A system for barcode primer retrieval and evaluation

Pappadà G[1], Santamaria M[2], Scioscia G[3], Quinto V[1]

Motivation

The "DNA Barcode" initiative aims to provide a rapid and standardized protocol to identify ideally all the living species belonging to a particular taxonomic domain on the basis of a short DNA fragment taken from a standardized portion of the genome. This simple methodology could be very useful in many practical fields including the industrial one, where such a rapid and cheap species identification system could simplify the monitoring of food quality and traceability. One of the most important steps in building an effective barcode protocol to discriminate most of the species in a wide taxonomic range is the design of universal primers for the PCR amplification of the selected genetic marker. In order to meet this requirement, we propose a new bioinformatic tool able to predict the best primer pairs for the PCR amplification of the barcode region in a given taxonomic group.

Methods

At present, several barcoding species assignments have been performed in a broad range of taxa. The sequences of the primers used in these experiments are often stored in public databases. The primer finder we developed is based on the retrieval of the barcode primers already known and available in the public database GenBank. In details, a GenBank query is performed to extract all the entries matching with the taxonomic group of interest. Then, the query results are restricted to those entries which contain the keyword "BARCODE". These entries often contain sequence data from both barcode region and forward and reverse PCR primers used for its amplification. In the following step, the primer sequences are extracted from all the entries belonging to the searched taxa and analysed in order to obtain the best candidate universal primers pair: forward and reverse primer sequences are separately aligned in order to identify conserved and variable positions. The consensus sequences produced for both forward and reverse primers would represent good candidates for the barcode amplification in all the species belonging to the examined taxonomic group. Finally, the quality of the new primers is evaluated by calculating the physical-chemical properties, which could affect the PCR reaction. Moreover, a test is performed to understand if, beyond the barcode region, the primer pair amplifies other regions too. The figure summarizes the workflow used for this approach.

[1] Exhicon I.C.T. S.r.l., Bari, Via avv. V. Malcangi 254, 70059 Trani, Italy [2] CNR - Istituto di Tecnologie Biomediche, Sede di Bari, Via Amendola 122/D, Bari, 70126, Italy. [3] IBM Italia S.p.A, Sede di Bari, Via P.L. Laforgia 14, 70125 Bari, Italy

Results

The resulting web application, which is equipped with instruments for the analysis of biosequences such as multiple alignment and consensus sequence finder, can support the researcher in obtaining an effective set of barcode primers in a specific taxonomic group whose size can be defined by the user himself. Indeed, he can select any taxonomy level (e.g. genus, family, order, etc.) as a search criterion. If a certain taxonomic group includes several species already barcoded, the developed system will be able to support the design of effective barcode primers in the other species belonging to the same group. The exploitation of primers already annotated increases the robustness of the system but, at the same time, it excludes from the analysis all the primer sequences that have not been annotated. A Barcode Primer Retrieval service prototype is available on the webpage: http://213.26.249.182/services/BarcodePrimers.html

Availability

http://213.26.249.182/services/BarcodePrimers.html

Contact e-mail

g.pappada@exhicon.it

Image