Contents lists available at ScienceDirect

Acta Haematologica Polonica

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

Original research article/Praca oryginalna

Genomic landscape of human erythroleukemia K562 cell line, as determined by next-generation sequencing and cytogenetics

Małgorzata Kurkowiak [1,*], *Monika Pępek* [1,2], *Marcin M. Machnicki* [1,2], *Iwona Solarska* [3], *Katarzyna Borg* [3], *Małgorzata Rydzanicz* [4], *Piotr Stawiński* [4,5], *Rafał Płoski* [4], *Tomasz Stokłosa* [1,**]

[1] *Department of Immunology, Center of Biostructure Research, Medical University of Warsaw,*
[2] *Postgraduate School of Molecular Medicine, ,*
[3] *Department of Diagnostic Hematology, Institute of Hematology and Transfusion Medicine, ,*
[4] *Department of Medical Genetics, Medical University of Warsaw,*
[5] *Department of Genetics, Institute of Physiology and Pathology of Hearing, Kajetany/Warsaw, Poland*

ARTICLE INFO

Article history:
Received: 27.02.2017
Accepted: 07.06.2017
Available online: xxx

Keywords:
- Chronic myeloid leukemia
- K562
- Human erythroleukemia
- Next-generation sequencing
- Genomic instability

ABSTRACT

We have performed detailed analysis of the genomic landscape of commercially available K562 cells, employing targeted enrichment of nearly 1300 cancer-related genes followed by next-generation sequencing (NGS) and also classical cytogenetics. Deep sequencing revealed 88 variants of potentially biological significance. Among them we have detected alterations in genes already known to be mutated in K562, such as *TP53* but also in several other genes, which are implicated in tumorigenesis and drug resistance, such as *MLH1*, *ASXL1* and *BRCA1* as the most prominent examples. Fluorescence *in situ* hybridization (FISH) of interphases of K562 cells revealed multiplication of the *BCR* and *ABL1* gene copies, as well as the amplification of the *BCR-ABL1* fusion gene. Our results may help to better understand genomic instability of the blastic phase of CML represented by the K562 cell line and can help researchers who want to employ this cell line in various experimental settings.

* *Corresponding author at*: Department of Immunology, Center of Biostructure, Medical University of Warsaw, Banacha 1a, Bldg F, 02-097 Warsaw, Poland.
** *Corresponding author at*: Department of Immunology, Center of Biostructure, Medical University of Warsaw, Banacha 1a, Bldg F, 02-097 Warsaw, Poland.
E-mail addresses: mkurkowiak@wum.edu.pl (M. Kurkowiak), tomasz.stoklosa@wum.edu.pl (T. Stokłosa).
http://dx.doi.org/10.1016/j.achaem.2017.06.002

Please cite this article in press as: Kurkowiak M, et al. Genomic landscape of human erythroleukemia K562 cell line, as determined by next-generation sequencing and cytogenetics. Acta Haematol Pol. (2017), http://dx.doi.org/10.1016/j.achaem.2017.06.002

Introduction

Chronic myeloid leukemia (CML) is a clonal hematopoietic stem cell disorder characterized by the increased and deregulated growth and maturation of myeloid cells in the bone marrow. Disease is caused by reciprocal chromosomal translocation t(9;22)(q34;q11) which results in the genetic abnormality called the Philadelphia chromosome (Ph), encoding fusion gene [1]. *BCR-ABL1* encodes a constitutively active oncogenic tyrosine kinase BCR-ABL1, which transforms hematopoietic stem cells by activating several proliferative and antiapoptotic pathways, but also by increasing genomic instability. One of the most frequently used CML cell line model is the K562 BCR-ABL1-positive human erythroleukemia cell line, which was derived in 1970 from a pleural effusion of a female patient with CML in blastic phase (CML-BP, also known as blast crisis) [2], decades before the era of targeted therapy with tyrosine kinase inhibitors. K562 cells do not express MHC molecules on their surface and serve also as one of the typical target cells for measuring activity of NK cells [3, 4]. Therefore, K562 represents an important tool for the studies of malignant hematopoiesis as well as for the studies on the molecular pathogenesis of leukemia and human cancer in general. This is reflected by the fact that a number of publications mentioning K562 available in PubMed exceeds 800 per year in recent years and totals in more than 16.000 publications since 1975 with K562 among the key words. The cell line was cytogenetically characterized many times giving different results. K562 karyotype was described as hypodiploid in short term cultures and near triploid in long time cultures[2] but also Ph-positive hyperdiploid karyotypes [5], Ph positive and near triploid[6] or Ph-negative and near triploid [7, 8]. Possible reasons for such discrepancy include genomic instability of K562 cell line, especially in the long-term culture, amplification *BCR-ABL1* oncogene but also new mutations in DNA repair genes (such as *MLH1*) described in this work. However, despite so numerous research employing K562 cells a detailed genomic analysis of this cell line is not available and so far there are no publications describing genomic landscape of K562 cell line. To accurately characterize genetic features of K562 cell line that is currently used in leukemia research, we performed next-generation sequencing (NGS) of K562 cell line DNA from the early passage. We also investigated the copy number of *BCR*, *ABL1* and fusion gene *BCR-ABL1* using fluorescence *in situ* hybridization (FISH).

Materials and methods

Cell line

The K562 cell line was purchased from Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ, Braunschweig, Germany). Cells were grown in RPMI 1640 medium (Gibco, Life Technologies) with 10% fetal bovine serum (Invitrogen) at 37 °C in 5% CO_2. K562 cells were passaged 5 times before isolating genetic material with the use of Gentra Puregene Cell Kit (QIAGEN) for NGS sequencing and FISH experiments. The cell line was confirmed at each passage to be mycoplasma free (PCR-based test).

Next-generation and Sanger sequencing

Mutational analysis of K562 cell line was performed using SeqCap EZ Choice (Roche NimbleGen) custom enrichment. Coding sequences of almost 1300 genes, selected on the basis of literature review and major commercial cancer gene panels (Supplementary Table I), were sequenced on Illumina HiSeq 1500, as described previously [9, 10]. The mean depth of coverage was 146×, 96.5% of our target was covered at least 10× and 93.3% of the target was covered at least 20×. Paired, 100 bp reads were trimmed, quality-filtered and aligned to hg19 genome, followed by duplicate removal, variant calling and variant annotation (GATK, SnpEff [11]).

All variants that did not pass quality check were excluded from further analysis. Protein sequence-changing variants, splicing regions variants and start or stop codon-gain variants were then subject to filtering based on their frequencies in 1000 genomes [12] and Exome Sequencing Projects [Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA (URL:http://evs.gs.washington.edu/EVS/)]. All variants more common than 1% in those databases (also considering frequencies in European population) were excluded. Variants more common than 10% in our internal database (more than 2000 sequenced samples, representing various diseases, including cancer, inherited genetic disorders and normal tissue samples from affected individuals) were also removed. This filtering step was applied to remove frequent non-pathogenic variants unique to the Polish population as well as sequencing artifacts (false positives) generated during sequencing process, which cannot be eliminated using publicly available databases. The existence of selected mutations/variants identified by the NGS was confirmed by Sanger sequencing. Briefly, DNA of a particular coding region surrounding the mutation was amplified in PCR reaction using HotStarTaq Plus DNA Polymerase (QIAGEN). The PCR primer sequences were as follows: TP53-F: tgttcacttgtgccctgact, TP53-R: ttaacccctcctcccagaga, ASXL1e12F: tgtatgccatgacccttaagct, ASXL1e12R: cctcaccaccatcaccactg. The PCR products were purified using Agencourt AMPure XP (Beckman Coulter), labeled with BigDye Terminator v.3.1 (Applied Biosystems) according to the manufacturer's instructions and sequenced by an external sequencing facility.

CADD [13], PolyPhen 2 [14], SIFT [15], fathmm [16], MutationTaster [17] and CHASM [18, 19] methods were used to predict functional consequences of identified variants. CHASM was run on default parameters using acute myeloid leukemia passenger mutation rate table.

Fluorescence in situ hybridization (FISH)

K562 cells were harvested according to standard cytogenetical procedures. After cell synchronization by colcemid for 20 min at 37 °C (10 μg/ml, Biosera, France), pelleted cells underwent a hypotonic treatment using 0.075 M KCl solution (Merck, Germany) for 20 min at 37 °C to swell the cells. The cells were then fixed in cold Carnoy's fixative solution composed of 3:1 methanol and 100% acetic acid (Merck) and

Please cite this article in press as: Kurkowiak M, et al. Genomic landscape of human erythroleukemia K562 cell line, as determined by next-generation sequencing and cytogenetics. Acta Haematol Pol. (2017), http://dx.doi.org/10.1016/j.achaem.2017.06.002