Introduction

Leukemia is a kind of blood and bone marrow malignancy characterized by the fast development of abnormal white blood cells [1]. Based on the kind of stem cell involved and whether the leukemia is acute or chronic, there are four primary subtypes of leukemia [2]. Acute myeloid leukemia (AML) is the most frequent kind of leukemia in adults and is associated with a high number of annual deaths in the United States [3]. It is distinguished by an overabundance of undifferentiated myeloid cells in the bone marrow, resulting in a lack in normal blood cell synthesis [4]. The symptoms may include fever, weight loss, fatigue, breathlessness, frequent infections, and abnormal bleeding[5]. The causes of AML are diverse and can include exposure to therapeutic or environmental agents that damage DNA, although the exact etiology is often unclear. AML is caused by a combination of hereditary and ecological factors, such as genetic mutations, age, radiation exposure, chemical exposure, and previous treatments [6]. Genomic profiling advances have thrown some insight on the function of genetics in AML, but further research is needed to fully understand the mechanisms involved. Diagnosing AML involves analyzing peripheral blood or bone marrow samples for the presence of abnormal myeloid cells. It should be diagnosed when marrow or blood has > 20% blasts of myeloid lineage [7]. The prognosis of AML depends on factors such as the patient's age and subtype of the disease. It is mostly diagnosed at the median age of 68 years. Treatment...
typically involves chemotherapy, sometimes combined with targeted therapy drugs, and may be followed by a stem cell transplant. Recent studies have focused on the molecular pathogenesis of AML, identifying genetic variations that influence the prognosis and altering the classification of the disease. One of the commonly altered genes in AML is DNMT3A (~20%), which plays a role in DNA methylation [8]. It consists of 35 exons and encodes a protein of 912 amino acid. It was mapped to chromosome 2p23.3. Mutations in DNMT3A are linked to cytogenetically normal acute myeloid leukemia (CN-AML). These mutations impair the enzyme’s ability to methylate DNA fully, leading to changes in gene activity and the production of abnormal white blood cells [9]. In the field of research, bioinformatics methods and in silico techniques have revolutionized the study of life sciences. These computational approaches help categorize proteins based on structure and function and assist in the development of servers for molecular sorting using machine learning methods [10]. In silico methods also facilitate screening potential therapeutics against molecular targets, reducing the need for extensive laboratory work and conserving resources [11].

**M E T H O D S**

To perform in silico analysis, 18 computational tools were employed (Table 1).

**Table 1: Tools applied for analysis**

<table>
<thead>
<tr>
<th>In silico tools</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCBI</td>
<td>For retrieval of SNPs</td>
</tr>
<tr>
<td>SIFT, Align GVGD, FATHMM, PANTHER</td>
<td>For the identification of deleterious SNPs</td>
</tr>
<tr>
<td>Mupro, I-Mutant Suite, mCSM</td>
<td>For protein stability analysis</td>
</tr>
<tr>
<td>ConSurf</td>
<td>For estimation of the conservation profile</td>
</tr>
<tr>
<td>HOPE Project</td>
<td>For analysis of structural effects of DNMT3A gene mutation</td>
</tr>
<tr>
<td>SWISS-MODEL</td>
<td>For protein modeling</td>
</tr>
<tr>
<td>SOPMA</td>
<td>For secondary structure analysis</td>
</tr>
<tr>
<td>STRING</td>
<td>In order to anticipate protein-protein interactions</td>
</tr>
<tr>
<td>Vienna package</td>
<td>For the prediction of effect of mutations on RNA secondary structure</td>
</tr>
<tr>
<td>GEPIA, UALCAN, cBioPortal, Cytoscape</td>
<td>For analysis of post-translational modifications</td>
</tr>
</tbody>
</table>

**R E S U L T S**

From NCBI, five SNPs were reported to be found in AML. To identify the damaging and deleterious effects of SNPs that could interfere with the structure and function of DNMT3A gene, four in silico tools (SIFT, Align GVGD, FATHMM, and PANTHER) were used. All the SNPs were predicted as damaging and deleterious by these computational algorithms (Table 2). To analyze the effects of point mutations on the stability of protein structure, three software (Mupro, mCSM, and I-Mutant) were used. The SNPs were shown to reduce the protein structural stability (Table 3).

**Table 2: Analysis of damaging effects of DNMT3A mutations on structure and function of gene using in silico tools**

<table>
<thead>
<tr>
<th>DBSNP RS#</th>
<th>SNPs</th>
<th>SIFT</th>
<th>Align GVGD</th>
<th>FATHMM</th>
<th>PANTHER</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs147001633</td>
<td>R882P</td>
<td>Decrease</td>
<td>C65</td>
<td>Damaging</td>
<td>Probably damaging</td>
</tr>
<tr>
<td>rs147001633</td>
<td>R882L</td>
<td>Decrease</td>
<td>C65</td>
<td>Damaging</td>
<td>Probably damaging</td>
</tr>
<tr>
<td>rs377577594</td>
<td>R882S</td>
<td>Decrease</td>
<td>C65</td>
<td>Damaging</td>
<td>Probably damaging</td>
</tr>
<tr>
<td>rs377577594</td>
<td>R882G</td>
<td>Decrease</td>
<td>C65</td>
<td>Damaging</td>
<td>Probably damaging</td>
</tr>
<tr>
<td>rs377577594</td>
<td>R882C</td>
<td>Decrease</td>
<td>C65</td>
<td>Damaging</td>
<td>Probably damaging</td>
</tr>
</tbody>
</table>

To analyze the evolutionary conservation, ConSurf interpret the results, according to which all the SNPs were predicted as highly conserved (Figure 1).

**Figure 1: ConSurf results of DNMT3A gene amino acid sequence on a multi-colored bar sheet with conservation scale below**

- **Damaging**
  - A predicted functional residue (highly conserved and essential).
- **Deleterious**
  - A predicted functional residue highly conserved and essential.
- **Destabilizing**
  - A predicted structural residue highly conserved and essential.
- **Probably damaging**
  - A predicted functional residue conserved and essential.
- **Decrease stability**
  - A predicted functional residue conserved and essential.
- **Stabilizing**
  - A predicted functional residue conserved and essential.
- **Increase stability**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
- **Decrease**
  - A predicted functional residue conserved and essential.
residues were smaller than the wild-type residue (Table 4 and Figure 2). External interactions will be lost by the smaller size.

Table 4: HOPE’s interpretation of the effect of amino acid changes on DNMT3A protein structure and stability

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Size</th>
<th>Change of Charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>R882P</td>
<td>W&gt;M</td>
<td>+ve to neutral</td>
</tr>
<tr>
<td>R882L</td>
<td>W&gt;M</td>
<td>+ve to neutral</td>
</tr>
<tr>
<td>R882S</td>
<td>W&gt;M</td>
<td>+ve to neutral</td>
</tr>
<tr>
<td>R882G</td>
<td>W&gt;M</td>
<td>+ve to neutral</td>
</tr>
<tr>
<td>R882C</td>
<td>W&gt;M</td>
<td>+ve to neutral</td>
</tr>
</tbody>
</table>

Figure 2: The native (left) and mutant (right) amino acid residues of DNMT3A gene mutations are shown in schematic form.

For the homology modeling of the wild and mutant variants, SWISS MODEL was used. It is designated to construct 3D structure of protein (Figure 3). Different parameters (GMQE, QMEAN) were obtained using this tool (Table 5).

Table 5: Representing the different QMEAN Z score, GMQE, Identity, and template query number of wild type and mutant proteins via SWISS-Model

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Total number of amino acids</th>
<th>Number of amino acids in model</th>
<th>Template query number</th>
<th>Sequence identity</th>
<th>GMQE</th>
<th>QMEAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wildtype</td>
<td>912</td>
<td>613-912</td>
<td>6pa7.1.K</td>
<td>100.00%</td>
<td>0.39</td>
<td>-1.14</td>
</tr>
<tr>
<td>R882P</td>
<td>912</td>
<td>613-912</td>
<td>6pa7.1.K</td>
<td>99.85%</td>
<td>0.39</td>
<td>-1.07</td>
</tr>
<tr>
<td>R882L</td>
<td>912</td>
<td>613-912</td>
<td>6pa7.1.K</td>
<td>99.85%</td>
<td>0.39</td>
<td>-1.19</td>
</tr>
<tr>
<td>R882S</td>
<td>912</td>
<td>613-912</td>
<td>6pa7.1.K</td>
<td>99.85%</td>
<td>0.39</td>
<td>-1.19</td>
</tr>
<tr>
<td>R882G</td>
<td>912</td>
<td>613-912</td>
<td>6pa7.1.K</td>
<td>99.85%</td>
<td>0.39</td>
<td>-1.26</td>
</tr>
<tr>
<td>R882C</td>
<td>912</td>
<td>613-912</td>
<td>6pa7.1.K</td>
<td>99.85%</td>
<td>0.39</td>
<td>-1.23</td>
</tr>
</tbody>
</table>

Figure 3: Protein structures of wildtype and mutants using SWISS-MODEL Self-Optimized Prediction Method from Alignment abbreviated as SOPMA, was used to interpret the secondary structure of protein. It predicted different characteristics of secondary structure i.e., alpha helix, 310 helix, Pi helix, beta bridge, ambiguous states, and other states (Figure 4).

Figure 4: Presenting secondary structure of DNMT3A protein using SOPMA
To create protein interaction network of DNMT3A protein, STRING Database was used. It predicted that DNMT3A is functionally associated with five other proteins (EZH2, DNMT3L, MYC, DNMT1 and HIST2H3PS2). Any change in the structure of DNMT3A protein can affect the functions of these proteins related to it (Figure 5).

![Figure 5: Network of protein-protein interactions of DNMT3A using STRING Database](image)

For the assessment of secondary structure of mRNA, RNA fold webserver on the Vienna package was used. All the mutations resulted in anomalous RNA folding, thus affecting mRNA localization and influencing the translation of protein (Figure 6).

![Figure 6: Impacts of SNPs on RNA secondary structure by Vienna package](image)

For comprehensive and interactive analysis of expression and methylation levels of DNMT3A gene in AML, UALCAN was used. In Figure 8 (A), a significant variation of DNMT3A methylation was seen between male and female. In Figure 8 (B), methylation levels in different age groups of AML patients were shown. Individuals with age of 21 to 80 years were found to have significant methylation levels.

![Figure 7: (A) Overview of DNMT3A mRNA expression in AML from TCGA obtained from GEPIA2 (B) Kaplan-Meier plots showing overall survival between high and low DNMT3A groups in AML](image)

For check the PTM sites, cBioPortal was used. It predicted that 23% (altered / profiled = 131 / 560) of genetic alterations were reported in DNMT3A including in-frame, missense, splice and other mutations etc. It displayed all Post Translational Modifications (PTMs) available for the transcript. Different PTM types were shown with varying color codes such as green for phosphorylation, red for ubiquitination and purple for sumoylation (Figure 9).

![Figure 9: Representing genetic alterations of DNMT3A and PTM](image)
DISCUSSION

In silico methods are preferred in many scientific disciplines because they are cost-effective, time-efficient, flexible, reduce ethical concerns, and can make predictions about complex systems so they are used to better understand how mutations might disrupt protein structure and function [12]. DNMT3A is a de novo DNA methyltransferase that has lately gained attention as a result of its common mutation in a variety of immature and mature hematologic neoplasms. DNMT3A mutations occur early in cancer formation and tend to be associated with a poor prognosis in persons with acute myeloid leukemia (AML), making this gene an intriguing target for innovative therapies [13]. Disease-causing SNPs are frequently detected in evolutionarily conserved areas. Five SNPs of DNMT3A were found to be involved in AML. All the five SNPs (R882P, R882L, R882S, R882G, R882C) were retrieved from dbSNP of NCBI database. For the assessment of function, different prediction tools (SIFT, Align GVGD, FATHMM, and PANTHER) were used. These tools predicted that all the mutations were present in exposed region and are highly conserved, having high conservation scores. Therefore, increases the risk of tumorigenesis[20]. HOPe was used to assess the consequences of amino acid replacements on the protein's physical and chemical characteristics, spatial structure, hydrophobicity, size, charge, and function. All of the mutant residues were anticipated to be smaller than the wild-type residue. The lower size will eliminate external interactions. The variations in hydrophobicity and size between mutant and wildtype residues may cause protein framework disruption by disrupting H-bonding connections with adjacent residues [21]. The mapping of amino acid substitutions can be accomplished using 3D protein structure analysis. To create 3D models of the mutated residues, SWISS-MODEL was used. Chakravarty explained the changes in secondary structure during the transition suggest that helices and strands are likely to be extended at the expense of turns and coils. The decrease in binding factors of the interface residues as they transition from the unbound to the bound form reflects a loss in flexibility during complex formation [22]. For the analysis of secondary structure, SOPMA was used. This software predicted that all the mutations were present in the exposed region in the form of helix. Helices can tolerate more mutations than strands without change, because they have more inter-residue interactions [23]. Mutations that alter secondary structure inside the protein core are more likely to produce proteins that do not fold correctly, making their structures more difficult to crystallize. Hence, anomalous proteins are formed[24]. STRING database was used to analyze how DNMT3A is associated with other proteins. It showed the association of DNMT3A gene with five other genes (EZH2, DNMT3L, MYC, DNMT1 and
DNMT1 maintains methylation during DNA replication. Trowbridge et al., demonstrated that DNMT1 haploinsufficiency impaired leukemia stem cell (LSC) activity by depressing bivalent chromatin domains [25]. DNMT3A can reside in the nucleus as dimers, tetramers, and larger oligomeric complexes. The oligomers are made of either homo-dimeric DNMT3A molecules or heterodimeric DNMT3A–DNMT3L molecules [26]. Apart from programmed changes in oligomerization, such as those caused by developmental changes in DNMT3L expression and differential DNMT3A/3B isoform usage, a number of pathologic changes, such as mutations at DNMT3A binding interfaces have been shown to influence oligomerization and alter cell behavior [27]. For the prediction of changes in RNA secondary structure, RNA fold webserver in Vienna Package was used, according to which all the mutants of DNMT3A led to the abnormal folding of RNA, thus influencing RNA localization and affecting the protein translation. In the current study, it was elaborated that in addition to its known involvement in HSC differentiation, DNMT3A has been linked to the preservation of RNA splicing and genomic integrity, both of which are dramatically altered when DNMT3A is mutated [28]. The loss of DNMT3A resulted in the downregulation of spliceosome genes and aberrant RNA splicing. Mutations induce abnormalities of DNMT3A splicing, likely through changing exonic splicing silencers [29]. GEPIA2 was utilized to examine the expression of DNMT3A in AML and overall survival analysis. It was illustrated by GEPIA2 that mRNA expression of DNMT3A gene was significantly higher in AML as compared to normal tissues. DNMT3A expression was shown to be higher in AML in a prior study [30] and its mutations were independently linked with poor outcome in AML patients with an intermediate-risk cytogenetic profile or CN-AML [31]. UALCAN was employed for a thorough and interactive investigation of the expression and methylation levels of the DNMT3A gene in AML. A statistically significant overrepresentation of DNMT3A methylation status was reported in patients ≥ 50 years old in recent research. There was no evident relationship between DNMT3A methylation status and gender. While the prevalence of AML is increasing, no difference in frequency has been seen between males and females [32]. For prediction of post translational modification sites, cBioPortal was used. Large-scale studies have recently revealed that overlap between PTMs and SNPs results in damaged PTMs, which severely influence both gene and protein function and are linked to human cancer [33]. Radivojac et al., also discovered a link between phosphorylation site disruptive variations and somatic cancer mutations [34]. Cytoscape was used to visualize molecular interaction networks of human DNMT3A protein with other associated proteins [35]. Maintaining protein interactions is critical for maintaining system homeostasis [36]. Any change in the gene leads to disruption of functioning of correlated genes. Thus, various studies have found a high relationship between DNMT3A genetic variants and prognosis in AML patients, with mutations predicting a markedly bad prognosis in AML patients[37].

CONCLUSIONS
In this study, in silico tools were used to analyze the impact of DNMT3A mutations, which are associated with hematologic neoplasms, particularly AML. The findings shed light on the potential mechanisms underpinning DNMT3A’s function in cancer formation and emphasize its therapeutic potential. All the analyzed SNPs were found to be deleterious and damaging, destroying the DNMT3A structure and function. These SNPs may reduce the DNMT3A capacity to fully methylate DNA which abrupt its activity preventing the normal differentiation, ultimately leading to AML. Overall, this study demonstrates the importance of in silico methods in elucidating the complex molecular processes involved in cancer development and progression, particularly AML.

Authors Contribution
Conceptualization: AMA
Methodology: SA, KJ, AT
Formal analysis: SA, KJ, AT
Writing-review and editing: SA, AMA

All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest
The authors declare no conflict of interest.

Source of Funding
The authors received no financial support for the research, authorship and/or publication of this article.

REFERENCES
Single Nucleotide Changes in DNMT3A Gene in AML Patients

DOI: https://doi.org/10.54393/pbmj.v6i04.882


