The Potential of DNA Barcode-Based Delineation Using Seven Putative Candidate Loci of the Plastid Region in Inferring Molecular Diversity of Cowpea at Sub-Species Level.

Abstract

The novelty and suitability of the mitochondrial gene *CO*1 in DNA barcoding as a reliable identification tool in animal species are undisputed. This is attributed to its standardized sequencing segment of the mitochondrial cytochrome c oxidase-1 gene (*CO*1) which has the necessary universality and variability making it a generally acceptable barcode region. *CO*1 is a haploid single locus that is uniparentally-inherited. Protein-coding regions are present in high-copy numbers making it an ideal barcode. The mitochondrial oxidase subunit I (COI) gene is a robust barcode with a suitable threshold for delineating animals and is not subject to drastic length variation, frequent mononucleotide repeats or microinversions. However, a low nucleotide substitution rate of plant mitochondrial genome [mtDNA] precludes the use of *CO*1 as a universal plant DNA barcode and makes the search for alternative barcode regions necessary. Currently, there exists no universal barcode for plants. The plastid region reveals leading candidate loci as appropriate DNA barcodes yet to be explored in biodiversity studies in Kenya. Four of these plastid regions are portions of coding genes (*matK, rbcL, rpoB*, and *rpoC*1), and three noncoding spacers (*atpF-atpH, trnH-psbA*, and *psbK-psbL*) which emerge as ideal candidate DNA loci. While different research groups propose various combinations of these loci, there exists no consensus; the lack thereof impedes progress in getting a suitable universal DNA barcode. Little research has attempted to investigate and document the applicability and extend of effectiveness of different DNA regions as barcodes to delineate cowpea at subspecies level. In this study we sought to test feasibility of the seven putative candidate DNA loci singly and in combination in order to establish a suitable single and multi-locus barcode regions that can have universal application in delineating diverse phylogeographic groups of closely related Kenyan cowpea variants. In this study, our focus was based on genetic parameters including analyses of intra- and infra-specific genetic divergence based on intra- and infra-specific K2P distances; calculation of Wilcoxon signed rank tests of intra-specific divergence among loci and coalescence analyses to delineate independent genetic clusters. Knowledge of DNA candidate loci that are informative will reveal the suitability of DNA barcoding as a tool in biodiversity studies. Results of this study indicate that: *matK, trnH-psbA, psbK-psbL*, and *rbcL* are good barcodes for delineating intra and infraspecific distances at single loci level. However, among the combinations, *matK* + *trnH-psbA*, *rpoB* + *atpF-atpH* + *matK* are the best barcodes in delineating cowpea subvariants. *rbcL* gene can be a suitable barcode marker at single locus level, but overall, multi locus approach appears more informative than single locus approach. The present study hopes to immensely contribute to the scanty body of knowledge on the novelty of DNA barcoding in cataloguing closely related cowpea variants at molecular level and hopes to open up future research on genomics and the possibility of use of conserved regions within DNA in inferring phylogenetic relationships among Kenyan cowpea variants.

Authors.

Patrick Okoth[1], John Muoma[2], Mulaya Emmanuel[3], Wekesa Clabe[4], Dennis O. Omayio[2], Paul O. Angienda[1]