

## Comparison of MLL Fusion Genes Expression among the Cytogenetics Abnormalities of Acute Myeloid Leukemia and Their Clinical Effects

Senol Dogan<sup>1\*</sup>, Amina Kurtovic-Kozaric<sup>2</sup>, Albenita Hajrovic<sup>1</sup>, Muhamed Lisic<sup>1</sup> and Ercan Gokgoz<sup>3</sup>

<sup>1</sup>Department of Genetics and Bioengineering, Faculty of Engineering and Information Technologies, International Burch University, Sarajevo, Bosnia and Herzegovina

<sup>2</sup>Department of Clinical Pathology, University of Sarajevo, Sarajevo, Bosnia and Herzegovina

<sup>3</sup>Department of Information Technologies, Faculty of Engineering and Information Technologies, International Burch University, Sarajevo, Bosnia and Herzegovina

\*Corresponding author: Senol Dogan, Department of Genetics and Bioengineering, Faculty of Engineering and Information Technologies, International Burch University, Sarajevo, Bosnia and Herzegovina, Tel: +387 33 944 400; Fax: +387 33 944 500; E-mail: senol1dogan3@gmail.com

Rec date: March 21, 2016; Acc date: April 14, 2016; Pub date: April 21, 2016

Copyright: © 2016 Dogan S, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

### Abstract

Mixed-lineage leukemia (MLL) is a subtype of acute myeloid leukemia with more aggressive prognosis than other subtypes. Translocations of MLL gene with other partner genes, forming the MLL-fusion proteins (MLL-FPs), are the main characteristics of MLL leukemia. Many studies have demonstrated that MLL-FPs such as: MLL-AF4, MLL-AF6, MLL-AF9, MLL-AF10, MLL-ENL, MLL-ELL, MLL-EPS15, as well as partial tandem duplication are the most common abnormalities that play significant role in MLL-rearranged leukemia. Gene expression profiles from 197 patients and 180 clinical data were downloaded from TCGA database. R statistical program has classified clinical and genomic data simultaneously according to cytogenetic abnormalities. As a result of this analysis, the most frequent 47 MLL-FPs genes expression have been detected and compared with other cytogenetic abnormalities such as t(4;11), t(9;11), t(8;21), t(15;17), complex, inversion 16, trisomy 8 and cytogenetically normal AML. 35 out of 46 MLL-FPs genes presented with abnormal gene expression profile. This study showed that MLL-FPs are not just active and related with MLL, but also with other subtypes of AML.

**Keywords:** AML; MLL; Data mining; Cytogenetic abnormalities; Fusion protein; Gene expression

### Introduction

Acute myeloid leukemia (AML) is a clonal hematopoietic disorder that may be derived from either a hematopoietic stem cell or a lineage-specific progenitor cell [1]. Hematopoietic tissues have a potential to produce various types of malignancies, such as acute lymphocytic leukemia (ALL), acute myelogenous leukemia (AML), chronic myelogenous leukemia (CML), chronic lymphocytic leukemia (CLL) [2]. Acute leukemia is one of the most common types of leukemia among adults and constitutes 97% of all childhood malignancies, which show clonal expansion and changing specific stages of normal myeloid and lymphoid hematopoiesis [3]. Although cytogenetic analysis has been used to identify the pathogenesis of acute myeloid leukemia (AML) for more than decades, clinically defined subtypes are very heterogeneous diseases and difficult to characterize in a same group [4,5]. The cytogenetic of the subtypes and chromosome translocation are the main key properties to distinguish the disease and prognostic factors of AML [1,6,7]. The AML and ALL cytogenetic reports have revealed that many non-random chromosome abnormalities included specific genes that implicated in the process of leukemogenesis [8].

Lastly, the new term mixed-lineage leukemia was added to the literature [9]. MLL gene plays a positive regulator of global gene expression in early embryonic development and hematopoiesis [10]. The gene encodes a very important epigenetic transcription factors, such as HOX genes [11]. In addition to that, MLL fusion proteins, main properties of MLL, are produced by chromosomal translocations, which affect the MLL gene at 11q23 [12]. If AML includes MLL

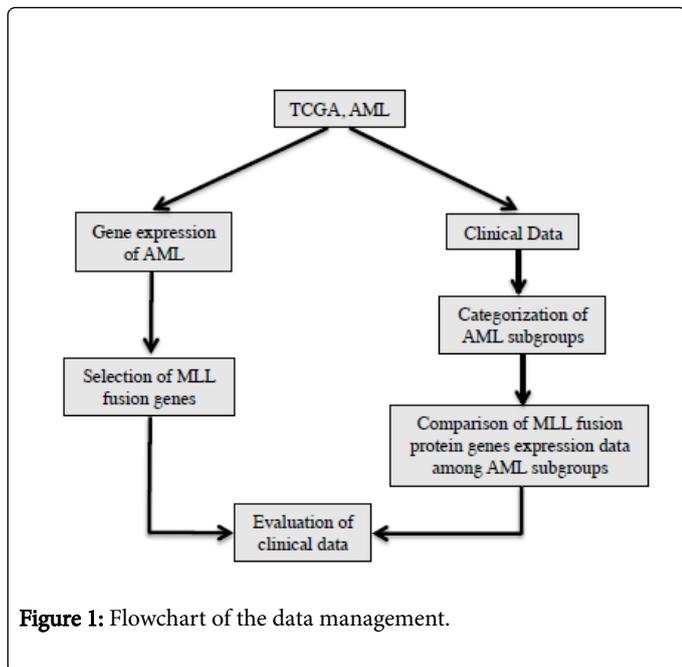
chromosomal rearrangements and produces fusion proteins, it is assumed to be followed by poor prognosis and aggressive for infants [13-15]. The fusion proteins are observed in both hematological and solid tumors as breaking within genes on each chromosome [7]. The translocation genes involved in AML might be transcription regulators, which determine the cellular development and cell fate [16]. There are more than 80 different partner genes, but the most commonly observed MLL-FPs are MLL-AF4 in ALL, MLL-AF9 and MLL-AF10 in AML and MLL-ELL in MLL [17].

Although it is well known that the fusion proteins activity are specific reasons for The previous studies show that statistical analysis of genomic data and clinical research should have been done together to understand better of new type of leukemia [18-21]. Development of MLL, the research has been done to show the potential correlation between the proteins and other subtypes. Finding the expression activity of the fusion genes is the main target of the research. The main idea of the research is to see the genes expression profile of MLL-FPs in other subtypes of AML and potential relation between them.

### Materials and Methods

#### The cancer genome atlas, R, HCE 3.5, Genevestigator

Gene expression and clinical data of acute myeloid leukemia are the main material for the research, which are downloaded from TCGA database, <https://tcga-data.nci.nih.gov/tcga/>. The data have been mined by different computational programs and web tools (Figure 1).



The statistical analyses and data comparison have been done by R statistical program (<https://www.r-project.org>), which is created for this particular research to analyze the data. The program categorizes genes which represent different expression profiles in the subtype of AML. HCE 3.5 (<http://www.cs.umd.edu/hcil/hce/>) has been used to cluster gene expressions. According to the clinical data, the different subtypes have been found and abnormally expressed MLL fusion genes were detected. Synchronization of the clinical and expression data is the key part of the method.

### Clinical data

AML clinical data have been used to find subtypes which are made of distinct cytogenetic abnormalities. Patients have different chromosomal abnormalities which are detected, identified and separated into different types: t(4;11), t(9;11), t(8;21), t(15;17), complex, inversion 16, trisomy 8 and AML. TCGA ID numbers of each patient is very useful for matching with their gene expression value. Therefore, same IDs have been used to find the correct gene expression value in the subtype of AML. The clinical data were prepared to find their gene expression values.

### Gene expression data according to AML subtypes

197 AML patients' expression values and 19798 genes were compared within each subtype. According to cytogenetic abnormalities which were derived from the clinical data, patients have been separated and categorized into subgroups of AML (Figure 1).

The IDs of each subgroup member is used to find their gene expression value. Then the abnormally high and low expressed genes of the subtypes have been compared among each other and average value (Table 1).

Abnormality	LEG	Expression	HEG	Expression	Average
t(4;11)	MLLT7	229.2	MLLT2	20.248	3.536

t(9;11)	ARHGEF1 2	208.2	MLLT2	30.241	4.271
t(8;21)	ArgBP2	242.7	MLLT2	16.497	3.975
t(15;17)	LASP1	64.0	ARHGEF1 2	74.948	3.276
Complex	ArgBP2	307.6	MLLT2	25.618	4.393
Inv16	ARHGEF1 2	222.3	MLLT2	23.813	4.406
Trisomy	ARHGEF1 2	115.5	MLLT2	36.988	4.424
AML	ArgBP2	269.1	MLLT2	23.846	4.372

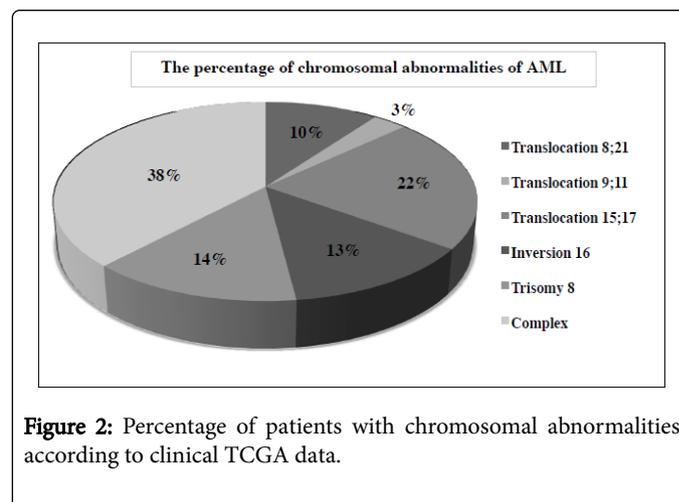
**Table 1:** Selected high and low expressed genes in the subgroups. LEG: Low Expressed Genes, HEG: High Expressed Genes. Average shows the gene expression values of all patients.

Although there are more than 80 MLL fusion encoded genes [15], the 46 most frequently found have been analyzed. The R program categorized and selected the genes depending on their expression abnormalities of AML cytogenetic disorders.

## Results

### Clinical data analysis

69 cytogenetic abnormalities patients have been extracted from 197 AML as subtypes which include t(8;21) (10%), t(15;17) (22%), complex (38%), MLL, inversion 16 (13%), trisomy 8 (14%) and MLL (3%). The following step was to find the MLL-FP genes expression profile of the subtypes and compare them among each other (Figure 2).



### Clinical outcome of the subtypes

The patients of the subtypes show different survival rate (Figure 3). While t(15;17) and inv16 have the highest survival rate, t(9;11) and t(4;11) have the lowest survival rate. It is obvious that patients with the subtypes involving MLL translocation on 11q23, have the worst prognosis.

### Gene expression result among subtypes

The comparison of 46 mostly common MLL-FPs gene expression shows different profile and abnormal deviations (Table 2).

	AVG	AML	t(4;11)	t(9;11)	t(8;21)	t(15;17)	Complex	Inversion 16	Trisomy 8
MLLT2	21945.8	23846	20248.8	30241.4	16497.8	212.2	25618.8	23813.7	36988
LASP1	13034.2	17052.2	14414.8	13910.3	16251.6	64	16301.3	17757.3	12539.8
FOXO3A	8659.5	11759	12371.2	5682.6	11351.2	127	12697.9	9441.6	8945.2
PICALM	10142.1	14368.3	9036.1	11011.2	11205.8	1238.5	13839.2	13086.2	11577.9
CREBBP	8020.4	9493.4	7986.9	9701.6	9430.9	1139.3	9945.3	9956.9	7982.1
MIFL	7627.6	8627.3	6424.6	7559.8	8206.8	1423.6	9061.8	8358.5	12358.3
GMPS	4416.2	5379.6	5605.7	6211.5	4718.9	247.1	4647.9	6337.5	3144.6
CASC5	2301.1	2548.7	3861.9	2660	2015.7	87.4	2456.3	2839.2	2187
EP300	3166.5	3312.2	3189.5	4507.7	3076.4	536.6	3723.9	2877	4254.9
MAML2	1424.2	1680.4	2960.4	496.3	1549.7	205.2	2286	2249	222.4
LPP	3290.6	4443.1	2728.7	5231.2	3063.8	291.3	3969.8	4503	3246
SMAP1	3634.3	5139.7	2673.7	4660	4466	243.4	4703.9	5112.1	3581.3
MLLT6	3385.8	3898.9	2271.4	3983.2	3189.3	156.2	4269.8	3708.4	6122.3
GAS7	4971	7917.8	2235.9	7763.8	7880.8	774.1	5250.9	6934.7	3956.6
AFF3	3981	3905.3	2087.3	6345.4	3564.9	504.1	3271.5	5386.7	6706.8
ARHGAP26	2363	2117.7	1651.8	2908.6	1126	4982.7	1451.3	2580.4	1839.8
MLLT11	4439.1	4338.1	1472.1	1766.3	7518.2	497.2	5269.1	5832.7	8718.2
PBX1	698	902.7	1471.5	587.3	332.5	344.6	1338.4	326.3	485.5
EEN	1592.3	1989.1	1401.1	2037.2	1699.3	178	2245.8	2108.2	1476.7
RARA	1428	1770.4	1252.5	1324.1	2258.1	94.5	1854.4	1911.1	1301.5
MLLT3	1031.3	1473.1	1247.1	1317.9	861	173.4	2453.2	991.3	175.5
SEPT9	1443.1	1702.4	1142.1	1347.2	1551.4	291.5	1600.5	1782.9	2386
MPFYVE	1248.8	1245.7	1135.6	1889.8	1136	312.5	1210.1	1228.8	1829.2
NCKIPSD	1649.6	1504.1	1089.2	1874.2	2418.1	398.6	1751.5	1891.8	2123.8
ELL	773	745.1	992.2	580.1	912.8	461.3	752.8	904.4	807.2
TIRAP	1526.9	941.6	954.7	1052.6	956.8	5210.1	912.1	925.3	676.8
DAB2IP	416.9	437.5	549.1	382.9	503.6	81.4	556.7	421.3	423.5
TET1	1308	690.1	469.6	1034.8	1811.7	3251.6	981	946.7	660.4
ArgBP2	656.6	269.1	398.4	233.8	242.7	2972.6	307.6	199.2	241.7
ARHGEF12	10942.5	277.9	384.6	208.2	303.7	74948.2	414.8	222.3	115.5
MLLT4	307.21	239.84	229.60	146.4	345.25	481.82	477.25	239.25	230.85
MLLT7	670.1	439.6	229.2	339.2	523.4	2403.7	418.5	303.9	472.6
MLLT10	7718.3	7720.9	6137.4	9266.8	8128.1	6755.4	8030.5	7558.9	8151.2

MYOF1	2595.5	3397.7	5361.5	4787.5	935.9	1699	1894	3363.2	127.1
ACACA	3846.4	2989.4	2590.1	2932.4	4305.8	4407	3097.3	3382.2	6209.8
FNBP1	1195.6	1814	1405.3	774.5	922	1009.6	1918.8	2123.4	215.6
EPS15	1658.8	1418.5	1192.8	1293.9	1843.2	3034	1592.6	1447.8	1207.4

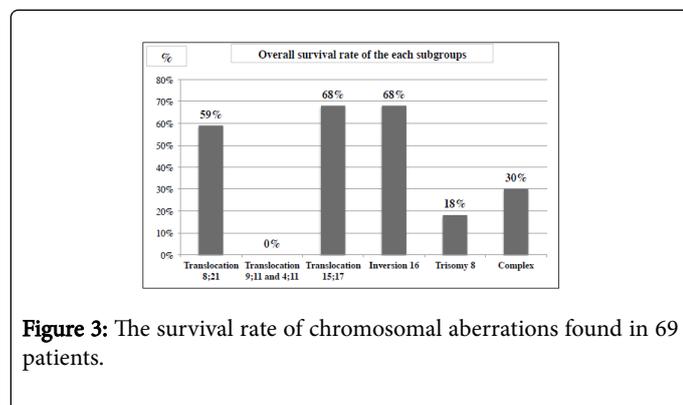
**Table 2:** MLL fusion genes expression in different subtypes of AML.

The black and gray cells represent the high expressed and low expressed values of the genes, respectively. The comparison is done between average of the gene expression and the subtypes.

MLLT2 is up-regulated in all translocations except in t(15;17). MLLT4 and MLLT7 are down-regulated in t(4;11) having a significant deviation when compared to average. For instance, ARHGEF12 is down-regulated in t(9;11), inv16, trisomy 8 and AML, but it is up-regulated in t(15;17). ArgBP2 is a quite good example of down-regulated gene expression in t(8;21), complex and AML. LASP1 expression fluctuates, it is down-regulated in t(15;17), but up-regulated in t(8;21). AF1Q and MLLT10 expression is high, but MYOF1 and GPHN expressions are low in int(8;21) and t(9;11) subtypes, respectively. t(15;17) has the highest gene expression diversity in both high and low directions if compared with other subtypes. t(15;17) and trisomy 8 have more diverse genes which represent up and down-regulated in the subgroups. For instance, EPS15, ArgBP2, AFX, ARHGAP26, CXXC6, ARHGEF12, TIRAP are highly expressed genes and MAPRE1, EP300, LASP1, MLLT6, RARA, SEPT9, ELL, EEN, CCDC94, MPFYVE, AF15q14, CREBBP, GAS7, CIP29, CBL, PICALM, MAML2, ABI1, AF9-MLLT3, AF9q34, FOXO3A, SMAP1, AFF3, NCKIPSD, EEFSEC, GMPS, LPP, AF4P12, MIFL, AF4, AF1Q are low expressed in t(15;17). Trisomy 8 as a subtype includes AF1Q, AF4p12, MIFL, AF4, GPHN, ACACA, MLLT6 as highly expressed genes and MYOF1, ARHGEF 12, TIRAP, MAML2, FNBP1 as low expressed ones. Interestingly, inv16 has no significant MLL-FPs gene expression variations. The results support our theory that MLL-FPs genes are not just active in MLL, but also in other subtypes.

### MLL-FP genes and clinical outcome

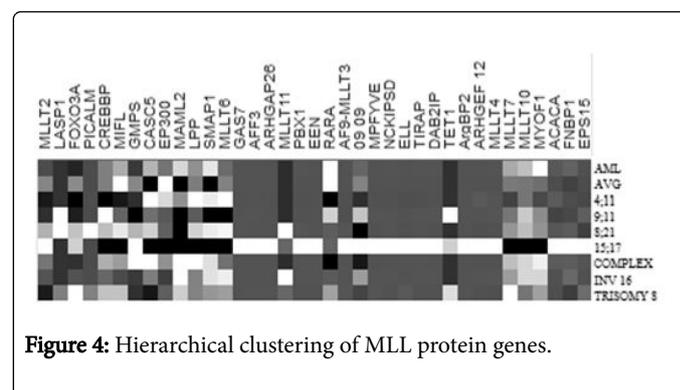
According to the AML subtypes, clinical outcome shows high heterogeneity. While t(15;17) and inv16 have the highest survival rate and relatively good prognosis, t(9;11) is lethal and leads to very bad prognosis (Figure 3).



**Figure 3:** The survival rate of chromosomal aberrations found in 69 patients.

### Hierarchical clustering of MLL protein genes

MLL-FP genes are correlated with each other depending on their expression and clustered to determine their relation. Some of the genes show strong correlation with each other in different subtypes (Figure 4). The first group genes are ACACA, FNBP1 and EPS15 and strongly correlated to all the subtypes. MLLT7, MLLT10, MYOF1 are the second group, ArgBP2, ARGHEF12 and MLLT4 are the third highly correlated genes. TET1 is distinct and has a unique expression profile in all the groups. MPFYVE, NCKIPSD, ELL, TIRAP, DAB2IP, PBX1, EEN, ARHGAP26, AFF3 and GAS7 are the fourth and the maximum correlated genes in all the subtypes. The fifth group of genes demonstrate diverse expression value and are slightly correlated with each other (Figure 4).



**Figure 4:** Hierarchical clustering of MLL protein genes.

### Discussion

However, it is well known that the genes responsible for MLL-FPs are important key factor for MLL leukemogenesis, the research has been done to investigate the expression profile in other subtypes of AML. 46 most commonly found MLL-FPs genes have been compared among the subtypes that the gene expression profile is presented as high and low (Table 2). Analysis of variations in gene expression in different cytogenetic abnormalities showed that significant number of genes is down-regulated in t(15;17). Furthermore, there are some genes which have opposite direction expressions. Therefore, MLL-FPs show diverse expression profiles in different subtypes of AML.

Despite of the fact that MLL is a new subtype of AML leukemia, there is no significant and efficient individual therapy so far. MLL is a result of crucial change in genetic conformation by the translocation and MLL protein complex that can lead to leukemogenesis. The analysis of MLL-FPs might be prospective target to find common properties and patterns between MLL type and other subtypes of AML. Moreover, we can find the correlation between abnormal expression of the genes and clinical outcome among all the subtypes.

## Conclusion

MLL is very aggressive type and there is no strong therapy. The bad prognosis and clinical outcome may be correlated with the fusion proteins activity. That's why the gene expression profile might be used to find therapeutic target genes and help us to see the relation between the subtypes of AML and MLL-FPs. The major contribution of the research is to find the reason of different clinical outcome and survival rate among the subtypes of AML and MLL-FP activity.

## Acknowledgements

We are thankful for Bosna Sema Education Institutions for their kind help and support.

## References

1. Kumar C, Chandra (2011) Genetic abnormalities and challenges in the treatment of acute myeloid leukemia. *Genes & cancer* 2: 95-107.
2. Campo E, Swerdlow SH, Harris NL, Pileri S, Stein H, et al. (2011) The 2008 WHO classification of lymphoid neoplasms and beyond: evolving concepts and practical applications. *Blood* 117 19: 5019-5032.
3. Paguirigan AL, Smith J, Meshinchi S, Carroll M, Maley C, et al. (2015) Single-cell genotyping demonstrates complex clonal diversity in acute myeloid leukemia. *Science translational medicine* 7: 281-282.
4. Bacher U, Schnittger S, Haferlach T (2010) Molecular genetics in acute myeloid leukemia. *Current opinion in oncology* 22: 646-655.
5. Grimwade D, Hills RK (2009) Independent prognostic factors for AML outcome. *ASH Education Program Book* 1: 385-395.
6. Rowley JD (2008) Chromosomal translocations: revisited yet again. *Blood* 112: 2183-2189.
7. Rabbitts TH (1994) Chromosomal translocations in human cancer. *Nature* 372: 143-149.
8. Mrozek K, Heerema NA, Bloomfield CD (2004) Cytogenetics in acute leukemia. *Blood reviews* 18: 115-136.
9. Mirro J, Zipf (1985) Acute mixed lineage leukemia: clinicopathologic correlations and prognostic significance. *Blood* 66: 1115-1123.
10. Kourlas PJ, Strout (2000) Identification of a gene at 11q23 encoding a guanine nucleotide exchange factor: evidence for its fusion with MLL in acute myeloid leukemia. *Proceedings of the National Academy of Sciences* 97: 2145-2150.
11. Shen Y, Zhu (2011) Gene mutation patterns and their prognostic impact in a cohort of 1185 patients with acute myeloid leukemia. *Blood* 118: 5593-5603.
12. Slany RK (2009) The molecular biology of mixed lineage leukemia. *Haematologica* 94: 984-993.
13. Behm FG, Raimondi (1996) Rearrangement of the MLL gene confers a poor prognosis in childhood acute lymphoblastic leukemia, regardless of presenting age. *Blood* 87: 2870-2877.
14. Krivtsov AV, Armstrong SA (2007) MLL translocations, histone modifications and leukaemia stem-cell development. *Nature Reviews Cancer* 7: 823-833.
15. Li Z, Lu (2008) Distinct microRNA expression profiles in acute myeloid leukemia with common translocations. *Proceedings of the National Academy of Sciences* 105: 15535-15540.
16. Cleary ML (1991) Oncogenic conversion of transcription factors by chromosomal translocations. *Cell* 66: 619-622.
17. Daser A, Rabbitts TH (2004) Extending the repertoire of the mixed-lineage leukemia gene MLL in leukemogenesis. *Genes & development* 18: 965-974.
18. Senol Dogan, Kurtovic-Kozaric A (2015) Changes of Molecular, Cellular and Biological Activities According to microRNA-mRNA Interactions in Ovarian Cancer. *Computational Molecular Biology* 5: 1-8.
19. Kurtovic-Kozaric A (2015) PRPF8 defects cause misplacing in myeloid malignancies. *Leukemia* 29: 126-36.
20. S Dogan, Kurtovic-Kozaric A, Karlı G (2016). The Detection of Extremely High and Low Expressed Genes by EGEF Algorithm in Invasive Breast Cancer. *J Biom Biostat* 286: 1-2.
21. Kurtovic-Kozaric A, Hasic A, Jerald PR (2016) The reality of cancer treatment in a developing country: the effects of delayed TKI treatment on survival, cytogenetic and molecular responses in chronic myeloid leukaemia patients. *British journal of haematology* 172: 420-427.