World Academy of Science, Engineering and Technology
International Journal of Mathematical and Computational Sciences
Vol:14, No:12, 2020

Whole Exome Sequencing Data Analysis of Rare Diseases: Non-Coding Variants and Copy Number Variations

Authors : S. Fahiminiya, J. Nadaf, F. Rauch, L. Jerome-Majewska, J. Majewski

Abstract : Background: Sequencing of protein coding regions of human genome (Whole Exome Sequencing; WES), has demonstrated a great success in the identification of causal mutations for several rare genetic disorders in human. Generally, most of WES studies have focused on rare variants in coding exons and splicing-sites where missense substitutions lead to the alternation of protein product. Although focusing on this category of variants has revealed the mystery behind many inherited genetic diseases in recent years, a subset of them remained still inconclusive. Here, we present the result of our WES studies where analyzing only rare variants in coding regions was not conclusive but further investigation revealed the involvement of non-coding variants and copy number variations (CNV) in etiology of the diseases. Methods: Whole exome sequencing was performed using our standard protocols at Genome Quebec Innovation Center, Montreal, Canada. All bioinformatics analyses were done using in-house WES pipeline. Results: To date, we successfully identified several disease causing mutations within gene coding regions (e.g. SCARF2: Van den Ende-Gupta syndrome and SNAP29: 22q11.2 deletion syndrome) by using WES. In addition, we showed that variants in non-coding regions and CNV have also important value and should not be ignored and/or filtered out along the way of bioinformatics analysis on WES data. For instance, in patients with osteogenesis imperfecta type V and in patients with glucocorticoid deficiency, we identified variants in 5'UTR, resulting in the production of longer or truncating non-functional proteins. Furthermore, CNVs were identified as the main cause of the diseases in patients with metaphyseal dysplasia with maxillary hypoplasia and brachydactyly and in patients with osteogenesis imperfecta type VII. Conclusions: Our study highlights the importance of considering non-coding variants and CNVs during interpretation of WES data, as they can be the only cause of disease under investigation.

Keywords : whole exome sequencing data, non-coding variants, copy number variations, rare diseases
Conference Title : ICSRD 2020 : International Conference on Scientific Research and Development
Conference Location : Chicago, United States
Conference Dates : December 12-13, 2020