World Academy of Science, Engineering and Technology
International Journal of Medical and Health Sciences
Vol:15, No:07, 2021

Identification of Candidate Congenital Heart Defects Biomarkers by Applying a Random Forest Approach on DNA Methylation Data

Authors : Kan Yu, Khui Hung Lee, Eben Afrifa-Yamoah, Jing Guo, Katrina Harrison, Jack Goldblatt, Nicholas Pachter, Jitian Xiao, Guicheng Brad Zhang

Abstract : Background and Significance of the Study: Congenital Heart Defects (CHDs) are the most common malformation at birth and one of the leading causes of infant death. Although the exact etiology remains a significant challenge, epigenetic modifications, such as DNA methylation, are thought to contribute to the pathogenesis of congenital heart defects. At present, no existing DNA methylation biomarkers are used for early detection of CHDs. The existing CHD diagnostic techniques are time-consuming and costly and can only be used to diagnose CHDs after an infant was born. The present study employed a machine learning technique to analyse genome-wide methylation data in children with and without CHDs with the aim to find methylation biomarkers for CHDs. Methods: The Illumina Human Methylation EPIC BeadChip was used to screen the genome-wide DNA methylation profiles of 24 infants diagnosed with congenital heart defects and 24 healthy infants without congenital heart defects. Primary pre-processing was conducted by using RnBeads and limma packages. The methylation levels of top 600 genes with the lowest p-value were selected and further investigated by using a random forest approach. ROC curves were used to analyse the sensitivity and specificity of each biomarker in both training and test sample sets. The functionalities of selected genes with high sensitivity and specificity were then assessed in molecular processes. Major Findings of the Study: Three genes (MIR663, FGF3, and FAM64A) were identified from both training and validating data by random forests with an average sensitivity and specificity of 85% and 95%. GO analyses for the top 600 genes showed that these putative differentially methylated genes were primarily associated with regulation of lipid metabolic process, protein-containing complex localization, and Notch signalling pathway. The present findings highlight that aberrant DNA methylation may play a significant role in the pathogenesis of congenital heart defects.

Keywords : biomarker, congenital heart defects, DNA methylation, random forest
Conference Title : ICGPCHD 2021 : International Conference on General Pediatrics and Congenital Heart Diseases
Conference Location : Berlin, Germany
Conference Dates : July 22-23, 2021