Clonal evolution in leukemia

Adolfo A Ferrando[1–3] & Carlos López-Otín[4,5]

Human leukemias are liquid malignancies characterized by diffuse infiltration of the bone marrow by transformed hematopoietic progenitors. The accessibility of tumor cells obtained from peripheral blood or through bone marrow aspirates, together with recent advances in cancer genomics and single-cell molecular analysis, have facilitated the study of clonal populations and their genetic and epigenetic evolution over time with unprecedented detail. The results of these analyses challenge the classic view of leukemia as a clonal homogeneous diffuse tumor and introduce a more complex and dynamic scenario. In this review, we present current concepts on the role of clonal evolution in lymphoid and myeloid leukemia as a driver of tumor initiation, disease progression and relapse. We also discuss the implications of these concepts in our understanding of the evolutionary mechanisms involved in leukemia transformation and therapy resistance.

It has been more than 150 years since Charles Darwin imagined that "whilst this planet has gone cycling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved." Now, we know that tumor cell populations, including leukemia cells, also undergo Darwinian evolution during their malignant progression[1–4]. In fact, recent genetic and epigenetic studies of the life history of leukemia and other hematological neoplasias have provided important insights into the role of clonal evolution as a driver of tumor initiation, disease progression and relapse[5–7]. These works have revealed a more complex view of leukemia evolution than the linear architecture originally proposed by Peter Nowell[1]. Thus, a new model has emerged that describes highly branched clonal architectures in leukemogenesis from early stages of disease to relapse. These studies have also validated the idea that subpopulations of cells with self-renewing properties generate, sustain and propagate the disease, and, therefore, that they are both key units for evolutionary selection and the main targets for directed therapies[8,9].

This review will highlight a series of recent works describing genomic and epigenomic evolution during leukemia initiation and relapse, and as a driver of therapy resistance. We will discuss both lymphoid leukemias, such as acute lymphoblastic leukemia (ALL) and chronic lymphocytic leukemia (CLL), and myeloid leukemias, including acute myeloid leukemia (AML) and chronic myeloid leukemia (CML). We will also discuss the epiallelic heterogeneity and dynamics in the context of leukemia progression and resistance. Finally, we will address the clinical implications of clonal evolution in disease prevention, risk stratification, evaluation of therapeutic response and management of therapy resistance.

[1]Department of Pediatrics, Columbia University, New York, New York, USA. [2]Department of Pathology and Cell Biology, Columbia University, New York, New York, USA. [3]Institute for Cancer Genetics, Columbia University, New York, New York, USA. [4]Departamento de Bioquímica y Biología Molecular, Facultad de Medicina, Instituto Universitario de Oncología (IUOPA), Universidad de Oviedo, Oviedo, Spain. [5]Centro de Investigación Biomédica en Red de Cáncer, Spain. Correspondence should be addressed to A.A.F. (af2196@columbia.edu) and C.L.-O. (clo@uniovi.es).

Received 18 November 2016; accepted 26 July 2017; published online 6 October 2017; doi:10.1038/nm.4410

CLONAL EVOLUTION DURING LEUKEMIA INITIATION

Pre-leukemic mutations and development of pre-leukemic clones

Tumors are an aggregate of diverse populations of cells. Tumor heterogeneity originates from and is shaped by clonal evolution, a multistep process by which random mutations create genetic and epigenetic diversity that is then the subject of natural selection. In leukemia, clonal selection is driven by competition between normal hematopoietic stem cells (HSCs) and early progenitors for resources in the microenvironment during tumor initiation, and later, during disease progression, selection is driven by competition between leukemic clones.

In the case of familial cancer predisposition syndromes that are associated with increased risk of leukemia, germline mutations are the first event contributing to tumor initiation. Prominent examples include mutations in *TP53* in leukemias from individuals with Li–Fraumeni syndrome[10]; mutations in the transcription factor *RUNX1* in AML occurring in individuals with familial platelet disorder and who have a predisposition to acute myelogenous leukemia (FPD/AML)[11]; mutations in *CEBPA* (transcription factor), *GATA2* (transcription factor), *ANKRD26* (ankyrin repeat protein) and *DDX41* (RNA helicase) in familial AML[12–15]; and mutations in *PAX5* and *ETV6* (transcription factors) in familial ALL[16–18].

Nonfamilial leukemias also have shared mutations that may initiate the cancer. Early studies of ALL occurring in monozygotic twins revealed the presence of shared disease-causing chromosomal alterations and markers indicative of a common fetal cell of origin for both leukemias[19,20]. Yet, analysis of the unaffected monozygotic twin sibling of an individual with ALL carrying the *ETV6–RUNX1* fusion oncogene demonstrated the presence of an immature population of hematopoietic progenitors with the same *ETV6–RUNX1* fusion and

BOx 1 Clonal hematopoiesis, aging and leukemia

The development of multiple blood clones associated with increased age was originally identified in individuals without disease and otherwise unremarkable hematopoietic parameters, but demonstrated skewed X inactivation in the hematopoietic system of females[140]. However, it has been only with the recent advent of genomic approaches that the underlying mechanism and clinical significance of these findings have come to the fore. First, single-nucleotide polymorphism (SNP)-array studies of blood from healthy individuals detected clonal chromosomal copy-number alterations and uniparental disomy in about 3% of elderly individuals and about 0.5% of young adults[23,24]. Most notably, age-related clonal hematopoiesis (also known as idiopathic clonal hematopoiesis) is associated with an increased risk of subsequently developing hematologic malignancies[23–26]. The identification of *TET2* mutations in clonal hematopoiesis cells in about 6% of women older than 65 with nonrandom X inactivation established a common prevalent genetic mechanism in age-induced clonal hematopoiesis and myeloid malignancies[27]. In addition, whole-exome sequencing analyses have revealed the presence of clonal mutations in blood in about 10% of individuals over 65 years old, again in association with increased risk of hematologic cancer, and with frequent involvement of cancer- and leukemia-associated genes—primarily *DNMT3A* and *TET2*, but also *ASXL1*, *TP53*, *SF3B1*, *JAK2* and *SRSF2* (refs. 25,26). Similarly, about 17% of cases with idiopathic cytopenia of undetermined significance (ICUS) without dysplasia show clonal hematopoiesis[141]. Recent studies using improved sequencing and analysis methods have suggested that clonal hematopoiesis is much more frequent than thought in healthy middle-age individuals, thereby making it necessary to distinguish benign clonal hematopoiesis from malignant clonal hematopoiesis that could drive leukemia transformation[142].

immunoglobulin rearrangement as her sibling[21]. Notably, this early clonal population was devoid of additional disease-causing leukemia-associated genetic lesions present in her sister's leukemia lymphoblasts[21]. These observations support the notion that the presence of *ETV6–RUNX1* is not sufficient to induce leukemia transformation; rather, the acquisition of additional genetic and/or epigenetic alterations is required. In this context, an inflammatory environment triggered by infection during infancy can drive disease progression by favoring the expansion of *ETV6–RUNX1*–positive preleukemic cells and the accumulation of mutations induced by enzymes involved in immunoglobulin gene rearrangement and class-switch recombination, which occur as a result of the immune function of normal cells in response to infection (**Box 1**)[22]. In myeloid malignancies, age-related clonal hematopoiesis may represent a premalignant state, because this condition can precede the development of a myeloid tumor and frequently involves shared mutations in driver genes responsible for leukemia transformation[23–27].

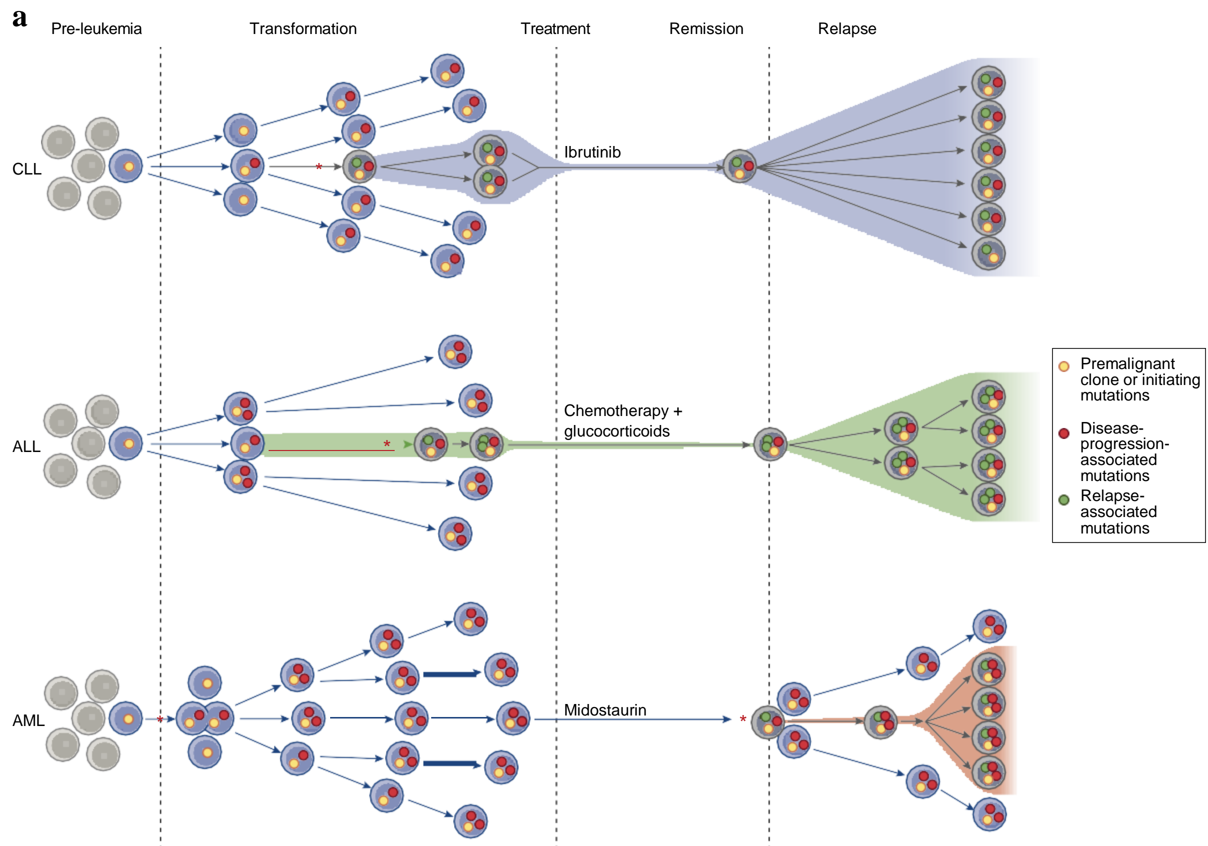
The shape of evolution: branched versus linear

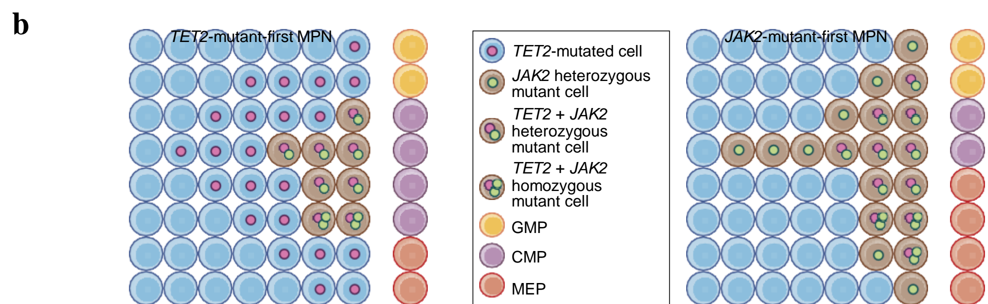
Clonal evolution is an active process that continually shapes the genomic landscape through the dynamic interplay between emerging genetic and epigenetic alterations and changing evolutionary pressures. Genomic analyses of the clonal composition of *ETV6-RUNX1*-positive leukemias at diagnosis indicate a branched evolutionary pattern, with multiple coexisting clones[28]. The identification of shared genetic alterations points to convergent evolution between the clones, but the genetic alterations could be acquired at variable stages by the clones, with no fixed determined order[28]. Similarly, early work tracking subclonal immunoglobulin heavy-chain rearrangements in *ALL* revealed a complex and branched pattern of clonal evolution[29,30]. Moreover, copy-number analysis of lymphoblastic leukemia xenografts with the *BCR-ABL1* fusion demonstrated marked genetic diversity between clones in the leukemia-initiating cell compartment, and subsequent branched, multiclonal architecture[31]. Single-cell and deep-sequencing analyses of AML have also shown a mix of heterogeneous clonal populations generated by the continuous acquisition and selection of convergent and divergent clonal mutations[32–34]. Recent studies in other hematological neoplasias, including CLL and myelodysplastic syndromes (MDS), have confirmed the relevance of a marked genetic heterogeneity in the generation of complex, dynamic and branched evolutionary routes during leukemogenesis[35–38] (**Fig.1**).

Initiating versus secondary mutations

Given the requirement of multiple genetic and epigenetic alterations for a fully transformed leukemia phenotype, it has been proposed that leukemia must originate from long-lived HSCs, because it would take time to accumulate these mutations. In fact, normal lymphoid and myeloid cells were shown to harbor the *RUNX1-RUNX1T1* fusion oncogene in an individual with AML in remission 15 years after treatment[39], which supports the presence of the mutation in the HSC compartment or in a multilineage progenitor with acquired self-renewal capacity. Moreover, genetic analysis of normal HSCs from AML samples has demonstrated the presence of preleukemic, nontransformed, immunophenotypically normal HSCs that contain pathogenic mutations in the methyltransferase gene *DNMT3A*, but without the nucleophosmin (*NPM1*) mutations found in association with the same *DNMT3A* mutations in AML blasts[40]. However, it is also possible that not all leukemias originate from the HSC compartment, because some genetic events—in particular, mutations in epigenetic regulators—may confer self-renewal properties to more committed progenitors, facilitating the acquisition of subsequent secondary mutations and the development of a fully transformed leukemia population.

In CLL, analysis of normal HSCs revealed genetic alterations in oncogenes and tumor suppressors outside the known neoplastic B cell lymphoid compartment[41]. In addition, data from xenograft experiments showed that HSCs from individuals with CLL are abnormally differentiated toward B cells[42]. In all, these results support the idea that clonal evolution from a preleukemic early hematopoietic progenitor with altered generation of B cells may underlie the origin of CLL. Similar studies in hairy cell leukemia (HCL)—a chronic lymphoproliferative disorder characterized by *BRAF*^{V600E} mutations—further suggested that chronic lymphoid malignancies are initiated by aberrant HSCs[43]. Moreover, analysis of intratumoral heterogeneity and mutational trajectories in MDS—characterized by ineffective hematopoiesis and increased risk of progression to AML—shows that there is subclonal evolution in the tumor population, with continuous acquisition of genetic alterations[38]. Further analysis of AML revealed that mutations involving epigenetic regulators (such as TET2 or ASXL1) and splicing factors (such as SF3B1 or SRF2) are predominant initiating lesions in AML, although they can also be present as secondary-hit events. Mutations in nucleolar proteins (NPM1), signaling factors (such as FLT3, JAK2 or CBL), transcription factors



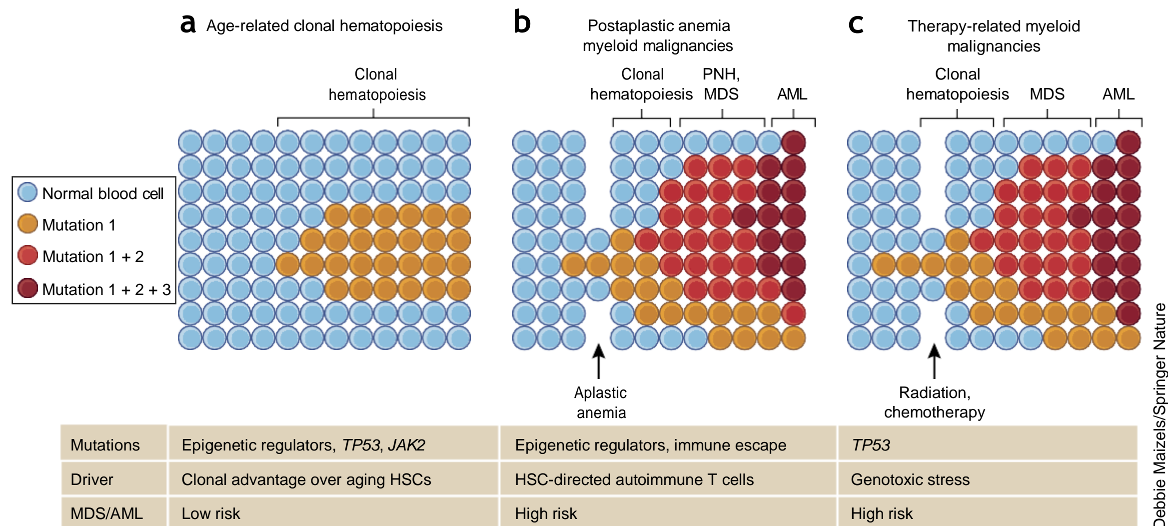


Founder mutation	*TET2*
TET2 populations	Expanded HSC and progenitor cells
TET2 + *JAK2* populations	Increased erythroid and megakaryocyte production
Clinico-biological associations	Older age Lower risk of thrombosis

Founder mutation	*JAK2*
JAK2 populations	Erythroid and megakaryocyte progenitors
TET2 + *JAK2* populations	Erythroid and megakaryocyte population expansion
Clinico-biological associations	Younger age Higher frequency of polycythemia vera

Debbie Maizels/Springer Nature

Figure 1 Different modes of clonal evolution in leukemias. (**a**) In CLL, the dominant clone after relapse is present in pre-treatment samples, where it can be detected at low frequency. This relapse-driven clone derives from the original tumor. It shares some early mutations involved in tumor transformation, but it has also acquired additional mutations in key genes conferring resistance to treatment (red asterisks, late mutations). Frequently in ALL, the relapse-driven clone is not directly derived from cells in the major clone at diagnosis, but from an ancestral tumor clone. In AML, a linear pattern of clonal evolution with secondary acquisition of resistance-driving mutations not detectable in the original diagnostic sample has been described as a mechanism of relapse after FLT3-inhibitor therapy with midostaurin. (**b**) Half of MPD cases with co-occurring mutations in *JAK2* (red circle) and *TET2* (green circle) result from an early *JAK2*^{V617F} mutation, whereas in the remaining cases, the *TET2* alteration occurs first. In *TET2*-first cases, the loss of *TET2* induces HSC and progenitor cell expansion. The secondary acquisition of *JAK2*^{V617F} then induces an excess of megakaryocyte and erythroid cell production. In *JAK2*-mutant-first cases, this mutation increases the output of megakaryocyte and erythroid cells, which expand upon the acquisition of a *TET2* secondary lesion. Brown shadowed cells indicate progenitors with increased megakaryocyte and erythroid cell output. GMP, granulocyte-monocyte progenitor. CMP, common myeloid progenitor. MEP, megakaryocyte-erythroid progenitor.



Mutations	Epigenetic regulators, *TP53*, *JAK2*	Epigenetic regulators, immune escape	*TP53*
Driver	Clonal advantage over aging HSCs	HSC-directed autoimmune T cells	Genotoxic stress
MDS/AML	Low risk	High risk	High risk

Figure 2 Clonal hematopoiesis in aging, bone marrow failure syndromes and therapy-related leukemias. (**a**) In aging, clonal hematopoiesis emerges from mutations in epigenetic regulators (*DNMT3A*, *TET2* and *ASXL1*), splicing factors (*SF3B1* and *SRSF2*) and other genes, such as *TP53* and *JAK2*. The selection driver is the clonal advantage of mutant cells over aging HSCs, and the risk to progression to myelodysplastic syndrome (MDS) and transformation to acute myeloid leukemia (AML) is low. (**b**) In individuals with aplastic anemia, clonal hematopoiesis emerges from mutations in epigenetic regulators plus specific cytogenetic abnormalities (MDS) or from genetic events linked with immune escape in patients with paroxysmal nocturnal hemoglobinuria (PNH). Both MDS and PNH are clonal disorders and represent independent complications that arise during aplastic anemia evolution, but they rarely coexist. The clonal-selection driver is the autoimmune attack by cytotoxic T cells of HSCs, and the risk of transformation to AML is high. (**c**) Therapy-related myeloid malignancies are characterized by high prevalence of *TP53* mutations, and the selection driver for clonal hematopoiesis is genotoxic stress. In this case, clonal hematopoiesis increases the risk of therapy-related MDS/AML, but there are cases in which direct clonal evolution does not occur.

(such as RUNX1 or ETV6) and chromosomal alterations (monosomy 7, trisomy 8 or del(5q)) are acquired as secondary events in the natural history of the disease[38,44].

Whole-genome sequencing analysis of acute leukemia cases originating from progression from MDS clonal hematopoiesis (**Box 1**) to overt secondary AML (sAML) demonstrated an oligoclonal composition at both stages of the disease, and shows that progression to sAML is characterized by the persistence of an MDS founding clone, followed by the expansion of new dominant clones with additional mutations[45]. Of note, although most mutations in both MDS and AML were transitions at CpG dinucleotides, there were more transversions in patients with sAML who were treated with the DNA methyltransferase inhibitor decitabine during the MDS phase before sAML diagnosis. These observations suggest that DNA methylation biases the type of mutations that can be acquired, and that exposure to chemotherapy shapes mutational patterns[45]. Clonal-architecture analyses in other hematological neoplasias, such as follicular lymphoma, have also confirmed the relevance of mutations in early progenitors for disease development[46].

Collectively, these observations suggest that, as HSCs accumulate mutations over time, genetic alterations that provide a clonal advantage can drive the emergence of clonal hematopoiesis (**Fig. 1** and **Box 1**). Should additional oncogenic mutations accumulate, these preleukemic clones can transform into fully malignant populations. Typical early events in AML include the loss of *TET2* or *DNMT3A* or the expression of the mutant R132H IDH1, which disrupt DNA methylation and promote self-renewal and the expansion of HSCs[40,47,48].

Evolutionary pressures in initiation

Therapy-associated AML and MDS following chemotherapy for other malignancies are characterized by a high prevalence of mutations in *TP53*, which were originally attributed to the mutagenic effects of chemotherapy[49,50]. However, analyses of normal hematopoietic progenitor samples collected from these patients before their original chemotherapy indicate the presence of preleukemic cells harboring these leukemia-associated *TP53* lesions[51]. TP53 induces programmed cell death following genotoxic stress, and loss of both alleles of *TP53* induces chemoresistance, which supports a role for selective pressure from chemotherapy-induced DNA damage acting on pre-existing preleukemic cells in the pathogenesis of therapy-associated myeloid malignancies. These results and data derived from the analysis of other genes, mechanistically linked with the acquisition of clonal dominance (driver genes) in these leukemias, suggest that chemotherapy contributes to the development of therapy-related myeloid tumors by promoting the expansion of pre-existing preleukemic clones under the selective pressure of genotoxic stress[51–53] (**Fig. 2**).

Additional examples of the role of clonal selection in leukemia initiation are provided by studies of the increase in the number of cellular clones in blood associated with aging (**Box 1**) and in bone marrow failure syndromes. In acquired aplastic anemia, autoreactive cytotoxic T cells mediate the destruction of HSCs and trigger hematopoietic failure, requiring immunosuppressive therapy or bone marrow transplantation[54]. Yet, following immunosuppression, individuals with aplastic anemia frequently have clonal hematopoiesis and are at increased risk of developing myeloid malignancies and paroxysmal nocturnal hemoglobinuria (PNH), a bone marrow failure syndrome with clonal hematopoiesis and hemolytic anemia[55].

Clonal hematopoiesis in aplastic anemia emerges from pre-existing HSCs that are present at the time of diagnosis and contain mutations in genes also implicated in the development of MDS and AML, including *BCOR*, *BCoRL1*, *ASXL1* and *DNMT3A*[56]. The emergence of multiple independent mutant clones with different mutations in the same gene suggests that they arise as a result of Darwinian selection, and not by genetic drift from a reduced pool of surviving

BOx 2 Epigenetic marks

Epigenetic mechanisms allow genetically identical cells to achieve different stable phenotypes by facilitating the transcription of specific genome regions, through changes in chromatin organization[91]. Epigenetic marks include DNA and histone modifications, which form an intricate network of mutually reinforcing or counteracting signals. The most prominent DNA epigenetic mark involves CpG cytosine-5 methylation (5mC), but there are also other changes, such as hydroxylation, formylation and carboxylation. Several methods are available to map 5mC patterns on a genome-wide scale[143]. The most comprehensive coverage of 5mC at single-base level is obtained by shotgun sequencing of bisulphite-treated DNA. Other methods include enzymatic digestion with methylation-sensitive restriction enzymes and 5mC capture by methylated DNA-binding proteins, followed by DNA sequencing. Methyl-DNA immunoprecipitation is another approach in which extracted DNA is cleaved, denatured and precipitated using an antibody to 5mC, and then the precipitated fragments are sequenced.

Chromatin structure is also heavily influenced by histone modifications that delimitate functional elements in mammalian genomes. Large-scale mapping of histone-modification patterns by using methods such as chromatin immunoprecipitation followed by sequencing (ChIP-seq) has allowed the characterization of the chromatin structure determinants across the genome, in diverse cell types and under different conditions. Thus, DNA is packaged into nucleosomes, which vary in histone composition and histone modifications, such as histone H3 lysine 9 trimethylation (H3K9me3). Active promoters are commonly marked by H3K4me2, H3K4me3 and H3K4 acetylation (H3K4ac). Transcribed regions are enriched for H3K36me3 and H3K79me2 and active enhancers are relatively enriched for H3K4me1, H3K4me2 and H3K27ac, whereas repressed genes are usually associated with H3K9me2, H3K9me3 and H3K27me3 marks. All these epigenetic marks involving DNA methylation and histone modifications are dynamically adapted to the changing conditions during clonal evolution of leukemia, and they directly contribute to the development of this process.

HSCs. Progression to MDS and AML is associated with accumulating cytogenetic abnormalities—such as monosomy 7 and trisomy 8—and accelerated telomere attrition[57,58]. In some cases, cells that have undergone clonal hematopoiesis harbor mutations in *PIGA* and uniparental disomy for the short arm of chromosome 6 (6pUPD), two genetic events linked with escape from autoimmunity, rather than with progression to malignant transformation. Loss of specific HLAs involved in the presentation of antigens driving autoimmune response against the HSC compartment allows the clones with the disomy for 6pUPD to escape from the attack of cytotoxic T cells[59,60]. In agreement with immune pressure driving the evolution of these clones, and although the precise mechanism for immune escape of *PIGA*-mutated clones in aplastic anemia is unclear, it seems that the loss of some glycosylphosphatidylinositol-anchored proteins may impair immune recognition of HSCs[61,62] (**Fig. 2**).

In Fanconi anemia—the most frequent cause of bone marrow failure syndrome—the chromosomal fragility arising from the defect in the *FA*/*BRCA* DNA-repair pathway confers a very high predisposition to MDS and AML. In this case, transformation into MDS and AML is associated with the onset of somatic chromosomal translocations involving, frequently, 1q, 7q, 3q (involving *EVI1*) and 21q (involving *RUNX1*)[63]. Another example of clonal evolution with leukemia predisposition in the context of defective hematopoiesis is provided by neutropenia syndromes caused by mutations in *ELANE, HAX1* and *WAS* (encoding neutrophil elastase, an apoptosis inhibitor and an actin polymerization activator, respectively)[64–66]. Treatment with granulocyte-colony-stimulating factor (CSF3) is effective for these diseases[67]. However, about 20% of patients eventually develop a myeloid malignancy harboring mutations in the CSF3 receptor gene[68], which suggests that increased CSF3 signaling provides a selective advantage driving clonal expansion[69]. Of note, these mutations are present in many cases without evidence of transformation, and they can antecede by months or years the development of a myeloid malignancy[70].

Order of mutation and clonal evolution

The order of mutation acquisition may strongly influence the disease phenotype, its clinical characteristics and the response to therapy[71,72]. Studies in MDS and related hematological neoplasms suggested that early driver mutations affecting specific genes dictate future evolutionary trajectories of disease, with distinct clinical outcome[71]. Further studies in myeloproliferative neoplasms—such as polycythemia vera and essential thrombocytemia—confirmed that the order in which mutations are acquired influences clonal evolution and clinical features. These myeloproliferative neoplasms show high prevalence of the *JAK2*^{V617F} mutation, which, in 10% of cases, co-occurs with a *TET2* mutation[72]. In half of these cases, the *JAK2* mutation can be detected at an early stage, whereas in the remaining samples, *TET2* is the earlier initiating mutation and, yet, in both situations, both single- and double-mutant clones could be detected, indicating that in both situations, double-mutant clones do not effectively outcompete single-mutant populations. *TET2*-first samples have increased numbers of common myeloid progenitors, whereas *JAK2*-first cases show a predominance of megakaryocyte and erythrocyte progenitors[72]. *TET2* mutations in *TET2*-first cases induce the expansion of HSCs and progenitor cells, but do not contribute to excess megakaryocyte and erythroid cells in the absence of a cooperating *JAK2* mutation[72]. By contrast, *JAK2*^{V617F}-mutated HSCs and progenitors generate increased numbers of erythroid and megakaryocyte cells, but expand only upon the acquisition of a *TET2* secondary mutation[72]. With regard to the influence of the mutation on clinical outcome, patients in whom the *JAK2*^{V167F} mutation is acquired first more often presented with polycythemia vera rather than essential thrombocytemia, and had increased risk of thrombosis and increased sensitivity to JAK2 inhibition with ruxolitinib in *in vitro* assays[72].

In angioimmunoblastic T cell lymphoma—a tumor of transformed follicular T cells—*TET2* and *DNMT3A* mutations emerge as early events and can be found in non–T cell hematopoietic populations, whereas the G17V mutation in the small GTPase RHOA, which is characteristic of these aggressive lymphomas, is restricted to the T cell compartment, supporting an instructive role in lineage specification[73]. Detailed longitudinal reconstruction of the evolutionary histories of other blood neoplasias will help to further define the order and constraints in which early driver events are acquired, and how they influence the characteristics of subsequent mutations, as well as disease pathogenesis.

The interplay between genetic and epigenetic evolution

Epigenetic changes substantially contribute to the dynamics of the evolutionary process, which drives the pervasive diversification of

the nascent leukemic cells[74–76] (**Fig. 3** and **Box 2**). These epigenetic modifications are somatically heritable and, similar to gene-coding mutations, are subjected to selection forces in a Darwinian fashion that finally results in the expansion of the most fit cell variants. Notably, DNA-methylation alterations in leukemia and related blood cell malignancies are several orders of magnitude more frequent than genetic changes, and thus have enormous potential to contribute to the acquisition of clonal heterogeneity[77–82]. Individuals with AML have substantial epiallele (specific DNA-methylation pattern of a genetic locus) diversity at diagnosis and at relapse when compared to normal controls[81]. Likewise, whole-methylome analysis of childhood B-precursor ALL samples has shown that there is a tendency toward hypomethylation in relapsed tumors[83], whereas higher epiallele burden in AML, CLL and CML is associated with more aggressive disease and worse clinical outcomes[76,84–86]. These findings suggest that leukemogenesis is associated with broad, dynamic methylation reprogramming on which evolutionary pressure can act during the establishment and progression of the disease, or during disease treatment.

The broad diversity of epigenetic regulators with genetic alterations in virtually all leukemia types and subtypes[87–91] supports the prominent role of epigenetic changes in promoting tumor initiation, and in clonal evolution during disease progression and in response to therapy. The precise molecular mechanisms underlying this epigenomic rewiring linked to clonal evolution of leukemia cells are unclear. Several studies have suggested that epigenetic reprogramming may be connected with the genomic mutational landscape of leukemias. Thus, high levels of methylation heterogeneity in CLL correlate with genetic subclonal complexity[84,92], and *BCR-ABL1* expression in CML cells is sufficient to trigger aberrant DNA methylation and thus potentially contributes to leukemia evolution[93].

Further evidence for genetic–epigenetic cross-talk during clonal evolution in leukemia is provided by the fact that somatic mutations acquired at disease initiation in epigenetic regulators—such as *TET2*—cooperate with genetic alterations linked to progression (such as *FLT3* mutations) to synergistically reprogram the DNA methylome[94]. However, recent comparative genomics and epigenomics studies in AML have revealed that genetic and epigenetic diversification occur with distinct kinetics and often follow independent pathways[81]. Some individuals with AML show high epiallele diversity and low somatic mutation burden in their leukemia cells at diagnosis, whereas others present with high somatic mutation and lower epiallele burdens, suggesting the occurrence of epigenetically driven as compared to genetically driven modes of tumor heterogeneity in this disease, with some intermediate situations[81].

Global epigenetic heterogeneity in leukemias may be due to a stochastic and dynamic response of tumor cells to a variety of environmental stresses occurring during leukemogenesis or after chemotherapy treatment. These stochastic changes in DNA methylation affect large regions of the epigenome of cancer cells, destabilizing the nuclear architecture and facilitating selection for traits that provide survival and growth advantages to these cells against the selective pressure imposed by the changing microenvironment[87]. However, it is also possible that certain genetic mutations could trigger epigenome reorganization and disordered DNA methylation at specific loci. For example, in AML, epigenetic heterogeneity may be mediated through localized expression of lineage-specific transcription factors in HSCs[95], whereas in B cell malignancies, it might be associated with the function of activation-induced cytosine deaminase[96].

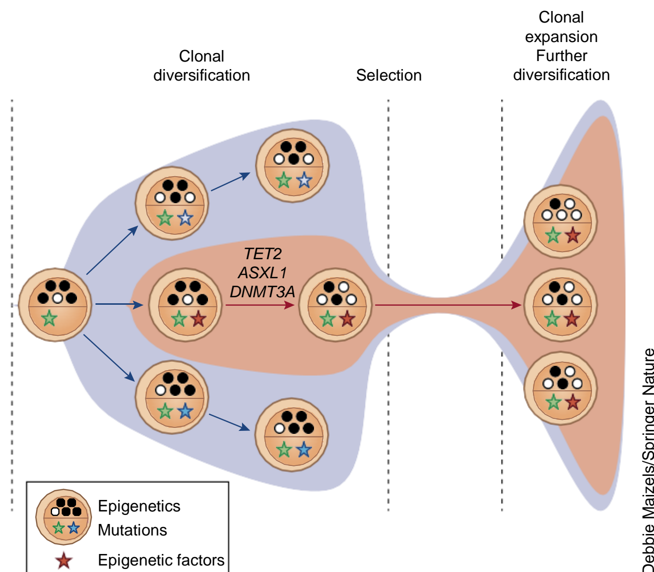


Figure 3 Contribution of epigenetics to clonal evolution in leukemia. Each subclone has both a mutational and an epigenetic signature, depicted using stars and circles, respectively. During clonal diversification, epigenetic signatures can evolve, perhaps driven by mutations in epigenetic-remodeling factors. After treatment (selection), the epigenetic signature associated with the fittest clone is expanded and may evolve.

In summary, and regardless of the involved mechanisms, it seems that epigenetic heterogeneity has an important clinical and functional effect on leukemia and is an important component of nongenetic mutational strategies used by leukemia cells during their evolutionary trajectories in search of the maximum fitness.

Clonal selection and targeted therapies

A turning point in the natural history of disease occurs with the initiation of treatment. Cytotoxic chemotherapy and targeted treatments impose a major selection pressure on leukemia clones, override most other evolutionary selection factors and become a major determinant of clonal evolution (**Fig. 1**). The effects of therapy in clonal evolution are most evident in the analyses of samples from individuals who have relapsed after treatment with specific targeted agents that inhibit a single factor.

Leukemia lymphoblasts from patients with CML who relapse after treatment with imatinib—a tyrosine-kinase inhibitor targeting the BCR-ABL1 oncoprotein—typically harbor mutations in the *BCR-ABL1* oncogene that abrogate kinase inhibition[97]. Notably, imatinib-resistance driving mutations can be detected in subclones pretreatment in most cases, supporting the notion that the selection of pre-existing resistance populations, and not ongoing acquisition of kinase-domain mutations, drives disease progression and relapse[98]. Treatment with second (dasatinib and nilotinib)- and third (bosutinib, ponatinib)-generation kinase inhibitors can overcome many of the imatinib-resistance-driving mutations[99]. Yet, treatment with these inhibitors can select for a different set of resistance-driving alleles[100].

Similarly, individuals with CLL treated with ibrutinib—a Bruton tyrosine kinase (BTK) inhibitor that abrogates B cell receptor signaling—have *BTK* mutations in their leukemic cells upon relapse that specifically block ibrutinib binding[101], or *PLCG2* mutations, which activate BCR-signaling independently of BTK[102]. Similarly to kinase-inhibitor resistance in CML, *BTK*-resistance-associated mutations can be detected as minor subclones present before treatment initiation,

	Preleukemia →	Diagnosis →	Clinical remission →	Relapse
Clinical objectives	Early detection and risk assessment	Prognostic evaluation and tailored treatment	MRD quantitation and risk assessment	Prognostic evaluation and tailored salvage therapy
Open questions	Can we predict the risk of leukemia progression in patients with clonal hematopoiesis?	Do the mutational profile and the level of genetic and epigenetic clonal heterogeneity predict relapse risk?	Do the mutational profile or the pattern of clonal evolution in MRD predict relapse risk?	Do the mutational profile or the clonal evolutionary pattern at relapse correlate with outcomes after salvage therapy?
Genomic and epigenomic profiling goals	• Clonal hematopoiesis mutational profile • Clonal heterogeneity • Clonal expansion • Clonal evolution	• Leukemia mutational profile • Clonal heterogeneity	• MRD mutational profile • Clonal heterogeneity • Clonal dominance-expansion	• Relapse mutational profile • Clonal heterogeneity • Pattern of clonal evolution

Debbie Maizels/Springer Nature

Figure 4 Clinical implications of the clonal evolution of leukemia. The flow charts show clinical aspects and open questions of clonal evolution in relation to disease prevention, risk stratification, evaluation of therapeutic response and management of therapy resistance in leukemia. MRD, minimal residual disease.

supporting a role for Darwinian selection of pre-existing resistant subclones in disease progression[103,104]. In addition, the treatment of *FLT3*-mutated AML with the multi-target kinase inhibitor midostaurin often results in disease progression, owing to emerging clones harboring FTL3-kinase-domain mutations that impair drug binding and kinase inhibition[105]. Similarly, the treatment of acute promyelocytic leukemia—a type of AML driven by expression of the *PML-RARA* fusion gene—with arsenic trioxide, which induces degradation of the PML-RARA oncoprotein, results in the emergence of cells with mutations in the arsenic-binding B2 domain of PML, in the ligand-binding domain of RARA or in both[106].

Mechanisms of escape and resistance to immunotherapy

Cellular therapy using adoptively transferred T cells armed with chimeric antigen receptors (CAR-T) targeting CD19 has strong antitumor activity in a variety of B cell malignancies[107], with high rates of complete remission in relapse and B-precursor ALL that is refractory to treatment[108–110], and yet, relapse and resistance are an emerging clinical problem[111]. Under the selective pressure of cytotoxic T cells recognizing a single surface antigen, leukemia cells could escape by downregulating expression of the target antigen. However, CD19 is required for B-precursor leukemia growth and imposes a barrier for immunotherapy escape by mere antigen deletion[112]. Relapse after anti-CD19-directed CAR-T cell therapy is often driven by leukemia lymphoblasts expressing alternatively spliced forms of CD19, which are still functional in promoting proliferation but are devoid of the exon encoding the CAR-binding epitope[113]. Persistent CD19 CAR-T cell immune pressure can also induce an epigenetic lineage switch coupled with the loss of CD19 expression as a mechanism of resistance. This seems to be particularly relevant in mixed-lineage leukemia-rearranged (MLLr) infant lymphoblastic leukemias, in which CD19 CAR-T therapy can induce selection of populations with myeloid features accompanied by loss of CD19 expression[114,115].

A similar phenotypic switch to AML has been observed in an individual with MLLr relapsing after CD19 targeted therapy with a CD3–CD19 bispecific antibody[116], and is probably favored by the characteristic mixed lymphoid–myeloid transcriptional program of MLLr ALL[117]. A related reprogramming mechanism of immune escape has been reported in a patient with Richter syndrome, a form of high-grade lymphoma transformed from CLL, who progressed following CAR-T cell therapy with the development of a plasmablastic lymphoma with the loss of CD19 and emergence of a CD19-negative CLL population[118].

Mechanisms of resistance to cytotoxic combination chemotherapy

In the context of multiagent combination chemotherapy, the patterns of mutations selected at relapse vary in different diseases. Analysis of paired diagnostic and relapse ALL samples indicates that relapses emerge primarily from an ancestral clone related to, but different from, the main leukemia population present at diagnosis[6,28,31,119–121]. Moreover, relapsed ALL is highly genetically heterogeneous, probably reflecting the complex regimen of combination chemotherapy used in the treatment of this disease, which includes glucocorticoids, antimetabolites, microtubule-spindle poisons and DNA-damaging agents. Yet, analysis of matched diagnosis and relapsed ALL samples shows frequent mutations in *CREBBP* and other epigenetic regulators[121–123], as well as in oncogenic signaling factors with a particularly high prevalence of activating mutations in the neuroblastoma RAS viral oncogene homolog (*NRAS*), and kirsten rat sarcoma viral oncogene homolog (*KRAS*) genes[121,123].

Notably, some mutations are clearly linked to chemotherapy resistance at the mechanistic level. With regards to treatment with glucocorticoids, mutations in the glucocorticoid-receptor gene *NR3C1* or in *TP53*, which would impair glucocorticoid response and DNA-damage-induced apoptosis, respectively, are commonly found in ALL at relapse. Nevertheless, the most prominent genetic event in ALL at relapse is the presence of gain-of-function mutations in *NT5C2* present in some B-precursor ALL and T-ALL cases at relapse[124,125]. *NT5C2* encodes a cytosolic nucleotidase responsible for the dephosphorylation and export of purine nucleoside monophosphates, an activity that antagonizes the effects of 6-mercaptopurine, a central drug used in the maintenance phase of ALL therapy. Consistently, relapse-associated *NT5C2* mutations are linked to early relapse and progression under therapy and induce resistance to 6-mercaptopurine when expressed in ALL cells[124,125]. In addition, activating mutations in *PRPS1*—which encodes the enzyme responsible for the entry step in purine and pyrimidine synthesis—drive resistance to 6-mercaptopurine by blocking the incorporation of this drug to the salvage pathway of purine biosynthesis[126]. Mutations in *NT5C2, PRPS1* and *CREBBP* are selected for at relapse[121,123]. By contrast, most other mutations associated with ALL at relapse show a heterogeneous pattern of clonal evolution. Thus, activating mutations in *KRAS* and *NRAS* —present in about 40% of relapsed ALL cases—are sometimes retained or acquired at the time of relapse, whereas, in other cases, they are present at diagnosis but lost at relapse. This phenomenon is probably related to the frequent subclonal nature of these mutations, the context-specific

effects on their interaction with other mutations, and their variable influence in response to chemotherapy, with increased sensitivity to spindle-poison drugs and increased resistance to methotrexate and glucocorticoids[121,127,128].

The pattern of evolution at relapse is less well defined in other hematologic tumors than in ALL (**Fig. 1**). In MDS, analysis of the cells associated with disease progression after treatment has indicated a role for both linear and branching evolution. In this case, treatment with nonmyeloablative drugs—such as lenalidomide, 5-azacitidine or temsirolimus—induced fluctuations in the bone marrow oligoclonal composition, yet hematopoiesis remained clonal, as a result of the expansion of either the main clonal population, a minor subclone or a new clonal population not detected before the onset of therapy[38].

In AML, early cytogenetic studies of paired samples at diagnosis and relapse demonstrated that most leukemia cases with chromosomal alteration remained cytogenetically stable or acquired additional chromosomal alterations at relapse; only a few patient samples at relapse presented with karyotypes unrelated to that of their corresponding diagnostic sample[129]. This model was further supported by analysis of copy-number alterations, which verified the close relationship between AML cells at relapse with the major leukemia population detected at diagnosis[130]. Whole-genome sequencing of diagnosis–relapse pairs further established that AML relapses are the result of clonal persistence and linear evolution in most cases, with only a fraction of AML relapses originating from an ancestral yet highly related subclonal population[32]. In this context, some mutations (*NPM1*, *IDH1*, *IDH2*, *DNMT3A*, *ASXL1*, *TET2* and *MLL*-partial tandem duplication) and most acquired copy-number alterations and copy-neutral loss of heterozygosity events are retained at relapse, which supports a role at the early stages of disease development in the pathogenesis of AML. Other mutations that are present at diagnosis, such as those in *CEBPA*, *FLT3*, *RUNX1*, *BCORL1* or *KRAS*, can occasionally be lost at relapse[131–133].

Finally, in CLL, most driver mutations are subclonal at diagnosis[35], and the progression of CLL after therapy has been associated with linear evolution (35/59 cases), branched clonal evolution (21/59) cases and, exceptionally (2 cases), with no apparent genetic evolution evident in whole-exome-sequencing mutation data[36]. In this disease, early driver chromosomal alterations (tri(12), del(13q) and del(11q)) remained stably clonal, and *TP53* mutations, del(17p) and *IKZF3* mutations showed increases in variant frequency indicative of positive selection, whereas mutations in *SF3B1* and *ATM* were equally likely to show increased or decreased clonal frequency following therapy[36].

Clinical implications

The recognition of genetic and epigenetic heterogeneity in leukemia and the realization of an important role of clonal hematopoiesis and clonal evolution in tumor initiation, disease progression and relapse have profound implications for the diagnosis and treatment of these blood malignancies (**Fig. 4**). Increased risk of developing a hematological neoplasia in individuals with age-related clonal hematopoiesis may support the need for close monitoring (**Box 1**). Nevertheless, the cumulative risk of developing leukemia resulting from these factors is relatively low, which highlights the need for identifying improved biomarkers for the detection of those cases at higher risk of transformation. In this regard, we need to precisely assess whether mutations associated with the size of the mutant clone and the dynamics of clonal hematopoiesis correlate with a higher risk of leukemia transformation. Notably, the presence of age-related clonal hematopoiesis with a candidate driver mutation may not confer an elevated risk of developing leukemia over the presence of clonal hematopoiesis per se. By contrast, when clonal hematopoiesis is found in the context of aplastic anemia, the presence of *BCOR-*, *BCORL1-* or *PIGA-* mutated clones is associated with favorable response to immunosuppressive therapy, whereas clones with epigenetic-factor mutations (*DNMT3A* and *ASXL1*) are associated with worse response to immunosuppression, show higher risk of progression to MDS and AML and confer inferior overall survival[56]. Detailed evaluation of the predictive value of quantitative (kinetics of clonal expansion) and qualitative (increased clonal heterogeneity or acquisition of new genetic alterations) changes in clonal composition in the context of age-associated clonal hematopoiesis and aplastic anemia will be required to establish their relevance in prognosis. Therapeutic interventions beyond 'watch and wait' in patients with high risk of leukemia transformation may include chemoprevention using drugs that promote enhanced myeloid differentiation or even treatment with targeted therapies specifically designed to antagonize the effect of driving mutations and curtail the selective advantage of the expanding hematopoietic clone.

Then, upon transformation, we need to know whether increased levels of genetic and epigenetic heterogeneity are associated with outcome. Clonal heterogeneity in AML has not been related to prognosis, yet increased epigenetic heterogeneity may be associated with poorer outcomes[81]. In CLL, high levels of locally disordered methylation at promoters—a measure of both epigenetic heterogeneity and clonal evolution—are associated with adverse clinical outcomes[84,134]. Thus, it is likely that the prognostic value of specific genetic alterations may depend on their clonal or subclonal representation, which opens up the possibility of developing an algorithm for predicting a patient's relapse risk from the analysis of subclonal dynamics of the disease before treatment. However, perhaps the most powerful strategy for improved prognostic evaluation might be the implementation of new tools for the analysis of minimal residual disease (MRD). In ALL, quantitative differences in MRD levels are strongly associated with prognosis and may correlate with relapse risk[135]. Accordingly, genomic profiling and evaluation of clonal composition and dynamics in serial MRD samples may improve the predictive value of these analyses. Likewise, the identification of MRD clonal populations harboring genetic variants associated with chemotherapy resistance may facilitate the development of tailored therapies aimed at curtailing the emergence of relapse.

Conclusions, controversies and future directions

In the present era of genomic revolution, clonal evolution represents an inconvenient truth in leukemia and other human malignancies. We can efficiently sequence the genomes of patients with leukemia, discover specific mutations and propose personalized therapies, but this cartography-based genomic approach is seriously compromised by the multilayered plasticity and tireless adaptive behavior of tumor cells. Arguably, if we move from an impressionist view of the leukemia landscape to a close-up portrait of individual cells captured by leading resolution methods, we will likely demonstrate that every leukemic cell is unique and distinct from all the others that make up that particular hematological neoplasia.

Fortunately, recent progress in different areas suggests that, among this apparent chaos, there is some order waiting to be deciphered. The advent of new technologies, such as single- cell sequencing methods, to analyze intratumor heterogeneity[136], together with *in vivo* and *in vitro* functional analyses of genes associated with this process, is facilitating the definition of the molecular determinants and underlying mechanisms of clonal evolution in leukemia[6]. The currently available global picture of leukemia indicates that multi-clonal heterogeneity is

virtually universal in all blood malignancies. The clonal architecture of each leukemia is the result of the continuous emergence of genetic and epigenetic variants under selection by competition for microenvironment interactions and by administered therapies.

However, beyond this general overview, our knowledge of the dynamics of clonal diversification and the cooperative or competitive forces that finally determine the rise and fall of the evolving subclones is limited. One can envision this process as parallel to that of quasi-species evolution of unicellular infectious microorganisms that struggle for their existence under limited resources in a dynamic environment, and under selection by the host immune system. Likewise, our ability to predict the evolutionary trajectories and the relative biological and clinical relevance of different subclones is limited, thereby hampering our opportunities to develop anticipation-based chemotherapy strategies[5].

Deep-sequence analysis of multiple longitudinal samples collected during disease course and patient treatment could help clinicians to evaluate the comparative dynamics of different subpopulations, and their respective abilities to contribute to the effective repopulation of leukemic niches. Further understanding of the clonal evolutionary process that occurs in individuals with leukemia will also contribute to clarifying the role of therapy as a driver of clonal diversification, as well as to allow the development of innovative treatments that could limit the Darwinian-selection-driven resilience of the disease. The development of these therapies will have to address questions on what is the more appropriate scheduling for combinatorial approaches that target both early clonal genetic lesions and branch mutations arising during the diversification process. Immunotherapy offers much promise as a transversal therapeutic approach covering multiple clonal populations. Indeed, CAR-T cell therapies, bispecific T cell–engager antibodies—such as blinatumomab—and immune-checkpoint inhibitors—such as nivolumab—have already offered promising results for the treatment of different relapsed and/or refractory hematological neoplasias[137]. Lastly, new immunotherapy approaches aimed at targeting the adaptive neoantigen landscape generated by the evolving clones may offer additional therapeutic opportunities[107].

Many of the studies discussed in this article are based on exome data, but this approach fails to capture the complete clonal architecture of leukemias, and it lacks the comprehensiveness of whole-genome sequencing and the depth of targeted sequencing approaches. Furthermore, although epigenetic changes add another even more variable layer to the molecular complexity of clonal evolution, it is yet unclear whether epigenetic alterations are a mere epiphenomenon secondary to mutations and structural chromosomal alterations, or whether they play an active part themselves in the evolutionary process. Likewise, the cell of origin and the relevant cellular compartment for clonal diversification and selection—stem cells as opposed to early or committed precursors—remain to be fully clarified for many leukemias. Although in some cases founder mutations may occur in HSCs, it is also possible that some genetic events—in particular, mutations in epigenetic regulators—may confer self-renewal properties to more committed progenitors. Yet, this is an area for which direct experimental data are limited. Additional caution in defining and answering this question is also warranted by recent studies redefining the hierarchy of hematopoietic differentiation and the process of lineage commitment[138,139]. Larger studies using single-cell analysis, whole-genome sequencing with substantial coverage, and deep, targeted resequencing of large series of leukemia samples are needed to best define both the cell of origin and the dynamics of malignant evolution, as well as the precise structure of evolutionary trees. Additional, large-scale and highly sensitive studies will also be required to evaluate the prevalence of age-related clonal hematopoiesis in healthy individuals and to identify robust markers predictive of leukemia transformation.

In summary, leukemia cells struggle for their existence and explore their particular 'garden of forking paths' through clonal evolutionary processes that are, in essence, Darwinian. However, there is no grandeur in this view of life, because the fierce competition for space and resources of leukemic cells under multiple environmental pressures continuously generates new entities that seriously compromise the life of patients with leukemia. Hopefully, the pervasive advance in the understanding of these complex evolutionary histories occurring in the course of a human lifetime may finally lead to new clinical strategies for a more effective control of blood malignancies.

Acknowledgments
We thank M. Mittelbrunn (CBM-Hospital 12 de Octubre, Madrid, Spain), X.S. Puente (Universidad de Oviedo, Oviedo, Spain), P. Menéndez (J. Carreras Leukemia Research Institute, Barcelona, Spain), R. Rabadán (Columbia University, New York, New York, USA), J. Soulier (Université Paris Diderot, Paris, France) and all members of our labs for their helpful comments on the manuscript. A.A.F. is supported by grants from the National Cancer Institute (NCI) of the National Institutes of Health (NIH), the Leukemia & Lymphoma Society, the Chemotherapy Foundation and the Rally Foundation. C.L.-O. is supported by grants from European Union (DeAge, ERC-Advanced Grant), Ministerio de Economía y Competitividad SAF2014-52413-R, Instituto de Salud Carlos III (RTICC), CIBERONC, Plan Feder, and EDP Foundation. The generous support by J.I. Cabrera is also acknowledged.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Nowell, P.C. The clonal evolution of tumor cell populations. *Science* **194**, 23–28 (1976).
2. Merlo, L.M., Pepper, J.W., Reid, B.J. & Maley, C.C. Cancer as an evolutionary and ecological process. *Nat. Rev. Cancer* **6**, 924–935 (2006).
3. Burrell, R.A., McGranahan, N., Bartek, J. & Swanton, C. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* **501**, 338–345 (2013).
4. Greaves, M. Evolutionary determinants of cancer. *Cancer Discov.* **5**, 806–820 (2015).
5. Puente, X.S. & López-Otín, C. The evolutionary biography of chronic lymphocytic leukemia. *Nat. Genet.* **45**, 229–231 (2013).
6. Landau, D.A., Carter, S.L., Getz, G. & Wu, C.J. Clonal evolution in hematological malignancies and therapeutic implications. *Leukemia* **28**, 34–43 (2014).
7. Greaves, M. Leukaemia 'firsts' in cancer research and treatment. *Nat. Rev. Cancer* **16**, 163–172 (2016).
8. Dick, J.E. Stem cell concepts renew cancer research. *Blood* **112**, 4793–4807 (2008).
9. Jan, M. & Majeti, R. Clonal evolution of acute leukemia genomes. *Oncogene* **32**, 135–140 (2013).
10. Malkin, D. *et al.* Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science* **250**, 1233–1238 (1990).
11. Song, W.J. *et al.* Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. *Nat. Genet.* **23**, 166–175 (1999).
12. Smith, M.L., Cavenagh, J.D., Lister, T.A. & Fitzgibbon, J. Mutation of CEBPA in familial acute myeloid leukemia. *N. Engl. J. Med.* **351**, 2403–2407 (2004).
13. Hahn, C.N. *et al.* Heritable GATA2 mutations associated with familial myelodysplastic syndrome and acute myeloid leukemia. *Nat. Genet.* **43**, 1012–1017 (2011).
14. Noris, P. *et al.* ANKRD26-related thrombocytopenia and myeloid malignancies. *Blood* **122**, 1987–1989 (2013).
15. Polprasert, C. *et al.* Inherited and somatic defects in DDX41 in myeloid neoplasms. *Cancer Cell* **27**, 658–670 (2015).
16. Shah, S. *et al.* A recurrent germline PAX5 mutation confers susceptibility to pre-B cell acute lymphoblastic leukemia. *Nat. Genet.* **45**, 1226–1231 (2013).
17. Zhang, M.Y. *et al.* Germline ETV6 mutations in familial thrombocytopenia and hematologic malignancy. *Nat. Genet.* **47**, 180–185 (2015).

18. Moriyama, T. *et al*. Germline genetic variation in ETV6 and risk of childhood acute lymphoblastic leukaemia: a systematic genetic study. *Lancet Oncol*. **16**, 1659–1666 (2015).
19. Greaves, M.F., Maia, A.T., Wiemels, J.L. & Ford, A.M. Leukemia in twins: lessons in natural history. *Blood* **102**, 2321–2333 (2003).
20. Sanjuan-Pla, A. *et al*. Revisiting the biology of infant t(4;11)/MLL-AF4+ B-cell acute lymphoblastic leukemia. *Blood* **126**, 2676–2685 (2015).
21. Hong, D. *et al*. Initiating and cancer-propagating cells in TEL-AML1-associated childhood leukemia. *Science* **319**, 336–339 (2008).
22. Greaves, M. Infection, immune responses and the aetiology of childhood leukaemia. *Nat. Rev. Cancer* **6**, 193–203 (2006).
23. Jacobs, K.B. *et al*. Detectable clonal mosaicism and its relationship to aging and cancer. *Nat. Genet*. **44**, 651–658 (2012).
24. Laurie, C.C. *et al*. Detectable clonal mosaicism from birth to old age and its relationship to cancer. *Nat. Genet*. **44**, 642–650 (2012).
25. Genovese, G. *et al*. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N. Engl. J. Med*. **371**, 2477–2487 (2014).
26. Jaiswal, S. *et al*. Age-related clonal hematopoiesis associated with adverse outcomes. *N. Engl. J. Med*. **371**, 2488–2498 (2014).
27. Busque, L. *et al*. Recurrent somatic TET2 mutations in normal elderly individuals with clonal hematopoiesis. *Nat. Genet*. **44**, 1179–1181 (2012).
28. Anderson, K. *et al*. Genetic variegation of clonal architecture and propagating cells in leukaemia. *Nature* **469**, 356–361 (2011).
29. Li, A.H., Rosenquist, R., Forestier, E., Lindh, J. & Roos, G. Detailed clonality analysis of relapsing precursor B acute lymphoblastic leukemia: implications for minimal residual disease detection. *Leuk. Res*. **25**, 1033–1045 (2001).
30. de Haas, V. *et al*. Quantification of minimal residual disease in children with oligoclonal B-precursor acute lymphoblastic leukemia indicates that the clones that grow out during relapse already have the slowest rate of reduction during induction therapy. *Leukemia* **15**, 134–140 (2001).
31. Notta, F. *et al*. Evolution of human BCR-ABL1 lymphoblastic leukaemia-initiating cells. *Nature* **469**, 362–367 (2011).
32. Ding, L. *et al*. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* **481**, 506–510 (2012).
33. Jan, M. *et al*. Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. *Sci. Transl. Med*. **4**, 149ra118 (2012).
34. Paguirigan, A.L. *et al*. Single-cell genotyping demonstrates complex clonal diversity in acute myeloid leukemia. *Sci. Transl. Med*. **7**, 281re2 (2015).
35. Puente, X.S. *et al*. Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature* **526**, 519–524 (2015).
36. Landau, D.A. *et al*. Mutations driving CLL and their evolution in progression and relapse. *Nature* **526**, 525–530 (2015).
37. Makishima, H. *et al*. Dynamics of clonal evolution in myelodysplastic syndromes. *Nat. Genet*. **49**, 204–212 (2017).
38. Mossner, M. *et al*. Mutational hierarchies in myelodysplastic syndromes dynamically adapt and evolve upon therapy response and failure. *Blood* **128**, 1246–1259 (2016).
39. Miyamoto, T., Weissman, I.L. & Akashi, K. AML1/ETO-expressing nonleukemic stem cells in acute myelogenous leukemia with 8;21 chromosomal translocation. *Proc. Natl. Acad. Sci. USA* **97**, 7521–7526 (2000).
40. Shlush, L.I. *et al*. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature* **506**, 328–333 (2014).
41. Damm, F. *et al*. Acquired initiating mutations in early hematopoietic cells of CLL patients. *Cancer Discov*. **4**, 1088–1101 (2014).
42. Kikushige, Y. *et al*. Self-renewing hematopoietic stem cell is the primary target in pathogenesis of human chronic lymphocytic leukemia. *Cancer Cell* **20**, 246–259 (2011).
43. Chung, S.S. *et al*. Hematopoietic stem cell origin of BRAFV600E mutations in hairy cell leukemia. *Sci. Transl. Med*. **6**, 238ra71 (2014).
44. Sperling, A.S., Gibson, C.J. & Ebert, B.L. The genetics of myelodysplastic syndrome: from clonal haematopoiesis to secondary leukaemia. *Nat. Rev. Cancer* **17**, 5–19 (2017).
45. Walter, M.J. *et al*. Clonal architecture of secondary acute myeloid leukemia. *N. Engl. J. Med*. **366**, 1090–1098 (2012).
46. Green, M.R. *et al*. Mutations in early follicular lymphoma progenitors are associated with suppressed antigen presentation. *Proc. Natl. Acad. Sci. USA* **112**, E1116–E1125 (2015).
47. Xie, M. *et al*. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat. Med*. **20**, 1472–1478 (2014).
48. Corces-Zimmerman, M.R., Hong, W.J., Weissman, I.L., Medeiros, B.C. & Majeti, R. Preleukemic mutations in human acute myeloid leukemia affect epigenetic regulators and persist in remission. *Proc. Natl. Acad. Sci. USA* **111**, 2548–2553 (2014).
49. Horiike, S. *et al*. Distinct genetic involvement of the TP53 gene in therapy-related leukemia and myelodysplasia with chromosomal losses of Nos 5 and/or 7 and its possible relationship to replication error phenotype. *Leukemia* **13**, 1235–1242 (1999).
50. Side, L.E. *et al*. RAS, FLT3, and TP53 mutations in therapy-related myeloid malignancies with abnormalities of chromosomes 5 and 7. *Genes Chromosom. Cancer* **39**, 217–223 (2004).
51. Wong, T.N. *et al*. Role of TP53 mutations in the origin and evolution of therapy-related acute myeloid leukaemia. *Nature* **518**, 552–555 (2015).
52. Takahashi, K. *et al*. Preleukaemic clonal haemopoiesis and risk of therapy-related myeloid neoplasms: a case-control study. *Lancet Oncol*. **18**, 100–111 (2017).
53. Gibson, C.J. *et al*. Clonal hematopoiesis associated with adverse outcomes after autologous stem-cell transplantation for lymphoma. *J. Clin. Oncol*. **35**, 1598–1605 (2017).
54. Young, N.S., Calado, R.T. & Scheinberg, P. Current concepts in the pathophysiology and treatment of aplastic anemia. *Blood* **108**, 2509–2519 (2006).
55. Socié, G., Rosenfeld, S., Frickhofen, N., Gluckman, E. & Tichelli, A. Late clonal diseases of treated aplastic anemia. *Semin. Hematol*. **37**, 91–101 (2000).
56. Yoshizato, T. *et al*. Somatic mutations and clonal hematopoiesis in aplastic anemia. *N. Engl. J. Med*. **373**, 35–47 (2015).
57. Maciejewski, J.P., Risitano, A., Sloand, E.M., Nunez, O. & Young, N.S. Distinct clinical outcomes for cytogenetic abnormalities evolving from aplastic anemia. *Blood* **99**, 3129–3135 (2002).
58. Dumitriu, B. *et al*. Telomere attrition and candidate gene mutations preceding monosomy 7 in aplastic anemia. *Blood* **125**, 706–709 (2015).
59. Katagiri, T. *et al*. Frequent loss of HLA alleles associated with copy number-neutral 6pLOH in acquired aplastic anemia. *Blood* **118**, 6601–6609 (2011).
60. Afable, M.G. II *et al*. SNP array-based karyotyping: differences and similarities between aplastic anemia and hypocellular myelodysplastic syndromes. *Blood* **117**, 6876–6884 (2011).
61. Hillmen, P., Lewis, S.M., Bessler, M., Luzzatto, L. & Dacie, J.V. Natural history of paroxysmal nocturnal hemoglobinuria. *N. Engl. J. Med*. **333**, 1253–1258 (1995).
62. Ogawa, S. Clonal hematopoiesis in acquired aplastic anemia. *Blood* **128**, 337–347 (2016).
63. Quentin, S. *et al*. Myelodysplasia and leukemia of Fanconi anemia are associated with a specific pattern of genomic abnormalities that includes cryptic RUNX1/AML1 lesions. *Blood* **117**, e161–e170 (2011).
64. Horwitz, M., Benson, K.F., Person, R.E., Aprikyan, A.G. & Dale, D.C. Mutations in ELA2, encoding neutrophil elastase, define a 21-day biological clock in cyclic haematopoiesis. *Nat. Genet*. **23**, 433–436 (1999).
65. Klein, C. *et al*. HAX1 deficiency causes autosomal recessive severe congenital neutropenia (Kostmann disease). *Nat. Genet*. **39**, 86–92 (2007).
66. Devriendt, K. *et al*. Constitutively activating mutation in WASP causes X-linked severe congenital neutropenia. *Nat. Genet*. **27**, 313–317 (2001).
67. Bonilla, M.A. *et al*. Effects of recombinant human granulocyte colony-stimulating factor on neutropenia in patients with congenital agranulocytosis. *N. Engl. J. Med*. **320**, 1574–1580 (1989).
68. Rosenberg, P.S. *et al*. Stable long-term risk of leukaemia in patients with severe congenital neutropenia maintained on G-CSF therapy. *Br. J. Haematol*. **150**, 196–199 (2010).
69. Germeshausen, M., Ballmaier, M. & Welte, K. Incidence of CSF3R mutations in severe congenital neutropenia and relevance for leukemogenesis: Results of a long-term survey. *Blood* **109**, 93–99 (2007).
70. Skokowa, J. *et al*. Cooperativity of RUNX1 and CSF3R mutations in severe congenital neutropenia: a unique pathway in myeloid leukemogenesis. *Blood* **123**, 2229–2237 (2014).
71. Papaemmanuil, E. *et al*. Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood* **122**, 3616–3627, quiz 3699 (2013).
72. Ortmann, C.A. *et al*. Effect of mutation order on myeloproliferative neoplasms. *N. Engl. J. Med*. **372**, 601–612 (2015).
73. Cortés, J.R. & Palomero, T. The curious origins of angioimmunoblastic T-cell lymphoma. *Curr. Opin. Hematol*. **23**, 434–443 (2016).
74. Shaknovich, R., De, S. & Michor, F. Epigenetic diversity in hematopoietic neoplasms. *Biochim. Biophys. Acta* **1846**, 477–484 (2014).
75. Guièze, R. & Wu, C.J. Genomic and epigenomic heterogeneity in chronic lymphocytic leukemia. *Blood* **126**, 445–453 (2015).
76. Li, S., Mason, C.E. & Melnick, A. Genetic and epigenetic heterogeneity in acute myeloid leukemia. *Curr. Opin. Genet. Dev*. **36**, 100–106 (2016).
77. Kulis, M. *et al*. Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. *Nat. Genet*. **44**, 1236–1242 (2012).
78. Figueroa, M.E. *et al*. DNA methylation signatures identify biologically distinct subtypes in acute myeloid leukemia. *Cancer Cell* **17**, 13–27 (2010).
79. Milani, L. *et al*. DNA methylation for subtype classification and prediction of treatment outcome in patients with childhood acute lymphoblastic leukemia. *Blood* **115**, 1214–1225 (2010).
80. Geng, H. *et al*. Integrative epigenomic analysis identifies biomarkers and therapeutic targets in adult B-acute lymphoblastic leukemia. *Cancer Discov*. **2**, 1004–1023 (2012).
81. Li, S. *et al*. Distinct evolution and dynamics of epigenetic and genetic heterogeneity in acute myeloid leukemia. *Nat. Med*. **22**, 792–799 (2016).
82. Pan, H. *et al*. Epigenomic evolution in diffuse large B-cell lymphomas. *Nat. Commun*. **6**, 6921 (2015).
83. Sandoval, J. *et al*. Genome-wide DNA methylation profiling predicts relapse in childhood B-cell acute lymphoblastic leukaemia. *Br. J. Haematol*. **160**, 406–409 (2013).
84. Landau, D.A. *et al*. Locally disordered methylation forms the basis of intratumor methylome variation in chronic lymphocytic leukemia. *Cancer Cell* **26**, 813–825 (2014).

85. Oakes, C.C. *et al*. DNA methylation dynamics during B cell maturation underlie a continuum of disease phenotypes in chronic lymphocytic leukemia. *Nat. Genet.* **48**, 253–264 (2016).
86. Heller, G. *et al*. Next-generation sequencing identifies major DNA methylation changes during progression of Ph+ chronic myeloid leukemia. *Leukemia* **30**, 1861–1868 (2016).
87. Feinberg, A.P., Koldobskiy, M.A. & Göndör, A. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. *Nat. Rev. Genet.* **17**, 284–299 (2016).
88. Greenblatt, S.M. & Nimer, S.D. Chromatin modifiers and the promise of epigenetic therapy in acute leukemia. *Leukemia* **28**, 1396–1406 (2014).
89. Roberts, K.G. & Mullighan, C.G. Genomics in acute lymphoblastic leukaemia: insights and treatment implications. *Nat. Rev. Clin. Oncol.* **12**, 344–357 (2015).
90. Woods, B.A. & Levine, R.L. The role of mutations in epigenetic regulators in myeloid malignancies. *Immunol. Rev.* **263**, 22–35 (2015).
91. Shen, H. & Laird, P.W. Interplay between the cancer genome and epigenome. *Cell* **153**, 38–55 (2013).
92. Oakes, C.C. *et al*. Evolution of DNA methylation is linked to genetic aberrations in chronic lymphocytic leukemia. *Cancer Discov.* **4**, 348–361 (2014).
93. Amabile, G. *et al*. Dissecting the role of aberrant DNA methylation in human leukaemia. *Nat. Commun.* **6**, 7091 (2015).
94. Shih, A.H. *et al*. Mutational cooperativity linked to combinatorial epigenetic gain of function in acute myeloid leukemia. *Cancer Cell* **27**, 502–515 (2015).
95. Zhang, X. *et al*. DNMT3A and TET2 compete and cooperate to repress lineage-specific transcription factors in hematopoietic stem cells. *Nat. Genet.* **48**, 1014–1023 (2016).
96. Pasqualucci, L. *et al*. Expression of the AID protein in normal and neoplastic B cells. *Blood* **104**, 3318–3325 (2004).
97. Gorre, M.E. & Sawyers, C.L. Molecular mechanisms of resistance to STI571 in chronic myeloid leukemia. *Curr. Opin. Hematol.* **9**, 303–307 (2002).
98. Roche-Lestienne, C. *et al*. Several types of mutations of the Abl gene can be found in chronic myeloid leukemia patients resistant to STI571, and they can pre-exist to the onset of treatment. *Blood* **100**, 1014–1018 (2002).
99. Branford, S., Melo, J.V. & Hughes, T.P. Selecting optimal second-line tyrosine kinase inhibitor therapy for chronic myeloid leukemia patients after imatinib failure: does the BCR-ABL mutation status really matter? *Blood* **114**, 5426–5435 (2009).
100. Cortes, J. *et al*. Dynamics of BCR-ABL kinase domain mutations in chronic myeloid leukemia after sequential treatment with multiple tyrosine kinase inhibitors. *Blood* **110**, 4005–4011 (2007).
101. Woyach, J.A. *et al*. Resistance mechanisms for the Bruton's tyrosine kinase inhibitor ibrutinib. *N. Engl. J. Med.* **370**, 2286–2294 (2014).
102. Liu, T.M. *et al*. Hypermorphic mutation of phospholipase C, $\gamma 2$ acquired in ibrutinib-resistant CLL confers BTK independency upon B-cell receptor activation. *Blood* **126**, 61–68 (2015).
103. Burger, J.A. *et al*. Clonal evolution in patients with chronic lymphocytic leukaemia developing resistance to BTK inhibition. *Nat. Commun.* **7**, 11589 (2016).
104. Ahn, I.E. *et al*. Clonal evolution leading to ibrutinib resistance in chronic lymphocytic leukemia. *Blood* **129**, 1469–1479 (2017).
105. Smith, C.C. *et al*. Validation of ITD mutations in FLT3 as a therapeutic target in human acute myeloid leukaemia. *Nature* **485**, 260–263 (2012).
106. Goto, E. *et al*. Missense mutations in PML-RARA are critical for the lack of responsiveness to arsenic trioxide treatment. *Blood* **118**, 1600–1609 (2011).
107. Maus, M.V., Grupp, S.A., Porter, D.L. & June, C.H. Antibody-modified T cells: CARs take the front seat for hematologic malignancies. *Blood* **123**, 2625–2635 (2014).
108. Davila, M.L. *et al*. Efficacy and toxicity management of 19-28z CAR T cell therapy in B cell acute lymphoblastic leukemia. *Sci. Transl. Med.* **6**, 224ra25 (2014).
109. Lee, D.W. *et al*. T cells expressing CD19 chimeric antigen receptors for acute lymphoblastic leukaemia in children and young adults: a phase 1 dose-escalation trial. *Lancet* **385**, 517–528 (2015).
110. Maude, S.L. *et al*. Chimeric antigen receptor T cells for sustained remissions in leukemia. *N. Engl. J. Med.* **371**, 1507–1517 (2014).
111. Restifo, N.P., Smyth, M.J. & Snyder, A. Acquired resistance to immunotherapy and future challenges. *Nat. Rev. Cancer* **16**, 121–126 (2016).
112. Chung, E.Y. *et al*. CD19 is a major B cell receptor-independent activator of MYC-driven B-lymphomagenesis. *J. Clin. Invest.* **122**, 2257–2266 (2012).
113. Sotillo, E. *et al*. Convergence of acquired mutations and alternative splicing of CD19 enables resistance to CART-19 immunotherapy. *Cancer Discov.* **5**, 1282–1295 (2015).
114. Jacoby, E. *et al*. CD19 CAR immune pressure induces B-precursor acute lymphoblastic leukaemia lineage switch exposing inherent leukaemic plasticity. *Nat. Commun.* **7**, 12320 (2016).
115. Gardner, R. *et al*. Acquisition of a CD19-negative myeloid phenotype allows immune escape of MLL-rearranged B-ALL from CD19 CAR-T-cell therapy. *Blood* **127**, 2406–2410 (2016).
116. Rayes, A., McMasters, R.L. & O'Brien, M.M. Lineage switch in MLL-rearranged infant leukemia following CD19-directed therapy. *Pediatr. Blood Cancer* **63**, 1113–1115 (2016).
117. Armstrong, S.A. *et al*. MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nat. Genet.* **30**, 41–47 (2002).
118. Evans, A.G. *et al*. Evolution to plasmablastic lymphoma evades CD19-directed chimeric antigen receptor T cells. *Br. J. Haematol.* **171**, 205–209 (2015).
119. Mullighan, C.G. *et al*. Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science* **322**, 1377–1380 (2008).
120. Bardini, M. *et al*. Clonal variegation and dynamic competition of leukemia-initiating cells in infant acute lymphoblastic leukemia with MLL rearrangement. *Leukemia* **29**, 38–50 (2015).
121. Oshima, K. *et al*. Mutational landscape, clonal evolution patterns, and role of RAS mutations in relapsed acute lymphoblastic leukemia. *Proc. Natl. Acad. Sci. USA* **113**, 11306–11311 (2016).
122. Mullighan, C.G. *et al*. CREBBP mutations in relapsed acute lymphoblastic leukaemia. *Nature* **471**, 235–239 (2011).
123. Ma, X. *et al*. Rise and fall of subclones from diagnosis to relapse in pediatric B-acute lymphoblastic leukaemia. *Nat. Commun.* **6**, 6604 (2015).
124. Meyer, J.A. *et al*. Relapse-specific mutations in NT5C2 in childhood acute lymphoblastic leukemia. *Nat. Genet.* **45**, 290–294 (2013).
125. Tzoneva, G. *et al*. Activating mutations in the NT5C2 nucleotidase gene drive chemotherapy resistance in relapsed ALL. *Nat. Med.* **19**, 368–371 (2013).
126. Li, B. *et al*. Negative feedback-defective PRPS1 mutants drive thiopurine resistance in relapsed childhood ALL. *Nat. Med.* **21**, 563–571 (2015).
127. Ariës, I.M. *et al*. Towards personalized therapy in pediatric acute lymphoblastic leukemia: RAS mutations and prednisolone resistance. *Haematologica* **100**, e132–e136 (2015).
128. Jones, C.L. *et al*. MAPK signaling cascades mediate distinct glucocorticoid resistance mechanisms in pediatric leukemia. *Blood* **126**, 2202–2212 (2015).
129. Estey, E., Keating, M.J., Pierce, S. & Stass, S. Change in karyotype between diagnosis and first relapse in acute myelogenous leukemia. *Leukemia* **9**, 972–976 (1995).
130. Raghavan, M. *et al*. Segmental uniparental disomy is a commonly acquired genetic event in relapsed acute myeloid leukemia. *Blood* **112**, 814–821 (2008).
131. Parkin, B. *et al*. Clonal evolution and devolution after chemotherapy in adult acute myelogenous leukemia. *Blood* **121**, 369–377 (2013).
132. Sood, R. *et al*. Somatic mutational landscape of AML with inv(16) or t(8;21) identifies patterns of clonal evolution in relapse leukemia. *Leukemia* **30**, 501–504 (2016).
133. Krönke, J. *et al*. Clonal evolution in relapsed NPM1-mutated acute myeloid leukemia. *Blood* **122**, 100–108 (2013).
134. Nadeu, F. *et al*. Clinical impact of clonal and subclonal TP53, SF3B1, BIRC3, NOTCH1, and ATM mutations in chronic lymphocytic leukemia. *Blood* **127**, 2122–2130 (2016).
135. Pui, C.H. *et al*. Clinical utility of sequential minimal residual disease measurements in the context of risk-based therapy in childhood acute lymphoblastic leukaemia: a prospective study. *Lancet Oncol.* **16**, 465–474 (2015).
136. Kim, J.Y. & Gatenby, R.A. Quantitative clinical imaging methods for monitoring intratumoral evolution. *Methods Mol. Biol.* **1513**, 61–81 (2017).
137. Batlevi, C.L., Matsuki, E., Brentjens, R.J. & Younes, A. Novel immunotherapies in lymphoid malignancies. *Nat. Rev. Clin. Oncol.* **13**, 25–40 (2016).
138. Notta, F. *et al*. Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science* **351**, aab2116 (2016).
139. Sun, J. *et al*. Clonal dynamics of native haematopoiesis. *Nature* **514**, 322–327 (2014).
140. Busque, L. *et al*. Nonrandom X-inactivation patterns in normal females: lyonization ratios vary with age. *Blood* **88**, 59–65 (1996).
141. Kwok, B. *et al*. MDS-associated somatic mutations and clonal hematopoiesis are common in idiopathic cytopenias of undetermined significance. *Blood* **126**, 2355–2361 (2015).
142. Young, A.L., Challen, G.A., Birmann, B.M. & Druley, T.E. Clonal haematopoiesis harbouring AML-associated mutations is ubiquitous in healthy adults. *Nat. Commun.* **7**, 12484 (2016).
143. Jones, P.A. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **13**, 484–492 (2012).