World Academy of Science, Engineering and Technology
International Journal of Biotechnology and Bioengineering
Vol:10, No:04, 2016

Isolate-Specific Variations among Clinical Isolates of Brucella Identified by Whole-Genome Sequencing, Bioinformatics and Comparative Genomics

Authors : Abu S. Mustafa, Mohammad W. Khan, Faraz Shaheed Khan, Nazima Habibi

Abstract : Brucellosis is a zoonotic disease of worldwide prevalence. There are at least four species and several strains of Brucella that cause human disease. Brucella genomes have very limited variation across strains, which hinder strain identification using classical molecular techniques, including PCR and 16 S rDNA sequencing. The aim of this study was to perform whole genome sequencing of clinical isolates of Brucella and perform bioinformatics and comparative genomics analyses to determine the existence of genetic differences across the isolates of a single Brucella species and strain. The draft sequence data were generated from 15 clinical isolates of Brucella melitensis (biovar 2 strain 63/9) using MiSeq next generation sequencing platform. The generated reads were used for further assembly and analysis. All the analysis was performed using Bioinformatics work station (8 core i7 processor, 8GB RAM with Bio-Linux operating system). FastQC was used to determine the quality of reads and low quality reads were trimmed or eliminated using Fastx_trimmer. Assembly was done by using Velvet and ABySS softwares. The ordering of assembled contigs was performed by Mauve. An online server RAST was employed to annotate the contigs assembly. Annotated genomes were compared using Mauve and ACT tools. The QC score for DNA sequence data, generated by MiSeq, was higher than 30 for 80% of reads with more than 100x coverage, which suggested that data could be utilized for further analysis. However when analyzed by FastQC, quality of four reads was not good enough for creating a complete genome draft so remaining 11 samples were used for further analysis. The comparative genome analyses showed that despite sharing same gene sets, single nucleotide polymorphisms and insertions/deletions existed across different genomes, which provided a variable extent of diversity to these bacteria. In conclusion, the next generation sequencing, bioinformatics, and comparative genome analysis can be utilized to find variations (point mutations, insertions and deletions) across different genomes of Brucella within a single strain. This information could be useful in surveillance and epidemiological studies supported by Kuwait University Research Sector grants MI04/15 and SRUL02/13.

Keywords : brucella, bioinformatics, comparative genomics, whole genome sequencing
Conference Title : ICBBBCB 2016 : International Conference on Bioinformatics, Biomedicine, Biotechnology and Computational Biology
Conference Location : Boston, United States
Conference Dates : April 25-26, 2016