

Bayesian Classification of DNA Barcodes.

Mike Anderson,

Abstract:

DNA barcodes are short strands of nucleotide bases taken from the cytochrome c oxidase subunit 1 (COI) of the mitochondrial DNA (mtDNA). It has been proposed that these barcodes may be used as a method of differentiating between biological species (Hebert, Ratnasingham, and deWaard 2003). Current methods of species classification and clustering utilize distance measures that are heavily dependent on both evolutionary model assumptions as well as a clearly defined "gap" between intra- and interspecies variation (Meyer and Paulay 2005). We point out the limitations of such distance measures and propose a character-based method of species classification which utilizes an application of Bayes' rule to overcome these deficiencies. The proposed method is shown to provide accurate species-level classification and provide answers to important questions not addressable with current methods.