Genetic markers as tools for individualized disease monitoring in pediatric acute lymphoblastic leukemia

PhD thesis book

Bálint Egyed, MD

Károly Rácz Doctoral School of Clinical Medicine
Semmelweis University

| Supervisors: | Gábor Kovács, MD, D.Sc |
| | Ágnes Félné Semsei, Ph.D |

| Official reviewers: | Zsolt Nagy, MD, Ph.D |
| | Beáta Soltész, Ph.D |

Head of the Complex Examination Committee:
Judit Demeter, MD, D.Sc

Members of the Complex Examination Committee:
Gergely Kriván, MD, Ph.D
Hajnalka Andrikovics, MD, Ph.D

Budapest
2022

I. INTRODUCTION

Contrary to most cancer types, acute lymphoblastic leukemia (ALL) is predominantly presented in childhood. Although, ALL has become a curable disease in the last decades, the survival rate in some groups (e.g. patients with resistant central nervous system [CNS] involvement) is far behind the average, and many patients have severe therapeutic side effects (e.g. cardiotoxicity). Therefore, the two major current aims in the management of ALL are (*i*) detecting refractory subclones very early and (*ii*) predicting toxic effects of the chemotherapy which limit treatment efficacy or quality of life.

To date, for the follow-up of treatment response and prediction of relapse-free survival, there is no more sensitive method than estimating the measurable residual disease (MRD). However, MRD evaluation requires regular invasive procedures (e.g. bone marrow biopsy, spinal tap), and tumor cell detection in hidden body compartments, such as the CNS niche, remains a challenge. The cytologic examination of the cerebrospinal fluid (CSF) still yields a high proportion of false-negative reports due to the paucity of cells in this body fluid, therefore, the CNS leukemia is often underdiagnosed. This leads us to the investigation of tumor-derived, soluble, subcellular biomarkers which may indicate leukemia MRD in the CSF. The emerging role of conserved small non-coding RNAs, particularly microRNAs (miRs), in regulatory processes associated with cancer proliferation

and differentiation makes them good candidates for ALL MRD indicators. Furthermore, as the production of important miR carrier particles, so called extracellular vesicles (EVs), increases in the tumor microenvironment, these leukemia-derived bodies may also facilitate proper 'cancer cell biopsy'.

During the first 30 years after cancer diagnosis, pediatric ALL survivors are 15 times as likely to have heart failure as their siblings without cancer. This progressive and irreversible cardiotoxicity-related disorder mainly roots in anthracycline use in the treatment of ALL. The anthracycline-induced cardiotoxicity (ACT) is unpredictable: there are patients with cardiac problems who received very low doses of anthracyclines while others were administered with high doses and escaped the side effect. This proposes the key role of inter-individual genetic variability in ACT development. Single nucleotide polymorphisms (SNPs) are inherited differences in the germline DNA sequence with a low rate of recurrent mutations, which make them stable genetic markers. SNPs were associated with individual risk of diseases and various adverse drug reactions earlier. Application of SNP genotyping in ACT risk prediction may force a way of safer individualized medication in pediatric ALL.

We hypothesized that genetic markers (miRs and SNPs) can be utilized as personalized residual disease indicator and cardiotoxicity predictor in pediatric ALL.

II. OBJECTIVES

Newly identified genetic markers guiding the physician's decision in a progressive disease course or a chemotherapy-induced toxic event may contribute to better survival rates and/or quality of life after recovery from childhood ALL. We aimed to identify (*i*) a miR pattern in various body fluids (CSF, bone marrow, peripheral blood) indicating the leukemic burden in the CNS niche, and (*ii*) SNPs or interacting SNP networks associated with a pathological decrease in the heart function. Our main questions were:

1. Are there measurable leukemia-specific miRs in the CSF and can we distinguish between patients with and without unambiguous CNS leukemia based on miR expression measurements?
2. Do the peripheral blood or the bone marrow miR expression patterns correlate with the stage of CNS involvement, thus, provide less invasive CNS leukemia markers?
3. As EVs are significant reservoirs of miRs, can we demonstrate a difference in CSF small EV density between CNS-positive and -negative patients?
4. Can we identify SNPs which predispose to or defend from left ventricular function decline in a long-term follow-up analysis of pediatric ALL survivors?
5. How does the interaction network or priority order of SNPs relevant in chemotherapy-induced cardiotoxicity look?

III. METHODS

Briefly, nucleic acids (miR-rich RNA or DNA fractions) were isolated from patient samples collected at different type points after the diagnosis of acute leukemia. In our CNS leukemia biomarker study, we measured the expression of leukemia-associated miRs in peripheral blood (PB), bone marrow (BM) and CSF samples from the first period of the chemotherapy. In the pharmacogenetic analysis of ACT, SNP genotyping was performed in germline DNA separated from leukemia-free PB collected in disease remission. Array-based high-throughput qPCR measurements were applied to screen miR expressions and SNPs in the samples.

Patient cohorts and samples

Liquid biopsies (PB, BM and CSF samples) in the CNS leukemia study were processed within 2 hours after collection in two pediatric hematology centers in Budapest. Platelet-free plasma (PFP) were separated from PB and BM. CNS status of patients was defined as follows: CNS-involved (CNS$^+$; nontraumatic lumbar puncture with unambiguously identified blasts on cytospin preparation of CSF) and CNS-negative (CNS$^-$; blast-free CSF). A discovery screening step with diagnostic PB (de novo ALL, n=15; relapsed ALL, n=5) and CSF samples (matched CNS$^+$ vs. CNS$^-$ ALL patients, n=10) was performed. Then, 132 samples (PB, BM and CSF collected at different treatment checkpoints) of patients with ALL,

acute myeloid leukemia or mixed phenotype acute leukemia were analyzed for MRD in the CNS niche.

To reach long follow-up times in our ACT study, we enrolled those patients with pediatric ALL (n=622) retrospectively, who had undergone chemotherapy between 1989 and 2015 in 6 Hungarian pediatric hematology centers. Children with osteosarcoma (OSC, n=39) diagnosed in the same time period was also involved as a complementary cohort. The patients were treated with anthracyclines (ANTs) in the first year of chemotherapy. The patients were followed-up by echocardiography (ECHO) routinely in the clinical practice to monitor their left ventricular function. The left ventricular fractional shortening (LVFS) were determined in several follow-up time points (from the diagnosis to >15 years after the treatment). The worst heart function of each patient was used to define patients for our case-control study. Cases were those who had LVFS $\leq 28\%$ at any time point during the follow-up (n = 20); patients, who received the same chemotherapy but never had LVFS $\leq 28\%$ were regarded as controls (n = 641).

RNA isolation, cDNA synthesis and miR profiling

Total RNA was isolated from BM PFP, PB PFP and CSF samples using the miRNeasy Serum/Plasma Mini Kit (Qiagen, Hilden, Germany). Synthesis of cDNA by universal reverse transcription chemistry was performed with TaqMan Advanced miRNA cDNA Synthesis Kit

(Thermo Fisher Scientific, Waltham, MA, USA). A discovery screening of 47 leukemia-related, candidate miR expressions was carried out on high-throughput Custom TaqMan Advanced Low-Density Array (TLDA) microfluidic cards (Thermo Fisher Scientific). As a validation step, differentially expressed miRs identified in the discovery cohort were quantified by qPCR with TaqMan Advanced miRNA Assays (Thermo Fisher Scientific) on an extended sample pool.

Determination of CSF EV density and characteristics

As EVs are the main reservoir and transporter of miRs, we examined EV fractions in CSF samples. Firstly, $CD63^+$ (i.e. exosome type) EVs were captured and magnetically separated in CSF by anti-CD63-coated beads (Exosome-Human CD63 Isolation/Detection Reagent and DynaMag Spin Magnet, Thermo Fisher Scientific). Beads were then measured on a FACSCalibur (BD Biosciences, San Jose, CA, USA) instrument. Secondly, small EV-rich fractions (irrespective of the presence of CD63) were isolated by ultracentrifugation (UC) at 100,000g two times, and labeled with anti-CD63 and anti-CD81 antibodies. EVs were visualized by immunoelectron microscopy (JEOL 1011 TEM, Tokyo, Japan).

Germline DNA separation and SNP genotyping

DNA was isolated from PB using QIAmp DNA Blood Midi or Maxi Kits (Qiagen, Hilden, Germany). Seventy

SNPs in 26 genes encoding cardiomyocyte transporters involved in ANT import or elimination as well as enzymes in the hepatic metabolism of ANT were selected for further analysis. Genotyping was conducted using TaqMan® OpenArray™ Genotyping System (Thermo Fisher Scientific).

Statistical and bioinformatic analysis

MiR expression fold change values were calculated according to the comparative cycle threshold (Ct) algorithm for each sample. Reference miR selection and normalization of the expressions' (geNorm and NormFinder packages in *R*), principal component analysis after the discovery miR screen (FactoMineR), and linear model analysis to compare the miRNomic profile of different sample groups (limma) were performed.

To investigate the influence of SNPs on the left ventricular cardiac parameters in a long-term (>15 years) follow-up setting, a case-control study was conducted with univariate and multivariate logistic regression. Estimated haplotype frequency in cases and controls and the haplotype-specific odds ratio were calculated by the Haploview 4.1 software. Bayesian network-based Bayesian multilevel analysis of relevance (BN-BMLA) method was applied to test for potential interaction of the studied cofactors and SNPs.

IV. RESULTS

High-throughput miR profiling of liquid biopsy samples from a discovery cohort of children with ALL
TLDA cards proved to be applicable for miR expression quantification in cell-free PB and CSF samples of ALL patients. In diagnostic (day 0) PB PFP samples (*de novo* vs. relapsed disease), all of the leukemia-related miRs were measurably expressed. Expression levels measured in diagnostic CSF samples were evaluated by principal component (PC) analysis with a goal of identifying miRs that could potentially influence the classification of the patients by their CNS status (CNS^+ vs. CNS^-). Three members of the miR-181-family (miR-181a, miR-181b, miR-181c) had consistent and considerable positive contributions to all the first three PCs and defined >70% of the expression variability between CNS^+ and CNS^- samples. In CNS^+ CSF samples, among all tested RNAs, levels of miR-181a, miR-181b and miR-181c showed the strongest reductions by the 15th day of the treatment, as their expression fold changes (ΔFC) were 36.2, 299.4 and 39.8 times lower than at the diagnosis, respectively.

Diagnosis of CNS leukemia by measuring CSF miR-181a level in an extended patient cohort
We were able to confirm the high expression of miR-181a in diagnostic CSF samples in a set of 8 CNS^+ ALL patients compared with 10 CNS^- ALL patients by conventional qPCR. MiR-181a expression levels conferred a more than

52-fold increased risk for CNS leukemia (CNS$^+$ vs. CNS$^-$ patients: ΔFC=52.30, p=1.49E-4). This phenomenon was independent of B- or T-cell immunophenotypes of ALL. MiR-181a yielded a sensitivity of 90.0% and specificity of 87.5% (area under the receiver operating characteristic curve: 92.5%) as a CNS disease marker at a cut-off ΔFC value of 3.7. Put in context, conventional cytomorphology technique provides a lower sensitivity with approximately 30%. By the 33rd day of ALL treatment (after two intrathecal chemotherapy courses), the average level of miR-181a decreased by 96.7% among CNS$^+$ patients, while it remained around the expression detected in diagnostic sample at our later time points in patients without CNS leukemia.

Similar results were not found when analyzing miR-181b expressions or miR-181a levels in acute myeloid leukemia (AML) or mixed phenotype acute leukemia (MPAL) patients.

Assessment of BM and PB miR-181a/-b expression to indicate CNS status in ALL

We searched for less invasive markers of CNS disease by analyzing BM and PB samples from the same cohort. There was no difference in miR-181a expression at diagnosis between CNS$^+$ and CNS$^-$ ALL groups in PB. On the contrary, BM miR-181a expression distinguished those patients with CNS$^+$ and CNS$^-$ ALL who were characterized by precursor B-cell immunophenotype

(ΔFC=9.18, p=0.04). A substantial decrease in miR181a level from the diagnosis of CNS$^+$ ALL by the 33rd day of therapy was also detectable in BM samples (ΔFC= -187.96, p=0.006).

Small EVs in CSF samples of CNS$^+$ ALL patients

A proof-of-concept study was conducted with very limited patient numbers (CNS$^+$ vs. CNS$^-$ patients, n=2 vs. n=2). Immunomagnetic bead separation showed no difference in the proportion of CD63$^+$ elements (mean±SE: 1.0±0.2% vs. 1.1±0.1%). Yet, the small EV density contrast between CNS$^+$ and CNS$^-$ CSF samples was obvious by immunoelectron microscopy: CNS$^+$ samples contained high amounts of CD63$^-$/CD81$^-$ EVs missing in CNS$^-$ cases.

Pharmacogenetic factors behind ACT in a considerable Hungarian cohort

Risk of pathological LVFS was significantly influenced by SNPs in *CYP3A5* and *SLC28A3* genes. *CYP3A5* rs4646450 TT was 17% among cases and 3% in ALL patients without pathological FS (p = 5.60E-03; OR = 6.94 [1.76-27.39]). This effect was confirmed by completing the ALL cohort with independent OSC patients (p = 4.81E-03; OR = 7.25 [1.83–28.78]). In subgroup analyses, we found that male patients (p = 4.00E-03; OR = 13.45 [2.26-80.1]) and patients who received higher ANT doses

(p = 5.00E-03; OR = 20.83 [2.49−174.3]) bore a higher risk for ACT by carrying rare homozygous TT genotype. In ALL, *SLC28A3* rs7853758 AA was 12% in cases, while only 1% among controls (p = 6.50E-03; OR = 11.56 [1.98-67.45]). However, we were not able to reproduce this result with *SLC28A3* rs7853758 in the combined ALL and OSC cohort.

Haplotype analysis and interaction mapping in the variable's network determining ACT risk

Haplotype block rs4646450–rs776746 TA in gene *CYP3A5* seemed to influence the ACT in the ALL cohort with moderate linkage, but only nominal significance (D' = 0.9; r^2 = 0.48; p = 3.05E-2; OR = 14.26 [1.28−158.54]). BN-BMLA analysis revealed interaction effect between two SNPs (rs7853758 and rs885004) in gene *SLC28A3*. Both SNPs are potentially strongly relevant in the Bayesian network of ACT (*P* = 0.55 and 0.36, respectively). It may suggest the refinement of the previously reported predisposing role of *SLC28A3* rs7853758 A allele in ACT by their interaction with the rs7853758 polymorphic locus.

V. CONCLUSIONS

In the first part of my thesis, we reported that leukemia-specific miRs are measurable in cell-free BM, PB and CSF liquid biopsy samples of patients with childhood ALL and the expression of miR-181 family in the CSF may be capable of distinguishing patients with and without CNS involvement. CSF miR-181a expression measurements showed higher sensitivity in CNS leukemia detection than the conventional cytologic examination. The alteration of miR-181a levels in the CSF seemed to have the potential for personalized ALL MRD monitoring in the CNS compartment.

In the second part of my thesis, genotyping a substantial Hungarian pediatric ALL cohort, we found potential pharmacogenetic factors behind ACT. In a long-term follow-up analysis, *CYP3A5* rs4646450 TT genotype was significantly more frequent in those children with pathological LVFS (<28%). Furthermore, *CYP3A5* rs776746 seems to sit on the same haplotype as *CYP3A5* rs4646450, and rs776746 is part of the strongly relevant factor's network of ACT. An interaction between *SLC28A3* rs7853758 and *SLC28A3* rs885004 polymorphisms was also observed by our self-developed SNP network analysis method, which complicates the previously described role of *SLC28A3* rs7853758 in ACT by other authors.

VI. BIBLIOGRAPHY OF THE CANDIDATE'S PUBLICATIONS

Publications related to the thesis

Egyed B, Kutszegi N, Sági JC, …, Semsei ÁF. (2020) MicroRNA-181a as novel liquid biopsy marker of central nervous system involvement in pediatric acute lymphoblastic leukemia. *J Transl Med*, 18(1): 250.
IF: 5.531 (D1)

Sági JC*, **Egyed B***, Kelemen A, …, Semsei ÁF. (2018) Possible roles of genetic variations in chemotherapy related cardiotoxicity in pediatric acute lymphoblastic leukemia and osteosarcoma. *BMC Cancer*, 18(1): 704.
* Co-first authors.
IF: 2.933 (Q1)

Rzepiel A, Kutszegi N, Gézsi A, Sági JC, **Egyed B**, …, Erdélyi DJ. (2019) Circulating microRNAs as minimal residual disease biomarkers in childhood acute lymphoblastic leukemia. *J Transl Med*, 17: 372.
IF: 4.124 (D1)

Publications not directly related to the thesis

Egyed B, Horváth A, Semsei ÁF, …, Kovács GT. (2022) Co-detection of VEGF-A and its regulator, microRNA-181a, may indicate central nervous system involvement in pediatric leukemia. *Pathol Oncol Res*, 28:1610096.
IF (2021): 2.874

Sági JC, Gézsi A, **Egyed B**, …, Erdélyi DJ. (2021) Pharmacogenetics of the central nervous system - toxicity and relapse affecting the CNS in pediatric acute lymphoblastic leukemia. *Cancers*, 13:2333.
IF: 6.575

Egyed B, Kovács GT, Kutszegi N, …, Semsei ÁF. (2018) New and traditional directions in the biology and management of childhood acute lymphoblastic leukemia. *Orv Hetil*, 159(20), 786-797.
IF: 0.564

Alpár D*, **Egyed B***, Bödör C, Kovács GT. (2021) Single-cell sequencing: biological insight and potential clinical implications in pediatric leukemia. *Cancers*, 13:5658.
* Co-first authors.
IF: 6.575

Constatin T, Andrási N, Ponyi A, …, **Egyed B**, …, Szabó A. (2021) Diagnosis and treatment of paediatric inflammatory multisystem syndrome. *Orv Hetil*, 162(17):652-667.
IF: 0.707

Kiss R, Gángó A, Benard-Slagter A, **Egyed B**, …, Alpár D. (2020) Comprehensive profiling of disease-relevant copy number aberrations for advanced clinical diagnostics of pediatric acute lymphoblastic leukemia. *Mod Pathol*, 33:812-824.
IF: 7.842