



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

**ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ**

**ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
"ΤΕΧΝΟΛΟΓΙΕΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΣΤΗΝ ΙΑΤΡΙΚΗ ΚΑΙ ΤΗ ΒΙΟΛΟΓΙΑ"
ΒΙΟΠΛΗΡΟΦΟΡΙΚΗ**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Δημιουργία εργαλείου για αναζήτηση πρωτεομικής
ομοιότητας βασισμένη πάνω σε δευτερογενή δομή.**

Δημήτριος Γ. Κοντός

Επιβλέπουσα: Σοφία Κοσσίδα, Ερευνήτρια Γ'

**Αθήνα
Αύγουστος 2014**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Δημιουργία εργαλείου για αναζήτηση πρωτομικής ομοιότητας βασισμένη πάνω σε δευτερογενή δομή.

Δημήτριος Γ. Κοντός

A.M.:ΠΙΒ055

ΕΠΙΒΛΕΠΟΥΣΑ: Σοφία Κοσσίδα, Ερευνήτρια Γ'

**Αθήνα
Αύγουστος 2014**

ΠΕΡΙΛΗΨΗ

Έχουν παρατηρηθεί στη φύση πρωτεΐνες που έχουν την ίδια λειτουργία, ανήκουν στην ίδια οικογένεια αλλά έχουν χαμηλή ταυτότητα αλληλουχίας (λιγότερο από 20%). Το πιο ενδιαφέρον όμως είναι ότι οι δευτεροταγείς δομές τους όταν προσδιορίστηκαν με κρυσταλλογραφία ακτίνων Χ έδειξε ότι μοιράζονται τα ίδια δομικά χαρακτηριστικά. Έτσι προέκυψε η ανάγκη για μια αναζήτηση ομοιότητας πρωτεϊνών σε δευτεροταγείς δομικές πληροφορίες που σε συνδυασμό με την κλασική αναζήτηση BLAST ομοιότητα αλληλουχίας θα μπορούσε να βοηθήσει στην έρευνα για την αναδίπλωση της πρωτεΐνης στο χώρο.

Στη εργασία αυτή αναπτύξαμε ένα web rmi java applet εργαλείο μέσω του οποίου ο τελικός χρήστης έχει τη δυνατότητα να συνδυάσει τα αποτελέσματα στοίχισης της αμινοξικής του ακολουθίας αλλά και τα αποτελέσματα της αντίστοιχης δευτεροταγής της ακολουθίας με γνωστές PDBs ακολουθίες. Για τη δυνατότητα μετατροπής μιας αμινοξικής στην αντίστοιχη δευτεροταγή υπάρχουν πολλοί αλγόριθμοι πρόβλεψης. Στη παρούσα εργασία στηριχθήκαμε στον αλγόριθμο που χρησιμοποιεί το Strap εργαλείο.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Δευτεροταγής στοίχιση πρωτεϊνών

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: blast, sfs, dssp.

ABSTRACT

Herein, the novel Protein Secondary Structure Profile suite has been developed in order to perform protein similarity searches that do not rely solely on primary sequence rather than secondary structural information. There are large databases available nowadays, which contain huge structural and sequence information. However, in many cases the primary sequence comparisons do not cover similarity searches adequately, which makes the need for a new structure-based similarity search approach highly critical. Provided that the function of each biological molecule is dictated directly by its structure, it may be seen as a universal law that a protein's secondary structure is more conserved than its primary sequence of aminoacids. There are numerous cases of extremely low sequence identity (less than 20%) between homologous viral proteins, which however have the same function, belong to the same viral family and, more strikingly, when their structures were determined by X-ray crystallography, it was evident that they shared the same structural features as well. Taken together, we propose that our novel structural similarities approach presented here, in conjunction with classic sequence similarity BLAST searches would yield results of great interest to science.

In this work we developed a web rmi java applet tool through which the end user has the ability to combine the results of the alignment of the amino acid sequence but also lose the corresponding secondary sequence with sequences of known PDBs. For the conversion of an amino acid corresponding respectively to the secondary are many prediction algorithms. In the present study, we relied on the algorithm that uses the tool Strap

SUBJECT AREA: Secondary structure alignment.

KEYWORDS: blast, sfs, dssp.