Microarray Technique for Studying Genetic Diversity in Saudi Sickle Cell Patients

Awwad Alenezy, MD* Dalal S.AlShaya, Ph.D** Hayat Alafari, Ph.D** Amna Alotiby, Ph.D*** Talat Bukhari, Ph.D*** Abdulkarim S. BinShaya, Ph.D**** Hisham Waggiallah, Ph.D**** Wael Alturaiki, Ph.D***** Yasser Alnaam, Ph.D****** Arwa F. Alanazi, BchD******* Amani F. Alanazi, BRT******** Faris Q.B.Alenzi, Ph.D****

ABSTRACT

Introduction: Sickle cell anemia is one of the most common heritable hematologic diseases affecting humans. Approximately 3 million individuals had genetic blood diseases in Saudi Arabia, of whom 30% lived in Dammam Region. The aim of this study was to conduct complete gene survey studies using microarray technology.

Material and Methods: Blood samples from 90 unrelated sickle cell disease patients were obtained from the KKUH, Riyadh between from January 2017 and to June 2020. In this study, linkage disequilibrium has been determined between single nucleotides polymorphism loci in the same region of beta globin gene to identify which of them had a role for the unique variable appearance of the disease affect. To achieve such goal, the Haploview program was used.

Results: The obtained results revealed the region from 5246694 to 5251625 which contains 9323 bases, showing three single-nucleotide polymorphism (SNPs) in the beta globin gene region in chromosome number 11, besides the haplotypes that were appeared in the samples under investigation. This study also showed a significant correlation between SNP2 - SNP3 and between SNP 1 - SNP3, and a negative correlation between SNP1 - SNP2.

Conclusion: This study has used genome-wide association study (GWAS) in understanding the genetic diversity that explains the phenotypic shape of sickle cell disease (SCD) patients in Saudi Arabia. It is therefore important to conduct further studies at a large level in Saudi Arabia to confirm these important results, which will increase current understanding of the SCD’s nature.

Keywords: Genetic origin, Mutation, Sickle cell disease, Single nucleotide, Single phenotype

INTRODUCTION

Sickle cell disease (SCD) is a major health problem according to the World Health Organization, and there are more than half million new cases each year. SCD is featured by several clinical signs such as: stroke, chest pain, abnormal hemoglobin in red blood cells, and recurrent vasoconstriction. Chronic hemolysis usually leads to tuberculosis, delayed growth, anemia, and jaundice. Some SCD patients are sensitive for rheumatism, leg ulcers, pulmonary arterial hemorrhage and hypertension. At the national level, SCD represents about 0.4-8% while SCD carriers ranged from 2% to 27%. Currently, in Saudi Arabia, approximately 3 million individuals had genetic blood diseases in Saudi Arabia, of whom 30% lived in Dammam Region. Several publications showed that there are at least two different types of sickle cell anemia, the average type centralized in Dammam region, while the severe type based in southwestern region.

MATERIALS AND METHODS

Patients: The study was conducted on 90 of sickle cell anemia patients selected randomly from the attending the blood diseases clinic at at two regional hospitals based in Riyadh, Saudi Arabia: The King Khalid University Hospital (KKUH), and The King Faisal Specialist Hospital...
Research Center (KFSHRC), from January 2017 to June 2020. The study protocol was processed according to the Declaration of Helsinki. Written informed consent from all patients were obtained. Ethical approval is obtained from both KKUH, and KFSHRC, Saudi Arabia.

**Samples:** Ten ml of venous blood was withdrawn from each patient and distributed to two tubes (each containing 5 ml) of ethylenediamine tetra acetic acid (EDTA). All samples were processed in the KKUH, pathology department. The Adeva 2120 (Siemens Company) was used for measuring hematological markers and Architect from Abbott Company for measuring biochemical markers.

**Sequencing Analysis of PCR Results Using Biosystems Analyzer:** Sequencing analysis refers to the determination of the nucleotide arrangement (G, A, T, and C) along with the DNA. It is an important and useful process in biotechnology, since it allows discovery of mutations and accurate diagnosis of genetic diseases using the 3730 DNA Analyzer from Applied Biosystems (DNA Sequencing by Capillary Electrophoresis Applied Biosystems Chemistry Guide Second Edition).

**Microarray Technology:** The protocol used in this study was applied as advised by the manufacture. GeneChip-Human Mapping 10K Array and Manual kit – Microarray Affymetrix (California, USA). Haploview, an important software program in bioinformatics, was used to analyze data contained in Hap Map, and to evaluate the quality of genetic data for a particular disease.22,23

**Statistical Analysis:** The data obtained was subjected to a statistical analysis using Window Excel and SPSS v17 statistical tools. ANOVAs tests for multiple comparisons and significant analysis (p < 0.05) were carried out.

**RESULTS**

The study was conducted on 90 patients with sickle cell anemia from the outpatients clinic of blood diseases at three major regional hospitals in Saudi Arabia. A comprehensive database was created including the patient's name, patient history, and biochemical and blood biometrics tests. The patient's data was also studied to determine the disease severity, the frequency of clinical symptoms and pain. Genome-wide association study (GWAS) was also carried out to find genes/genomic regions that contribute to diseases and genetic variation. In this study, the sites of a single-nucleotide polymorphism (SNPs) on SCD patients were examined. For this purpose, one of the most important programs in the field of bioinformatics was the haploview, which provides a comprehensive set of tools for analyzing large areas of the genome in many ways. In this study, genotype was performed using the AffymetrixChip technique. Three SNPs were found in beta region of chromosome 11. The haplotypes presented in the research samples are shown in Table 3 and Figures 2 and 3.

**Clinical and Blood Data in the Research Sample:** Table-1 presents a summary of hematological and biochemical measurements from patients with sickle cell anemia that have been followed for more than 12 months. Table-2 also shows the statistical analysis of different parameters with their occurrence of in certain mutations.

**Table 1:** Hematological and biochemical measurements from SCA patients

<table>
<thead>
<tr>
<th>Parameter</th>
<th>WBC (Mean)</th>
<th>Hemoglobin</th>
<th>Platelets</th>
<th>LDH</th>
<th>Bilirubin</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>White Blood Cells (WBC)</strong></td>
<td>11.5</td>
<td>95</td>
<td>371</td>
<td>440</td>
<td>51</td>
</tr>
<tr>
<td><strong>Std. Deviation</strong></td>
<td>5</td>
<td>23</td>
<td>152</td>
<td>201</td>
<td>32</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mutation</th>
<th>Significance Level (P-Value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High WBC count</td>
<td>Mutated HBBF8_3</td>
<td>0.03</td>
</tr>
<tr>
<td>High Hemoglobin level</td>
<td>HBBF4_3</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>HBBF5_1</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>HBBF8_2</td>
<td>0.02</td>
</tr>
<tr>
<td>High LDH level</td>
<td>Mutated HBBF9_1</td>
<td>0.009</td>
</tr>
<tr>
<td>High the bilirubin level</td>
<td>Mutated HBBF9_1</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Genome-wide association studies (GWAS) are used to find genes or genomic regions that contribute to diseases and genetic variation. In this study, the sites of a single-nucleotide polymorphism (SNPs) on SCD patients were examined. For this purpose, one of the most important programs in the field of bioinformatics was the haploview, which provides a comprehensive set of tools for analyzing large areas of the genome in many ways. In this study, genotype was performed using the AffymetrixChip technique. Three SNPs were found in beta region of chromosome 11. The haplotypes presented in the research samples are shown in Table 3 and Figures 2 and 3.

**Table 3:** Percentage of individual patterns of haplotypes found in the genotype of beta-globin for the samples

<table>
<thead>
<tr>
<th>Percent</th>
<th>Haplotype</th>
<th>HapID</th>
</tr>
</thead>
<tbody>
<tr>
<td>77.1%</td>
<td>AAA</td>
<td>Hap 1</td>
</tr>
<tr>
<td>12.7%</td>
<td>ACA</td>
<td>Hap 2</td>
</tr>
<tr>
<td>6.8%</td>
<td>ACC</td>
<td>Hap 3</td>
</tr>
<tr>
<td>3.4%</td>
<td>CCC</td>
<td>Hap 4</td>
</tr>
</tbody>
</table>

**Figure 1:** The nature of the patients

**Figure 2:** Single-nucleotide polymorphism (SNPs) identified in the beta region
Microarray Technique for Studying Genetic Diversity in Saudi Sickle Cell Patients

**DISCUSSION**

The relationship between disease and genetic variation was examined using GWAS to determine which unaffected genes that contribute to a particular disease. The SNPs are important markers that have been extensively examined in many correlative studies. Many studies have confirmed that these markers control a particular disease with different effects due to different genetic variants.

Important genetic links can be explained by a direct correlation, where SNP is associated with one or several of its multiple aspects. Therefore, the genetic association of the SNPs with the linkage disequilibrium (LD) determines the true causal variables. Another physiological study has confirmed the role of the presumed mutation in a disease.

In this study, three SNPs were identified in the beta-globin generegion of the SCD patients (SNP1: rs11036364), (SNP2: rs 16911905), and (SNP3: rs1003586). This study also showed a significant correlation between SNP2 - SNP3 and between SNP 1 - SNP3, and a negative correlation between SNP1 - SNP2. Previous study showed that SNPs in a single phenotype can be affected by the phenotype of the whole group.

SCD patients with the genotype AA had a higher risk of clinical distress than those with AC or CC pattern. The single nucleotide form detected in this study is different in its location and may be due to the differences in the genetic origin of our samples. A mutation in the beta-globin gene may be associated with a specific individual pattern.

Hattori et al. showed that on 98 SCD studies, 54% had Benin haplotype, while Bantu haplotype existed on 27%. Miller demonstrated that Arabian/Indian haplotype had 13% fetal hemoglobin, while African haplotype had lower HbF. It was implied that HbF had a theoretical importance. The SC mutation exists in Africa with several genetic haplotypes. Four haplotypes have been associated with HbS in Africa and the 5th is in India and/or Arabian Peninsula and are linked with disease severity. These SCD haplotypes in Africa could increase the current understanding of genetic factors that shape the SCD phenotype. Currat et al study showed the exposure of a specific mutation (rs782144 SNP) across west African populations. They concluded that geographical distribution of known SC haplotypes is still not established in many African countries. Nogbri’s group demonstrated that Arab Indian haplotype is one of the major HBB haplotypes showing different clinical and hematological parameters compared to other haplotypes such as: Senegal, Cameroon and Benin.

**CONCLUSION**

In this study, linkage disequilibrium has been determined between single nucleotides polymorphism loci in the beta globin gene to identify which of them had a link to the disease appearance and affect. Promising data was obtained, however, larger and generalized samples from different locations from Saudi Arabia and Arab world are needed to confirm these results.

**Authorship Contribution:** All authors share equal effort contribution towards (1) substantial contributions to conception and design, acquisition, analysis and interpretation of data; (2) drafting the article and revising it critically for important intellectual content; and (3) final approval of the manuscript version to be published. Yes.

**Potential Conflict of Interest:** None.

**Acceptance Interest:** None.

**Funding:** This work was supported by the PSAU (grant no:2019/03/10211)

**REFERENCES**