



NATIONAL AND KAPODISTRIAN UNIVERSITY OF ATHENS

**SCHOOL OF SCIENCE
FACULTY OF INFORMATICS AND TELECOMMUNICATIONS**

**INTERDISCIPLINARY POSTGRADUATE PROGRAMME
"INFORMATION TECHNOLOGIES IN MEDICINE AND BIOLOGY"**

POSTGRADUATE THESIS

**Comparative Study of Target Prediction Algorithms, utilizing
artificial and experimental data on Kaposi's sarcoma-
associated herpes & Epstein–Barr virus**

Sotiris L. Konstantinidis

SUPERVISOR: Artemis Hatzigeorgiou, Researcher

ATHENS

MARCH 2014



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

**ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ**

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

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Συγκριτική Μελέτη Αλγορίθμων Πρόβλεψης Στόχων, με την
χρήση τεχνητών και πειραματικών δεδομένων στους ιούς
Kaposi's sarcoma-associated herpes και Epstein Barr**

Σωτήρης Α. Κωνσταντινίδης

Επιβλέποντες: Άρτεμις Χατζηγεωργίου, Ερευνήτρια

ΑΘΗΝΑ

ΜΑΡΤΙΟΣ 2014

POSTGRADUATE THESIS

Comparative Study of Target Prediction Algorithms, utilizing artificial and experimental data on Kaposi's sarcoma-associated herpes & Epstein–Barr virus

Sotiris L. Konstantinidis

A.M.: 063

SUPERVISOR: Artemis Hatzigeorgiou, Researcher

THESIS COMMITTEE: Artemis Hatzigeorgiou, Researcher
Vangelis Karkaletsis, Senior Researcher
Ilias Maglogiannis, Assistant Professor

MARCH 2014

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Συγκριτική Μελέτη Αλγορίθμων Πρόβλεψης Στόχων, με την χρήση τεχνητών και πειραματικών δεδομένων στους ιούς Kaposi's sarcoma-associated herpes και Epstein Barr

Σωτήρης Λ. Κωνσταντινίδης

A.M.: 063

ΕΠΙΒΛΕΠΟΥΣΑ: Άρτεμις Χατζηγεωργίου, Ερευνήτρια

ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ: Άρτεμις Χατζηγεωργίου, Ερευνήτρια
Βαγγέλης Καρκαλέτσης, Α' Ερευνητής
Ηλίας Μαγκλογιάννης, Επίκουρος Καθηγητής

MΑΡΤΙΟΣ 2014

ABSTRACT

miRNAs are small non-coding molecules 20-22 nt long. These molecules are expressed endogenously and their role is to regulate gene expression. Furthermore miRNA's are not only expressed through the DNA of the human organism, but also viruses express their miRNA's. To this end, a major challenge is the development and use of reliable computational tools in order to support the ongoing research of life scientists, in deciphering the exact mechanisms of miRNAs gene regulation and how this affects the human organism functionality as well as the occurrence of severe diseases like cancer. The goal of this work is firstly to choose the appropriate biology experiment regarding the interactions between viral miRNAs and human genes, in order to use it as a test set. Secondly, is to use this test set in order to evaluate several target prediction algorithms and finally to use a reliable way to compare the outcome of these algorithms.

SUBJECT AREA: Bioinformatics

KEYWORDS: computational approaches, machine learning, microRNA, target prediction, gene regulation

ΠΕΡΙΛΗΨΗ

Τα μικρά RNA είναι μικρού μήκους μη-κωδικά μόρια, μεγέθους 20-22 νουκλεοτιδίων. Τα μόρια αυτά εκφράζονται ενδογενώς και ο ρόλος τους είναι να ρυθμίζουν την γονιδιακή έκφραση. Ακόμα περισσότερο μικρά RNA δεν εκφράζονται μόνο από το ανθρώπινο DNA αλλά επίσης ξένοι οργανισμοί εκφράζουν τα δικά τους μικρά RNA. Με δεδομένα αυτά μια μεγάλη πρόκληση είναι η ανάπτυξη και χρήση αξιόπιστων υπολογιστικών εργαλείων, ώστε να υποστηρίξουν την τρέχουσα έρευνα των επιστημόνων υγείας στην προσπάθεια τους να αποκωδικοποιήσουν τον ακριβή ρόλο των μικρών RNA και πως αυτός επηρεάζει την λειτουργία του ανθρώπινου οργανισμού, όπως επίσης και την εμφάνιση σοβαρών ασθενειών. Ο στόχος της εργασίας αυτής είναι αρχικά να δημιουργηθεί ένα σύνολο δεδομένων έλεγχου από την βιβλιογραφία. Εν συνεχεία, στόχος είναι να εξεταστούν αλγόριθμοι πρόβλεψης στόχων σε σχέση με το σύνολο δεδομένων ελέγχου.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Βιοπληροφορική

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: υπολογιστικά εργαλεία, μηχανική μάθηση, πρόβλεψη στόχων, μικρά RNA, ρύθμιση γονιδίων