Genetic prognostic factors in acute myeloid leukemia

Thesis of Ph.D. dissertation

Koszarska Magdalena Renata

Eötvös Loránd University, Faculty of Sciences
Doctoral School of Biology
Classical and Molecular Genetics PhD Program

Leader of the PhD School: Anna Erdei DSc, member of HAS
Leader of PhD Program: Tibor Vellai DSc
Supervisors: Hajnalka Andrikovics M.D., PhD and Váradi András DSc

Hungarian National Blood Transfusion Service
Budapest, 2014.
Introduction

Acute myeloid leukemia (AML) is a malignant disorder characterized by increased proliferation and blocked differentiation of a hematopoietic stem cell clone. AML has a highly heterogeneous genetic background with more than two hundred different cytogenetic and molecular genetic aberrations. Acquired cytogenetic abnormalities are the most important leukemia-related prognostic factors for treatment outcome and survival and are divided into 3 prognostic groups: favorable, intermediate and adverse. The favorable cytogenetic risk group covers the structural changes of genes affecting the subunits of the core-binding factor (CBF): t(8;21), inv 16, t(16;16) in detail, or the RARA gene: t(15;17). The adverse cytogenetic risk group includes: deletions of 5, 5q, 7, 7q, 17p, inv(3)(q21;q26), t(3;3)(q21;q26), rearrangements of (11q23) except t(9;11) and complex karyotypes (defined as either ≥three or ≥five abnormalities). The intermediate cytogenetic risk group consists of other abnormalities, which do not belong either to the favorable or to adverse cytogenetic groups. Testing for molecular acquired genetic mutations helps for the risk stratification in different cytogenetic subgroups, especially in patients with normal karyotype (NK).

A common mutation occurring in about 25% of AML and 35% of NK-AML patients is the internal tandem duplications (ITD) of the fms-like tyrosine kinase 3 (FLT3) gene, which encodes for a receptor tyrosine kinase. FLT3-ITDs occur in the juxtamembrane domain (JMD) or in the tyrosine kinase domain 1 (TKD1), resulting in constitutive activation of this receptor. FLT3-ITD AML accounts for a distinct subgroup of AML with well-known clinical features, and impaired treatment outcome due to increased relapse rate. The most frequent acquired, recurrent mutation identified in approximately 35% of AML and 55% of NK-AML patients is a 4 basepair long insertion in the nucleophosmin 1 gene (NPM1). The mutation alters the nuclear localization signal of the NPM1 protein resulting in a dominant cytoplasmic localization. AML patients harboring this mutation have a better prognosis.

A new type of recently identified, acquired genetic alteration in AML affects mutational hotspots of isocitrate dehydrogenase 1 and 2 (IDH1, codon 132 and IDH2 codons 140 and 172). Under normal circumstances, IDH enzymes catalyze the conversion of isocitrate to α–ketoglutarate (αKG). The presence of IDH mutant enzymes results in aberrant production of 2-hydroxyglutarate (2HG), resulting in the block of αKG dependent enzymes such as tet methylcytosine dioxygenases 1 and 2 (TET1 and 2), or histone demethylases causing global, aberrant DNA and histone methylation, altered gene expression profiles and consecutively
impaired stem cell differentiations. $IDH^{mut}$ AML also characterized by distinct clinical features, but data on its prognostic impact were controversial.

The clinical manifestation and the treatment outcome of AML depend not only on the acquired mutations in the malignant clone, but also on the inherited genetic variations of the individual patient affecting eg. pharmacokinetics. The ATP-binding cassette, sub-family G, member 2 ($ABCG2$) multidrug transporter is preferentially expressed in pharmacological barriers, (liver, kidney, placenta, blood-brain barrier). This protein modulates the absorption, metabolism and toxicity of numerous drugs and xenobiotics, and causes multidrug resistance in cancer. Interestingly, $ABCG2$ was also found to be expressed in stem cells (including haematopoietic and leukemic stem cells in AML). $ABCG2$ expression on stem cells presumably plays a protective role against various endo- and exotoxins. $ABCG2$ expression may be responsible for the development of multidrug resistance during chemotherapy. Many SNPs causing mis-trafficking and early degradation were identified in the coding region of $ABCG2$ gene. To date, among the most studied variants the following two nonsynonymous single nucleotide polymorphisms (SNP) are represented: V12M (c.34G>A), and Q141K (c.421C>A). V12M has not yet been associated with any deviation in protein expression or function. On the other hand, Q141K, with a variable allele frequency between 5-30% in various ethnic groups but preferentially occurring in Asian populations, was found to decrease protein expression in vitro, and to reduce drug efflux and ATPase activity.

Detailed characterization of acquired and inherited genetic background in malignant disorders may help in establishing a targeted and a personalized treatment strategy for improving outcome.
Aims of the study

**Impact of acquired mutations (FLT3, NPM1, IDH1/2) in AML**

Recent studies suggest that personalized medicine is an important tool in order to choose the most optimal therapy for cancer patients. Detailed characterization of acquired genetic background in malignant disorders, such as AML, is essential in risk stratification and in selection of the most optimal therapy, for each patient. Our aims were the following in the context of acquired mutations:

1. To set up and optimize a PCR followed by a capillary electrophoresis in order to calculate the frequencies of FLT3-ITDs and NPM1 mutations in our AML cohort.
2. To study the association between FLT3-ITD load, size and clinical characteristics or treatment outcome in AML.
3. To set up three different detection methods (allele specific PCR, high resolution melting and sequencing) in order to calculate the frequencies of IDH1/2 mutations in AML.
4. To verify the clinical and the prognostic relevance of different IDH1/2 mutations in AML.

**ABCG2 genetic variations in healthy individuals and in AML patients**

Inherited genetic variations have also been found to play a great role in the clinical manifestation and the treatment outcome of AML patients. They account for risk stratification and choice of therapy in order to achieve the best outcome. In connection, our aims were the following:

1. To develop an antibody-based quantitative flow cytometric assay for the determination of membrane proteins including ABCG2 on red blood cell (RBC) membranes.
2. To test the effect of common ABCG2 polymorphic variants (V12M and Q141K) on ABCG2 protein expression on RBC membrane in healthy individuals.
3. To investigate the effect of genetic variants on disease susceptibility, clinical manifestation and treatment outcome in AML patients.
Methods

Subjects
1) Genotyping was completed on 389 AML patients (FLT3-ITD n=324, IDH n=376, ABCG2 SNP n=389) [183 males/206 females; median age: 50; range: (16-93 years)] diagnosed and followed between 2001 and 2009. Treatment outcome was investigated in patients younger than 65 years old and treated with curative intention (n=324, n=314, n=307 in FLT3-ITD, IDH and ABCG2 studies respectively).
2) The determination of ABCG2 protein expression level on red blood cells (RBC) was performed in 61 healthy volunteers (47 unrelated individuals and 14 family members of two probands selected from the volunteers).
3) A control group, consisting of 202 healthy unrelated individuals was screened for ABCG2 SNPs (V12M, Q141K).

Molecular genetic methods
All mutation analyses (FLT3-ITD, NPM1, IDH1/2), were performed from genomic DNA isolated from bone marrow or peripheral blood samples at the time of diagnosis. ABCG2 expression level was also investigated at the time of diagnosis, but from mRNA isolated from the bone marrow. ABCG2 SNPs testing was performed from genomic DNA isolated from bone marrow or peripheral blood samples at diagnosis or any time during follow-up.

FLT3-ITD and NPM1 mutational status was defined according to the detected fluorescently labeled PCR product on capillary electrophoresis (fragment analysis). The FLT3-ITD load was quantified as the ratio of the peak height of the FLT3-mutant allele divided by the sum of the mutant and the wild type peak heights (ITD/ITD+WT).

IDH1 and IDH2 mutations were screened by high resolution melting (HRM) and by allele-specific PCR (AS-PCR). Samples, which showed IDH1 positivity by AS-PCR and HRM, were sequenced in order to identify the exact type of mutation. For IDH2, those samples were sequenced, which gave contradictory results by AS-PCR and HRM.

The most common SNPs in ABCG2 (V12M) and (Q141K) were genotyped using the LightCycler480 (Roche Diagnostics, Basel, Switzerland) allelic discrimination system. The coding region of ABCG2 gene, (exons 2-16) and the exon-intron boundaries were sequenced in order to identify novel mutations.
Flow cytometry (FACS)

FACS analysis was performed from freshly drawn peripheral blood using monoclonal antibodies specifically recognizing ABCG2 (BXP21 and BXP34-recognizing an intracellular ABCG2 epitope and 5D3 recognizing an ABCG2 extracellular epitope).

Statistical analysis

Continuous variables are presented as median and range. The Mann-Whitney or the Kruskal-Wallis tests were used to compare continuous variables in subgroups according to FLT3-ITD mutation status, different ITD load and size and IDH mutation status. Fischer’s exact test and the χ2 tests were performed to compare dichotomous variables. Genotype and allele frequencies with 95% confidence interval (AF ± 95%CI) are presented. Estimated haplotype frequencies (EHF) were calculated by the web-based SNPStats web tool. The log-rank test was used to compare disease free survival (DFS) and overall survival (OS) between groups separated by ITD load and size or IDH mutation status or ABCG2 genotypes. A Cox proportional hazards model was computed for multivariate analysis of OS and DFS with the calculation of hazard ratio (HR) and 95% confidence interval (95%CI). The statistics were performed using Statistical Package for the Social Science (SPSS) version 13.0.
Results

Impact of acquired mutations (FLT3, NPM1, IDH1/2) in AML

FLT3-ITD mutations were detected in 82/388 (21.1%) cases. The presence of ITD was related to de novo AML etiology (p<0.001). It associated more frequently with myelomonocytic morphology (p=0.0004). FLT3-ITD^{pos} patients presented with significantly higher white blood cell (WBC) counts at diagnosis (p<0.001). FLT3-ITD^{pos} associated with intermediate risk cytogenetics (p<0.001) and NPM1 mutation (p<0.001). In our entire patient cohort, FLT3-ITD^{pos} itself did not associate with adverse prognosis, only relapse rate seemed to be increased (p=0.08). The combined FLT3-ITD^{neg} and NPM1^{pos} subgroup showed better OS and DFS in the intermediate and normal karyotype subgroups.

The median FLT3-ITD load was 26%. Using 50% as a cut-off, clinical characteristics at presentation (sex, etiology of AML, morphology, blood counts, cytogenetics and NPM1), were not different between FLT3-ITD^{pos} individuals with low ITD load (FLT3-ITD^{<50%}; n=56) vs. high ITD load (FLT3-ITD^{>50%}; n=15). Treatment outcome was investigated in 43 FLT3-ITD^{<50%} and 15 FLT3-ITD^{>50%} patients. Regarding the outcome parameters, presence of higher FLT3-ITD load resulted in worse DFS and OS (p=0.010, 0.038 respectively).

The size of FLT3-ITD varied from 6 to 210 base pair (bp). 22/81 patients (27.2%) displayed more than one mutant amplicons (size calling was possible in 81/82 FLT3-ITD positive cases). FLT3-ITD mutants with insertion length shorter or longer than 48bp showed similar clinical characteristics at diagnosis, but profoundly different treatment outcome in the entire AML group. Longer ITDs were associated with higher early death ratio (p=0.01); a tendency toward higher relapse rate (p=0.06); resulting in worse OS (p=0.02) and DFS (p=0.005). Patients with FLT3-ITD^{>48bp} had worse OS and DFS compared to ITD negative patients (n=255; p=0.04; p=0.013 respectively). Interestingly individuals with longer than 60bp insertions (FLT3-ITD^{>60bp}; n=15) showed better OS and DFS compared to patients with insertion between 48 and 60bp (FLT3-ITD^{48-60bp}; n=15, p=0.014; p=0.019 respectively), indicating that medium sized ITD confers the worst prognosis among ITD positive patients.

IDH1 and IDH2 mutations were mutually exclusive, detected in 32 (8.5%) and 28 cases (7.5%). All mutations were in heterozygous form. In IDH1, R132C (n=14, 43.7%) and R132H (n=10, 31.2%) substitutions were the most frequent. In IDH2, R140Q (n=20, 71.4%) and R172K (n=8; 28.6%) were identified. IDH mutations combined were associated with: older age (p=0.001), higher platelet counts (p=0.001), intermediate karyotype (p<0.0001),
Results

*NPM1* mutation (*p*=0.022) as well as with a lower mRNA expression level of *ABCG2* gene (*p*=0.006) at diagnosis. Remission rate, relapse rate, overall survival (OS) were not different in *IDH1* mutant (n=24) vs. *IDH1 and 2* wild type (n=269) or in *IDH2* mutant (n=21) vs. *IDH1 and 2* wild type. Marked differences in the clinical presentation were observed between *IDH1* R132H and R132C mutant AML patients. R132H mutant AML cases were more likely to have de novo origin (90%), while R132C positive AMLs were secondary to myelodysplastic syndrome (MDS), or therapy related in 50% of cases (*p*=0.08). FAB M1 was more common in R132C (*p*=0.02). PLT count at diagnosis was higher in R132H (*p*=0.050). Interestingly, particular *IDH* mutations differed in association with *NPM1* mutations: concomitant *IDH* and *NPM1* mutations occurred in 14.3% (2/14) of *IDH1* R132C vs. 70% (7/10) R132H carriers (*p*=0.02) and in 47.4% (9/20) of *IDH2* R140Q vs. 0% (0/8) R172K carriers (*p*=0.02). *IDH1* R132H (n=9) negatively influenced OS compared to *IDH1 and 2* wild type (*p*=0.02) or to R132C (n=9, *p*=0.019). No differences were found comparing various *IDH2* mutations with respect to treatment outcome.

**ABCG2 genetic variations in healthy individuals and in AML patients**

*ABCG2 expression on RBC*

ABCG2 protein expression on RBC did not associate with age or sex. Among the 47 donors, there were 11 individuals with heterozygous Q141K variant (allele frequency: 11.7±6.6%), and 3 individuals with heterozygous V12M variant (allele frequency: 3.2±3.6%) genotypes. Individuals carrying the heterozygous ABCG2 Q141K variant exhibited significantly lower expression of ABCG2 (5.27±1.19) on RBC, as compared to homozygous wild-type individuals (6.13±0.61, *p*=0.011). There was no significant difference between homozygous wild-type individuals and heterozygous V12M carriers.

*ABCG2 nonsense mutations*

Two unrelated individuals out of 47 exhibited lower RBC ABCG2 expression (2.65±0.29). Sequencing of the *ABCG2* gene in these cases revealed two different heterozygous nonsense mutations: (1) an arginine to stop codon change in exon 7 (c.706C>T, p.R236X, rs140207606) and (2) a small deletion also in exon 7 (c.791_792delTT, p.L264HfsX14). Low ABCG2 expression phenotypes were segregated in both families with the presence of the respective mutations.
**ABCG2 V12m and Q141K in AML**

Genotype and allele frequencies (AF) were not different in AML patients as a whole cohort compared to controls. On the other hand, AML subgroup analysis revealed that *RUNXI-RUNXIT1* t(8;21)(q22;q22) translocation] positive patients (n=18) showed higher V12M AF compared to control individuals (p=0.02). Among AML patients younger than 60 years receiving daunorubicin based induction chemotherapy (n=256), ABCG2 141K carriers displayed longer overall (p=0.072) and disease free survival (p=0.048) compared to individuals carrying the major allele in homozygous form (QQ genotype). Early death due to toxicity, resistant disease, remission and relapse rates were not different between ABCG2 141K carriers and homozygous major allele carriers.
Conclusions:

Impact of acquired mutations (FLT3, NPM1, IDH1/2) in AML

Our results suggest that the length of FLT3-ITD mutations may influence disease outcome, in a way that medium-sized ITDs (48-60 bp) may confer worse prognosis than shorter or longer mutations. As a possible explanation, ITD was reported to lead to the constitutive activation of the kinase domain by disrupting the autoinhibitory interaction between the juxtamembrane domain and activation loop. In line with this hypothesis, longer ITDs could lead to more profound destabilization of the inactive kinase. On the other hand, longer ITDs were reported to be inserted more C-terminally in the FLT3 protein, which may cause increased interference with the kinase activity, leading to a reduced transforming capacity.

IDH1 and 2 mutations are associated with distinct clinical characteristics. Alterations in NPM1 associations between IDH1 R132C and R132H or between IDH2 R140Q and R172K suggest differences between particular mutations. IDH substitutions are generally considered as a weak prognostic factors influencing survival in an inconsistent way in different AML subgroups, which may originate from different amino acid changes affecting outcome in opposite ways. Our observations may suggest that specific mutations might differ in the level of 2HG production or they might act differently on other pathways besides blocking αKG-dependent enzymes.

ABCG2 genetic variations in healthy individuals and in AML patient

Based on our results, FACS may be a rapid and reliable assay for the quantitative determination of membrane protein expression in human erythrocytes. We found significant differences in ABCG2 expression levels between homozygous wild-type and heterozygous 141K polymorphic variant carriers. Our further observations suggest that the ABCG2 genetic variants associated with the absence of high frequency blood group antigens (p.R236X, p.L264HfsX14) may be more common than previously described. Our results also suggest that common ABCG2 polymorphisms might influence AML predisposition and treatment outcome differentially depending on AML genetic background or chemotherapy applied.

In summary, we conclude that that mutational profiling can be important and useful for risk stratification and therapeutical decisions. Therefore detailed characterization of acquired and inherited genetic background in malignant disorders, such as AML is necessary in establishing a targeted and a personalized treatment strategy for improving outcome.
Summary

AML has a highly heterogeneous acquired genetic background. Each of AML patients harbors an unique set of mutations and molecular features characterizing the disease. Mutational profiling of acquired and inherited variants can be important and useful for risk stratification guided therapeutical decisions. One of the most frequent mutations affects FLT3 gene: the internal tandem duplication (ITD) in FLT3 autoinhibitory juxtamembrane region causing a permanent receptor signal transduction and uncontrolled proliferation of leukemic cells. Another recently identified, acquired aberration affects mutational hotspots of isocitrate dehydrogenase 1 and 2 (IDH1 codon 132 and IDH2 codons 140 and 172). The inherited polymorphic variants of ATP-binding cassette multidrug transporter, sub-family G, member 2 (ABCG2 V12M and Q141K) may influence drug transport during chemotherapy.

In this work I focused on the impact of acquired mutations (FLT3-ITD, IDH1/2) and inherited polymorphisms (ABCG2) on the clinical manifestation and treatment outcome in AML and on impact of ABCG2 inherited variations on ABCG2 protein expression in healthy individuals.

FLT3-ITD was detected in 21% of AML cases. Patients with higher ITD load or medium size insertions (between 48 and 60bp) showed worse treatment outcome and survival compared to patients with lower load or insertions with other size. IDH1 and IDH2 mutations were mutually exclusive, detected in 8% and 7% of AML cases respectively. IDH1 R132H negatively influenced survival compared to IDH1/2 wild type or to R132C.

Healthy individuals carrying the heterozygous ABCG2 Q141K variant exhibited significantly lower expression of ABCG2 protein in red blood cell membrane, as compared to homozygous wild-type individuals. Sequencing of the ABCG2 gene of two unrelated individuals showing lower erythrocyte ABCG2 expression revealed heterozygous mutations resulting in premature termination (R236X, L264HfsX14). ABCG2 Q141K positively influenced survival in patients with daunorubicin-based treatment regimens.

Detailed characterization of acquired and inherited genetic variants in malignant disorders, such as AML helps to predict prognosis and to apply risk adapted treatment strategies aiming at improving outcome.
Publications mentioned in the Ph.D. dissertation:


Other publications:


Acknowledgements

At first I wish to thank my boss Attila Tordai for giving me the opportunity to work in his laboratory and perform my PhD study. This dissertation would not have been possible without the chance he gave me 5 years ago.

I would like to express the deepest appreciation to my advisers Hajnalka Andrikovics and András Váradi who continually supported me. I cannot find words to express my gratitude to Hajnalka Andrikovics for answering all of my "short questions", lasting forever.

I share the credit of my work with my colleagues; András Bors, who helped me many times with his good advices and excellent sense of humor; Nóra Meggyesi who supported me whenever I needed it; Julia, Magdi, Klári and Csilla for their help in the laboratory, "Hungarian language lessons" and "cook lessons"; Tünde, Brigi and Martina, for their optimism, great jokes and excellent atmosphere. I am indebted to all my colleagues for their support and acceptance.

I would like to express my gratitude to everyone from St. Istvan and St. Laszlo Hospital with whom I had a pleasure to work; Arpad Batai, Judit Csomor, Nora Lovas, Andrea Sipos, Eva Karaszi, Janos Dolgos, Sandor Fekete, Judit Reichardt, Eniko Lehoczky, Otto Csaczsovszki, Peter Remenyi, Kozma Andras, Adam Emma and Tamas Masszi.

A special note of thank to Ildikó Kasza, György Várady, George Scheffer, Adrienn Németh, Gergely Szakács, Edit Szabó and Balázs Sarkadi for their collaboration.

I owe my profound gratitude to groups of Balázs Sarkadi, András Varádi and Gergely Szakac for the real scientific discussions during the Tuesday meetings.

It gives me great pleasure in acknowledging the support and help of Professor Anna Erdei and Laszlo Orosz.

I owe a huge debt of gratitude to Viola Pomozi, Tőkés Natália and Mellie Pieraks for providing me with delicious coffee, cakes and "scientific discussions".

I wish to thank, my amazing friends; Asia Łukasik, Asia Kogut, Dzufa, Marta Baranska, Paulina Fros-Pasz, Łukasz Duszak, Kasia Kulas who make my world a happy place.

This PhD would not have been possible without my family; especially without my mum who always believes in me, my father, who I believe, takes care of me, my crazy sisters, brothers and their families who cause that I never feel lonely.

Above all I would to thank my husband Béla Kakuk for his endless support during my PhD. I love you, you make me happy every day!