Seminars in Cancer Biology xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Seminars in Cancer Biology

journal homepage: www.elsevier.com/locate/semcancer

Review

Epigenetics in pediatric acute lymphoblastic leukemia

Jessica Nordlund*, Ann-Christine Syvänen

Department of Medical Sciences and Science for Life Laboratory, Uppsala University, Sweden

ARTICLE INFO

Keywords:
Acute lymphoblastic leukemia (ALL)
DNA methylation
Epigenetics
Epigenomics
Subtyping

ABSTRACT

Acute lymphoblastic leukemia (ALL) is the most common malignancy in children. ALL arises from the malignant transformation of progenitor B- and T-cells in the bone marrow into leukemic cells, but the mechanisms underlying this transformation are not well understood. Recent technical advances and decreasing costs of methods for high-throughput DNA sequencing and SNP genotyping have stimulated systematic studies of the epigenetic changes in leukemic cells from pediatric ALL patients. The results emerging from these studies are increasing our understanding of the epigenetic component of leukemogenesis and have demonstrated the potential of DNA methylation as a biomarker for lineage and subtype classification, prognostication, and disease progression in ALL. In this review, we provide a concise examination of the epigenetic studies in ALL, with a focus on DNA methylation and mutations perturbing genes involved in chromatin modification, and discuss the future role of epigenetic analyses in research and clinical management of ALL.

1. Acute lymphoblastic leukemia (ALL)

Acute lymphoblastic leukemia (ALL) is the most common form of pediatric cancer. Structural chromosomal rearrangements, which can lead to expressed fusion genes, are together with clinical features such as white blood cell count at diagnosis and minimal residual disease status, the main basis for diagnosis, risk stratification, and prognosis of pediatric ALL [1]. Patients with ALL are classified into genetic subtypes based on the occurrence of recurrent chromosomal abnormalities detected by karyotyping (G-banding), fluorescent in situ hybridization (FISH), and/or polymerase chain reaction (PCR) amplification. ALL arises from hematopoietic cells in either the B-cell precursor (BCP-ALL) or T-cell lineages (T-ALL). Both the BCP-ALL and T-ALL immunophenotype groups comprise multiple subtypes defined by chromosomal alterations that are believed to be the leukemia-initiating lesions [1]. In most protocols, the important subtypes for prognosis of ALL are T-ALL and the BCP-ALL subtypes high hyperdiploidy (HeH), t(12;21)*ETV6-RUNX1*, t(1;19)*E2A-PBX1*, t(9;22)*BCR-ABL1*, dic(9;20), iAMP21, hypodiploidy (< 45 chr), and *KMT2A* (also known as *MLL1*) rearrangements.

However, 20% of newly diagnosed BCP-ALL cases do not belong to any of the known genetic subtypes. These patients, whose subtype is referred to as B-other, are limited to clinical and minimal residual disease data for informing treatment decisions and also typically lack leukemia-specific genetic changes that can be used for disease monitoring during treatment [2]. During 2016–2017, several previously unknown recurrent genomic rearrangements involving the *DUX4*, *ZNF384*, or *MEF2D* genes were discovered in the B-other subgroup [3,4,5,6]. These and other emerging subtypes of ALL have been reviewed in detail elsewhere [1].

Normal hematopoietic cell development requires tightly controlled regulation of DNA methylation, chemical modification of histones, and expression of non-coding RNAs, all of which may be deregulated during leukemic transformation. DNA methylation is by far the most well characterized epigenetic modification, and is involved in the regulation of gene expression, maintenance of genome stability, and cellular differentiation. Many studies have implicated aberrant epigenetic regulation in the pathogenesis, treatment outcome and recurrence of ALL. This review will summarize and discuss the functions and consequences of epigenetic alterations with a focus on DNA methylation and somatic mutational signatures in epigenetic regulating genes in pediatric ALL as elucidated by recent studies.

Abbreviations: 5mc, 5-methyl cytosine; 5hmc, 5-hydroxymethylcytosine; ALL, Acute Lymphoblastic Leukemia; ASE, Allele-specific gene expression; B-All, B-cell precursor lineage ALL; CpG, CG dinucleotide; CGI, CpG island; CIMP, CpG island methylator phenotype; DMCs, Differentially methylated CpG sites; DMRs, Differentially methylated regions; DNAm, DNA methylation; EFS, Event free survival; ETP-ALL, Early T-cell precursor ALL; FISH, Fluorescent in situ hybridization; HELP, HpaII tiny fragment Enrichment by Ligation-mediated PCR; NGS, Next generation sequencing; OS, Overall survival; PCR, Polymerase chain reaction; RRBS, Reduced representation bisulfite sequencing; T-ALL, T-cell lineage ALL; WGBS, Whole genome bisulfite sequencing

* Corresponding author at: Molecular Medicine, Box 1432, BMC, 75144 Uppsala, Sweden.
E-mail address: Jessica.Nordlund@medsci.uu.se (J. Nordlund).

http://dx.doi.org/10.1016/j.semcancer.2017.09.001
Received 12 June 2017; Received in revised form 21 August 2017; Accepted 2 September 2017

2. Technology for DNA methylation analysis

The human genome contains 28 million CpG sites that are targets for DNA methylation. The discovery of epigenetic biomarkers has significantly expanded owing to the development of methods for interrogating DNA methylation on a genome-wide scale using hybridization microarrays and next generation sequencing (NGS). Methods for detecting DNA methylation can be divided into three broad classes; those involving (I) enrichment of methylated genomic positions, (II) digestion with methylation- sensitive restriction enzymes, and (III) utilization of sodium bisulfite treatment, which converts unmethylated cytosine residues into thymine, while methylated cytosines are protected against conversion. These approaches vary in many aspects, such as required DNA input, degree of genomic resolution and coverage, and ability of quantification. Thus, when comparing results from different DNA methylation studies the method(s) used for determination of DNA methylation levels should be considered. Here, we will briefly summarize methodologies applied for genome-wide DNA methylation analysis in studies of ALL patient cohorts.

The earliest genome-wide methods implemented in ALL involved enrichment of methylated DNA fragments, either by a combination of methylation specific restriction enzyme digestion and PCR [7,8] or immunoprecipitation of methylated DNA fragments [9], followed by reading out the results using hybridization microarrays. The relative abundance of the enriched DNA regions is an estimate of the amount of cytosine methylation at any given region of the genome that is represented by probes on the array. The strength of these methods is that they cover relatively large genomic regions, but a limitation is that enrichment methods do not provide single-base resolution of DNA methylation at individual CpG sites.

Sodium bisulfite treatment of DNA enables measurement of the methylation status of individual cytosine residues at a single-base resolution. Targeted analysis of DNA methylation in bisulfite-treated DNA can in principle be performed by any SNP genotyping method. The MassArray system (Agena, formerly Sequenom) has been used during more than two decades because it allows robust quantification of the DNA methylation levels at targeted CpG sites [10]. A custom designed Golden Gate assay (Illumina) for 1536 CpG sites in candidate gene promoter regions was used in an early study of allele-specific regulation of gene expression by DNA methylation in primary ALL cells [11]. Today the most frequently used methods with capacity of single-base resolution of bisulfite-converted DNA with high-throughput analysis of many CpG sites and samples in parallel are the Infinium BeadChip assays (Illumina). The BeadChip assays interrogate the methylation status of cytosine residues by genotyping cytosine or thymine (methylated vs unmethylated cytosine residues) using a predetermined set of probes in a microarray format. The Infinium assays offer quantitative measurement of DNA methylation and have been launched with increasing numbers of target CpG sites over the last decade, starting with the HumanMethylation 27 K BeadChip (27k array) that mostly targeted CpG islands [12], followed by the HumanMethylation 450 K (450k array) [13] and Infinium MethylationEPIC (850k array) BeadChips [14], which in addition to CpG islands and genes, also assay CpG island shores, gene bodies, enhancers, and other non-coding genomic regions. The BeadChips provide a user friendly and straightforward approach for analyzing hundreds of thousands of CpG sites in many samples at a relatively low cost.

Although the targeted approaches using BeadChip assays for DNA methylation analysis offer advantages for analysis of large patient cohorts, they target only up to 3% of the 28 million CpG sites in the genome, while complete genome-wide DNA methylome maps can only be achieved by whole-genome sequencing of bisulfite–treated DNA (WGBS) [15]. Several approaches for creating WGBS libraries from genomic DNA in combination with bisulfite conversion followed by sequencing by next generation sequencing (NGS) have been described. The different WGBS library construction approaches may affect the genomic coverage and the accuracy of the methylation calling, as shown in a recent bench-marking study [16]. Although the reagent costs for whole-genome sequencing (WGS) are decreasing, the large amount of DNA required in the first generation WGBS methods together with the high cost and limited availability of user-friendly methods for analysis of the WGBS data have so far limited the size of the patients cohorts subjected to WGBS in ALL [17,18].

A potential source of error in interpretation of DNA methylation levels using bisulfite conversion is that both 5-methyl cytosine (5mC) and 5-hyroxylmethyl cytosine (5hmC) are read as cytosine, and thus cannot be discriminated. This drawback of bisulfite treatment can be circumvented by introduction of an oxidation step in the bisulfite treatment procedure [19]. From a biological point of view discrimination between 5hmC and 5mC at single-base resolution in DNA is desirable because 5hmC may have opposite functions to 5mC, for example by how it affects gene expression. Demethylation of 5hmC occurs as a result of oxidation of 5mC by enzymes of the TET (ten-eleven translocation) family [20]. TET mutations are rare, but have been detected in about 1% of ALL cases [21,22]. Mutations in the *TET2* enzyme are associated with reduced 5hmC levels in acute myeloid leukemia [23], however it is unclear whether 5hmC occurs in ALL cells and hence it is not known to what extent 5hmC may influence the results of the bisulfite conversion-based assays commonly used to study DNA methylation in ALL.

There are several factors that can affect the interpretation of quantitative DNA methylation in the analyses of ALL cells, such as the sample number in the studies, tumor heterogeneity due to presence of DNA from subclonal populations of leukemic blasts, and presence of DNA from normal hematopoietic or blood cells within the leukemic DNA samples. It is difficult to experimentally identify and define epigenetic states within a population of bulk cells because of the heterogeneity within the leukemic cell population. Methods are needed to ascertain the uniqueness of the ALL methylome in the presence of normal cells when bone marrow or peripheral blood samples from ALL patients are analyzed in bulk [24,25]. In combination with recent reductions in DNA sequencing costs, single-cell sequencing offers a breakthrough for future analysis of DNA methylation in single cells from heterogeneous tumor populations [26,27,28].

3. DNA methylation in ALL cells

Methylation of cytosine residues in CpG dinucleotides plays a pivotal role in the establishment of cellular identity by influencing gene expression [29] and is a widespread and common feature of all human cancers, including leukemias [30,31]. The DNA methylomes of cancer cells have been found to contain large hypomethylated blocks and to display deregulation of the tightly controlled boundaries between methylated and unmethylated genomic regions [32]. Hypermethylation of CpG dense regions also known as CpG islands (CGIs) is the most systematically studied type of aberrant methylation across human cancers [31]. Accumulating evidence suggests that the pathogenesis and phenotypic characteristics of leukemic cells are the results of a combination of specific targeted and genome-wide alterations of DNA methylation [33,34].

The earliest studies to investigate aberrant DNA methylation in primary ALL cells analyzed CGIs nearby or in individual candidate genes. Although the small number of CGIs and patients included in these initial studies limits the interpretation of their results, they suggest that aberrant promoter methylation is associated with prognosis [35], cytogenetic alterations [36], cytogenetic subtype [37], and relapse [38]. These early findings combined with technological advances outlined in Section 2 above, have spurred several groups to study aberrant changes in DNA methylation of ALL cells on a genome-wide scale. The design of these studies in terms of methods for DNA methylation analysis, number of ALL samples, type of control cells used, and their main results are summarized in Table 1.