna, Austria

<sup>4</sup>Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

### Abstract

Over the past few years, massively parallel sequencing technologies have revealed with high resolution the tremendous genetic and epigenetic heterogeneity in chronic lymphocytic leukemia (CLL). We have learned how the molecular architecture differs not only between affected individuals but also within samples and over time. These insights have catalyzed our understanding of the pathobiology of CLL and point to critical signaling pathways in the development and progression of the disease. Several key driver alterations have been identified, which serve to refine prognostic schemata but also inspire the development of new therapeutic strategies. Ongoing advances in technology promise to further elucidate the molecular basis of CLL, and this knowledge is anticipated to aid us in understanding and addressing the clinical challenge presented by the vast variability in the clinical course of patients with CLL.

---

A hallmark of chronic lymphocytic leukemia (CLL) is its tremendously variable clinical course. As many as 80% of CLL patients are asymptomatic at diagnosis, but many progress to extensive lymphadenopathy, hepatosplenomegaly, and life-threatening cytopenias within only a few years. Others, however, remain asymptomatic over decades with 20–30% having a life expectancy not significantly different from the general population.<sup>1,2</sup>

An enduring goal of CLL studies has been to better understand the basis of this clinical variability. Of note, because of its high prevalence, relatively slow progression, and the ready availability of leukemia samples from patient peripheral blood, CLL has been

---

© 2014 Elsevier Inc. All rights reserved.

**Corresponding Author:** Catherine J. Wu, MD, Dana-Farber Cancer Institute, Department of Medical Oncology, 450 Brookline Avenue, Dana 540B, Boston, MA 02215, Phone: (617) 632-5943, Fax: (617) 632-6380, cwu@partners.org.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

### Conflicts of Interest

The authors declare that they have no conflicts of interest or competing financial or personal relationships that could inappropriately influence the content of this article.

continuously at the forefront of genomic research. Thus, while the first prognostic schema, established in the 1970s,<sup>3,4</sup> was based on clinical features, newer studies have focused on the role of somatic genomic alterations in the pathogenesis of CLL and in turn, have examined their impact on clinical outcome. For example, mutational status of the immunoglobulin heavy chain variable-region gene (*IGHV*) divides CLL into two genetically distinct groups, which likely reflects different cells of origin and has emerged as an important disease-associated prognostic factor.<sup>5,6</sup> In a separate landmark study, presence of four common CLL-associated cytogenetic aberrations, deletions of chromosomes 11q, 13q and 17q as well as trisomy of chromosome 12, could stratify CLL patients into prognostically distinct groups and, for the first time, was linked with clinical course and survival.<sup>7</sup>

Remarkably, the introduction of next-generation sequencing has led to a breathtakingly exponential increase in the knowledge of the molecular underpinnings of CLL over the last three years (Figure 1). As described in this review, these recent studies have revealed several notable, even surprising and paradigm-shifting insights such that our perception about this disease has greatly evolved in a relatively short span of time. First, whole-exome sequencing (WES) of large sample cohorts have clearly shown the high degree of genetic variability amongst CLL patients (*Section IIA*), with the discovery of novel common gene mutations that likely play a role in the pathobiology of CLL. Furthermore, a startlingly high degree of intra-sample clonal heterogeneity was recently uncovered, based on the detection and quantification of leukemia-specific alterations that mark various cell subpopulations within a CLL sample (*Section IIB*). At the same time, genome-wide approaches to examine the DNA methylome have revealed epigenetic heterogeneity amongst patients (*Section IIIA*) and within individual samples (*Section IIIB*). Collectively, these observations demonstrate the complex interrelationship between genetic and epigenetic features of each sample. These new discoveries are undergoing evaluation as features to incorporate into new prognostic and therapeutic schema that promise to improve the clinical care of patients with CLL (*Section IV*).

## II. Profiling Genetic Heterogeneity in CLL

Massively-parallel sequencing (MPS) techniques have provided the ability to rapidly sequence millions of DNA fragments with relatively low sample input. As an alternative to whole genome sequencing (WGS), selective restriction of reads to the coding regions of the genome by WES has drastically reduced the costs of sequencing per sample. This latter approach has facilitated the rapid sequencing of large sample cohorts, which in turn has enabled the drawing of associations between genetic alterations and clinical features. Over the last few years, the results of up to a dozen whole CLL genomes<sup>8-10</sup> and more than 300 whole CLL exomes<sup>11-14</sup> across different centers worldwide have been reported.

### IIa. Unbiased Discovery of Key CLL Drivers

A major goal of the large-scale cancer sequencing studies has been the identification of key alterations that drive malignancy. The development of massively parallel sequencing has led the parallel development of advanced computational algorithms for analyzing these large datasets. In general, these algorithms detect cancer-specific alterations with a high

probability of being cancer drivers on the basis of whether they are present at a significantly higher than expected rate given the known background mutation rate of the cancer. In CLL, these efforts have corroborated known CLL-associated alterations (i.e. mutations in *TP53* and *ATM*), but importantly have identified numerous previously unknown somatic changes, the majority of which have been confirmed across independent sample cohorts. The first studies of DNA sequencing in CLL found mutations in *MYD88*, *NOTCH1*, *XPO1*, as well as in *BIRC3*.<sup>8,11</sup> Subsequently, results of WES of two well-powered cohorts of approximately 100 patients each further detected several novel somatic alterations in CLL (i.e. in *FBXW7*, *POT1*, *CHD2*). Strikingly, both studies identified the novel finding of recurrent mutations in the splicing machinery cofactor *SF3B1* in 10–15% of patients.<sup>9,15</sup> Most recently, the largest single CLL sequencing cohort to date was reported, comprising 160 patients, in which numerous lower frequency mutations (i.e. in *NRAS*, *KRAS*, *HIST1H1E*, *SAMHD1* and *MED12*) were identified.<sup>13</sup>

These somatic alterations are present in critical components of a number of cellular pathways, and include DNA damage and cell cycle control (*TP53*, *ATM*, *POT1*, *BIRC3*), mRNA processing (*XPO1*, *SF3B1*), NOTCH signaling (*NOTCH1*), inflammatory pathways (*MYD88*), as well as chromatin modification (*CHD2*) (Table 1).<sup>9,13,15</sup> Several of the significantly mutated genes display a clustering of mutations in hotspots within highly evolutionarily conserved gene regions, and strongly support the idea that they are positively selected gain-of-function alterations. For example, mutation in *MYD88*, a critical adaptor molecule of the interleukin-1 receptor–toll-like receptor (TLR) signaling pathway, has been found almost exclusively at position L265P, which is localized within the interleukin-1 receptor–TLR domain. It is noteworthy that this particular mutation appears to occur selectively in B cell malignancies, most prominently in Waldenström's macroglobulinaemia and DLBCL.<sup>16,17</sup> In another example, *NOTCH1*, a transmembrane protein central to the Notch signaling pathway, is recurrently affected by a 2 bp frameshift deletion CLL in the C-terminal PEST domain,<sup>11</sup> leading to pathway activation, increased cell survival and resistance against pro-apoptotic stimuli.<sup>18</sup> *SF3B1* encodes the core catalytic subunit of the spliceosome complex and its mutations localize to 900 basepairs within the C-terminal region<sup>9,15,19</sup> and have been noted to impact splicing at 3' splice sites.<sup>20,21</sup> Another recurrently mutated gene affecting RNA processing is the nuclear transport gene *XPO1* with mutations clustering at a highly conserved site at residue E571K<sup>8,9,22,23</sup> Finally, the shelterin *POT1* encodes a protein essential for telomere function, of which recurrent mutations in CLL affect key residues required to bind telomeric DNA and lead to substantial telomeric dysfunction associated with increased genomic instability and numerous chromosomal abnormalities.<sup>14</sup>

The significantly mutated CLL genes also include examples of tumor-suppressor genes (*TP53*, *BIRC3*, *ATM*). *TP53* is furthermore involved in the region of chromosome 17p, and *BIRC3* and *ATM* at 11q, which are often found deleted in CLL and which correspond to poor prognosis.<sup>9,13,24</sup>

Further clues on the functional role of alterations can be inferred based on the patterns of co-segregation and mutual exclusivity (Table 1). Interestingly, the significantly mutated genes in CLL seem to be differentially represented between the *IGHV* mutated and unmutated

CLLs. While the former appears to be associated with del(13q) and mutations in *MYD88*, the latter commonly display mutations affecting *NOTCH1*, *SF3B1* and *ATM* and *TP53* and associations with trisomy 12, del(11q) and del(17p), respectively.<sup>9,13,22,25–27</sup> Lesions of *BIRC3* and *TP53* have been noted to occur in a mutually exclusive fashion. Likewise, mutations in *SF3B1*, *NOTCH1* and *MYD88* also appear to be exclusive of each other. These patterns suggest possible distinct evolutionary paths, whereby certain subclonal alterations may confer advantage when occurring in the genomic context of particular ancestor lesions. Alternatively, mutual exclusivity could indicate that alterations have highly similar downstream effects, thus, functionally redundant secondary mutations do not provide any further advantage to the tumor cell. On the other hand, consistent co-occurrence suggests synergistic effects between alterations that enhance fitness of the malignant clone and promote selection of driver combinations.

As the numbers of studies examining the incidence of these key mutations in CLL have grown, it has become also clear that their frequency in patient groups largely depends on the composition of the sequenced cohort. Thus, while mutation frequency in *SF3B1* ranges between 4–12% in early CLL, it rises to 17–24% of patients by the time of disease progression.<sup>9,24,25</sup> Similarly, mutations in *TP53*, *ATM*, and *NOTCH1* have higher incidence in cohorts with advanced disease across independent studies.<sup>9,27–29</sup> The mutation rate of *NOTCH1* is further markedly increased in patients with lymphomatous transformation.<sup>12</sup> By contrast, mutation rates of *MYD88* appear stable throughout the course of CLL (Figure 2).<sup>9,30</sup>

Altogether, these patterns of association between distinct genetic alterations suggest the presence of different potential trajectories in the development of the CLL genome. Moreover, the patterns of genetic alterations are distinct in each individual CLL, and demonstrate the tremendous inter-individual heterogeneity in CLL (Figure 3A).

## IIb. Intraleukemic Genetic Heterogeneity and Clonal Evolution

Even as we have gained greater understanding of variation in the spectrum of genetic alterations amongst patients with CLL, analysis of WES data has also revealed the extensive degree of genetic heterogeneity within individual samples. The existence of genetically distinct subpopulations was already suggested by studies using fluorescence *in situ* hybridization, in which many chromosomal alterations were observed only in subfractions of cells.<sup>7</sup> Moreover, the number of genomic alterations was shown to increase throughout the course from newly diagnosed to progressive and further to relapsed CLL.<sup>31–33</sup> Application of whole-exome and whole-genome sequencing methodologies however have provided more precise quantitation and more global assessment of this phenomenon.

Using these technologies, a series of recent studies have indicated that a single time point genetic profile of a CLL sample represents a snapshot of multiple different tumor cell populations that are related to each other and changing over time. By integrating information on the allelic frequencies of mutations together with local copy number and purity information, Landau et al recently demonstrated the possibility of inferring proportions of cell subpopulations harboring a genetic alteration (Figure 3B). Interestingly, certain somatic mutations were detected preferentially in clonal or subclonal fashion, suggesting an order of

alterations corresponding to earlier and later drivers. Presence of a subclonal driver by itself also indicated poor prognosis and more rapid disease progression. Furthermore, of the 149 patients study, a subset of samples assessed longitudinally showed clear patterns of clonal evolution commonly following therapy, which was also associated with worse overall outcome. Importantly, aggressive subclones representing a majority of the tumor cell population at relapse were often already detectable in pre-treatment samples.<sup>13</sup> Complex and highly heterogeneous evolutionary dynamics with linear and branching subclones as well as marked clonal shifts over time and multiple cycles of therapies were also seen in smaller series of 22 patients studied by DNA arrays as well as in 3 patients repeatedly assessed by NGS in a detailed multi time-point analysis over up to 7 years.<sup>10,34</sup>

### III. Profiling Epigenetic Heterogeneity in CLL

The recent explosion in the number of mutations in known epigenetic regulatory genes identified across human cancers has underscored the importance of epigenetic control in tumor suppression.<sup>35</sup> It is highly likely that these epigenetic changes cooperate with genetic mutations to mold the evolutionary tumor landscape. The best-studied epigenetic modification to date is CpG methylation, which occurs by conversion of cytosine in DNA to 5-methylcytosine by addition of methyl groups to CpG sites regulating gene expression. In cancer, global genome-wide hypomethylation is accompanied by localized hypermethylation and an increase in expression of DNA methyltransferase.<sup>36</sup>

#### IIIa. Epigenetic Differences Amongst Patients

CpG methylation profiles clearly differ among prognostic subcategories of CLL.<sup>37,38</sup> These aberrantly methylated loci have been shown to include genes involved in CLL pathobiology, such as *BCL2*,<sup>39</sup> *TCL1*,<sup>40</sup> death-associated protein kinase 1 (*DAPK1*),<sup>41</sup> *LPL*, *ZAP70* and *NOTCH1*, as well as gene regulators and pathways involved in B cell signaling.<sup>42</sup> Also aberrantly methylated in CLL are several microRNA (miRNA) promoters, and this change can impact the expression of those miRNAs with known roles in CLL.<sup>43</sup> Finally, two lincRNA isoforms were found deregulated in CLL through changes in their promoter methylation and degree of histone modification. They map to the frequently deleted region of chromosome 13q14 and seem to play a role in controlling transcription of multiple other genes at this locus.<sup>44</sup>

Highly comprehensive genome-wide methylation profiling has been performed using the Illumina 450k arrays. These DNA hybridization chips evaluate 485,000 methylation sites per sample, covering over 95% of CpG islands, but also miRNA promoters, and CpG sites outside of CpG islands. In a study analyzing samples from 139 CLL patients with this technology, significant methylation differences were found in association with CLL genotypes (i.e. mutations in *SF3B1* and *NOTCH1*; trisomy 12 or del(11q)). Furthermore, methylation profiling could categorize patients into 3 distinct clinicopathologic groups (Figure 3C). Specifically, CLLs with unmutated *IGHV* were strongly related to naïve B cells, whereas CLLs with predominantly mutated *IGHV* aligned with mature B cells. These results suggest a methylation imprint corresponding to the putative cell of origin. Unexpectedly, a third group was uncovered with mainly mutated *IGHV*, but with a methylation signature more similar to naïve B cells. Compared to the other two groups,

patients of the third group had an intermediate clinical outcome. It was noted that methylation of CpGs in the gene bodies outside CpG islands showed strongest correlation with gene expression.<sup>45</sup>

### IIIb. Intraleukemic Epigenetic Heterogeneity

Given the existence of genetically distinct subclones within individual patient cancer cells, it might be expected that epigenetic alterations could also show complex intratumoral differences with changes over time. Indeed, recent analysis of 450k array data revealed high heterogeneity within individual CLL samples compared to healthy B cells (Figure 3D).<sup>46</sup> Also using this platform, Cahill et al examined methylation in paired diagnostic and follow-up samples. While they found over 2000 sites differentially methylated between *IGHV* mutated and unmutated CLL, they did not observe significant differences in methylation patterns over time nor between peripheral blood and lymph nodes. These data support the idea that altered methylation is an earlier rather than later leukemogenic event.<sup>42</sup> A recent 450k array analysis on 28 longitudinally followed CLL cases, however, showed that the majority of genetically evolved cases also showed epigenetic changes but that epigenetic evolution was not observed in the absence of genetic changes.<sup>47</sup> These results suggest a temporal hierarchy in which genetic alterations precede marked epigenetic changes and yet co-operate together. In the same study, Oakes et al additionally applied next generation targeted bisulfite sequencing on 28 selected regions. In contrast to array-based detection, massively parallel sequencing technologies provide the opportunity to study methylation at base pair resolution and with sequence context. Thus, allele-specific methylation in CLL could be identified, showing a stochastic pattern with random distribution between neighboring CpGs, which contrasted starkly to physiologically imprinted regions that showed methylation consistently occurring on the same allele. Reduced-representation bisulfite sequencing (RRBS) is a recently introduced NGS technology, which enables genome-scale methylation examination restricted to CG-rich sites, and hence is cost-effective. A recent RRBS study of 104 primary CLL samples uncovered the high degree of intra-tumoral methylation heterogeneity in CLL compared to normal B cells. This heterogeneity was evaluated to stem primarily from stochastic variation in DNA methylation, termed “locally disordered methylation”. Hence the high heterogeneity in methylation status within patient samples was assessed to relate not to ordered but differing methylation states arising from an admixture of cell subpopulations but rather from an increased proportion of cells with disordered methylation state. This high degree of “noise” in the CLL methylome was also linked with altered transcriptional regulation, clonal evolution and adverse clinical outcome.<sup>46</sup>

## IV. Clinical Implications of Genomic Discoveries in CLL

Many of the recent insights on CLL gained by the various large-scale sequencing studies have substantial clinical implications. Overall, they provide new markers that serve to refine prognostic information but also point to promising novel therapeutic targets and new strategies for rationally approaching CLL treatment.



#### IVa. Refining CLL Prognostic Schemata

Multiple studies on unselected cohorts comprising mainly untreated patients have demonstrated the prognostic value of four driver alterations - mutations in *SF3B1*, *NOTCH1*, *BIRC3* and *TP53*.<sup>9,11,19,22,24,48</sup> Other studies assessed the impact of these lesions in more selected cohorts, such as patients enrolled on therapeutic clinical trials. For example, analysis of the German CLL4 trial cohort, which compared fludarabine with or without cyclophosphamide, revealed the poor prognostic impact of *TP53* mutations as being equivalent to chromosomal del(17p), and that these lesions together were associated with extraordinary poor response to chemotherapy.<sup>49</sup> Similarly, in the UK LRF CLL4 trial, comparing different fludarabine-containing regimens, integration of mutations in *TP53* together with those in *SF3B1* and *NOTCH1* confirmed patients with *TP53* alterations as the group with shortest survival and poorest therapeutic response. However, *SF3B1* and *NOTCH1* mutations, albeit not associated with fludarabine response, both had an independent negative impact on overall survival.<sup>27</sup> Finally, in the German CLL2H cohort, which received alemtuzumab in the setting of fludarabine-refractoriness, *NOTCH1* mutations also had longer progression-free survival, suggesting that this subgroup might particularly benefit from the anti-CD52 antibody treatment. In contrast, *SF3B1* mutation had neither impact on response rates nor overall and progression-free survival.<sup>50</sup> These contrasting results indicate that different drivers may have a variable role, depending on the context.

Given the explosion of studies that have genotyped patients across unselected and a limited set of clinical trial cohorts, Rossi et al sought integrate this mutational information (on *TP53*, *NOTCH1*, *SF3B1* and *BIRC3*) together with conventional FISH cytogenetic data to an effort to improve the predictive ability of clinical prognostic schema.<sup>2</sup> Indeed, when this was done, predictive accuracy was significantly enhanced in a large cohort of 637 newly diagnosed CLL cases, and was validated in an independent group of 370 patients. Patients with *TP53* or *BIRC3* lesions consistently had the worst prognosis, followed by patients with mutations in *SF3B1* and *NOTCH1* and del(11q). Notably, these high-risk lesions developed in about a 20% of initially low-risk individuals during the course of the disease, among which del(11q) and abnormalities of *TP53*, *NOTCH1*, *SF3B1* as well as *BIRC3* again represented a major fraction.<sup>2</sup> Another large study assessing mutations of *TP53*, *NOTCH1*, *SF3B1*, *FBXW7*, *MYD88* and *XPO1* together with cytogenetics and *IGHV* mutational status in 1160 patients confirmed the strong prognostic value of *NOTCH1*, *TP53* and *SF3B1* alterations with independent impact on overall survival of the latter two.<sup>22</sup>

While the above studies address the question of whether presence or absence of a driver mutation in a CLL sample is prognostic, it has become evident that addressing the fraction of CLL cells bearing a driver mutation within a clonally heterogeneous population and which may participate in clonal evolution may present daunting challenges to physicians treating patients with this disease. Clearly, therapy imposes a strong selective pressure that impacts the relative proportions of subclones harboring driver mutations within the malignant population. Thus, the dominant driver mutations in relapse samples are frequently already present in minor subclones before treatment.<sup>13</sup> These observations support the well-established “watch and wait” strategy of withholding treatment for asymptomatic CLL. On

the other hand, a more thorough understanding of these interclonal dynamics might anticipate which subpopulations could become problematic in the future and thereby suggest approaches to personalize antecedent treatment.<sup>51</sup> Early detection of driver lesions in small subclones may potentially require surveillance using ultra-deep sequencing approaches, although validation of this strategy is still lacking. In support of this idea, however, a recent study demonstrated that presence of minor subclones harboring *TP53* mutation comprised of as little as 2% of the cancer cells within a population in early CLL pre-dated its emergence as a dominant subclone at relapse and the development of a chemorefractory phenotype.<sup>52</sup>

#### IVa. Novel Therapeutic Targets

To date, the only genetic features that have driven therapeutic decisions are deletions of 17p and *TP53* alterations.<sup>48</sup> However, the relative high frequency of somatic mutations in particular genes indicates that certain pathways are likely essential to CLL and that targeting these gene pathways could provide effective therapy. Indeed, a number of agents targeting these genes and pathways are under development, and include inhibitors of Notch signaling (Y-secretase inhibitors)<sup>53</sup>; inhibitors of the spliceosome<sup>54</sup>, of nuclear RNA export<sup>23</sup>, and of telomerase.<sup>55</sup> Another potential new target in CLL includes activation-induced cytidine deaminase (*AID*), which drives somatic hypermutation in normal B cell development by the generation of point mutations and initiation of double strand breaks.<sup>56,57</sup> More than 40% of CLLs have marked expression of *AID*, and in these cases inhibition of homologous recombination *in vitro* preferentially induced apoptosis in leukemia cells.<sup>58</sup>

Besides pointing to potentially vulnerable pathways in CLL, recurrently mutated genes in CLL have also been recently used to define the subclonal architecture of CLL. These studies have strongly suggested a specific temporal hierarchy in the acquisition of genetic and epigenetic changes, with clonal events (i.e. “earlier”) enriched for alterations, which more selectively affect B cells.<sup>13</sup> These observations suggest that elimination of CLL early in disease course may be particularly effective, prior to its genetic diversification. Consistent with the idea that targeting the common “trunk”<sup>59</sup> might be an effective therapeutic strategy, clinical studies targeting B cell-specific pathways such as through using anti-CD20 antibodies or novel kinase inhibitors of B cell receptor signaling pathways have demonstrated striking efficacy in CLL.<sup>60</sup>

#### V. Future Directions

Within the same time frame that studies using massively parallel sequencing in CLL have been reported, analysis of other cancers using these technologies have been also completed. These datasets together with CLL data have yielded on one hand, information regarding the fundamental mechanisms underlying somatic alteration in cancer and have also identified processes that are common or unique to CLL compared to other tumors.<sup>56,57</sup> Recent NGS efforts have examined the mutation spectrum of other mature B cell lymphomas as well; not surprisingly, similar alterations appear across the malignancies, but CLL also has its own unique constellations of abnormalities (Figure 4).<sup>8,11,13,15,17,61–69</sup>

Overall, these studies across cancers have yielded a number of further important insights. First, overall mutation rates of different malignancies have been established, and it is clear



that CLL is a low mutation-rate tumor, similar to other leukemias, and is mutated up to 10-fold less than the carcinogen-induced cancers.<sup>56,57,70</sup> Second, examination of the spectrum of mutations and the organization of affected gene regions have revealed evidence that cancers occasionally undergo catastrophic shattering events (i.e. so-called “chromothripsis”) as well as coordinated genomic rearrangements across different chromosomes.<sup>71,72</sup> Finally, comparison of CLL in relation to other cancers has revealed the spectrum of mutations in CLL to be consistent with a footprint of somatic hypermutation which is conventionally catalyzed by enzymes of the AID/APOBEC family of cytidine deaminases.<sup>56,57</sup>

At this juncture, the continued advances in technology and analytic tools promise the ability to gain answers to a number of critical questions in the near term. First, by integration of genetic and transcriptional data, we can understand genotype and phenotype relationships. Distinct RNA expression modules have already been linked to CLL genotypes, and over 60% of these associations were corroborated in an independent validation cohort yielding defined hypotheses for experimental studies.<sup>73</sup> Second, the low mutation rate of CLL also suggests that we can likely gain comprehensive analysis of the mutation landscape in the near future with a saturating number of exomes.<sup>70</sup> Third, single cell sequencing technologies promise to yield opportunities to dissect genetic heterogeneity in CLL, further trace the evolutionary tree of individual cases and strengthen understanding of genotype and phenotype relationships in this disease.<sup>74</sup> Lastly, investigating CLL development by comparison between monoclonal B cell lymphocytosis (MBL) and CLL as well as by learning how germline variants predispose to CLL might help us to further understand evolutionary trajectories in the disease. Recent meta-analyses of genome-wide population studies to examine CLL susceptibility have assessed on the order of magnitude of thousands of patients and controls. These efforts have corroborated 13 previously known risk loci for CLL and have newly found another 13 loci associated with inherited disease susceptibility. Interestingly, some of the identified genes suggest an overlap with regions of somatic driver alterations, involving *BCL2* or the 3' UTR of *POT1*.<sup>75,76</sup> Striking reports on CLL cases with germline mutations in putative drivers exist for *DAPK1*<sup>41</sup>, for *SAMHD1*, a nuclease involved in DNA damage response,<sup>77</sup> and for genes of the microRNA precursor miR-16-1-miR-15a.<sup>78</sup>

In closing, the recent technological developments have yielded a breathtaking increase in our knowledge about the CLL genome. As we have reviewed herein, the insights gained from dissecting the nature and basis of clonal evolution in CLL fuel a myriad of further lines of investigation that can have concrete impact on the clinical practice and treatment of CLL. From a more basic biological standpoint, future studies will likely address whether different genomic alterations provide distinct roles at defined stages in development of the disease, from predisposition to initial transformation of a clonal B cell population and to progression with relapsed disease following treatment and even transformation.

## Acknowledgments

M.G. acknowledges support by a Marie-Curie International Outgoing Fellowship from the European Union (PIOF-2013-624924). C.J.W. acknowledges support from the Blavatnik Family Foundation, AACR (SU2C Innovative Research Grant), NHLBI (1R01HL103532-01) and NCI (1R01CA155010-01A1).

## References

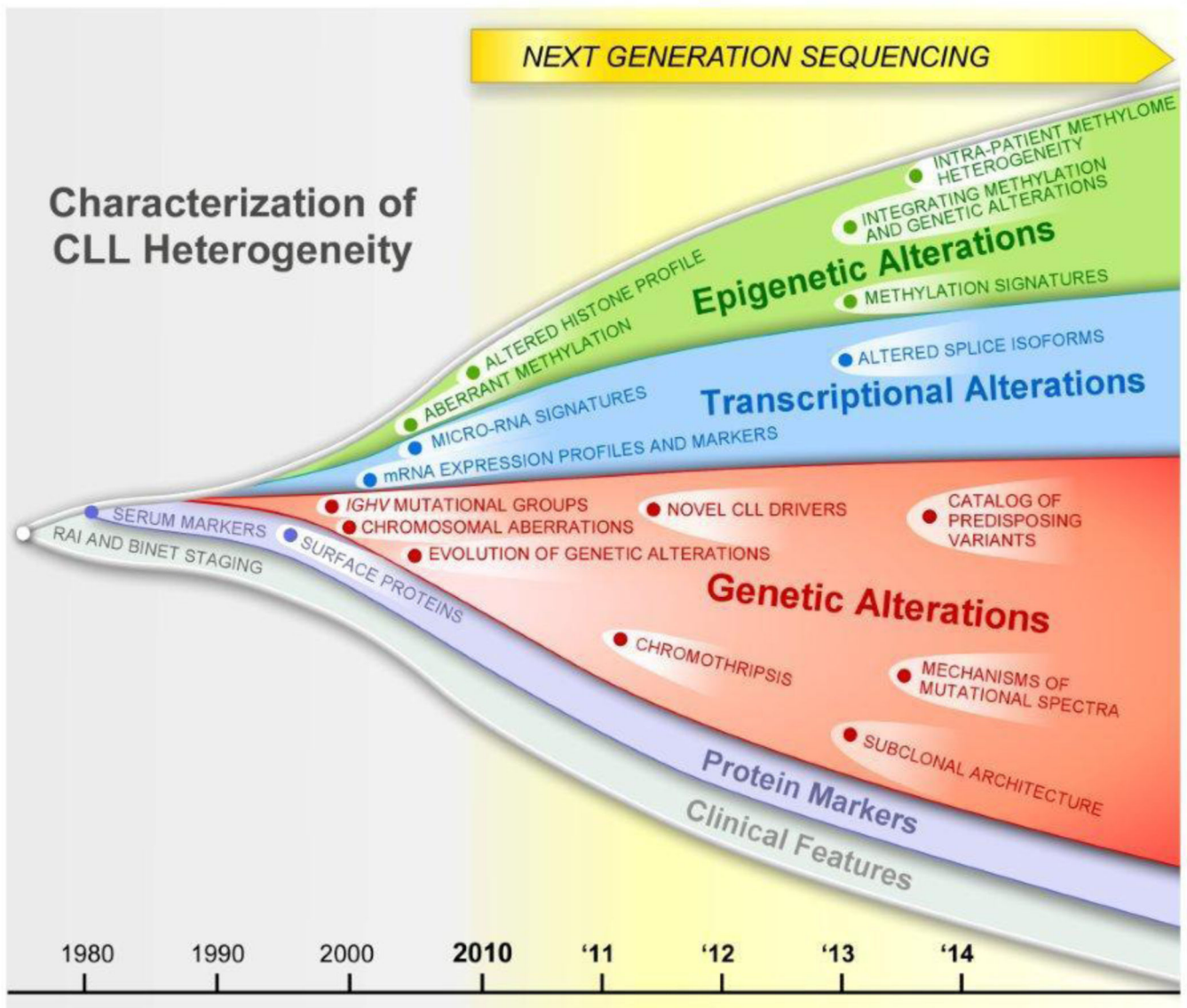
1. Oscier D, Dearden C, Erem E, et al. Guidelines on the diagnosis, investigation and management of chronic lymphocytic leukaemia. *Br J Haematol*. 2012; 159:541–564. [PubMed: 23057493]
2. Rossi D, Rasi S, Spina V, et al. Integrated mutational and cytogenetic analysis identifies new prognostic subgroups in chronic lymphocytic leukemia. *Blood*. 2013; 121:1403–1412. [PubMed: 23243274]
3. Rai KR, Sawitsky A, Cronkite EP, Chanana AD, Levy RN, Pasternack BS. Clinical staging of chronic lymphocytic leukemia. *Blood*. 1975; 46:219–234. [PubMed: 1139039]
4. Binet JL, Auquier A, Dighiero G, et al. A new prognostic classification of chronic lymphocytic leukemia derived from a multivariate survival analysis. *Cancer*. 1981; 48:198–206. [PubMed: 7237385]
5. Damle RN, Wasil T, Fais F, et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood*. 1999; 94:1840–1847. [PubMed: 10477712]
6. 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; 94:1848–1854. [PubMed: 10477713]
7. Dohner H, Stilgenbauer S, Benner A, et al. Genomic aberrations and survival in chronic lymphocytic leukemia. *New Engl J Med*. 2000; 343:1910–1916. [PubMed: 11136261]
8. Puente XS, Pinyol M, Quesada V, et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature*. 2011; 475:101–105. [PubMed: 21642962]
9. Wang L, Lawrence MS, Wan Y, et al. SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. *New Engl J Med*. 2011; 365:2497–2506. [PubMed: 22150006]
10. Schuh A, Becq J, Humphray S, et al. Monitoring chronic lymphocytic leukemia progression by whole genome sequencing reveals heterogeneous clonal evolution patterns. *Blood*. 2012; 120:4191–4196. [PubMed: 22915640]
11. Fabbri G, Rasi S, Rossi D, et al. Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation. *J Exp Med*. 2011; 208:1389–1401. [PubMed: 21670202]
12. Fabbri G, Khiabani H, Holmes AB, et al. Genetic lesions associated with chronic lymphocytic leukemia transformation to Richter syndrome. *J Exp Med*. 2013; 210:2273–2288. [PubMed: 24127483]
13. Landau DA, Carter SL, Stojanov P, et al. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell*. 2013; 152:714–726. [PubMed: 23415222]
14. Ramsay AJ, Quesada V, Foronda M, et al. POT1 mutations cause telomere dysfunction in chronic lymphocytic leukemia. *Nat Genet*. 2013; 45:526–530. [PubMed: 23502782]
15. Quesada V, Conde L, Villamor N, et al. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nat Genet*. 2012; 44:47–52. [PubMed: 22158541]
16. Ngo VN, Young RM, Schmitz R, et al. Oncogenically active MYD88 mutations in human lymphoma. *Nature*. 2011; 470:115–9. [PubMed: 21179087]
17. Treon SP, Xu L, Yang G, et al. MYD88 L265P somatic mutation in Waldenstrom's macroglobulinemia. *New Engl J Med*. 2012; 367:826–833. [PubMed: 22931316]
18. Arruga, F.; Gizdic, B.; Serra, S., et al. Functional impact of NOTCH1 mutations in chronic lymphocytic leukemia. *Leukemia*. Epub 2013 November 19, 10.1038/leu.2013.319. Available from [www.nature.com/leu](http://www.nature.com/leu)
19. Wan Y, Wu CJ. SF3B1 mutations in chronic lymphocytic leukemia. *Blood*. 2013; 121:4627–4634. [PubMed: 23568491]
20. Wan Y, Brooks A, Wang L, et al. SF3B1 mutation alters the selection of 3' RNA splice sites in chronic lymphocytic leukemia. *Blood*. 2013; 122:117.
21. Ferreira PG, Jares P, Rico D, et al. Transcriptome characterization by RNA sequencing identifies a major molecular and clinical subdivision in chronic lymphocytic leukemia. *Genome Res*. 2014 Feb; 24(2):212–226. [PubMed: 24265505]

22. Jeromin S, Weissmann S, Haferlach C, et al. SF3B1 mutations correlated to cytogenetics and mutations in NOTCH1, FBXW7, MYD88, XPO1 and TP53 in 1160 untreated CLL patients. *Leukemia*. 2014; 28:108–117. [PubMed: 24113472]
23. Lapalombella R, Sun Q, Williams K, et al. Selective inhibitors of nuclear export show that CRM1/XPO1 is a target in chronic lymphocytic leukemia. *Blood*. 2012; 120:4621–4634. [PubMed: 23034282]
24. Rossi D, Fangazio M, Rasi S, et al. Disruption of BIRC3 associates with fludarabine chemorefractoriness in TP53 wild-type chronic lymphocytic leukemia. *Blood*. 2012; 119:2854–2862. [PubMed: 22308293]
25. Cortese D, Sutton LA, Cahill N, et al. On the way towards a ‘CLL prognostic index’: focus on TP53, BIRC3, SF3B1, NOTCH1 and MYD88 in a population-based cohort. *Leukemia*. 2014 Mar; 28(3):710–713. [PubMed: 24217197]
26. Weissmann S, Roller A, Jeromin S, et al. Prognostic impact and landscape of NOTCH1 mutations in chronic lymphocytic leukemia (CLL): a study on 852 patients. *Leukemia*. 2013; 27:2393–2396. [PubMed: 23860447]
27. Oscier DG, Rose-Zerilli MJ, Winkelmann N, et al. The clinical significance of NOTCH1 and SF3B1 mutations in the UK LRF CLL4 trial. *Blood*. 2013; 121:468–475. [PubMed: 23086750]
28. Rossi D, Brusca A, Spina V, et al. Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. *Blood*. 2011; 118:6904–6908. [PubMed: 22039264]
29. Rossi D, Rasi S, Fabbri G, et al. Mutations of NOTCH1 are an independent predictor of survival in chronic lymphocytic leukemia. *Blood*. 2012; 119:521–529. [PubMed: 22077063]
30. Messina M, Del Giudice I, Khiabani H, et al. Genetic lesions associated with chronic lymphocytic leukemia chemo-refractoriness. *Blood*. epub 2014 February 18, 10.1182/blood-2013-10-534271. Available from [bloodjournal.hematologylibrary.org](http://bloodjournal.hematologylibrary.org).
31. Bea S, Lopez-Guillermo A, Ribas M, et al. Genetic imbalances in progressed B-cell chronic lymphocytic leukemia and transformed large-cell lymphoma (Richter’s syndrome). *Am J Pathol*. 2002; 161:957–968. [PubMed: 12213724]
32. Shanafelt TD, Hanson C, Dewald GW, 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 Clin Oncol*. 2008; 26:e5–e6. [PubMed: 18467710]
33. Shanafelt TD, Witzig TE, Fink SR, et al. Prospective evaluation of clonal evolution during long-term follow-up of patients with untreated early-stage chronic lymphocytic leukemia. *J Clin Oncol*. 2006; 24:4634–4641. [PubMed: 17008705]
34. Braggio E, Kay NE, VanWier S, et al. Longitudinal genome-wide analysis of patients with chronic lymphocytic leukemia reveals complex evolution of clonal architecture at disease progression and at the time of relapse. *Leukemia*. 2012; 26:1698–1701. [PubMed: 22261920]
35. Shen H, Laird PW. Interplay between the cancer genome and epigenome. *Cell*. 2013; 153:38–55. [PubMed: 23540689]
36. Ziller MJ, Gu H, Muller F, et al. Charting a dynamic DNA methylation landscape of the human genome. *Nature*. 2013; 500:477–481. [PubMed: 23925113]
37. Rahmatpanah FB, Carstens S, Hooshmand SI, et al. Large-scale analysis of DNA methylation in chronic lymphocytic leukemia. *Epigenomics*. 2009; 1:39–61. [PubMed: 20495622]
38. Kanduri M, Cahill N, Goransson H, et al. Differential genome-wide array-based methylation profiles in prognostic subsets of chronic lymphocytic leukemia. *Blood*. 2010; 115:296–305. [PubMed: 19897574]
39. Hanada M, Delia D, Aiello A, Stadtmauer E, Reed JC. bcl-2 gene hypomethylation and high-level expression in B-cell chronic lymphocytic leukemia. *Blood*. 1993; 82:1820–1828. [PubMed: 8104532]
40. Yuille MR, Condie A, Stone EM, et al. TCL1 is activated by chromosomal rearrangement or by hypomethylation. *Genes Chromosomes Cancer*. 2001; 30:336–341. [PubMed: 11241786]
41. Raval A, Tanner SM, Byrd JC, et al. Downregulation of death-associated protein kinase 1 (DAPK1) in chronic lymphocytic leukemia. *Cell*. 2007; 129:879–890. [PubMed: 17540169]

42. Cahill N, Bergh AC, Kanduri M, et al. 450K–array analysis of chronic lymphocytic leukemia cells reveals global DNA methylation to be relatively stable over time and similar in resting and proliferative compartments. *Leukemia*. 2013; 27:150–158. [PubMed: 22922567]
43. Baer C, Claus R, Frenzel LP, et al. Extensive promoter DNA hypermethylation and hypomethylation is associated with aberrant microRNA expression in chronic lymphocytic leukemia. *Cancer Res*. 2012; 72:3775–3785. [PubMed: 22710432]
44. Garding A, Bhattacharya N, Claus R, et al. Epigenetic upregulation of lncRNAs at 13q14.3 in leukemia is linked to the In Cis downregulation of a gene cluster that targets NF-κB. *PLoS Genet*. 2013; 9:e1003373. [PubMed: 23593011]
45. Kulis M, Heath S, Bibikova M, et al. Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. *Nat Genet*. 2012; 44:1236–1242. [PubMed: 23064414]
46. Landau DA, Clement K, Boyle P, et al. Increased Local Disorder of DNA Methylation Forms the Basis of High Intra-Leukemic Epigenetic Heterogeneity and Enhances CLL Evolution. *Blood*. 2013; 122:596.
47. Oakes CC, Claus R, Gu L, et al. Evolution of DNA methylation is linked to genetic aberrations in chronic lymphocytic leukemia. *Cancer Discov*. 2014 Mar; 4(3):348–361. [PubMed: 24356097]
48. Pospisilova S, Gonzalez D, Malcikova J, et al. ERIC recommendations on TP53 mutation analysis in chronic lymphocytic leukemia. *Leukemia*. 2012; 26:1458–1461. [PubMed: 22297721]
49. Zenz T, Eichhorst B, Busch R, et al. TP53 mutation and survival in chronic lymphocytic leukemia. *J Clin Oncol*. 2010; 28:4473–4479. [PubMed: 20697090]
50. Schnaiter A, Paschka P, Rossi M, et al. NOTCH1, SF3B1, and TP53 mutations in fludarabine-refractory CLL patients treated with alemtuzumab: results from the CLL2H trial of the GCLLSG. *Blood*. 2013; 122:1266–1270. [PubMed: 23821658]
51. Puente XS, Lopez-Otin C. The evolutionary biography of chronic lymphocytic leukemia. *Nat Genet*. 2013; 45:229–231. [PubMed: 23435090]
52. Rossi D, Khiabani H, Spina V, et al. Clinical impact of small TP53 mutated subclones in chronic lymphocytic leukemia. *Blood*. epub 2014 February 5, 10.1182/blood-2013-11-539726. Available from [bloodjournal.hematologylibrary.org](http://bloodjournal.hematologylibrary.org).
53. Groth C, Fortini ME. Therapeutic approaches to modulating Notch signaling: current challenges and future prospects. *Semin Cell Dev Biol*. 2012 Jun; 23(4):465–472. [PubMed: 22309842]
54. Bonnal S, Vigevani L, Valcarcel J. The spliceosome as a target of novel antitumour drugs. *Nat Rev Drug Discov*. 2012 Nov; 11(11):847–859. [PubMed: 23123942]
55. Ruden M, Puri N. Novel anticancer therapeutics targeting telomerase. *Cancer Treat Rev*. 2013; 39:444–456. [PubMed: 22841437]
56. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. *Nature*. 2013; 500:415–421. [PubMed: 23945592]
57. Lawrence MS, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*. 2013; 499:214–218. [PubMed: 23770567]
58. Lamont KR, Hasham MG, Donghia NM, et al. Attenuating homologous recombination stimulates an AID-induced antileukemic effect. *J Exp Med*. 2013; 210:1021–1033. [PubMed: 23589568]
59. Greaves M, Maley CC. Clonal evolution in cancer. *Nature*. 2012; 481:306–313. [PubMed: 22258609]
60. Niemann CU, Jones J, Wiestner A. Towards targeted therapy of chronic lymphocytic leukemia. *Adv Exp Med Biol*. 2013; 792:259–291. [PubMed: 24014301]
61. Bea S, Valdes-Mas R, Navarro A, et al. Landscape of somatic mutations and clonal evolution in mantle cell lymphoma. *Proc Natl Acad Sci U S A*. 2013 Nov 5; 110(45):18250–18255. [PubMed: 24145436]
62. Okosun J, Bodor C, Wang J, et al. Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. *Nat Genet*. 2014; 46:176–181. [PubMed: 24362818]
63. Love C, Sun Z, Jima D, et al. The genetic landscape of mutations in Burkitt lymphoma. *Nat Genet*. 2012; 44:1321–1325. [PubMed: 23143597]

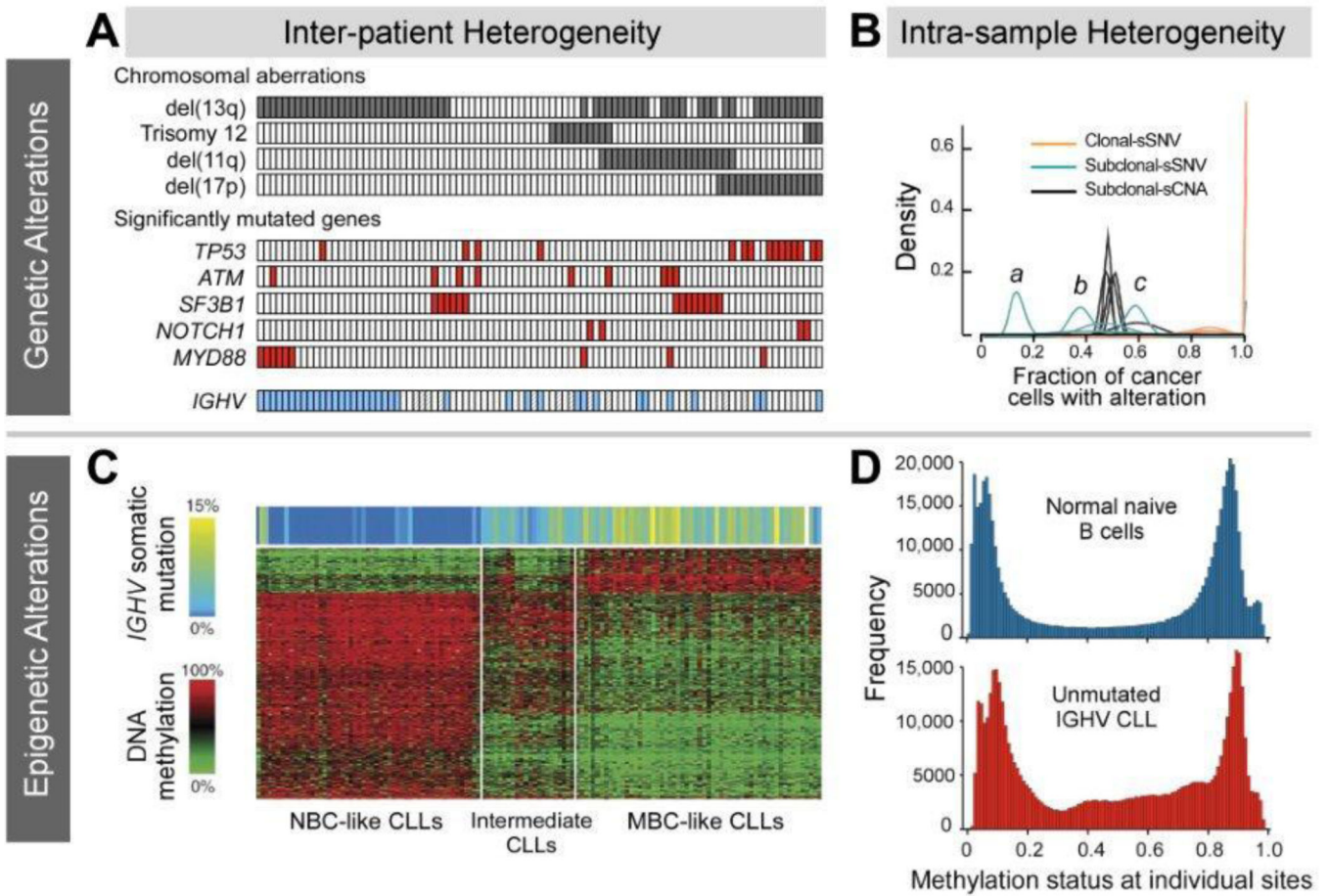
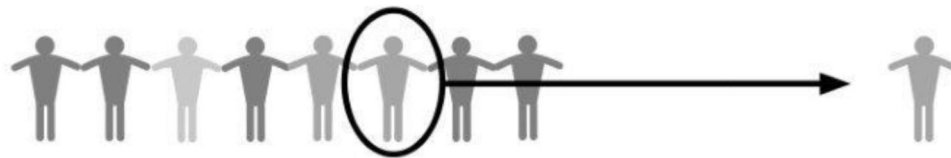
64. Richter J, Schlesner M, Hoffmann S, et al. Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing. *Nat Genet.* 2012; 44:1316–1320. [PubMed: 23143595]
65. Schmitz R, Young RM, Ceribelli M, et al. Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. *Nature.* 2012; 490:116–120. [PubMed: 22885699]
66. Pasqualucci L, Trifonov V, Fabbri G, et al. Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat Genet.* 2011; 43:830–837. [PubMed: 21804550]
67. Rossi D, Trifonov V, Fangazio M, et al. The coding genome of splenic marginal zone lymphoma: activation of NOTCH2 and other pathways regulating marginal zone development. *J Exp Med.* 2012; 209:1537–1551. [PubMed: 22891273]
68. Chapman MA, Lawrence MS, Keats JJ, et al. Initial genome sequencing and analysis of multiple myeloma. *Nature.* 2011; 471:467–472. [PubMed: 21430775]
69. Lohr JG, Stojanov P, Carter SL, et al. Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. *Cancer Cell.* 2014; 25:91–101. [PubMed: 24434212]
70. Lawrence MS, Stojanov P, Mermel CH, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature.* 2014; 505:495–501. [PubMed: 24390350]
71. Stephens PJ, Greenman CD, Fu B, et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell.* 2011; 144:27–40. [PubMed: 21215367]
72. Baca SC, Prandi D, Lawrence MS, et al. Punctuated evolution of prostate cancer genomes. *Cell.* 2013; 153:666–677. [PubMed: 23622249]
73. Kotliar D, Landau D-A, Litvin O, et al. Reconstructing a genotype-phenotype map in chronic lymphocytic leukemia. *Blood.* 2013; 122:2857.
74. Method of the year 2013. *Nat Methods.* 2014; 11:1. [PubMed: 24524124]
75. Berndt SI, Skibola CF, Joseph V, et al. Genome-wide association study identifies multiple risk loci for chronic lymphocytic leukemia. *Nat Genet.* 2013; 45:868–876. [PubMed: 23770605]
76. Speedy HE, Di Bernardo MC, Sava GP, et al. A genome-wide association study identifies multiple susceptibility loci for chronic lymphocytic leukemia. *Nat Genet.* 2014; 46:56–60. [PubMed: 24292274]
77. Clifford R, Louis T, Robbe P, et al. SAMHD1 is mutated recurrently in chronic lymphocytic leukemia and is involved in response to DNA damage. *Blood.* 2014; 123:1021–1031. [PubMed: 24335234]
78. Calin GA, Ferracin M, Cimmino A, et al. A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *New Engl J Med.* 2005; 353:1793–1801. [PubMed: 16251535]
79. Mansouri L, Cahill N, Gunnarsson R, et al. NOTCH1 and SF3B1 mutations can be added to the hierarchical prognostic classification in chronic lymphocytic leukemia. *Leukemia.* 2013; 27:512–514. [PubMed: 23138133]
80. Gonzalez D, Martinez P, Wade R, et al. Mutational status of the TP53 gene as a predictor of response and survival in patients with chronic lymphocytic leukemia: results from the LRF CLL4 trial. *J Clin Oncol.* 2011; 29:2223–2229. [PubMed: 21483000]
81. Zenz T, Habe S, Denzel T, et al. Detailed analysis of p53 pathway defects in fludarabine-refractory chronic lymphocytic leukemia (CLL): dissecting the contribution of 17p deletion, TP53 mutation, p53-p21 dysfunction, and miR34a in a prospective clinical trial. *Blood.* 2009; 114:2589–2597. [PubMed: 19643983]
82. Balatti V, Bottoni A, Palamarchuk A, et al. NOTCH1 mutations in CLL associated with trisomy 12. *Blood.* 2012; 119:329–331. [PubMed: 22086416]
83. Zenz T, Vollmer D, Trbusek M, et al. TP53 mutation profile in chronic lymphocytic leukemia: evidence for a disease specific profile from a comprehensive analysis of 268 mutations. *Leukemia.* 2010; 24:2072–2079. [PubMed: 20861914]



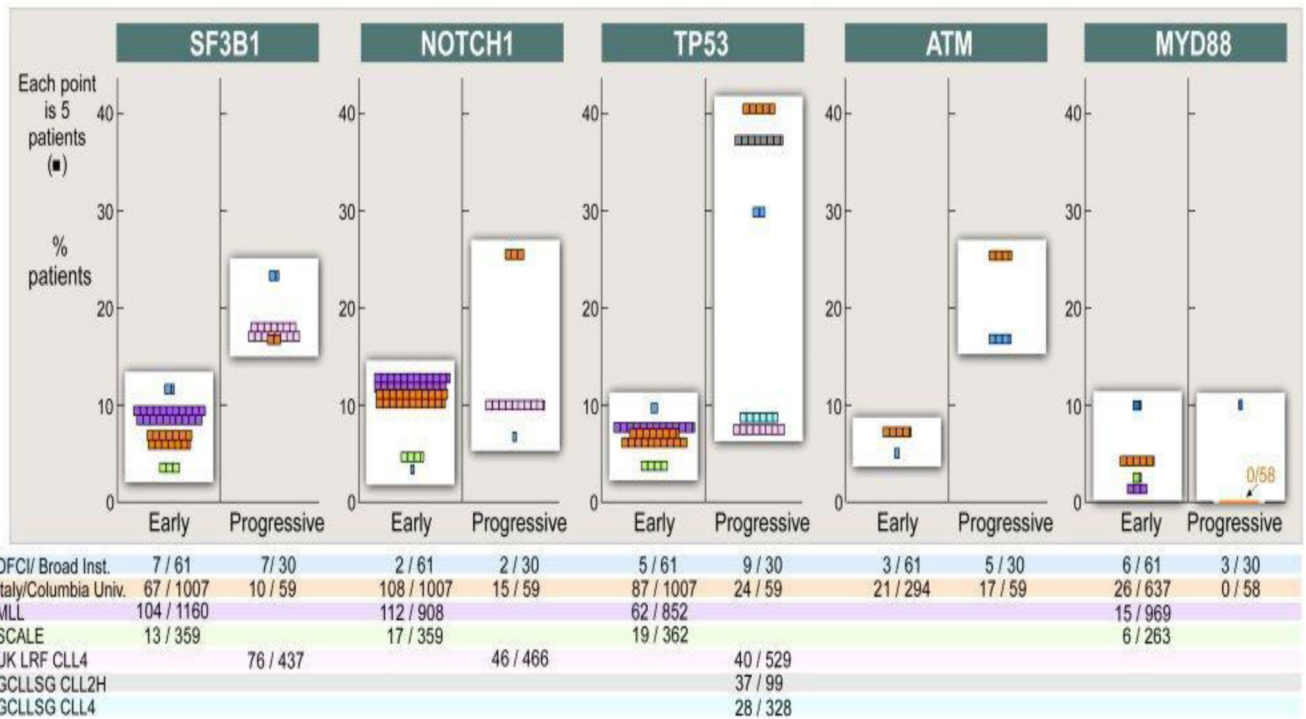


**Figure 1.** Evolution and growth in our understanding of CLL heterogeneity over time.





**Figure 2.** Inter- and intraleukemic genetic (A, B) and epigenetic (C, D) heterogeneity in CLL revealed by next-generation sequencing. Panel A, B, C were adapted with permission from Wang et al,<sup>9</sup> Landau et al,<sup>13</sup> Kulis et al,<sup>45</sup> respectively. Panel D was provided by Landau et al.<sup>46</sup>



**Figure 3.**

Frequency of genetic alterations in CLL depending on the cohort. The largest cohort per center has been taken into account, and the number of mutated cases over total size of the cohort is noted. Unselected cohorts have been included from: DFCI (Dana-Farber Cancer Institute)/Broad Institute,<sup>9,13</sup> Italy (Amedeo Avogadro University of Eastern Piedmont, Novara; Sapienza University, Rome)/Columbia University,<sup>2,11,24,28-30</sup> MLL (Munich Leukemia Laboratory),<sup>22,26</sup> SCALE (Scandinavian Lymphoma Etiology).<sup>25,79</sup> Included are also cohorts from the clinical trials UK LRF (UK Lymphoma Research Foundation) CLL4,<sup>27,80</sup> GCLLSG (German CLL Study Group) CLL4,<sup>49</sup> GCLLSG CLL2H.<sup>81</sup> “Early” - newly diagnosed and untreated patients without explicit evidence for progressive disease; “progressive” - patients with symptomatic CLL requiring treatment, or with relapse after treatment.

Platform (N)	MCL	FL	BL	DLBCL	SMZL	CLL	WM	MM
	WGS (4) WES (29) TS (172)	WGS (6) WES (4) TS (100)	WGS (4) WES (63) TS (149) RNAseq (41)	WES (100) RNAseq (80) SNParray (105)	WES (8) TS (109) RNAseq (6)	WGS (4) WES (270) TS (303)	WGS (30) TS (57)	WES (203) WGS (38)
<i>TP53</i>	++++		++++	+++	+++	+++		++
<i>NOTCH1</i>	++		++	++	++	+++		
<i>SF3B1</i>						+++		
<i>ATM</i>	++++					++		++
<i>MYD88</i>			++	+++	++	++	++++	
<i>XPO1</i>					++	++		+
<i>CHD2</i>						++		
<i>POT1</i>						++		
<i>SAMHD1</i>						++		+
<i>FBXW7</i>						++		
<i>BIRC3</i>	++				++	++		
<i>HIST1H1E</i>		+++				++	++++	
<i>LRP1B</i>			++++	++++		++		++
<i>KLHL6</i>		++				+		
<i>DDX3X</i>			++++	+		+		
<i>ITKB</i>					++	+		
<i>EGR2</i>					+	+		
<i>KRAS</i>						+		++++
<i>NRAS</i>						+		++++
<i>MED12</i>						+		
<i>ZMYM3</i>						+		
<i>BCOR</i>						+		
<i>RIPK1</i>						+		
<i>CCND1</i>	++++							++
<i>MLL2</i>	+++			++++	+++			+
<i>WHSC1</i>	+++							
<i>DCP1D</i>	+++							
<i>TRMP6</i>	+++							
<i>MEF2B</i>	++	+++		++				
<i>NOTCH2</i>	++				++++			
<i>SP140</i>	++							++
<i>CREBBP</i>		++++		++++				
<i>KDM2B</i>		++++		++				
<i>EZH2</i>		++++		+++				
<i>TNFRSF14</i>		++++						
<i>CARD11</i>		+++	++	+++	++			+
<i>TNFAIP3</i>		+++		+++	++			
<i>STAT6</i>		+++						
<i>EBF1</i>		+++						
<i>CD79B</i>		++		++				
<i>MYC</i>			++++	++				++
<i>ID3</i>			++++					
<i>MLL3</i>			++++	++++				++
<i>SMARCA4</i>			++++	+++				
<i>CCND3</i>			+++	++				
<i>GNA13</i>			+++	++				
<i>EP300</i>			++	+++	++			
<i>PIM1</i>			++	+++				+
<i>ARID1A</i>				+++	++		+++	+
<i>PRDM1</i>				+++				++
<i>BCL2</i>				+++				
<i>TBL1XR1</i>					+++			
<i>FAM46C</i>								+++
<i>DIS3</i>								+++
References	61	62	63, 64, 65	63, 65, 66	67	8, 11, 13, 15	17	68, 69

**Figure 4.** Recurrent putative driver alterations in mature B cell non-Hodgkin lymphomas. MCL - mantle cell lymphoma, FL - follicular lymphoma, BL - Burkitt’s lymphoma, DLBCL - diffuse large B cell lymphoma, SMZL - splenic marginal zone lymphoma, CLL, WM - Waldenström’s macroglobulinaemia and MM - multiple myeloma assessed by NGS (WGS - Whole Genome Sequencing, WES - Whole Exome Sequencing, TS - Targeted Sequencing [NGS or Sanger Sequencing], RNAseq, or SNP array). Shown are all alterations that were significant in CLL in at least one study. For alterations in other entities, selected ones have

been validated and occurred in a significant and high proportion (>10% of cases) in at least one study, or were identified across independent studies and were significant in at least one study.

**Table 1**

Evidence for co-segregation and mutual exclusivity of genetic alterations in CLL

<b>Alterations associated with M-IGHV<sup>8,9,15,22</sup></b>				
<b>Pathway</b>	<b>Alteration</b>	<b>Co-Segregation</b>	<b>Mutual Exclusivity</b>	<b>Location</b>
Chromatin modification	<i>CHD2</i>			15q26
Inflammatory pathways	<i>MYD88</i>	Del13q <sup>9,22,25</sup>	<i>SF3B1</i> , <sup>25</sup> <i>NOTCH1</i> <sup>25</sup>	3p22
<b>Alterations associated with U-IGHV<sup>8,9,11,14,15,22,25,27,28,82,83</sup></b>				
<b>Pathway</b>	<b>Alteration</b>	<b>Co-Segregation</b>	<b>Mutual Exclusivity</b>	<b>Location</b>
DNA damage response, cell cycle control	<i>ATM</i>	Del(11q) <sup>9</sup>		11q22-q23*
	<i>BIRC3</i>	Del(11q) <sup>2,25</sup>	<i>TP53</i> , <sup>2,24,30</sup> <i>NOTCH1</i> , <sup>24</sup> <i>SF3B1</i> <sup>24</sup>	11q22*
	<i>POT1</i>	<i>SF3B1</i> <sup>14</sup>		7q31.33
	<i>TP53</i>	Del(17p) <sup>2,9,25,49,80-82</sup>	<i>SF3B1</i> , <sup>28</sup> Tris(12), <sup>49</sup> <i>BIRC3</i> <sup>24</sup>	17p13.1*
mRNA processing	<i>SF3B1</i>	Del(11q) <sup>9,22</sup>	<i>MYD88</i> <sup>25</sup> , Tris(12) <sup>28</sup> <i>NOTCH1</i> <sup>22,25,27,79</sup> , <i>TP53</i> <sup>30</sup>	2q33.1
	<i>XPO1</i>	<i>NOTCH1</i> <sup>8</sup>		2p15
NOTCH signaling	<i>NOTCH1</i>	Tris(12), <sup>2,9,22,26,27,29</sup> <i>XPO1</i> , <sup>8</sup> <i>TP53</i> <sup>11,26</sup>	<i>SF3B1</i> , <sup>22,25,27,79</sup> <i>MYD88</i> , <sup>25</sup> <i>TP53</i> , <sup>29</sup> <i>BIRC3</i> , <sup>24</sup> Del(13q) <sup>26</sup>	9q34.3
	<i>FBXW7</i>	Tris(12) <sup>9,22</sup>	<i>SF3B1</i> , <sup>22</sup> <i>NOTCH1</i> <sup>9</sup>	4q31.3

\* involved in corresponding common chromosomal alterations