Chaber, R., Gurgul, A., Wróbel, G., Tomoń, A., Paszek, S., Potocka, N., Haus, O., Lejman, M., Łach, K., Szmatoła, T., Jasielczuk, I., Rybka, B., Ryczan-Krawczyk, R., Stąpor, S., Ciebiera, K., Arthur, C. J., & Zawlik, I. (2018). The distinguishable DNA whole genome methylation profile of 2 cases of pediatric precursor B acute lymphoblastic leukaemia (BCP ALL) with prodromal, preleukemic phase: A case report. *Medicine (Baltimore)*, 97(42), [e12763]. https://doi.org/10.1097/MD.00000000000012763

Publisher's PDF, also known as Version of record

License (if available): CC BY-NC

Link to published version (if available): 10.1097/MD.00000000000012763

Link to publication record in Explore Bristol Research

This is the final published version of the article (version of record). It first appeared online via Wolters Kluwer Health at https://doi.org/10.1097/MD.00000000000012763. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/
The distinguishable DNA whole genome methylation profile of 2 cases of pediatric precursor B acute lymphoblastic leukaemia (BCP ALL) with prodromal, preleukemic phase

A case report

Radosław Chaber, MD, PhD a,∗, Artur Gurgul, PhD b, Grażyna Wróbel c, Anna Tomań, MD a, Sylwia Paszek, MSc a, Natalia Potocka, MSc d, Olga Haus b, Monika Lejman, PhD d, Kornelia Lach, MSc a, Tomasz Szymatoł a, MSc a, Igor Jasielczuk, MSc b, Blanka Rybka, PhD d, Renata Ryczan-Krawczyk, MSc c, Sylwia Stępor, MSc a, Krzysztof Ciebie r, MSc h, Christopher J. Arthur, M.Chem., PhD i, Izabela Zawlik a,*, b, c, d, e, f, g, h, i

Abstract

Rationale: A prolonged, prodromal phase before definitive paediatric precursor B acute lymphoblastic leukaemia (BCP ALL) diagnosis is rarely observed.

Patients concerns: In the first, the patient presented with an aplastic preleukemic phase, whilst the second presented with a rheumatic-like preliminary phase.

Diagnoses: The case reports of two patients with BCP ALL with a prodromal phase lasting a few weeks are presented.

Interventions and outcomes: DNA whole genome profile methylation analysis of bone marrow cells obtained at diagnosis revealed a pattern of methylation that was readily distinguishable from both healthy and standard course BCP ALL bone marrow samples.

Lessons: The biological implication of this observation remains unclear, with many differentially methylated loci involved in many processes like neurogenesis, cell projection organization and adhesion along with leucocyte activation and apoptosis. The prevalence and clinical significance of these methylation changes is unknown but this data indicates that the epigenetic basis of BCP ALL with a prolonged, prodromal phase requires a more detailed assessment.

Abbreviations: ALL = acute lymphoblastic leukaemia, BM = bone marrow, CBC = complete blood count, DM = differentially methylated, ESR = erythrocyte sedimentation rate, PLT = platelet count, PP = prodromal phase, SC = standard course, WBC = white blood cells.

Keywords: BCP ALL, DNA methylation, pediatric, 2 cases

1. Introduction

Acute lymphoblastic leukaemia (ALL) onset is usually sudden with a short history. A prolonged, prodromal phase (PP) before definitive ALL diagnosis is rarely observed. ALL can be recognized by bone marrow (BM) biopsy/aspirate and/or peripheral blood smear examination, WHO criteria from 2008[1] state that to establish a diagnosis of ALL, more than 20% to 25% of the lymphoblasts in bone marrow (BM) must possess a diagnostic phenotype. In the unusual case, that a patient presents with fewer than 20% lymphoblasts in the BM and with no evidence of an extramedullary mass but demonstrates 1 of the known recurring cytogenetic abnormalities associated with ALL, the patient may be considered to have lymphoblastic leukaemia. The observation of fewer than 20% unequivocal lymphoblasts in
the BM should, however, also prompt a search for lymphoblastic lymphoma in an extramedullary location.[1,2]

Leukaemia initiation requires nucleotide sequence changes to occur, such as point mutations, amplifications or chromosome translocations. In concert with these genetic changes, disrupted epigenetic regulation (via DNA methylation, post-translational histone modifications, and interaction with non-coding RNA (miRNA or siRNA)) results in the abnormal expression of key genes responsible for cell proliferation and differentiation.[3]

The most common clinical symptoms of ALL in children are secondary to peripheral pancytopenia resulting from normal BM precursors. Sometimes a clinical view is complemented by hepatosplenomegaly, lymphadenopathy, symptoms of the central nervous system, or tests involving persistent fever, weight-loss and bone pain (particularly limb pain). Usually, examination of BM aspirates reveals up to 70% to 100% undifferentiated cells which correspond to the lymphoblast precursors, making ALL diagnosis proven. Definitive diagnosis of ALL thus normally takes no more than 2–3 weeks from the time of the first symptoms presenting.

In our previous work, hierarchical clustering of whole-genome DNA methylation profiles obtained from BM aspirates at diagnosis of a group of 38 patients with pediatric B-cell precursor (BCP) ALL revealed 2 cases with noticeably different methylation profiles.[4] Detailed clinical analysis of these patients showed that both had preceding clinical symptoms of leukaemia many weeks before ALL diagnosis was established. This pre-leukemic phase with incomplete features was misleading and delayed treatment. Herein we present the case reports of these 2 children. We, therefore, sought to fully analyze the differences in CpG methylation sites between BCP ALL patients with the standard course (SC) (BCP ALL SC) and the 2 with incomplete feature PP (BCP ALL PP).

2. Materials and methods

2.1. Ethical statement

Ethics Committee approval was obtained from the Institutional Review Board of the Medical University of Lodz (number, RNN/226/11/K/E). Informed consent has been obtained from parents/legal guardians of all the participating children. Furthermore informed written consent was obtained from the parents of the 2 patients presented in this case report for publication. The study protocol conforms to the ethical guidelines of the 1975 Declaration of Helsinki.

2.2. Whole genome DNA methylation profile analysis

In our previous study,[5] material comprised 38 samples of BM obtained from patients with pediatric precursor-B acute lymphoblastic leukaemia, at the time of the final ALL diagnosis, and 4 control non-leukemic samples. The analysis of cytogenetic ALL subtypes revealed that among BCP ALL SC patients there were 2 cases of triplody and 12 cases of hyperplody which might serve as a control for the analyzed BCP ALL PP patients. The previous results also showed, that factors such as gender and age are not important confounders for the global methylation profile differentiation of pediatric leukaemia.[5] DNA was purified using QIAamp DNA Blood Mini Kit (QIAGEN), assessed for fragmentation by agarose gel electrophoresis and quantified using Qubit 2.0 fluorimeter (Thermo Fisher Scientific). CpG methylation analysis was performed using Illumina (San Diego, CA) MethylationEPIC BeadChip, allowing analysis of 850K sites per sample.

2.3. Data quality control and analysis

The raw intensity data were checked for quality using the BeadArray Controls Reporter software (Illumina) and analyzed using the ChAMP package pipeline.[6] First, probes with detection of P value < .01 and with fewer than 3 beads in at least 5% of samples per probe were excluded. Additionally, non-CpG probes, SNP-related probes, multi-hit probes and probes located on chromosome X and Y were also removed. Then, the beta values (the proportion of DNA methylation at a CpG site) for 753,390 sites were calculated and assessed for quality by evaluation of beta multidimensional scaling (MDS) and density plots across the study groups. A beta value of 0 represents a completely unmethylated CpG site and a beta value approaching 1 represents a fully methylated CpG site. Beta values were normalized using the BMIQ method.[7] Singular value decomposition (SVD)[8] was used to identify the most significant components of variation, including technical variation. Differential methylation analysis between groups was performed using the champ (DMP) function which implements the limma package[9] to calculate the P value for differential methylation using a linear model. The obtained p-values were corrected for multiple testing using the Benjamini–Hochberg procedure.[10]

2.4. Functional genes annotation and analysis

The genes associated with specific differentially methylated (DM) sites were separated depending on CpG site location (promoter, gene body) and analyzed in terms of molecular functions, biological presses, cellular components, pathways and phenotypes using WebGestalt (WEB-based GEne SeT AnaLysis) toolkit,[11] exploiting information obtained from GO, KEGG, WikiPathways, Human Phenotype Ontology and PharmGKB databases. Over-representation tests were performed with respect to all known human genes (genome), identifying enriched categories with a corrected P value (false discovery rate [FDR]) lower than 0.05 and requiring at least 5 genes per enriched category.

2.5. Methylation assay performance and differential methylation analysis

The assay performance, as evaluated based on control probes and BeadArray Controls Reporter software, was satisfactory across all studied samples. After preliminary filtering, 753,390 probes were normalized and the batch effect was evaluated using the SWD method.

The comparison of methylation level of 753,390 sites between the 2 cases with BCP ALL PP and BCP ALL SC samples allowed the identification of 11,854 DM CpGs (adj P < .05). Of the sites, 7239 were hypermethylated in cases with PP ALL with an average delta beta between groups of 0.224 (±0.109). The remaining 4615 sites were hypomethylated with a slightly lower absolute delta beta value of 0.170 (±0.071). Both hyper- and hypomethylated CpGs were distributed on all 22 autosomes and the number of CpG per chromosome ranged from 1203 on HSA1 to 145 on HSA21.

CpG context analysis showed that 2090 of DM sites were located in gene promoter regions (TSS200 and TSS1500), 1433 in genes 5'UTR or first exon and 4770 within gene bodies. Most of the detected DM sites (7302, 61.6%) were located outside known CpG islands and remaining ones were positioned in islands (18.3%) or islands’ shelves and shores (20.1%). The detailed analysis of distribution of hyper- and hypomethylated sites showed that hypermethylation in BCP ALL PP is more common within promoter regions (22.7% versus 9.7% of all DM CpGs) and

2.3. Data quality control and analysis

The raw intensity data were checked for quality using the BeadArray Controls Reporter software (Illumina) and analyzed using the ChAMP package pipeline.[6] First, probes with detection of P value < .01 and with fewer than 3 beads in at least 5% of samples per probe were excluded. Additionally, non-CpG probes, SNP-related probes, multi-hit probes and probes located on chromosome X and Y were also removed. Then, the beta values (the proportion of DNA methylation at a CpG site) for 753,390 sites were calculated and assessed for quality by evaluation of beta multidimensional scaling (MDS) and density plots across the study groups. A beta value of 0 represents a completely unmethylated CpG site and a beta value approaching 1 represents a fully methylated CpG site. Beta values were normalized using the BMIQ method.[7] Singular value decomposition (SVD)[8] was used to identify the most significant components of variation, including technical variation. Differential methylation analysis between groups was performed using the champ (DMP) function which implements the limma package[9] to calculate the P value for differential methylation using a linear model. The obtained p-values were corrected for multiple testing using the Benjamini–Hochberg procedure.[10]

2.4. Functional genes annotation and analysis

The genes associated with specific differentially methylated (DM) sites were separated depending on CpG site location (promoter, gene body) and analyzed in terms of molecular functions, biological presses, cellular components, pathways and phenotypes using WebGestalt (WEB-based GEne SeT AnaLysis) toolkit,[11] exploiting information obtained from GO, KEGG, WikiPathways, Human Phenotype Ontology and PharmGKB databases. Over-representation tests were performed with respect to all known human genes (genome), identifying enriched categories with a corrected P value (false discovery rate [FDR]) lower than 0.05 and requiring at least 5 genes per enriched category.

2.5. Methylation assay performance and differential methylation analysis

The assay performance, as evaluated based on control probes and BeadArray Controls Reporter software, was satisfactory across all studied samples. After preliminary filtering, 753,390 probes were normalized and the batch effect was evaluated using the SWD method.

The comparison of methylation level of 753,390 sites between the 2 cases with BCP ALL PP and BCP ALL SC samples allowed the identification of 11,854 DM CpGs (adj P < .05). Of the sites, 7239 were hypermethylated in cases with PP ALL with an average delta beta between groups of 0.224 (±0.109). The remaining 4615 sites were hypomethylated with a slightly lower absolute delta beta value of 0.170 (±0.071). Both hyper- and hypomethylated CpGs were distributed on all 22 autosomes and the number of CpG per chromosome ranged from 1203 on HSA1 to 145 on HSA21.

CpG context analysis showed that 2090 of DM sites were located in gene promoter regions (TSS200 and TSS1500), 1433 in genes 5'UTR or first exon and 4770 within gene bodies. Most of the detected DM sites (7302, 61.6%) were located outside known CpG islands and remaining ones were positioned in islands (18.3%) or islands’ shelves and shores (20.1%). The detailed analysis of distribution of hyper- and hypomethylated sites showed that hypermethylation in BCP ALL PP is more common within promoter regions (22.7% versus 9.7% of all DM CpGs) and
known CpG islands (27.9% versus 3.2%) whereas hypomethylation occurs more frequently in gene bodies (48.4% versus 35.0%) and regions outside islands (open sea; 78.8% versus 50.6%). DM sites and their annotations are reported in Supplementary File 1, http://links.lww.com/MD/C352.

Hypothesizing that some of the sites DM between BCP ALL PP and BCP ALL SC may conform to the methylation profile found in healthy BM samples, an additional comparison was made with respect to control samples. This analysis allowed the identification 41,952 DM sites of which 1511 overlapped with sites differing with respect to BCP ALL SC (Supplementary File 2, http://links.lww.com/MD/C553). Most of these sites were hypermethylated with respect to the control (1094; 72.4%) with an average delta beta of 0.228. Most of the sites were also hypermethylated with respect to BCP ALL SC with samples (1008) with average delta beta of 0.234. The selected DM sites were scattered across all autosomes with the highest number located on the biggest chromosomes and in gene bodies (35.8%), intergenic regions (24.9%), and outside known CpG islands (50.6%).

Unsupervised hierarchical clustering of the samples based on probes differing between BCP ALL PP and both ALL BCP SC and the control samples shows a clear separation of the 2 BCP ALL PP methylation profiles from both remaining groups with higher similarity to BCP ALL SC cases (Fig. 1). Comparable results were obtained when principal component analysis (PCA) was applied to the same probes (Fig. 2). For further gene functions analysis, only CpGs differing in methylation level between BCP ALL SC and control were used. 0

3. Results

3.1. Patient 1

This patient is a 10-year-old male with a history of the growth hormone deficiency, at admission he was undergoing growth hormone supplementation therapy. The patient was initially referred to hematology due to abnormal complete blood count (CBC) values (initial values: hemoglobin (Hb) of 9.6 g/dL, leukocytes (WBC) of 0.69 × 10^9/L, neutrophils of 0.1 × 10^9/L, and a platelet count (PLT) of 121 × 10^9/L). Examination of peripheral blood smears revealed no blast cells. For 2 weeks before admission, the patient presented with a refractory fever of up to 39°C, neutrophils of 0.1 × 10^9/L. A peripheral blood smear revealed the presence of undifferentiated lymphoid cells (2%). The levels of CRP protein and ferritin as ESR were increased reflecting the active inflammatory process. Unfortunately, BM aspiration during this prodromal symptom phase was not performed, so BM cellularity and blast cell percentage before establishing the formal diagnosis is unknown.

Three weeks later, after the exclusion of rheumatic disorders, the patient was referred to haematology with persistent clinical symptoms and stable values of CBC, CRP, ferritin, and ESR. Physical examination revealed no symptoms of infection, no lymphadenopathy or hepatosplenomegaly, the neurological status, skin and testes were normal. The treatment according to ALL IC BFM 2009 protocol[6] for an intermediate risk group without central nervous system involvement was introduced. A good prednisone response was observed and he achieved haematological remission on the fifteenth day with minimal residual disease level 1.8 × 10^−3. Presently, the maintenance oral chemotherapy is being continued and the patient has stayed in complete remission for 22 months.

3.2. Patient 2

This patient is a 6-year-old male was presented to rheumatology unit with a 5 months history of bone pain in both legs. Irregular episodes of fever up to 39°C (with frequency once a week) with spontaneous remission began 1 month after the onset of the bone pains. A systemic connective tissue disease was suspected. Comprehensive diagnostics was implemented. Initial CBC included Hb of 11.7 g/dL, leukocytes (WBC) of 6.22 × 10^9/L, neutrophils of 1.95 × 10^9/L, and a PLT of 550 × 10^9/L. A peripheral blood smear revealed a slightly increased percentage of lymphocytes (66%) and monocytes (10%) with the presence of atypical lymphoid cells (2%). The levels of CRP protein and ferritin as ESR were increased reflecting the active inflammatory process. Unfortunately, BM aspiration during this prodromal symptom phase was not performed, so BM cellularity and blast cell percentage before establishing the formal diagnosis is unknown.

A good prednisone response was observed and haematological remission on the fifteenth day with minimal residual disease 7.1 × 10^−3 was achieved. Presently the patient has been in complete remission for 38 months after diagnosis without BM transplantation.
4. Discussion

During our study of whole genome methylation changes within leukaemia methylation pattern across genome showed visible separation of the 2 samples profiles with clear differences with respect to control and BCP ALL SC patient groups. Detailed phenotypic analysis showed that these samples were obtained from cases of BCP ALL with few weeks duration PP before the confirmation of a formal diagnosis according to the WHO criteria. The first patient presented with an aplastic preleukemic phase, the second one with rheumatic-like preliminary phase. In both cases, treatment was delayed until the WHO criteria were fulfilled, so the “watch and wait” strategy was implemented. The
potential acceleration of leukaemia with its serious complications can be a reason for some concern for both patients and practitioners.

Aplastic PP s are rare but well recognized in childhood BCP ALL.\textsuperscript{[13–15]} This neoplasm usually develops via a minimum of 2 discrete stages. First, the acquisition of chromosomal abnormalities (e.g., ETV6-RUNX1 fusion, high hyperdiploidy) predominately (but not exclusively) during fetal hemopoiesis which drives the expansion of a clinically silent, or covert, but persistent preleukemic clone. Second, in a relatively small fraction of such cases, the accrual, postnatally of further, secondary genetic changes promote or precipitates disclosure of clinical leukaemia.\textsuperscript{[13,16,17]} Occasionally, this form of ALL is initiated by infection, explaining the long-lasting remission due to induced, significant endogenous corticosteroids production.\textsuperscript{[6]} The frequency of BCP ALL with preleukemic phase is low, and up to 2\% of the ALL cases can be included in this group\textsuperscript{[14,18–20]} while full symptomatic ALL appears after a sudden recuperation of the CBC, with 95\% of cases progressing to ALL within 6 months of initial diagnosis.\textsuperscript{[21]}

Musculoskeletal complaints that depend upon lymphoblast expansion in BM cavities may be the initial dominating symptoms of newly diagnosed leukaemia as well as rheumatic diseases, for example, juvenile idiopathic arthritis. It is extremely rare, however, to diagnose ALL in children with such complaints who were originally referred to a pediatric rheumatology unit (<1\% of all cases).\textsuperscript{[22–25]} In these cases, the mean time between symptom onset and final diagnosis of leukaemia is about 3 months.\textsuperscript{[23]}

The DM sites between BCP ALL PP and both BCP ALL SC and control samples were associated with 816 different genes. The genes were enriched in many GO biological processes which were summarized with the software package Revigo\textsuperscript{[26]} to allow for the detection of their major categories (Table 1). Among the major classes of enriched biological processes are those associated, for example, with neurogenesis, cell projection organization and adhesion but also leukocyte activation and apoptosis.

The analyzed genes are also enriched in several KEGG pathways. Notably, these include T or B cell receptor signalling pathways, endometrial cancer or non-small cell lung cancer.

Figure 2. Principal component analysis based on probes differing in methylation level between B-cell precursor acute lymphoblastic leukaemia with incomplete feature prodromal phase (BCP ALL PP) and both BCP ALL with the standard course (BCP ALL SC) and control samples.\textsuperscript{[5]} BCP ALL PP = B-cell precursor acute lymphoblastic leukaemia with incomplete feature prodromal phase, SC = standard course.
The disease phenotypes enriched by the genes include acute myeloid leukaemia and Wiskott-Aldrich syndrome (Table 3).

Detailed analysis of the genes associated with the probes with the most significant differences in methylation level between BCP ALL PP and BCP ALL SC allowed the detection of 3 genes previously associated with the B cell phenotype, leukaemia or general tumour suppression and included: ENPP1, TCFL5, and LRRC3B genes.\cite{27–29} The DM sites associated with ENPP1 gene were mainly located in the promoter region (TSS200) and were predominantly hypermethylated. For TCFL5 gene, the single DM sites were situated in TSS1500 (hypomethylated) and 1st exon (hypermethylated). In the case of the LRRC3B gene, the hypermethylated sites were distributed across the whole gene and were annotated to TSS200, 1st exon and 5' UTR.

To gain a better insight into the affected processes, the gene list was subdivided into 2 categories depending on the location of the probe within the gene body or promoter region (TSS200, TSS1500) as methylation in these areas may have opposite effect on the genes' expression. We identified 336 probes associated with the promoter

### Table 1

| Term ID      | Description                                      | log10 P value |
|--------------|--------------------------------------------------|---------------|
| GO:0022008   | Neurogenesis                                      | -10.5784      |
| GO:0030030   | cell projection organization                      | -8.3665       |
| GO:006928    | movement of cell or subcellular component         | -7.5498       |
| GO:0040011   | Locomotion                                        | -6.3918       |
| GO:003085    | positive regulation of catalytic activity         | -6.4841       |
| GO:002610    | biological adhesion                               | -6.3605       |
| GO:007155    | cell adhesion                                     | -6.1739       |
| GO:0046834   | lipid phosphorylation                             | -5.7399       |
| GO:007726    | cell-cell signaling                               | -5.7055       |
| GO:006357    | regulation of transcription from RNA polymerase II | -5.3449       |
| GO:001775    | cell activation                                   | -5.266        |
| GO:0032970   | regulation of actin filament-based process        | -5.1361       |
| GO:0048870   | cell motility                                     | -5.1163       |
| GO:0051674   | localization of cell                              | -5.1163       |
| GO:0040007   | Growth                                            | -4.6968       |
| GO:009611    | response to wounding                              | -4.5482       |
| GO:006915    | apoptotic process                                 | -4.4661       |
| GO:003029    | actin filament-based process                      | -4.4202       |
| GO:008283    | cell proliferation                                | -4.2161       |
| GO:0018212   | peptidyl-tyrosine modification                    | -3.8072       |
| GO:0045321   | leukocyte activation                              | -3.8081       |
| GO:0016092   | vesicle-mediated transport                        | -3.4564       |
| GO:0014066   | phosphatidylinositol 3-kinase signaling           | -3.4381       |
| GO:004708    | single-organism behavior                          | -3.2898       |
| GO:004332    | adherens junction organization                    | -3.2697       |
| GO:0035411   | catenin import into nucleus                       | -3.2696       |
| GO:006897    | Endocytosis                                       | -3.1125       |

Major classes of biological processes enriched by genes with differentially methylated probes between B-cell precursor acute lymphoblastic leukemia with incomplete feature prodromal phase (BCP ALL PP) and both BCP ALL with the standard course (BCP ALL SC) and control samples. The numerous enriched biological processes found with WebGestalt were analyzed with Revigo system to summarize them by removing redundant GO terms.

### Table 2

| FDR           | Involved genes                                      |
|---------------|-----------------------------------------------------|
| 0.000855      | AKT3; MAP3K8; FYN; GRB2; LCK; NFATC1; NFATC2; PK3D; PK3R2; PPP3R1; MAP2K1; VAV1; ZAP70; NCK2; CO2B; GRAP2 |
| 0.006788      | AKT3; RAP5A; DGG6; DGKH; DGKQ; DM11; DNM2; FYN; GAB1; GRB2; CICR2; PK3D; PK3R2; MAP2K1; DGK2; DGK0; CYTH1 |
| 0.010435      | AKT3; CHRM2; CHRNA4; FYN; GNAS2; KCNQ2; ACHE; PK3G; PK3R2; PRKGC; MAP2K1; GRB2; CICR2; PK3D; CDK8 |
| 0.012196      | AKT3; ERBB2; GRB2; PK3D; PK3R2; MAP2K1; TGF; TGFI; CDH1 |
| 0.017972      | AKT3; ERBB2; GRB2; PK3D; PK3R2; PRKGC; MAP2K1; RARA; TGF |
| 0.020693      | AKT3; GRB2; LNK; NFATC1; NFATC2; PK3R2; PK3D; PK3R2; PPP3R1; MAP2K1; VAV1 |
| 0.020693      | AKT3; GRB2; FER; FYN; SNAI1; NLK; TGF; TLFI; ACTN1; CDH1 |
| 0.020693      | AKT3; ERBB2; GRB1; GRB2; PK3D; PK3R2; PRKGC; MAP2K1; TGF; NCK2; NR2B |

Top 10 KEGG pathways enriched by genes with differentially methylated probes between B-cell precursor acute lymphoblastic leukemia with incomplete feature prodromal phase (BCP ALL PP) and both BCP ALL with the standard course (BCP ALL SC) and control.
regions of 214 different genes. These genes did not enrich any of the GO biological processes, phenotypes or pathways. Analysis of the 445 genes associated with probes located in gene bodies was found to be enriched in a wide range of biological processes that could be reduced to major categories and are related to processes such as regulation of GTPase activity, positive regulation of cell communication or apoptotic processes (Table 4).

The genes with observed methylation differences within the gene bodies also enriched KEGG pathways associated with intercellular communication or apoptotic processes (Table 4).

The frequency and implications of these methylation changes cannot be answered herein. Our data indicates, however, that the epigenetic basis of BCP ALL with an incomplete, prolonged, PP phase requires a more detailed assessment with a larger number of patients.

| Disease phenotypes enriched by genes with differentially methylated probes | Involved genes |
|---|---|
| Adhesion | 0.002432 |
| COH4; COH6; ENA5; SDK1; FER; ITGA11; ARPSAP26; FYN; CADM2; RGS8; GRB2; ICM3; ITGA6; ITGAE; ITGAL; ITGB1; LAM1; LIMS1; MIGAT5; MYH9; NEO1; NRCAM; OLF; PKC2G; APRB1P; STAB2; RAOL; PKRC2; PCHSA4; PCHSA6; PCHSA2; RDX; RAPH1; SLCA2; SVIL; TGFMR; TVO; VOC; DR1; NCK2; ACTN1; ARHGGEF7; STARD13; CLD10; CYTH1; NRK2C2; CYTIP; PCHSA8; COH1 |
| Drug interaction with drug | 0.002432 |
| FNA5; AD11; HSPP6; ADORA2A; ERBB2; F3; FYN; GAB1; S1N3A; GB2; APP; ITGAL; ITGB1; KON4; KONMA1; LCK; SNA3D; MAP1A; ACHE; PKC330; PKG2; PKRCE; PKRCE; MAP2K1; RDX; RDX; RYR2; STOBA; VAV1; ACTN1; ARHGGEF7; ABG2; BAG3; NCK2 |
| Leukemia, Myeloid, Acute | 0.004978 |
| C0D6; DBS; CSFR; CUX1; DMT1A4; ETBR; ARHSAP26; S1N3A; GATA2; IRSF8; M1R55; MPO; PBX1; M11; PRD16; SECT11; ZBR16; ZNF496; ABG2; BRE; NCK2B; CSA |
| Ventricular Fibrillation | 0.028874 |
| DPPE; KON1; KON2; RYR2; CA1N1C; CA1N2D1; EKGA1; TANC1; KONH7; NOS1AP |
| Wiskott-Aldrich Syndrome | 0.044225 |
| AB12; WHAM4; DNM1; DNM2; FYN; GRB2; ITSN2; AMRA1; ITSN1; NCK2; ARHGGEF7; MTSS1; MTSS1 |
| Developmental disorder NOS | 0.049191 |
| DNM1; DRYKTA; AUS2; FOXP1; HADHB; APP; KONG2; MAPT; MARK1; CHD7; MAP2K1; ARD1B; DCDC2C; NUAK1 |

Disease phenotypes enriched by genes with differentially methylated probes located within gene bodies between B-cell precursor acute lymphoblastic leukemia with incomplete feature prodromal phase (BCP ALL PP) and both BCP ALL with the standard course (BCP ALL SC) and control samples.

**Table 4**

| Major classes of biological processes enriched by genes with differentially methylated probes located within gene bodies. |
|---|---|
| **Description** | **log10 P value** |
| GO:0030387 | regulation of GTPase activity |
| GO:0010647 | positive regulation of cell communication |
| GO:0046834 | lipid phosphorylation |
| GO:0030302 | lamellipodium assembly |
| GO:0034330 | cell junction organization |
| GO:0030312 | vehicle-mediated transport |
| GO:0052070 | regulation of actin filament-based process |
| GO:0045216 | cell-cell junction organization |
| GO:0022650 | regulation of anatomical structure morphogenesis |
| GO:0040011 | locomotion |
| GO:0007009 | plasma membrane organization |
| GO:0008915 | apoptotic process |

Major classes of biological processes enriched by genes with differentially methylated probes located within gene bodies between B-cell precursor acute lymphoblastic leukemia with incomplete feature prodromal phase (BCP ALL PP) and both BCP ALL with the standard course (BCP ALL SC) and control samples. The numerous enriched biological processes found with WebGestalt were analyzed with Revigo to summarize them by removing redundant GO terms.

**Author contributions**

**Conceptualization:** Radoslaw Chaber, Christopher J. Arthur, Izabela Zawlik.

**Formal analysis:** Artur Gurgul, Tomasz Szmatola, Igor Jasieczuk, Krzysztof Ciebiera, Christopher J. Arthur.

**Investigation:** Sylwia Paszek, Natalia Potocka, Monika Lejman, Kornelia Lach, Tomasz Szmatola, Igor Jasieczuk, Krzysztof Ciebiera.

**Methodology:** Radoslaw Chaber, Artur Gurgul, Grażyna Wróbel, Anna Tomoń, Sylwia Paszek, Natalia Potocka, Olga Haus, Monika Lejman, Kornelia Lach, Tomasz Szmatola, Igor Jasieczuk, Blanka Rybka, Renata Ryczan-Krawczyk, Sylwia Stapor, Krzysztof Ciebiera, Christopher J. Arthur, Izabela Zawlik.
Project administration: Izabela Zawlik.
Supervision: Izabela Zawlik.
Validation: Olga Haus.
Visualization: Artur Gurgul, Christopher J. Arthur.
Writing – original draft: Radoslaw Chaiber, Artur Gurgul, Sylvia Paszek, Natalia Potocka, Krzysztof Ciebiera, Izabela Zawlik.
Writing – review & editing: Olga Haus, Izabela Zawlik.

References

[1] Swerdlow SH, Campo E, Harris NL, et al. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues, Fourth Edition. IARC WHO Classification of Tumours, No 2. Lyon, France: IARC Press; 2008.
[2] Vardiman JW, Thiele J, Arber DA, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. Blood 2009;114:937–51.
[3] Chiaretti S, Zini G, Bassan R. Diagnosis and subclassification of acute lymphoblastic leukemia. Mediatr J Hematol Infect Dis 2014;6:2005;397–420.
[4] Dawson MA, Kouzarides T. Cancer epigenetics: from mechanism to therapy. Cell 2012;150:12–27.
[5] Chaiber R, Gurgul A, Wrobel G, et al. Whole-genome DNA methylation characteristics in pediatric precursor B cell acute lymphoblastic leukemia (BCP ALL). PLoS One 2017;12:e0187422.
[6] Campbell M, Castillo L, Riccheri C, et al. A Randomized Trial of the BFM-Allenic Leukemia Final Version of Therapy Protocol from August-14-2009. Available at: http://www.bialaczka.org/wp-content/uploads/2016/10/ALLIC_BFM_2009.pdf Accesses 2018/09/29.
[7] Morris TJ, Butcher LM, Teschendorff AE, et al. ChAMP: 450k chip analysis methodology. Bioinformatics 2014;30:428–30.
[8] Teschendorff AE, Marabita F, Lechner M, et al. A beta-mixture quantile normalization method for correcting probe design bias in illumina infinium 450 k dna methylation data. Bioinformatics 2013;29:189–96.
[9] Teschendorff AE, Menon U, Gentry-Maharaj A, et al. An epigenetic signature in peripheral blood predicts active ovarian cancer. PLoS One 2009;4:e8274.
[10] Smyth GK. Gentleman R, Carey V, Dudoit S, et al. LinMA: Linear models for microarray data. Bioinformatics and Computational Biology Solutions Using R and Bioconductor New York, NY, Inc, Springer-2005;397–420.
[11] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Methodol 1995;57:289–300.
[12] Zhang B, Kirov SA, Snoddy JR. WebGestalt: an integrated system for exploring gene sets in various biological contexts. Nucleic Acids Res 2005;33:W741–748.
[13] Horsley SW, Colman S, McKinley M, et al. Genetic lesions in a preleukemic aplasia phase in a child with acute lymphoblastic leukemia. Genes Chromosomes Cancer 2008;47:333–40.
[14] Breathach F, Chesselis JM, Greaves MF. The aplastic presentation of childhood leukemia: a feature of common-ALL. Br J Haematol 1981;49:387–93.
[15] Hasle H, Heim S, Schroeder H, et al. Transient pancytopenia preceding acute lymphoblastic leukemia (pre-ALL). Leukemia 1995;9:605–8.
[16] Greaves MF, Wiemels JL. Origins of chromosome translocations in childhood leukemia. Nat Rev Cancer 2003;3:639–49.
[17] Greaves MF, Maia AT, Wiemels JL, et al. Leukemia in twins: lessons in natural history. Blood 2003;102:2521–33.
[18] Sills RH, Stockman JA3rd. Preleukemic states in children with acute lymphoblastic leukemia. Cancer 1981;48:110–2.
[19] Schaison G. Acute curable preleukemic bone marrow aplasia in children. Biomed Pharmacother 1982;36:74–6.
[20] Armata J, Grzeskowiak-Melanowska J, Balwierz W, et al. Prognosis in acute lymphoblastic leukemia (ALL) in children preceded by an aplastic phase. Leuk Lymphoma 1994;13:517–8.
[21] Matloub YH, Brunning RD, Arthur DC, et al. Severe aplastic anemia preceding acute lymphoblastic leukemia. Cancer 1993;71:264–8.
[22] Zombori L, Kovacs G, Csoka M, et al. Rheumatic symptoms in childhood leukaemia and lymphoma-a ten-year retrospective study. Pediatr Rheumatol Online J 2013;11:20.
[23] Trapani S, Grisolia F, Simonini G, et al. Incidence of occult cancer in children presenting with musculoskeletal complaints: a 10-year survey in a pediatric rheumatology unit. Semin Arthritis Rheum 2000;29:348–59.
[24] Gonçalves M, Terreir MT, Barbosa CM, et al. Diagnosis of malignancies in children with musculoskeletal complaints. Sao Paulo Med J 2005;123:21–3.
[25] Dorronsoro Martin I, Merino Muñoz R, Sastre-Urguellés A, et al. Malignant disease presenting as rheumatic manifestations. An Pediatr (Barc) 2004;61:593–7.
[26] Supek F, Bosnjak M, Škunca N, et al. REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One 2011;6:e21800.
[27] Yoon J, Wang H, Kim YC, et al. Plasma cell alloantigen ENPP1 is expressed by a subset of human B cells with potential regulatory functions. Immunol Cell Biol 2016;94:719–28.
[28] Silveira VS, Scrideli CA, Moreno DA, et al. Gene expression pattern contributing to prognostic factors in childhood acute lymphoblastic leukemia. Leuk Lymphoma 2013;54:310–4.
[29] Haraldson K, Kashuba VI, Dmitriev AA, et al. LRRC3B gene is frequently epigenetically inactivated in several epithelial malignancies and inhibits cell growth and replication. Biochimie 2012;94:1151–7.
[30] Kulis M, Merkel A, Heath S, et al. Whole-genome fingerprint of the DNA methylome during human B-cell differentiation. Nature Genet 2015;47:746–56.