Abstract

Sequences Signature and Genome Rearrangements in Mitogenomes

Marwa Al Arab

(Dissertation, 2018)

During the last decades, mitochondria and their DNA have become a hot topic of research due to their essential roles which are necessary for cells survival and pathology. In this study, multiple methods have been developed to help with the understanding of mitochondrial DNA and its evolution. These methods tackle two essential problems in this area: the accurate annotation of protein-coding genes and mitochondrial genome rearrangements. Mitochondrial genome sequences are published nowadays with increasing pace, which creates the need for accurate and fast annotation tools that do not require manual intervention. In this work, an automated pipeline for fast de-novo annotation of mitochondrial protein-coding genes is implemented. The pipeline includes methods for enhancing multiple sequence alignment, detecting frameshifts and building protein profiles guided by phylogeny. The methods are tested on animal mitogenomes available in RefSeq, the comparison with reference annotations highlights the high quality of the produced annotations. Furthermore, the frameshift method predicted a large number of frameshifts, many of which were unknown. Additionally, an efficient partially-local alignment method to investigate genomic rearrangements in mitochondrial genomes is presented in this study. The method is novel and introduces a partially-local dynamic programming algorithm on three sequences around the breakpoint region. Unlike the existing methods which study the rearrangement at the genes order level, this method allows to investigate the rearrangement on the molecular level with nucleotides precision. The algorithm is tested on both artificial data and real mitochondrial genomic sequences. Surprisingly, a large fraction of rearrangements involve the duplication of local sequences. Since the implemented approach only requires relatively short parts of genomic sequence around a breakpoint, it should be applicable to non-mitochondrial studies as well.