Comprehensive Next Generation Sequencing Based Genomic Analysis of The *HBB* Locus

Background: Hereditary diseases are common in Saudi Arabia. Such diseases include Sickle Cell Anemia and Thalassemia. Recent epidemiological studies report an alarming increase in the prevalence of hereditary blood disorders in Saudi Arabia. The majority of these positive results were in carriers of sickle-cell disease and β-thalassemia. Most of these diseases causing variants are located on chromosome 11 (p15.4) in the *HBB* gene.

Aim: We aim to tackle the issue of haemoglobinopathy genetic testing by identifying the most frequent and rare pathogenic variant mutations in the Saudi population utilizing next generation sequencing (NGS).

Methods: To identify the rare thalassemia variants, we have collected blood samples and extracted DNA of 183 transfusion dependent beta thalassemia patients. We conducted TaqMan genotyping tests of 6 frequent variants; NM_000518.4:c.92+1G>A, c.92+5G>C, c.93-21G>A, c.118C>T, c.20A>T and c.27dupG. After that, NGS was performed on samples with negative/heterozygous results for these variants. In addition, some samples were Sanger sequenced to validate NGS results and eliminate any possible false negative outcome.

Results: After initially genotyping 183 DNA samples with the 6 most frequent beta thalassemia mutations, 67 of these samples were sequenced using Ion Torrent next generation sequencing and 12 additional common and rare variants were identified, 8 of the 67 sequenced samples did not have any apparent mutation and failed to justify the disease designation. Sanger sequencing of 13 samples was performed for variant confirmation.

Conclusion: TaqMan assay genotyping is a fast, reliable and cost effective technique. Furthermore, TaqMan genotyping followed by NGS has successfully identified and characterized both common and rare pathogenic variations found in *HBB*. However, further studies are necessary to develop a unified approach towards SCD and β-thalassemia treatment applicable to all patients affected by these disorders utilizing currently available new technologies in gene editing.