

Genome Sequencing, Assembly and Annotation of *Gelidium Pristoides* from Kenton-on-Sea, South Africa

Authors : Sandisiwe Mangali, Graeme Bradley

Abstract : Genome is complete set of the organism's hereditary information encoded as either deoxyribonucleic acid or ribonucleic acid in most viruses. The three different types of genomes are nuclear, mitochondrial and the plastid genome and their sequences which are uncovered by genome sequencing are known as an archive for all genetic information and enable researchers to understand the composition of a genome, regulation of gene expression and also provide information on how the whole genome works. These sequences enable researchers to explore the population structure, genetic variations, and recent demographic events in threatened species. Particularly, genome sequencing refers to a process of figuring out the exact arrangement of the basic nucleotide bases of a genome and the process through which all the afore-mentioned genomes are sequenced is referred to as whole or complete genome sequencing. *Gelidium pristoides* is South African endemic Rhodophyta species which has been harvested in the Eastern Cape since the 1950s for its high economic value which is one motivation for its sequencing. Its endemism further motivates its sequencing for conservation biology as endemic species are more vulnerable to anthropogenic activities endangering a species. As sequencing, mapping and annotating the *Gelidium pristoides* genome is the aim of this study. To accomplish this aim, the genomic DNA was extracted and quantified using the Nucleospin Plank Kit, Qubit 2.0 and Nanodrop. Thereafter, the Ion Plus Fragment Library was used for preparation of a 600bp library which was then sequenced through the Ion S5 sequencing platform for two runs. The produced reads were then quality-controlled and assembled through the SPAdes assembler with default parameters and the genome assembly was quality assessed through the QUAST software. From this assembly, the plastid and the mitochondrial genomes were then sampled out using *Gelidiales* organellar genomes as search queries and ordered according to them using the Geneious software. The Qubit and the Nanodrop instruments revealed an A260/A280 and A230/A260 values of 1.81 and 1.52 respectively. A total of 30792074 reads were obtained and produced a total of 94140 contigs with resulted into a sequence length of 217.06 Mbp with N50 value of 3072 bp and GC content of 41.72%. A total length of 179281bp and 25734 bp was obtained for plastid and mitochondrial respectively. Genomic data allows a clear understanding of the genomic constituent of an organism and is valuable as foundation information for studies of individual genes and resolving the evolutionary relationships between organisms including Rhodophytes and other seaweeds.

Keywords : *Gelidium pristoides*, genome, genome sequencing and assembly, Ion S5 sequencing platform

Conference Title : ICPGB 2018 : International Conference on Plant Genomics and Biotechnology

Conference Location : Cape Town, South Africa

Conference Dates : November 15-16, 2018