

ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ Η/Υ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ ΚΑΙ ΔΙΚΤΥΩΝ

Υλοποίηση online εφαρμογής που θα επιτρέπει εξελιγμένες αναζητήσεις σε βάση δεδομένων

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

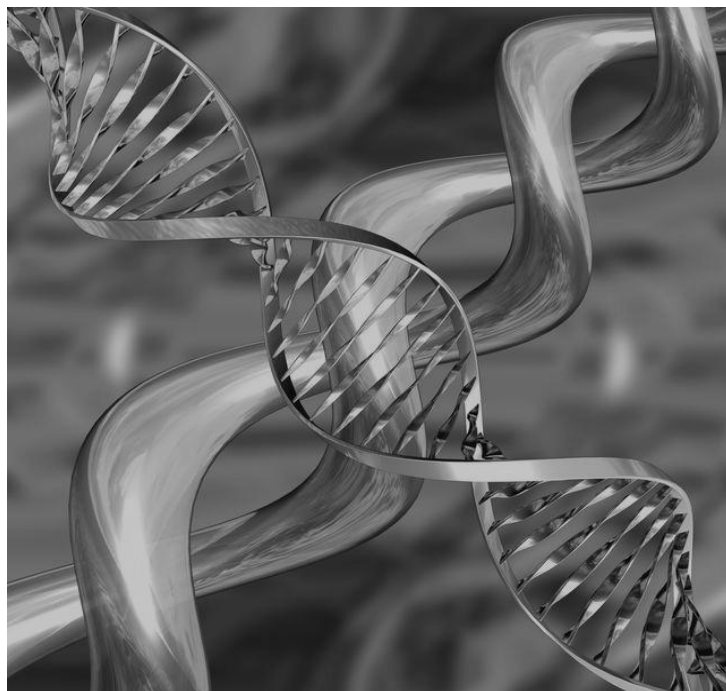
ΤΣΙΡΟΓΙΑΝΝΗ ΚΟΥΛΟΥΜΠΡΙΔΟΥ
ΙΩΑΝΝΑ

Επιβλέποντες καθηγητές:

Χούστης Ηλίας
Καθηγητής Π.Θ.

Χατζηγεωργίου Άρτεμις
Συμβασιούχος 407

Βόλος, Σεπτέμβρης 2011



Στην οικογένεια μου και στους φίλους μου

Ευχαριστίες

Με την περάτωση αυτής της διπλωματικής εργασίας, θα ήθελα να ευχαριστήσω τους επιβλέποντες καθηγητές της εργασίας κ. Χατζηγεωργίου Άρτεμις και κ. Χούστη Ηλία για την εμπιστοσύνη που έδειξαν στο πρόσωπό μου, τον κ. Αλεξίου Παναγιώτη για την πολύτιμη βοήθειά του στην υλοποίησή της παρούσας διπλωματικής, τις εύστοχες υποδείξεις και τις ουσιαστικές παρεμβάσεις του, καθώς και για την αποτελεσματική συνεργασία μας για την ολοκλήρωση της συγκεκριμένης εργασίας.

Επιπλέον, θα ήθελα να ευχαριστήσω ιδιαίτερα την οικογένεια μου και τους φίλους μου για την υποστήριξη τους και τη βοήθειά τους, καθώς βρίσκονταν πάντα δίπλα μου τόσο κατά τη διάρκεια των σπουδών μου όσο και κατά την εκπόνηση της παρούσας διπλωματικής εργασίας.

Τσιρογιάννη Κουλουμπρίδου Ιωάννα

Βόλος, 2011

Περίληψη

Σκοπός της διπλωματικής εργασίας ήταν η ανάπτυξη μιας εφαρμογής για την αναζήτηση και την παρουσίαση βιολογικών στόχων. Η εφαρμογή αυτή αφορά στην έρευνα γύρω από τα μόρια microRNA των οποίων η μελέτη διαδραματίζει πρωταγωνιστικό ρόλο στα πλαίσια της επιστημονικής κοινότητας. Συγκεκριμένα, οι ερευνητές - βιολόγοι ενδιαφέρονται για τη γνώση των βιολογικών στόχων των συγκεκριμένων μορίων πάνω στα μόρια που κωδικοποιούν τις πρωτεΐνες, ούτως ώστε να εξάγουν χρήσιμα συμπεράσματα τα οποία μπορούν να δώσουν λύσεις σε διάφορες παθήσεις. Η εκπόνηση της διπλωματικής πραγματοποιήθηκε σε συνεργασία με το Ερευνητικό Κέντρο Βιοϊατρικών Επιστημών «Αλέξανδρος Φλέμινγκ», το οποίο έχει αναπτύξει κατάλληλο αλγόριθμο για την πρόβλεψη των συγκεκριμένων στόχων. Η ανάπτυξη της εφαρμογής ακολούθησε σύγχρονες τεχνικές και μεθοδολογίες, όπως είναι η χρήση του αρχιτεκτονικού προτύπου Model - View - Controller .

Abstract

The purpose of this thesis was to develop an application for searching and presentation of biological targets. This application relates to research on microRNA molecules whose design plays a leading role within the scientific community. Specifically, the researchers - biologists interested in the knowledge of the biological targets of these molecules on the molecules that encode proteins in order to extract useful conclusions that can provide solutions to various diseases. The development of diploma held in collaboration with the Biomedical Sciences Research Center "Alexander Fleming", which has developed an appropriate algorithm for the prediction of specific targets. The application development has followed modern techniques and methodologies, such as the use of architectural pattern Model - View - Controller.

Πίνακας περιεχομένων

Ευχαριστίες.....	3
Περίληψη.....	4
Abstract	5
Πίνακας Εικόνων	8
1. Εισαγωγή	9
1.1 MicroRNA	9
2. DIANA LAB	11
2.1. DIANA-microT 3.0.....	11
2.2. Tarbase	12
2.3. MiRGen 2.0.....	12
2.4. DIANA-pathways.....	13
3. Βάση Δεδομένων.....	14
4. Αρχιτεκτονική	17
4.1. Model	18
4.2. View.....	18
4.3. Controller.....	19
5. Yii PHP Framework	20
5.1. Ταχύτητα	20
5.2. Κώδικας-Ποιότητα.....	20
5.3. Εισαγωγή δεδομένων από HTML φόρμες.....	20
5.4. Διαχείριση σφαλμάτων και καταγραφή συμβάντων.....	21
5.5. Τεκμηρίωση	21
5.6. Υποστήριξη	21
6. Υλοποίηση	22
6.1. Εισαγωγή στοιχείων	24
6.2. Αναζήτηση βιολογικών στόχων.....	24
6.3. Εμφάνιση αποτελεσμάτων.....	24
6.4. Η βασική ιδέα.....	24
6.5. Σκέψεις πριν την υλοποίηση... ..	25
6.6. Γλώσσες Διαδικτυακού Προγραμματισμού	29
6.7. Ανάλυση του Κώδικα.....	30
6.7.1. Περίπτωση Target	31

6.7.2. Περίπτωση Gene	38
6.7.3. Περίπτωση MatureMirna	45
7. Παραδείγματα τρεξίματος του κώδικα... ..	48
7.1. Επιλογή Target.....	48
7.2. Επιλογή Gene	52
7.3.Επιλογή Mature Mirna	55
8. Επίλογος	60
8.1. Σύνοψη και συμπεράσματα	60
8.2 Μελλοντικές Επεκτάσεις	61
9. Βιβλιογραφία.....	62

Πίνακας Εικόνων

Εικόνα 1: Ιστοσελίδα του Diana Lab (Πηγή: http://diana.cslab.ece.ntua.gr/)	11
Εικόνα 2: Συσχέτιση πινάκων στη βάση δεδομένων mirna2	15
Εικόνα 3: Μοντέλο συσχέτισης MODEL-VIEW-CONTROLLER (Πηγή: http://upload.wikimedia.org/wikipedia/commons/b/b5/ModelViewControllerDiagram2.svg)	18
Εικόνα 4: Διάγραμμα Ροής Δεδομένων	23
Εικόνα 5: Επίπεδο 1ο / Επιλογή	25
Εικόνα 6: Επιλογή Target	26
Εικόνα 7: Επιλογή Gene	27
Εικόνα 8: Επιλογή Mature Mirna	28
Εικόνα 9: Κώδικας σελίδας search.php	30
Εικόνα 10: Κώδικας συνάρτησης ActionSearch2 του Site Controller	31
Εικόνα 11: Κώδικας σελίδας search2.php	32
Εικόνα 12: Κώδικας συνάρτησης ActionResult του Site Controller	34
Εικόνα 13: Κώδικας για το query για την επιλογή Target	36
Εικόνα 14: Κώδικας της σελίδας results.php	38
Εικόνα 15: Κώδικας της σελίδας search2.php	39
Εικόνα 16: Κώδικας της συνάρτησης ActionResult1 του Site Controller	41
Εικόνα 17: Κώδικας για το query για την επιλογή Gene	42
Εικόνα 18: Κώδικας για τη σελίδα results1.php	44
Εικόνα 19: Κώδικας της σελίδας search3.php	45
Εικόνα 20: Κώδικας της συνάρτησης ActionResult2 του Site Controller	46
Εικόνα 21: Επιλογή Target	48
Εικόνα 22: Εισαγωγή δεδομένων για επιλογή Target	49
Εικόνα 23: Αποτελέσματα για επιλογή Target	50
Εικόνα 24: Αποτελέσματα για επιλογή Target	51
Εικόνα 25: Επιλογή Gene	52
Εικόνα 26: Εισαγωγή δεδομένων για επιλογή Gene	52
Εικόνα 27: Αποτελέσματα για την επιλογή Gene	53
Εικόνα 28: Αποτελέσματα για την επιλογή Gene	54
Εικόνα 29: Επιλογή Mature miRNA	55
Εικόνα 30: Εισαγωγή δεδομένων για επιλογή Mature miRNA	56
Εικόνα 31: Κώδικας query για επιλογή Mature miRNA	57
Εικόνα 32: Εισαγωγή δεδομένων για επιλογή Mature miRNA	58
Εικόνα 33: Αποτελέσματα για την επιλογή Mature miRNA	59

1. Εισαγωγή

1.1 MicroRNA

Τα microRNA(miRNA) είναι στενά συνδεδεμένα με τον άνθρωπο και τις λειτουργίες του ανθρώπινου οργανισμού. Το πόσο σημαντικά είναι μπορεί αρχικά να γίνει αντιληπτό από το γεγονός ότι στο ανθρώπινο γονιδίωμα κωδικοποιούνται περισσότερα από 1000 microRNA. Τι πραγματικά είναι όμως αυτά τα πολύ μικρά μόρια;

Τα microRNA είναι μικρά μόρια ριβονουκλεϊκού οξέος(μικρό-ριβονουκλεϊκό οξύ) που κατά μέσο όρο αποτελούνται από 22 νουκλεοτίδια και συναντώνται σε όλα τα ευκαρυωτικά κύτταρα. Είναι μετα-μεταγραφικοί ρυθμιστές που δένουν στις συμπληρωματικές ακολουθίες των αγγελιαφόρων RNA(mRNA messenger RNA) και πιστεύεται ότι ελέγχουν την γονιδιακή έκφραση. Αυτή η διαδικασία έχει ως αποτέλεσμα τη μεταγραφική καταστολή(translational repression) ή/και την σίγηση των γονιδίων(gene silencing).

Οι ιδιότητες των microRNA εμφανίζουν μεγάλες διαφορές ανάμεσα στους φυτικούς και τους ζωικούς οργανισμούς. Πιο συγκεκριμένα, στα φυτά η συμπληρωματικότητα των miRNA με τους στόχους τους που βρίσκονται στο αγγελιαφόρο RNA είναι σχεδόν τέλεια. Αντίθετα, στους ζωικούς οργανισμούς η συμπληρωματικότητα αυτή απέχει αρκετά από το τέλειο. Ένα microRNA μπορεί συνήθως να στοχεύσει σε πολλές διαφορετικές περιοχές πάνω στο ίδιο αγγελιαφόρο RNA ή και σε διαφορετικά αγγελιαφόρα RNA. Μια ακόμα διαφορά είναι το πού βρίσκονται αυτές οι περιοχές στόχοι στα αγγελιαφόρα RNA. Στα ζώα η περιοχή αυτή βρίσκεται συνήθως στα πρώτα τρία μη-μεταφρασμένα κομμάτια του messenger RNA (3'UTR). Από την άλλη πλευρά, στα φυτά οι στόχοι μπορεί να είναι σε αυτή την περιοχή (3'UTR), αλλά συνήθως βρίσκονται στην ίδια την κωδική περιοχή (coding sequence, CDS).

Η ανακάλυψη των microRNA είναι μια από τις πιο σημαντικές που έχουν γίνει στο τομέα της Βιολογίας τα τελευταία χρόνια. Αυτό δεν είναι καθόλου τυχαίο αν αναλογιστούμε τη σπουδαιότητα που παρουσιάζουν τα συγκεκριμένα μόρια σε πολλούς τομείς της ανθρώπινης υγείας. Χαρακτηριστικό παράδειγμα αποτελεί η ανακάλυψη μιας ομάδας ερευνητών από την Νέα Υόρκη ότι τα microRNA είναι στενά συνδεδεμένα με την μετάσταση του καρκίνου του μαστού. Συγκεκριμένα, ομάδα ερευνητών από το Αντικαρκινικό Κέντρο Memorial Sloan-Kettering της Νέας Υόρκης, εντόπισε αρκετά φυσικά ανθρώπινα microRNA που εμποδίζουν τον καρκίνο

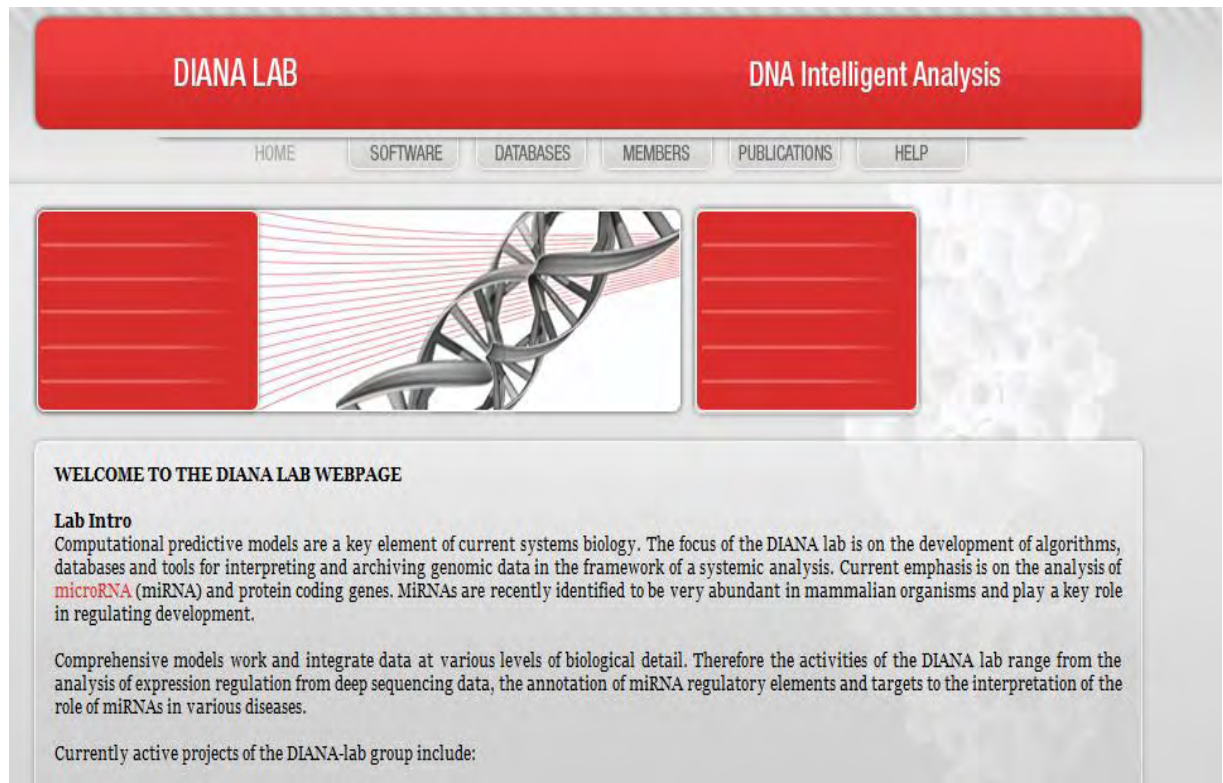
του μαστού να δώσει μεταστάσεις στον πνεύμονα και τα οστά, τα δύο βασικά σημεία μετάστασης. Ακόμα μια ερευνητική ομάδα με επιστήμονες από το πανεπιστήμιο του Τέξας MD Anderson Cancer Center ανέφερε ότι ένα μικρό κομμάτι RNA εμποδίζει τη μετάσταση του καρκίνου του προστάτη.

Διαπιστώνουμε λοιπόν ότι τα microRNA παίζουν ένα βασικό ρόλο ως ρυθμιστές πολλών κυτταρικών λειτουργιών μέσω της μέτα-μεταγραφικής καταστολής κωδικών γονιδίων. Η συνεχής αύξηση του ενδιαφέροντος της επιστημονικής κοινότητας τα τελευταία χρόνια για το ρόλο των microRNA έχει οδηγήσει στην ανάγκη σχεδίασης και ανάπτυξης τραπεζών δεδομένων microRNA και υπολογιστικών προγραμμάτων για την ανάλυση της λειτουργίας τους. Οι στόχοι της προσπάθειας αυτής είναι: (α) συνολικός λειτουργικός χαρακτηρισμός των miRNA και διαθεσιμότητα όλων των πληροφοριών μέσω μίας τράπεζας δεδομένων με διαχρονικό χαρακτήρα, (β) συσχετισμός της λειτουργίας των miRNA με ασθένειες μέσω μιας καινοτόμας οντολογίας και δημιουργία ενός υπολογιστικού εργαστηρίου για την ανάλυση λειτουργίας των miRNA, (γ) συσχετισμός πολυμορφισμών με στόχους και miRNA και πειραματική επαλήθευση.

Το site του DIANA LAB αποτελεί έναν ιστοχώρο που στοχεύει στην επίτευξη όλων των παραπάνω στόχων καθιστώντας εύκολη τη πρόσβαση των χρηστών σε μια μεγάλη γκάμα δεδομένων για τα microRNA. Σε αυτό το site μπορούμε να βρούμε ένα μεγάλο αριθμό προγραμμάτων για την αναζήτηση συγκεκριμένων πληροφοριών για τα microRNA. Τέτοια προγράμματα είναι τα Tarbase, DIANA-microT 3, miRGen 2 και DIANA-pathways.

2. DIANA LAB

Όπως αναφέραμε παραπάνω το site του DIANA LAB αποτελεί έναν από τους γνωστότερους ιστοχώρους στον τομέα της βιοπληροφορικής. Επικεντρώνεται κατά κύριο λόγο στην ανάπτυξη αλγορίθμων, βάσεων δεδομένων και εργαλείων που βοηθούν στην αρχειοθέτηση και στην ερμηνεία της γονιδιακής έκφρασης.



Εικόνα 1: Ιστοσελίδα του Diana Lab (Πηγή: <http://diana.cslab.ece.ntua.gr/>)

2.1. DIANA-microT 3.0

Ο DIANA-microT 3.0 είναι ένας αλγόριθμος που βασίζεται σε διάφορες παραμέτρους και αποσκοπεί στην πρόβλεψη γονιδιακών στόχων των microRNA. Πιο αναλυτικά, η εφαρμογή DIANA-microT 3.0 (Maragkakis, Alexiou et al., 2009) είναι ένα πρόγραμμα πρόβλεψης γονιδιακών στόχων το οποίο πέτυχε το μέγιστο precision στην αναγνώριση στόχων miRNA από πειράματα high-throughput στα οποία συγκρίθηκαν όλα τα ευρέως γνωστά αντίστοιχα προγράμματα. Ο web server του DIANA-microT 3.0

(Maragkakis, Reczko et al., 2009 Nucleic Acids Res) βασίζεται σε ένα περιβάλλον φιλικό προς τον χρήστη με συνδέσμους προς διάφορες βάσεις βιολογικών δεδομένων, οι οποίοι επιτρέπουν βαθύτερη κατανόηση των αποτελεσμάτων. Η αναζήτηση για αυτούς τους γονιδιακούς στόχους μπορεί να γίνει είτε με ένα ήδη γνωστό είτε με ένα νέο microRNA.

2.2. Tarbase

Η Tarbase είναι μια βάση δεδομένων που φιλοξενεί μία μεγάλη συλλογή στόχων των microRNA. Οι στόχοι αυτοί είναι πειραματικά αποδεδειγμένοι και αφορούν διάφορα είδη ζώων, φυτών και ιών που συγκεντρώνουν το ενδιαφέρον της επιστημονικής κοινότητας. Αν και υπάρχουν πολλά υπολογιστικά προγράμματα που μπορούν να προβλέψουν στόχους των microRNA θεωρήθηκε αναγκαίο να δημιουργηθεί μια βάση δεδομένων στην οποία θα αποθηκεύονται οι ήδη πειραματικά αποδεδειγμένοι στόχοι. Η τωρινή έκδοση της Tarbase περιέχει περισσότερους από 1300 στόχους και καθένας από αυτούς περιγράφεται από το microRNA το οποίο δένει πάνω του, το γονίδιο στο οποίο βρίσκεται, τα πειράματα τα οποία έγιναν για να ανακαλυφθεί, την ικανότητα του συγκεκριμένου στόχου να προκαλέσει μεταγραφική καταστολή και το paper από το οποίο πάρθηκε όλη αυτή η πληροφορία. Η συγκεκριμένη βάση ενώνεται λειτουργικά και με άλλες παρόμοιες και χρήσιμες βάσεις δεδομένων όπως οι Ensembl, Hugo, UCSC και SwissProt.

2.3. MiRGen 2.0

Στα τέλη του 2009 αναπτύχθηκε η εφαρμογή miRGen 2.0 (Alexiou, Vergoulis et al, 2009) η οποία χαρτογραφεί τις περίπλοκες αλληλεπιδράσεις των microRNA με μεταγραφικούς παράγοντες (Transcription Factors) που ρυθμίζουν την έκφραση των μεταγράφων microRNA. Επίσης περιέχονται χάρτες έκφρασης (expression maps) πολλών microRNA και πληροφορίες σχετικά με τα γονιδιωματικά χαρακτηριστικά των μεταγράφων των microRNA.

Η MiRGen2 είναι ουσιαστικά μια εφαρμογή, εύκολα προσβάσιμη και φιλική προς τον χρήστη που μας παρέχει πρόσβαση σε μια ολοκληρωμένη βάση δεδομένων που περιέχει αντίγραφα γονιδίων microRNA, περιοχές στις οποίες δένουν οι παράγοντες μεταγραφής, εκφράσεις microRNA και απλούς νουκλεοτιδιακούς πολυμορφισμούς που συνδέονται με τα microRNA. Η

MiRGen2 συνδέει τους παράγοντες μεταγραφής με τα microRNA μέσα από υποθετικές ή πειραματικά αποδεδειγμένες σχέσεις.

2.4. DIANA-pathways

Το DIANA-pathways είναι ένα υπολογιστικό εργαλείο που βασίζεται στο παγκόσμιο ιστό και σχεδιάστηκε για να αναγνωρίζει μοριακά μονοπάτια δυνητικά αλλαγμένα από την έκφραση απλών ή πολλαπλών microRNA. Το εργαλείο αυτό προσπαθεί να καλύψει την ανάγκη της επιστημονικής κοινότητας για μια γρήγορη και εύκολη ερμηνεία της ανάμειξης μιας σειράς γονιδίων σε διάφορες βιολογικές διαδικασίες. Όπως θα δούμε και στη συνέχεια, πραγματοποιεί μια διεξοδική ανάλυση πολλαπλών γονιδιακών στόχων των microRNA, συγκρίνοντας κάθε σετ από αυτούς τους στόχους με όλα τα γνωστά μονοπάτια KEGG. Η συνδυαστική δράση από τα συν-εκφραζόμενα microRNA στη διαμόρφωση ενός συγκεκριμένου μονοπατιού λαμβάνεται υπόψη με τη ταυτόχρονη ανάλυση πολλαπλών microRNA.

Το DIANA-pathways δέχεται ως είσοδο μια λίστα από ονόματα microRNA ή μια λίστα από Ensembl ID. Αν τα ονόματα των microRNA είναι ήδη ορισμένα, οι γονιδιακοί στόχοι θα ανασυρθούν αυτόματα από τη βάση δεδομένων του DIANA-microT 3.0. Επίσης ο χρήστης μπορεί να επιλέξει ανάμεσα σε ένα πιο αυστηρό(strict) ή σε ένα πιο χαλαρό(loose) σκορ κατωφλίου. Σε περίπτωση που η λίστα εισόδου του προγράμματος περιέχει ονόματα γονιδίων ή RefSeq ID τότε το πρόγραμμα τα μεταφράζει αυτόματα σε Ensembl ID. Στη συνέχεια η λίστα με τα γονίδια συγκρίνεται με τη βάση δεδομένων των KEGG μονοπατιών.

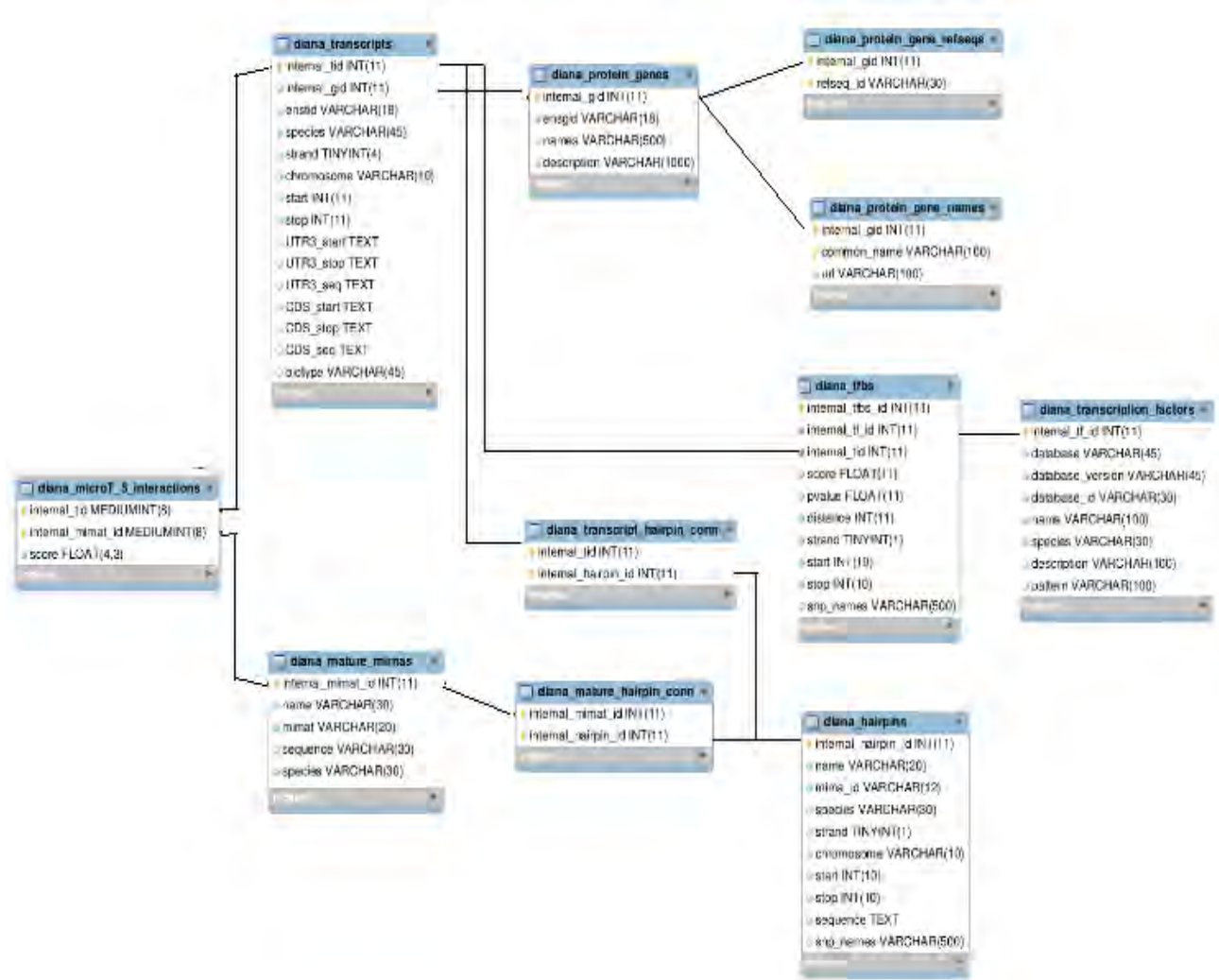
KEGG (Pathway): Είναι το όνομα του μονοπατιού, το οποίο είναι και σύνδεσμος πάνω στον οποίο αν κάνουμε κλικ, μας πάει σε μια σελίδα όπου φαίνεται γραφικά το μονοπάτι με τα γονίδια που υπήρχαν στη λίστα εισόδου του προγράμματος να είναι τονισμένα.

3. Βάση Δεδομένων

Με τον όρο βάση δεδομένων εννοούμε μία συλλογή από συστηματικά οργανωμένα σχετιζόμενα δεδομένα. Ας πάρουμε για παράδειγμα έναν τηλεφωνικό κατάλογο ο οποίος θεωρείται μια βάση δεδομένων, εφόσον αποθηκεύει αλλά και οργανώνει σχετιζόμενα τμήματα πληροφορίας, όπως είναι το όνομα και ο αριθμός τηλεφώνου. Ωστόσο, στον κόσμο των υπολογιστών, με την ορολογία βάση δεδομένων αναφερόμαστε σε μια συλλογή σχετιζόμενων δεδομένων τμημάτων πληροφορίας που αποθηκεύονται σε ηλεκτρονική μορφή. Πέρα από την ικανότητά της να αποθηκεύει δεδομένα, η βάση δεδομένων παρέχει βάσει του σχεδιασμού και του τρόπου ιεράρχησης των δεδομένων της σε προγράμματα ή συλλογές προγραμμάτων, τα ονομαζόμενα συστήματα διαχείρισης περιεχομένου, τη δυνατότητα να αντλούμε και να ανανεώνουμε γρήγορα δεδομένα.

Η συγκεκριμένη διπλωματική εργασία βασίστηκε πάνω στη βάση δεδομένων mirna2 του site DIANA LAB μέρος των πινάκων της οποίας παρουσιάζεται στην παρακάτω εικόνα.

Υλοποίηση online εφαρμογής που θα επιτρέψει εξελιγμένες αναζητήσεις σε βάση δεδομένων



Εικόνα 2: Συσχέτιση πινάκων στη βάση δεδομένων mirna2

Όπως φαίνεται στο παραπάνω σχήμα που αποτελεί ένα τμήμα της βάσης δεδομένων και συγκεκριμένα εκείνο πάνω στο οποίο χτίστηκε η παρούσα διπλωματική κάθε πίνακας αποτελείται από ένα πρωτεύον κλειδί.

Πιο συγκεκριμένα κάθε πίνακας αποτελείται από διάφορα πεδία. Για παράδειγμα ο πίνακας diana_microT5_interactions αποτελείται από 3 πεδία(internal_id, internal_mimat_id, score). Έχουμε επιλέξει ως κλειδιά για αυτόν τον πίνακα το internal_id και το internal_mimat_id. Τα κλειδιά χρησιμεύουν για τη σύνδεση των διαφόρων πινάκων μεταξύ τους. Πιο αναλυτικά, ο πίνακας diana_interactions συνδέεται με τον πίνακα diana_microT5_interactions με το κλειδί internal_id. Με τη σειρά του ο πίνακας diana_protein_genes συνδέεται με τον diana_interactions με το κλειδί internal_gid. Όμοια και οι υπόλοιποι πίνακες μεταξύ τους.

Η σύνδεση πινάκων με βάση τα κλειδιά χρησιμεύει στη γρηγορότερη αναζήτηση στη βάση δεδομένων για άντληση πληροφοριών από διαφορετικούς πίνακες.

Στη συνέχεια περιγράφονται αναλυτικά οι παραπάνω πίνακες της βάσης δεδομένων.

diana_microT5_interactions: Περιλαμβάνει ένα σκορ

diana_transcripts: Αναπαριστά ένα μετάγραφο ενός γονιδίου.

diana_mature_mirnas: Αναπαριστά ένα mature microRNA. Περιλαμβάνεται το όνομα και ένα σύνολο από επίπεδα έκφρασης.

diana_protein_genes: Αναπαριστά ένα γονίδιο που κωδικοποιεί μία πρωτεΐνη. Είναι εξειδίκευση του γονιδίου.

diana_protein_gene_names: Αναπαριστά ονόματα γονιδίων που χρησιμοποιούνται για την κωδικοποίηση μιας πρωτεΐνης.

diana_protein_gene_refseqs: Αναπαριστά μία εγγραφή στη δικτυακή βάση δεδομένων RefSeq. Κάθε γονίδιο μπορεί να διαθέτει ένα σύνολο από εγγραφές στη συγκεκριμένη βάση δεδομένων.

diana_hairpins: Αναπαριστά ένα microRNA με ιδιαίτερη δομή, δομή φουρκέτας.

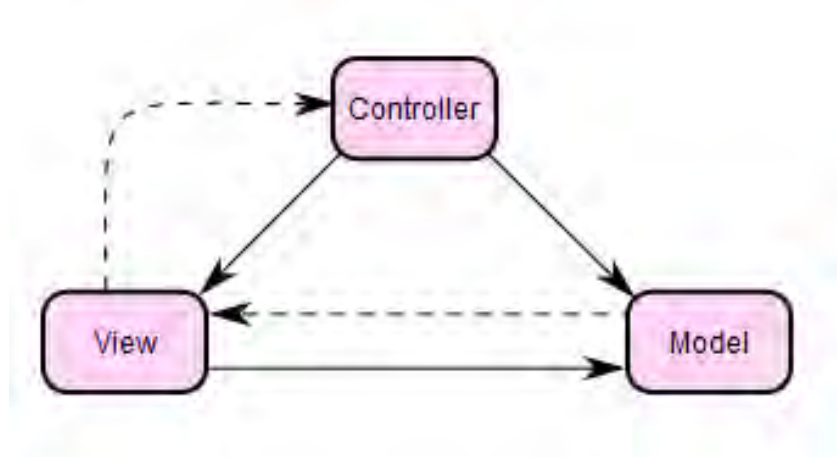
diana_mature_hairpin_conn: Περιέχει πεδίο-κλειδί για τη σύνδεση του πίνακα diana_hairpins με τον πίνακα diana_mature_mirnas

diana_transcript_hairpin_conn: Περιέχει πεδίο-κλειδί για τη σύνδεση του πίνακα diana_hairpins με τον πίνακα diana_transcripts

4. Αρχιτεκτονική

Η αρχιτεκτονική του συστήματος στηρίζεται στο αρχιτεκτονικό πρότυπο Model - View - Controller. Το συγκεκριμένο αρχιτεκτονικό πρότυπο βοηθάει στο διαχωρισμό της λογικής από την παρουσίαση στο υπό ανάπτυξη σύστημα με αποτέλεσμα να δίνει τη δυνατότητα για μεγαλύτερη ευελιξία τόσο στην ανάπτυξη όσο και στη συντήρηση του συστήματος αλλά και ευκολότερο έλεγχο των επιμέρους υποσυστημάτων. Μία εφαρμογή που αναπτύσσεται με βάση τη συγκεκριμένη τεχνολογία αποτελείται από ξεχωριστές δομικές μονάδες που υλοποιούν τους τρεις ξεχωριστούς ρόλους που ορίζονται από τη συγκεκριμένη αρχιτεκτονική. Τα μοντέλα (models) είναι κομμάτια κώδικα που αναπαριστούν τα δεδομένα της εφαρμογής. Οι όψεις (views) είναι κομμάτια κώδικα που στοχεύουν στην παρουσίαση του αποτελέσματος του συστήματος στο χρήστη. Οι ελεγκτές είναι τμήματα κώδικα που δέχονται και φέρουν εις πέρας τις εντολές του χρήστη. Στη σύγχρονη ανάπτυξη δικτυακών εφαρμογών το αρχιτεκτονικό πρότυπο Model - View -Controller βρίσκει ιδιαίτερη εφαρμογή και αποδεικνύεται σταδιακά ως ιδανική λύση για την ανάπτυξη πολύπλοκων εφαρμογών μιας και το μοντέλο ανάπτυξης δικτυακών συστημάτων επιβάλλει εγγενώς το διαχωρισμό της παρουσίασης, η οποία γίνεται μέσω HTML, και της λογικής, η οποία πραγματοποιείται στο διακομιστή της εφαρμογής. Σε αυτό το πλαίσιο ταιριάζει απόλυτα η λογική της λειτουργίας του συγκεκριμένου αρχιτεκτονικού προτύπου. Όπως φαίνεται στο σχήμα που ακολουθεί παρουσιάζονται διαγραμματικά τα επιμέρους στοιχεία της αρχιτεκτονικής και τονίζονται οι συσχετίσεις μεταξύ αυτών. Αξίζει να αναφέρουμε ότι ο ελεγκτής έχει πρόσβαση τόσο στις όψεις, τις οποίες δίνει προς απεικόνιση, όσο και στα μοντέλα, από τα οποία αντλεί πληροφορίες. Οι όψεις, με τη σειρά τους, έχουν τη δυνατότητα να άντλούν πληροφορίες από τα μοντέλα. Σε επόμενες ενότητες θα δούμε αναλυτικά την ανάλυση του αρχιτεκτονικού προτύπου Model - View - Controller στα επιμέρους τμήματα που το αποτελούν.

(Γεώργιος Πρέκας, 2009)



Εικόνα 3: Μοντέλο συσχέτισης MODEL-VIEW-CONTROLLER
(Πηγή:<http://upload.wikimedia.org/wikipedia/commons/b/b5/ModelViewControllerDiagram2.svg>)

4.1. Model

Αποτελεί την αναπαράσταση των δεδομένων που χρησιμοποιούνται και επεξεργάζονται από την εφαρμογή. Τα μοντέλα είναι βασισμένα στη λογική της πληροφορίας και όταν ένα μοντέλο αλλάζει την κατάσταση του ενημερώνει τα views που σχετίζονται άμεσα με αυτό ώστε να κάνουν ανανέωση της κατάστασής τους. Η πλειοψηφία των συστημάτων χρησιμοποιεί μηχανισμό μόνιμης αποθήκευσης για παράδειγμα μια σχεσιακή βάση δεδομένων. Αν και τα μοντέλα δεν είναι αναγκαίο να γνωρίζουν το μέσο αποθήκευσης των πληροφοριών είναι αρκετά πιο εύκολο αν το κάθε μοντέλο γνωρίζει πώς να αναγνώσει και να αποθηκεύσει τον εαυτό του στο μέσο μόνιμης αποθήκευσης της εφαρμογής. (Γεώργιος Πρέκας, 2009)

4.2. View

Προβάλλει ένα μοντέλο κατανοητό από το χρήστη και προσαρμοσμένο στο μέσο απεικόνισης. Είναι συνηθισμένο να παρέχονται διαφορετικά views για κάθε μοντέλο. Πέρα από τη διάκριση στον τρόπο που απεικονίζεται ένα μοντέλο, τα views διακρίνονται και με βάση το μέσο στο οποίο είναι σχεδιασμένα να προβάλλονται. Συχνά σε ένα μεγάλο σύστημα παρέχονται στους χρήστες ποικίλες δυνατότητες τόσο πρόσβασης όσο και απεικόνισης της πληροφορίας. Σε κάθε διαφορετικό μέσο απεικόνισης αντιστοιχεί και ένα διαφορετικό view το οποίο αντιλαμβάνεται τις ιδιαιτερότητες του

μέσου για να εμφανίσει τα στοιχεία του μοντέλου με τον καλύτερο δυνατό τρόπο. (Γεώργιος Πρέκας, 2009)

4.3. Controller

Ο controller αναλαμβάνει την ευθύνη για τη λήψη των μηνυμάτων και των εντολών εισόδου από το χρήστη και αποκρίνεται δρώντας πάνω στα μοντέλα ή εμφανίζοντας κάποιο view. Θα λέγαμε ότι είναι ο κρίκος που συνδέει το σύστημα και είναι υπεύθυνος για την περάτωση των εντολών του χρήστη κάθε εφαρμογής. Η κεντρική λειτουργία του είναι αρχικά η αναγνώριση των εντολών του χρήστη έπειτα η άντληση και η επεξεργασία της πληροφορίας από τα μοντέλα σύμφωνα με τις εντολές του χρήστη και τέλος η παρουσίαση της πληροφορίας που ζητήθηκε μέσω της κατάλληλης όψης. (Γεώργιος Πρέκας, 2009)

5. Yii PHP Framework

Το site του Diana Lab βασίστηκε πάνω στο αρχιτεκτονικό πρότυπο Model – View – Controller και χρησιμοποιήθηκε το Yii PHP framework, το οποίο κάνει πιο εύκολη την ανάπτυξη εφαρμογών σε PHP ακολουθώντας το παραπάνω αρχιτεκτονικό πρότυπο. Συνεπώς η συγκεκριμένη διπλωματική εργασία ακολούθησε τα παραπάνω μοντέλα για την ενσωμάτωση του κώδικα στο συγκεκριμένο site. Η επιλογή του framework αυτού βασίστηκε στις πολλές δυνατότητές του. Οι ενότητες που ακολουθούν περιγράφουν τις δυνατότητες άλλα και τα χαρακτηριστικά του Yii PHP framework. (Γεώργιος Πρέκας, 2009)

5.1. Ταχύτητα

Η ταχύτητα που παρέχει το Yii PHP framework στην εκτέλεση της εφαρμογής μέσω της τεχνικής της οκνηρής φόρτωσης των συστατικών του είναι αξιόλογη. Πιο συγκεκριμένα φορτώνει τα απαραίτητα συστατικά για την εκτέλεση της εφαρμογής καθώς αυτά απαιτούνται. Αντίθετα τα άλλα PHP frameworks για λόγους κυρίως ευκολίας φορτώνουν στη μνήμη όλα τα συστατικά πριν την εξυπηρέτηση του κάθε HTTP αιτήματος. Με αυτόν τον τρόπο αυξάνουν τη χρονική καθυστέρηση στην HTTP απάντηση. (Γεώργιος Πρέκας, 2009)

5.2. Κώδικας-Ποιότητα

Το Yii PHP framework αξιοποιεί τις αντικειμενοστραφείς δυνατότητες της PHP πράγμα που έχει ως αποτέλεσμα να αποτελείται από ευανάγνωστο κώδικα, εύκολο στη συντήρηση. Με τον τρόπο αυτό εγγυάται ότι τυχόν προβλήματα που ανακαλύπτονται στο framework επιλύονται εύκολα σε σύντομο χρονικό διάστημα. (Γεώργιος Πρέκας, 2009)

5.3. Εισαγωγή δεδομένων από HTML φόρμες

Το Yii PHP framework υποστηρίζει πλήρως τόσο την εισαγωγή όσο και την επαλήθευση δεδομένων από HTML φόρμες, κάνοντας τη χρήση απλή αλλά

και προσφέροντας μέγιστη ασφάλεια για το σύστημα. Επιπλέον, δίνει στο χρήστη μία σειρά από βοηθητικές μεθόδους και αντικείμενα με σκοπό να του προσφέρει έτοιμες λύσεις στα συχνότερα προβλήματα εισαγωγής δεδομένων από το χρήστη. (Γεώργιος Πρέκας, 2009)

5.4. Διαχείριση σφαλμάτων και καταγραφή συμβάντων

Το Yii PHP framework διαχειρίζεται τα σφάλματα με τον πιο δυνατό τρόπο παρέχοντας στον προγραμματιστή όλες τις τεχνικές πληροφορίες που απαιτούνται για να εντοπίσει και να επιδιορθώσει τις δυσλειτουργίες. Σε κάθε περίπτωση που αστοχεί παρέχονται πληροφορίες που σχετίζονται με το σφάλμα, το stack trace αλλά και τμήμα του κώδικα που προκάλεσε το πρόβλημα. (Γεώργιος Πρέκας, 2009)

5.5. Τεκμηρίωση

Μια άψογη τεκμηρίωση συνοδεύει το Yii PHP framework, η οποία μπορεί να βρεθεί τόσο σε επίπεδο κώδικα σε μορφή rhrdoc όσο και σε επίπεδο ξεχωριστής τεκμηρίωσης σε μορφή HTML αλλά και PDF. Με τον τρόπο αυτόν η ανάπτυξη εφαρμογών με βάση τη συγκεκριμένη πλατφόρμα διευκολύνεται εφόσον ο προγραμματιστής έχει στη διάθεσή του όλες τις πληροφορίες που απαιτούνται με κάθε λεπτομέρεια ώστε να υλοποιήσει το σύστημα σύμφωνα με τις προδιαγραφές που έχουν τεθεί. (Γεώργιος Πρέκας, 2009)

5.6. Υποστήριξη

Το Yii PHP framework αναπτύσσεται συνεχώς προβαίνοντας σε συχνές ανανεώσεις, οι οποίες έχουν στόχο τόσο τη διόρθωση τυχόν δυσλειτουργιών όσο και την προσθήκη καινούργιων χαρακτηριστικών. Μία ομάδα ανάπτυξης βρίσκεται στη διάθεση του προγραμματιστή, η οποία ανταποκρίνεται καθημερινά στα ερωτήματα των χρηστών μέσω ενός εξειδικευμένου forum συζητήσεων:

(<http://www.yiiframework.com/forum/>). (Γεώργιος Πρέκας, 2009)

6. Υλοποίηση

Το σύστημα που αναπτύχθηκε στο πλαίσιο της παρούσας διπλωματικής εργασίας είναι μία δικτυακή εφαρμογή αναζήτησης βιολογικών στόχων microRNA. Προσφέροντας μια σειρά από ξεχωριστά χαρακτηριστικά με στόχο τη βελτίωση της εμπειρίας του χρήστη προσπαθούμε μέσα από υπηρεσίες να προσαρμόσουμε την εφαρμογή ανάλογα με τις προτιμήσεις του.

Στα πλαίσια λοιπόν της διπλωματικής αυτής χρησιμοποιήθηκαν διάφοροι όροι για την αναζήτηση βιολογικών στόχων όπως : gene, protein gene, microRNA gene, transcript, hairpin microRNA, mature microRNA

Πιο αναλυτικά:

Gene^[1]: Αναπαριστά ένα γονίδιο. Το κάθε γονίδιο χαρακτηρίζεται από την αλληλουχία βάσεων του, από την περιγραφή του και από την τοποθεσία του, η οποία είναι σύνθετη ιδιότητα και περιλαμβάνει το είδος, το χρωμόσωμα, την αλυσίδα του χρωμοσώματος, την αρχή και το τέλος της αλληλουχίας βάσεων. Ένα γονίδιο μπορεί να είναι είτε protein gene είτε microRNA gene. Επίσης, μπορεί να σχετίζεται με ένα σύνολο από CPG, ένα σύνολο από transcription factors και ένα σύνολο από hairpin microRNAs.

CPG^[1]: Περιγράφει μία περιοχή του DNA όπου υπάρχουν διαδοχικά μία βάση C (κυτοσίνη) και μία βάση G (γουανίνη) στην ίδια αλυσίδα βάσεων και χαρακτηρίζεται από την τοποθεσία του, η οποία είναι σύνθετη ιδιότητα και ορίζεται όπως και παραπάνω.

Protein gene^[1]: Αναπαριστά ένα γονίδιο που χρησιμοποιείται για την κωδικοποίηση μίας πρωτεΐνης. Χαρακτηρίζεται από ένα μοναδικό αναγνωριστικό που αντιστοιχεί στην εγγραφή της δικτυακής βάσης δεδομένων, από ένα σύνολο ονομάτων και από σύνολο αναγνωριστικών για τη δικτυακή βάση δεδομένων Refseq. Ένα πρωτεϊνικό γονίδιο μπορεί να αντιστοιχίζεται σε ένα σύνολο από KEGG και σε ένα σύνολο από transcripts.

MicroRNA gene^[1]: Περιγράφει ένα γονίδιο που χρησιμοποιείται για την κωδικοποίηση ενός microRNA.

Transcript^[1]: Αναπαριστά ένα μετάγραφο ενός γονιδίου και χαρακτηρίζεται από ένα μοναδικό αναγνωριστικό που αντιστοιχεί στην εγγραφή της δικτυακής βάσης δεδομένων Ensembl.

Transcription factor^[1]: Αναπαριστά έναν παράγοντα μεταγραφής και χαρακτηρίζεται από το όνομά του και από το πρότυπό του.

Hairpin microRNA^[1]: Αναπαριστά ένα microRNA το οποίο παρουσιάζει μία ιδιαίτερη δομή, δομή φουρκέτας. Προκύπτει από τη μεταγραφή ενός γονιδίου και χαρακτηρίζεται από το όνομά του, από ένα μοναδικό αναγνωριστικό που αντιστοιχεί στην εγγραφή της δικτυακής βάσης δεδομένων miRBase και από την τοποθεσία του, η οποία είναι σύνθετη ιδιότητα και ορίζεται όπως και παραπάνω.

Mature microRNA^[1]: Αναπαριστά ένα τελικό microRNA, το οποίο προκύπτει από ένα hairpin microRNA και χαρακτηρίζεται από το όνομά του, από ένα μοναδικό αναγνωριστικό για τη δικτυακή βάση δεδομένων miRBase, από την αλληλουχία των βάσεων του και από ένα σύνολο επιπέδων έκφρασης.

Η διαδικασία αναζήτησης ακολουθεί το παρακάτω διάγραμμα:



Εικόνα 4: Διάγραμμα Ροής Δεδομένων

[1]: (Γεώργιος Πρέκας, 2009)

6.1. Εισαγωγή στοιχείων

Το συγκεκριμένο υποσύστημα λαμβάνει τις παραμέτρους της αναζήτησης που θέλει να πραγματοποιήσει ο χρήστης του συστήματος. Οι παράμετροι μπορεί να είναι μία ή περισσότερες όπως το όνομα ενός γονιδίου ή το όνομα ενός microRNA.

Ωστόσο τα αποτελέσματα που θα επιστραφούν μπορεί να είναι πάρα πολλά και έτσι ο χρήστης έχει τη δυνατότητα να τα περιορίσει εισάγοντας επιπλέον πληροφορία σε 2^ο επίπεδο ώστε να περιοριστούν τα αποτελέσματα. Θα το δούμε αναλυτικά στο κεφάλαιο της υλοποίησης.

6.2. Αναζήτηση βιολογικών στόχων

Το συγκεκριμένο υποσύστημα αναλαμβάνει να εντοπίσει τους βιολογικούς στόχους που ανταποκρίνονται στις παραμέτρους που έχει ορίσει ο χρήστης. Πιο αναλυτικά με την είσοδο των στοιχείων από το χρήστη το υποσύστημα της αναζήτησης ελέγχει τα δεδομένα και αναζητεί στη βάση δεδομένων για τις εγγραφές που έδωσε ο χρήστης.

6.3. Εμφάνιση αποτελεσμάτων

Το συγκεκριμένο υποσύστημα αναλαμβάνει την απεικόνιση των αποτελεσμάτων σε πίνακες. Είναι σχεδιασμένο έτσι ώστε να προσαρμόζεται ανάλογα με το είδος των αναζητούμενων στοιχείων.

Σε κάθε περίπτωση οι πίνακες που χρησιμοποιούνται αναδεικνύουν τη δομή των αποτελεσμάτων.

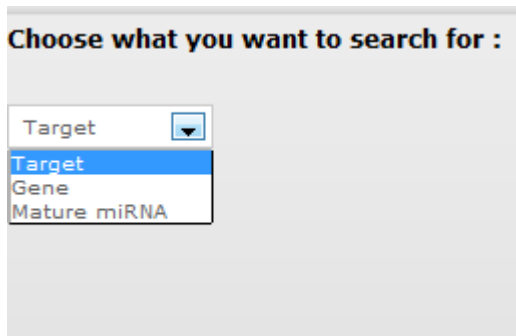
6.4. Η βασική ιδέα

Στο συγκεκριμένο υποκεφάλαιο παρουσιάζουμε τη βασική ιδέα αυτής της διπλωματικής. Συγκεκριμένα, ζητήθηκε η υλοποίηση μιας φόρμας αναζήτησης πάνω στη βάση δεδομένων mirna2. Ουσιαστικά δίνεται η δυνατότητα σε κάθε χρήστη να κάνει σύνθετες αναζητήσεις στη βάση

δεδομένων και να λαμβάνει τα αποτελέσματά του σε μορφή πινάκων. Τα αποτελέσματα αυτά είναι πεδία κάποιων πινάκων της βάσης miRNA2. Για την υλοποίηση βέβαια των σύνθετων αναζητήσεων απαιτούνται φίλτρα που περιορίζουν τον αριθμό των αναζητήσεων.

6.5. Σκέψεις πριν την υλοποίηση...

Πριν περάσουμε στην υλοποίηση σχεδιάσαμε τη δομή της αναζήτησης. Αρχικά σκεφτήκαμε ο χρήστης να μπορεί να επιλέγει μέσα από μια λίστα δεδομένων την αναζήτηση που επιθυμεί να κάνει στη βάση κάθε φορά. Δηλαδή να είναι σε θέση σε 1ο στάδιο να επιλέξει ανάμεσα σε:

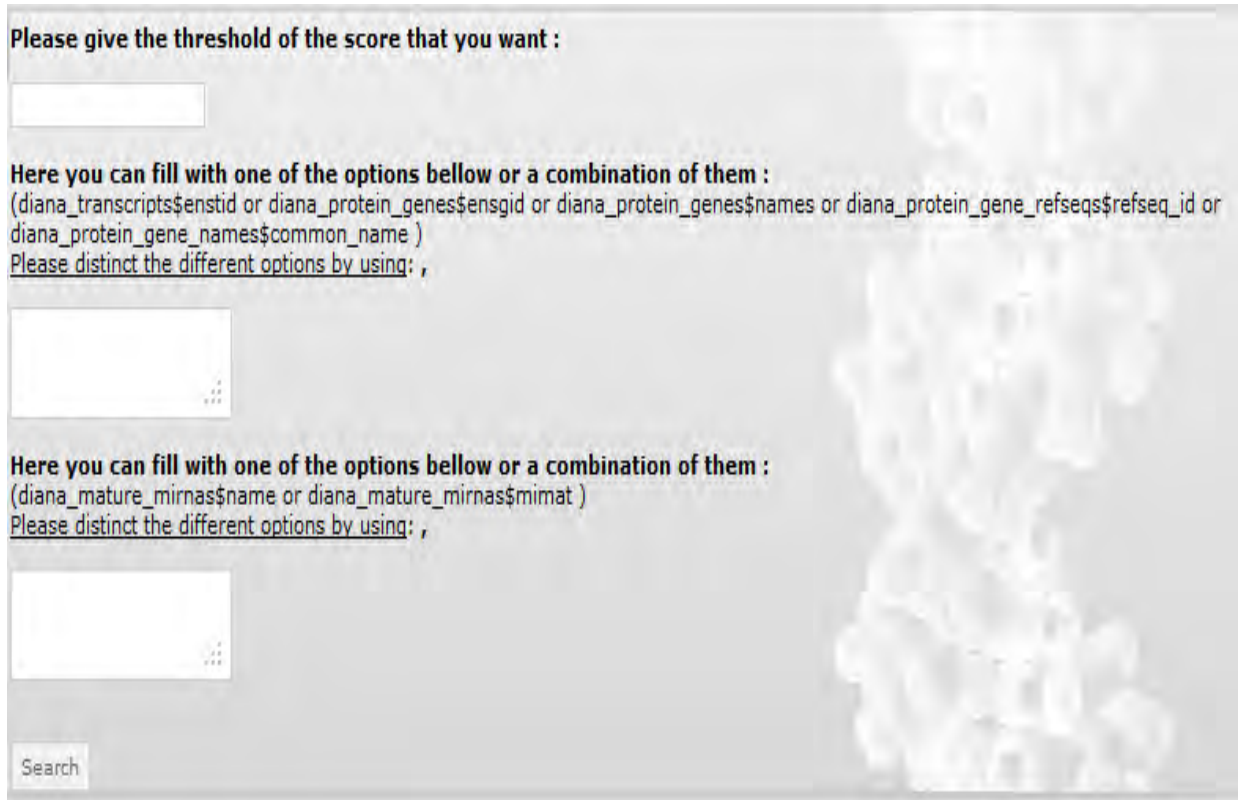


Εικόνα 5: Επίπεδο 1ο / Επιλογή

Συνεπώς σκεφτήκαμε για το 1ο στάδιο μια dropdown list από την οποία θα μπορεί ο χρήστης να επιλέξει αυτό που θέλει.

Έστω λοιπόν πως ο χρήστης επιλέξει Target. Ουσιαστικά δηλαδή ο χρήστης επέλεξε να ψάξει για στόχους. Προφανώς μέσα στη βάση δεδομένων miRNA2 υπάρχουν πάρα πολλά γονίδια στόχοι. Πρέπει λοιπόν να εισάγουμε σε 2ο επίπεδο κάποια φίλτρα που θα περιορίζουν τα αποτελέσματα της αναζήτησης. Σκεφτήκαμε λοιπόν ο χρήστης σε 2ο επίπεδο να μπορεί να εισάγει ο ίδιος στοιχεία με τα οποία να γίνεται η αναζήτηση στη βάση.

Πιο αναλυτικά, ο χρήστης μπορεί να δώσει ένα score για να περιορίσει τα αποτελέσματα αλλά και διάφορα πεδία πινάκων της βάσης που σχετίζονται με genes και microRNAs όπως φαίνεται στην εικόνα:



Please give the threshold of the score that you want :

Here you can fill with one of the options bellow or a combination of them :
(diana_transcripts\$enstid or diana_protein_genes\$ensgid or diana_protein_genes\$names or diana_protein_gene_refseqs\$refseq_id or diana_protein_gene_names\$common_name)
Please distinct the different options by using: ,

Here you can fill with one of the options bellow or a combination of them :
(diana_mature_mirnas\$name or diana_mature_mirnas\$mimat)
Please distinct the different options by using: ,

Search

Εικόνα 6: Επιλογή Target

Όπως φαίνεται παραπάνω ο χρήστης μπορεί να εισάγει ένα score όσες από τις επιλογές που εμφανίζονται στην εικόνα. Για παράδειγμα στο 1ο πεδίο τοποθετεί ένα score στο 2ο όσες από τις επιλογές,

diana_transcripts\$enstid or
diana_protein_genes\$ensgid or
diana_protein_genes\$names or
diana_protein_gene_refseqs\$refseq_id or
diana_protein_gene_names\$common_name

θέλει και όμοια στο 3ο από τις επιλογές,

diana_mature_mirnas\$name or diana_mature_mirnas\$mimat

Σε περίπτωση που ο χρήστης επιλέξει στο 1ο επίπεδο Gene τότε όμοια με τα παραπάνω χρησιμοποιούμε σε 2ο επίπεδο φίλτρα για να περιορίσουμε τα αποτελέσματα όπως φαίνεται στην παρακάτω εικόνα:

Here you can fill with one of the options bellow or a combination of them :
(diana_mature_mirnas\$name or diana_mature_mirnas\$mimat)
Please distinct the different options by using: ,

Here you can fill with one of the options bellow or a combination of them :
(diana_hairpins\$name or diana_hairpins\$mima_id)
Please distinct the different options by using: ,

Here you can fill with the option bellow :
(diana_transcription_factors\$name)
Please distinct the different options by using: ,

Search

Εικόνα 7: Επιλογή Gene

Όπως φαίνεται από την εικόνα, στο 1ο πεδίο ο χρήστης μπορεί να βάλει όσες από τις επιλογές θέλει από τις:

diana_mature_mirnas\$name or diana_mature_mirnas\$mimat

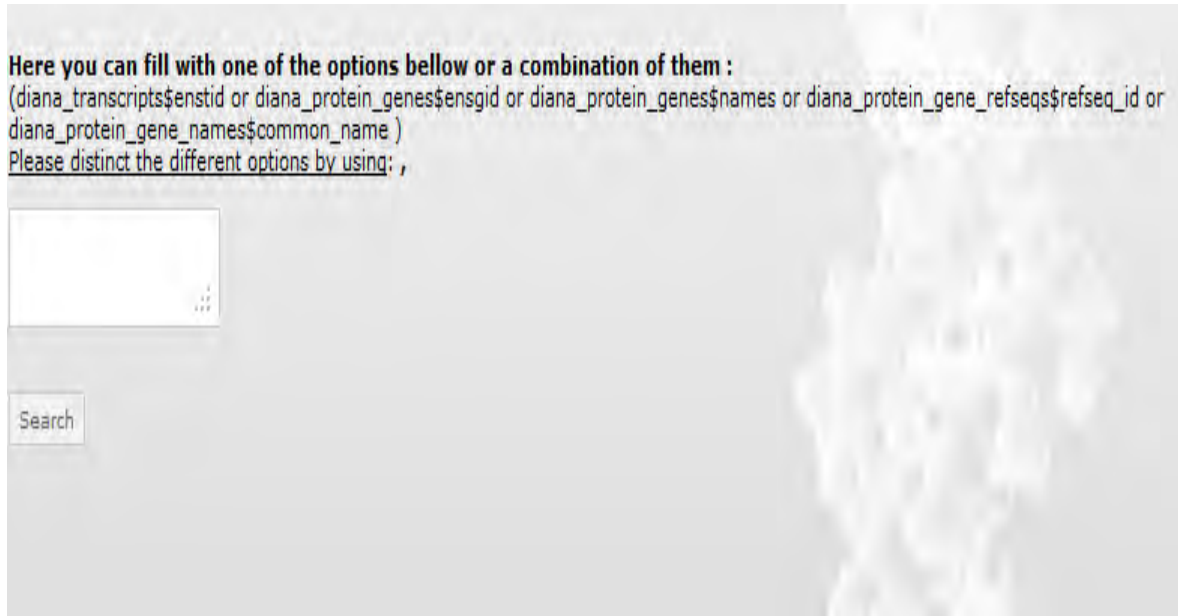
στο 2^ο όμοια με το 1^ο επιλέγει από τις:

diana_hairpins\$name or diana_hairpins\$mima_id

και στο 3ο επιλέγει από τις:

diana_transcription_factors\$name

Σε περίπτωση που ο χρήστης διαλέξει στο 1ο επίπεδο την επιλογή MatureMirna τότε αντίστοιχα με τα παραπάνω εμφανίζονται κατάλληλα στο 2ο επίπεδο φίλτρα για να γίνουν τα αποτελέσματα της αναζήτησης πιο συγκεκριμένα όπως φαίνεται στην εικόνα:



Εικόνα 8: Επιλογή Mature Mirna

Στο πεδίο που φαίνεται στην εικόνα ο χρήστης μπορεί να εισάγει όσες από τις παρακάτω επιλογές θέλει:

diana_transcripts\$enstid or

diana_protein_genes\$ensgid or

diana_protein_genes\$names or

diana_protein_gene_refseqs\$refseq_id or

diana_protein_gene_names\$common_name

Σε όλες τις επιλογές του 1ου επιπέδου που διαλέγει ο χρήστης, στόχος μας σε αυτή τη διπλωματική είναι να επιστρέψουμε τα σωστά αποτελέσματα που ζητήθηκαν κάνοντας ταυτόχρονα το σύστημα πιο αποδοτικό.

Ουσιαστικά λοιπόν οι σκέψεις μας πριν την υλοποίηση είχαν κυρίως στόχο την εύρεση των κατάλληλων φίλτρων ώστε να περιοριστούν και να γίνουν

πιο συγκεκριμένα τα αποτελέσματα της αναζήτησης στη βάση δεδομένων mirna2.

Στο σημείο αυτό όμως, αξίζει να μοιραστούμε τις σκέψεις μας για τη γλώσσα προγραμματισμού του κώδικα που χρησιμοποιήσαμε για την υλοποίηση όλων των παραπάνω πριν περάσουμε στην παρουσίαση του κώδικα που γράψαμε.

6.6. Γλώσσες Διαδικτυακού Προγραμματισμού

Ο προγραμματισμός στο διαδίκτυο είναι ένας πολύ μεγάλος τομέας της επιστήμης της πληροφορικής και των υπολογιστών. Οι γλώσσες προγραμματισμού με τις οποίες μπορεί κάποιος να δημιουργήσει ιστοσελίδες είναι πάρα πολλές και εξελίσσονται καθημερινά. Μερικά χαρακτηριστικά παραδείγματα είναι η Html, η Xml, η PHP, η Javascript, η Perl, η Python και πολλές άλλες. Φυσικά, δεν προσφέρουν όλες οι παραπάνω γλώσσες τις ίδιες δυνατότητες αλλά η καθεμία χρησιμοποιείται για ένα συγκεκριμένο σκοπό μέσα σε μια ιστοσελίδα.

Οι γλώσσες προγραμματισμού του διαδικτύου χωρίζονται σε διάφορες κατηγορίες. Η σημαντικότερη από αυτές είναι αυτή που διαχωρίζει τις γλώσσες που αφορούν κώδικα της πλευράς του πελάτη(client side coding) και αυτές που αφορούν κώδικα της πλευράς του εξυπηρετητή(server side coding). Στη κατηγορία client side coding ανήκουν γλώσσες όπως η Html, η Xml και η Javascript ενώ στη κατηγορία server side coding γλώσσες όπως η Python και η PHP.

Ο κώδικας που θα γράψουμε θέλουμε να τρέχει τόσο στην πλευρά του εξυπηρετητή όσο και στη πλευρά των πελατών που επισκέπτονται τον ιστοχώρο του DIANA LAB. Επίσης, θέλουμε κατά μια έννοια να είναι δυναμικός και όχι στατικός.. Η πιο κατάλληλη γλώσσα διαδικτυακού προγραμματισμού που μπορεί να κάνει όλα τα παραπάνω είναι η PHP σε συνδυασμό με την HTML.

Η PHP είναι μια γλώσσα προγραμματισμού για τη δημιουργία σελίδων web με δυναμικό περιεχόμενο. Μια σελίδα PHP περνά από επεξεργασία από ένα συμβατό διακομιστή του Παγκόσμιου Ιστού (π.χ. Apache), ώστε να παραχθεί σε πραγματικό χρόνο το τελικό περιεχόμενο, που θα σταλεί στο πρόγραμμα περιήγησης των επισκεπτών σε μορφή κώδικα HTML.

Ωστόσο όπως έχουμε περιγράψει σε προηγούμενα κεφάλαια το site του Diana Lab είναι «χτισμένο» πάνω στο framework yii. Συνεπώς για την υλοποίηση της παρούσας διπλωματικής χρησιμοποιήσαμε τις παραπάνω

γλώσσες προγραμματισμού προσαρμοσμένες πάνω στο συγκεκριμένο framework.

6.7. Ανάλυση του Κώδικα

Στο υποκεφάλαιο αυτό παρουσιάζεται η ανάλυση του κώδικα που γράφτηκε στα πλαίσια της παρούσας διπλωματικής εργασίας. Η ανάλυση αυτή ξεκινάει από την σελίδα που συναντάει ο χρήστης με την είσοδό του στη συγκεκριμένη εφαρμογή. Όπως έχει αναφερθεί παραπάνω, σε 1ο επίπεδο ο χρήστης πρέπει να διαλέξει από μια DropDown list την επιλογή που θέλει και να στείλει την επιλογή αυτή στο 2ο επίπεδο αναζήτησης. Ο κώδικας λοιπόν για αυτή τη σελίδα φαίνεται παρακάτω.

```
<?php
    //Create a list that performs the "actionsearch", of the "SiteController".

    echo CHtml::beginForm(array('site/search2'));
    echo '<br/>';
    echo '<h3><b>Choose what you want to search for : </b></h3>';
    echo '<br/>';echo '<br/>';
    echo CHtml::DropDownList('level1','',array(1=>'Target',2=>'Gene',3=>'Mature miRNA'));
    echo "<div class=\"row\">";
    echo '<br/>';
    echo CHtml::submitButton('Search');
    echo "</div>";
    echo CHtml::endForm();
?>
```

Εικόνα 9: Κώδικας σελίδας search.php

Όπως φαίνεται στον παραπάνω κώδικα αρχικά δίνουμε την εντολή beginForm για τη δημιουργία ενός πεδίου φόρμας. Στη συνέχεια δίνουμε τη δυνατότητα στο χρήστη μέσα από μια DropDown list να διαλέξει μία από τις επιλογές που του εμφανίζονται. Τέλος στέλνει την τιμή που διάλεξε στον site controller στη συνάρτηση Search2.

Ο site controller με τη σειρά του μέσα από την συνάρτηση,

```
public function actionSearch2()
```

λαμβάνει την τιμή που στάλθηκε και ανάλογα με την τιμή της ανακατευθύνει σε συγκεκριμένη ιστοσελίδα, όπως φαίνεται στην παρακάτω εικόνα.

```
public function actionSearch2 ()
{
    $level1=$_POST['level1'];

    if($level1==1) {
        $this->render('search2');
    }
    if($level1==2) {
        $this->render('search3');
    }
    if($level1==3) {
        $this->render('search4');
    }
}
```

Εικόνα 10: Κώδικας συνάρτησης ActionSearch2 του Site Controller

Συνεπώς, αν ο χρήστης επιλέξει Target ενεργοποιείται το 1ο if και κάνει ανακατεύθυνση στη σελίδα search2. Όμοια αν επιλέξει Gene ενεργοποιείται το 2ο if και κάνει ανακατεύθυνση στη σελίδα search3 και τέλος αν επιλέξει MatureMirna ενεργοποιείται το 3ο if και κάνει ανακατεύθυνση στη σελίδα search4.

6.7.1. Περίπτωση Target

Ας πάρουμε αρχικά την περίπτωση που επιλέγει ο χρήστης Target και συνεπώς ενεργοποιείται το 1ο if. Τότε ο κώδικας που θα εκτελεσθεί είναι ο παρακάτω.

Υλοποίηση online εφαρμογής που θα επιτρέπει εξελιγμένες αναζητήσεις σε βάση δεδομένων

```
<?php
//Create a list that performs the "actionsearch", of the "SiteController".

echo CHtml::beginForm(array('site/result'));
echo '<br/>';
echo '<b>Please give the threshold of the score that you want :</b> ';
echo '<br/>';
echo '<br/>';
echo CHtml::textField('score','');
echo '<br/>';
echo '<br/>';
echo '<b>Here you can fill with one of the options bellow or a combination of them : </b>';
echo '<br/>';
echo '(diana_transcripts$enstid or diana_protein_genes$ensgid or diana_protein_genes$names
    or diana_protein_gene_refseqs$refseq_id or diana_protein_gene_names$common_name )';
echo '<br/>';
echo '<u>Please distinct the different options by using</u>: <b>,</b> ';
echo '<br/>';
echo '<br/>';

echo CHtml::textArea('ch1','');
echo '<br/>';
echo '<br/>';
echo '<b>Here you can fill with one of the options bellow or a combination of them : </b>';
echo '<br/>';
echo '(diana_mature_mirnas$name or diana_mature_mirnas$mimat )';
echo '<br/>';
echo '<u>Please distinct the different options by using</u>: <b>,</b> ';
echo '<br/>';
echo '<br/>';
echo CHtml::textArea('ch2','');
echo "<div class=\"row\">";
echo '<br/>';
echo '<br/>';
echo CHtml::submitButton('Search');
echo "</div>";
echo CHtml::endForm();

?>
```

Εικόνα 11: Κώδικας σελίδας search2.php

Δημιουργούμε λοιπόν μια φόρμα μέσα στην οποία φτιάχνουμε 3 πεδία. 1 πεδίο textField για το score και 2 πεδία textArea μέσα στα οποία μπορεί ο χρήστης να εισάγει όσες από τις επιλογές θέλει και που του εμφανίζονται στα αντίστοιχα κουτάκια. Για παράδειγμα στο 1ο textArea μπορεί να εισάγει μόνο επιλογές από:

diana_transcripts\$enstid or
diana_protein_genes\$ensgid or
diana_protein_genes\$names or
diana_protein_gene_refseqs\$refseq_id or
diana_protein_gene_names\$common_name

Δηλαδή να εισάγει εγγραφές \$enstid από τον πίνακα diana_transcripts ή εγγραφές \$ensgid ή \$names από τον πίνακα diana_protein_genes ή εγγραφές \$refseq_id από τον πίνακα diana_protein_gene_refseqs ή εγγραφές \$common_name από τον πίνακα diana_protein_gene_names.

Όπως φαίνεται και στην εικόνα ο χρήστης μπορεί να εισάγει σε κάθε πεδίο textArea όσες εγγραφές θέλει αρκεί να χωρίζει τις εγγραφές μεταξύ τους με κόμμα.

Μόλις ο χρήστης εισάγει τις εγγραφές που θέλει στα κατάλληλα πεδία και αφού δώσει και ένα score για το κατώφλι τότε με το που πατήσει το submitButton στέλνονται τα δεδομένα στον sitecontroller στη συνάρτηση ActionResult.

Η συνάρτηση αυτή με τη σειρά της εκτελεί τον παρακάτω κώδικα.

Υλοποίηση online εφαρμογής που θα επιτρέπει εξελιγμένες αναζητήσεις σε βάση δεδομένων

```
public function actionResult()
{
    $score=$_POST['score'];
    $ch1=$_POST['ch1'];
    $ch2=$_POST['ch2'];

    $chars1 = preg_split('/', '/', $ch1, -1);
    $chars2 = preg_split('/', '/', $ch2, -1);
    $i=0;
    $j=0;

    while( $i != count($chars1) ){

        $query = "SELECT DISTINCT diana_microT_5_interactions.internal_tid, diana_microT_5_interactions.internal_mimat_id,
            diana_microT_5_interactions.score
        FROM diana_microT_5_interactions, diana_transcripts, diana_protein_genes, diana_protein_gene_refseqs,
            diana_protein_gene_names
        WHERE diana_microT_5_interactions.internal_tid = diana_transcripts.internal_tid
        AND diana_transcripts.internal_gid = diana_protein_genes.internal_gid
        AND diana_protein_genes.internal_gid = diana_protein_gene_refseqs.internal_gid
        AND diana_protein_genes.internal_gid = diana_protein_gene_names.internal_gid
        AND diana_transcripts.enstid = '$chars1[$i]'
        AND diana_microT_5_interactions.score < $score";

        $tempRow1[$i] = Yii::app()->getDb()->getCommandBuilder()->createSqlCommand($query)->query();
        $i=$i+1;
    }

    while( $j != count($chars2) ){
        $query = "SELECT DISTINCT diana_microT_5_interactions.internal_tid, diana_microT_5_interactions.internal_mimat_id,
            diana_microT_5_interactions.score
        FROM diana_microT_5_interactions, diana_mature_mirnas
        WHERE diana_microT_5_interactions.internal_mimat_id = diana_mature_mirnas.internal_mimat_id
        AND ( diana_mature_mirnas.name = '$chars2[$j]' OR diana_mature_mirnas.mimat = '$chars2[$j]' )
        AND diana_microT_5_interactions.score < $score";

        $tempRow2[$j] = Yii::app()->getDb()->getCommandBuilder()->createSqlCommand($query)->query();
        $j=$j+1;
    }

    $this->render('results',array('tempRow1'=>$tempRow1, 'tempRow2'=>$tempRow2, 'i'=>$i, 'j'=>$j ));
}
}
```

Εικόνα 12: Κώδικας συνάρτησης ActionResult του Site Controller

Η συνάρτηση αυτή παίρνει με τη μέθοδο POST τις μεταβλητές που της έστειλε η search2 και τις αποθηκεύει σε νέες μεταβλητές όπως φαίνεται

παραπάνω. Για παράδειγμα τη μεταβλητή `score` την αποθηκεύει: `$score=$_POST['score']` στη μεταβλητή `$score`.

Όμοια αποθηκεύει και τις διάφορες εγγραφές που εισήγαγε ο χρήστης στα πεδία `textArea` κρατώντας τες σ' ένα `string` για κάθε `textArea`.

Ωστόσο για να μπορέσει να γίνει η αναζήτηση στη βάση δεδομένων με βάση τις εγγραφές που εισήγαγε ο χρήστης πρέπει με κάποιο τρόπο το `string` να διαβαστεί και να τοποθετηθεί κάθε εγγραφή σε μια θέση ενός νέου πίνακα.

Όπως είχαμε πει παραπάνω ο χρήστης ξεχωρίζει κάθε εγγραφή από την επόμενη με ένα κόμμα. Συνεπώς χρειάζεται στο σημείο αυτό μια συνάρτηση που θα διαβάζει το `string` και θα το κόβει κάθε φορά που θα βρίσκει κόμμα εισάγοντας κάθε εγγραφή στη θέση ενός πίνακα.

Την παραπάνω διαδικασία την κάνει η συνάρτηση της `php`, `preg_split`

Για παράδειγμα έστω `$ch1=$_POST['ch1']`, στη μεταβλητή `$ch1` αποθηκεύεται το `string` που τραβάει με `POST` από το 1ο πεδίο `textArea` στη σελίδα `search2`. Για να τοποθετήσουμε μία μία τις εγγραφές σε έναν πίνακα εφαρμόζουμε την `preg_split` ως εξής:

```
$chars1 = preg_split('/',, $ch1, -1)
```

Σαν 1ο όρισμα παίρνει το στοιχείο στο οποίο χωρίζει τις εγγραφές μεταξύ τους, σαν 2ο το `string` και το 3ο όρισμα μας λέει πόσους χαρακτήρες πριν το στοιχείο χωρίσματος να κρατήσει. Δηλαδή αν ο χρήστης βάλει στο 1ο πεδίο `textArea` τα στοιχεία `ENST00000369829, F11C1.6a.2` τότε η συνάρτηση `preg_split` τα βάλει στον πίνακα `$char1` στην 1η θέση το `ENST00000369829` και στη 2η θέση το `F11C1.6a.2`

```
$chars1[0]= ENST00000369829
```

```
$chars[1]= F11C1.6a.2
```

Στο σημείο αυτό έχουμε έναν πίνακα με όλες τις εγγραφές που εισήγαγε ο χρήστης για αναζήτηση καθώς επίσης και το όριο για το κατώφλι.

Σύμφωνα λοιπόν με τον κώδικα εφαρμόζουμε ένα βρόχο επανάληψης για κάθε στοιχείο του πίνακα. Μέσα σε αυτό το βρόχο κάνουμε την επερώτηση στη βάση δεδομένων `mira2` όπως φαίνεται στην εικόνα παρακάτω:

```
while( $i != count($chars1) ){  
  
    $query = "SELECT DISTINCT diana_microT_5_interactions.internal_tid, diana_microT_5_interactions.internal_mimat_id,  
        diana_microT_5_interactions.score  
        FROM diana_microT_5_interactions, diana_transcripts, diana_protein_genes, diana_protein_gene_refseqs,  
        diana_protein_gene_names  
        WHERE diana_microT_5_interactions.internal_tid = diana_transcripts.internal_tid  
        AND diana_transcripts.internal_gid = diana_protein_genes.internal_gid  
        AND diana_protein_genes.internal_gid = diana_protein_gene_refseqs.internal_gid  
        AND diana_protein_genes.internal_gid = diana_protein_gene_names.internal_gid  
        AND diana_transcripts.enstid = '$chars1[$i]'  
        AND diana_microT_5_interactions.score < $score";  
  
    $tempRow1[$i] = Yii::app()->getDb()->getCommandBuilder()->createSqlCommand($query)->query();  
    $i=$i+1;  
}
```

Εικόνα 13: Κώδικας για το query για την επιλογή Target

Δηλαδή επιλέγουμε να μας επιστραφούν όλα τα πεδία του πίνακα diana_microT5_interactions με βάση τους πίνακες των πεδίων που εισάγει ο χρήστης και ενώνουμε όλους τους πίνακες μεταξύ τους με τα κλειδιά που είχαμε περιγράψει στην ενότητα 3. Βάση δεδομένων

Όπως παρατηρούμε στη συνθήκη WHERE έχουμε βάλει πέρα από τις ενώσεις των πινάκων με βάση τα κλειδιά του και μια συνθήκη για το score. Με αυτή τη συνθήκη ουσιαστικά περιορίζουμε τα αποτελέσματα της αναζήτησης. Για παράδειγμα αν ο χρήστης επιλέξει στη σελίδα search2.php μια τιμή score 0.407 τότε θα μας επιστραφούν όλες οι εγγραφές που πληρούν τις συνθήκες WHERE από τους παραπάνω πίνακες για τις οποίες το score είναι < 0,407 .

Με το πέρας της αναζήτησης έχει αποθηκευτεί στη μεταβλητή temprow[i] το αποτέλεσμα για την i-οστή θέση του πίνακα \$chars1.

Στη συνέχεια εφαρμόσουμε μια ίδια επερώτηση στη βάση δεδομένων για τα στοιχεία του πίνακα \$chars2 και αποθηκεύουμε τα αποτελέσματα σε μια μεταβλητή \$temprow2[i].

Ουσιαστικά στο σημείο αυτό έχουμε 2 πίνακες \$temprow1 \$temprow2 οι οποίοι περιέχουν τα αποτελέσματα της αναζήτησης για τις εγγραφές που εισήγαγε ο χρήστης στα πεδία textArea της σελίδας search2.php.

Για την εμφάνιση των αποτελεσμάτων κάνουμε ανακατεύθυνση σε μια νέα σελίδα την results.php με χρήση της εντολής:

```
$this->render('results',array('tempRow1'=>$tempRow1,  
'tempRow2'=>$tempRow2, 'i'=>$i, 'j'=>$j ))
```

Σε επόμενο λοιπόν βήμα για να εμφανίσουμε τα αποτελέσματα στην οθόνη του χρήστη τρέχουμε τον παρακάτω κώδικα:

```
<?php  
  
echo '<br/>';  
echo '<br/>';  
  
?><table border="2" >  
  
    <tr>  
    <td><u><b> <?php echo 'internal_tid';?></b></u></td>  
    <td><u><b> <?php echo 'internal_mimat_id';?></b></u> </td>  
    <td><u><b> <?php echo 'score'; ?></b></u> </td>  
    </tr>  
  
<?php  
    $z=0;  
while($z < $i){  
    while(($row1=$tempRow1[$z]->read())!==false) {  
  
        ?>  
        <tr>  
        <td> <?php echo $row1['internal_tid'];?></td>  
        <td> <?php echo $row1['internal_mimat_id'];?></td>  
        <td> <?php echo $row1['score'];?></td>  
        </tr>  
        <?php  
    }  
    $z=$z+1;  
}  
  
$z=0;
```

```
while($z < $j) {
    while(($row1=$tempRow2[$z]->read())!==false) {
        ?>
        <tr>
        <td> <?php echo $row1['internal_tid'];?></td>
        <td> <?php echo $row1['internal_mimat_id'];?></td>
        <td> <?php echo $row1['score'];?></td>
        </tr>
        <?php
    }
    $z=$z+1;
}
?>
</table>
```

Εικόνα 14: Κώδικας της σελίδας results.php

Πιο συγκεκριμένα, δημιουργούμε κατάλληλους πίνακες για την εμφάνιση των αποτελεσμάτων και στη συνέχεια με τις εντολές:

```
while(($row1=$tempRow1[$z]->read())!==false)
while(($row2=$tempRow2[$z]->read())!==false)
```

διαβάζουμε τα στοιχεία των πινάκων temprow1, temprow2 και τυπώνουμε τα πεδία του πίνακα diana_microT5_interactions.

6.7.2. Περίπτωση Gene

Σε αυτό το υποκεφάλαιο εξετάζουμε την περίπτωση όπου ο χρήστης επιλέγει στο 1ο επίπεδο Gene και συνεπώς ενεργοποιείται το 2ο if της συνάρτησης public function actionSearch2() του sitecontroller. Σε αυτήν την περίπτωση γίνεται ανακατεύθυνση στη σελίδα search3.php της οποίας ο κώδικας εμφανίζεται στην παρακάτω εικόνα.

```
<?php
//Create a list that performs the "actionsearch", of the "SiteController".

echo CHtml::beginForm(array('site/result1'));
echo '<br/>';
echo '<br/>';
echo '<b>Here you can fill with one of the options bellow or a combination of them : </b>';
echo '<br/>';
echo '(diana_mature_mirnas$name or diana_mature_mirnas$mimat)';
echo '<br/>';
echo '<u>Please distinct the different options by using</u>: <b>,</b> ';
echo '<br/>';
echo '<br/>';
echo CHtml::textArea('ch1','');
echo '<br/>';
echo '<br/>';
echo '<b>Here you can fill with one of the options bellow or a combination of them : </b>';
echo '<br/>';
echo '(diana_hairpins$name or diana_hairpins$mima_id)';
echo '<br/>';
echo '<u>Please distinct the different options by using</u>: <b>,</b> ';
echo '<br/>';
echo '<br/>';

echo CHtml::textArea('ch2','');
echo '<br/>';
echo '<br/>';
echo '<b>Here you can fill with the option bellow : </b>';
echo '<br/>';
echo '(diana_transcription_factors$name)';
echo '<br/>';
echo '<u>Please distinct the different options by using</u>: <b>,</b> ';
echo '<br/>';
echo '<br/>';

echo CHtml::textArea('ch3','');
echo "<div class=\"row\">";
echo '<br/>';echo '<br/>';
echo CHtml::submitButton('Search');
echo "</div>";
echo CHtml::endForm();

?>
```

Εικόνα 15: Κώδικας της σελίδας search2.php

Όπως βλέπουμε από τον κώδικα στη συγκεκριμένη σελίδα εμφανίζονται 3 πεδία textArea.

Στο 1ο από αυτά ο χρήστης μπορεί να εισάγει εγγραφές τύπου :

diana_mature_mirnas\$name or diana_mature_mirnas\$mimat

Στο 2ο εγγραφές τύπου :

diana_hairpins\$name or diana_hairpins\$mima_id

Στο 3ο εγγραφές τύπου :

diana_transcription_factors\$name

Όμοια με την ανάλυση που κάναμε για την επιλογή Target έτσι και εδώ μόλις πατήσει ο χρήστης το submitButton τότε στέλνονται όσα έχει εισάγει ο χρήστης στη συνάρτηση public function ActionResult1() του sitecontroller η οποία τραβάει τα δεδομένα με τη μέθοδο POST.

```
$ch1=$_POST['ch1']
```

```
$ch2=$_POST['ch2']
```

```
$ch3=$_POST['ch3']
```

Στη συνέχεια για να απομονώσει τις εγγραφές του χρήστη χρησιμοποιεί τη συνάρτηση preg_split και αποθηκεύει σε κάθε θέση των πινάκων \$chars1, \$chars2, \$chars3 τα δεδομένα που εισήγαγε ο χρήστης στα 3 πεδία textArea.

Ο κώδικας της συνάρτησης public function ActionResult1() είναι παρόμοιος σε λειτουργία με τον κώδικα της συνάρτησης public function ActionResult() και φαίνεται παρακάτω:


```
public function actionResult1()
{
    $ch1=$_POST['ch1'];
    $ch2=$_POST['ch2'];
    $ch3=$_POST['ch3'];

    $chars1 = preg_split('/', '/', $ch1, -1);
    $chars2 = preg_split('/', '/', $ch2, -1);
    $chars3 = preg_split('/', '/', $ch3, -1);

    $i=0;
    $j=0;
    $x=0;

    while( $i != count($chars1) ){

        $query = "SELECT DISTINCT diana_transcripts.internal_tid, diana_transcripts.internal_gid,
diana_transcripts.enstid, diana_transcripts.species, diana_transcripts.strand, diana_transcripts.chromosome,
diana_transcripts.start, diana_transcripts.stop
        FROM diana_mature_mirnas, diana_microT_5_interactions, diana_transcripts
        WHERE diana_mature_mirnas.internal_mimat_id = diana_microT_5_interactions.internal_mimat_id
        AND diana_microT_5_interactions.internal_tid = diana_transcripts.internal_tid
        AND ( diana_mature_mirnas.name = '$chars1[$i]' OR diana_mature_mirnas.mimat = '$chars1[$i]' ) ";

        $tempRow1[$i] = Yii::app()->getDb()->getCommandBuilder()->createSqlCommand($query)->query();
        $i=$i+1;
    }

    while( $j != count($chars2) ){

        $query = "SELECT DISTINCT diana_transcripts.internal_tid, diana_transcripts.internal_gid,
diana_transcripts.enstid, diana_transcripts.species, diana_transcripts.strand,
diana_transcripts.chromosome, diana_transcripts.start, diana_transcripts.stop
        FROM diana_hairpins, diana_transcript_hairpin_conn, diana_transcripts
        WHERE diana_hairpins.internal_hairpin_id = diana_transcript_hairpin_conn.internal_hairpin_id
        AND diana_transcript_hairpin_conn.internal_tid = diana_transcripts.internal_tid
        AND ( diana_hairpins.name = '$chars2[$j]' OR diana_hairpins.mima_id = '$chars2[$j]' ) ";

        $tempRow2[$j] = Yii::app()->getDb()->getCommandBuilder()->createSqlCommand($query)->query();
        $j=$j+1;
    }
}
```

Εικόνα 16: Κώδικας της συνάρτησης ActionResult1 του Site Controller

Υλοποίηση online εφαρμογής που θα επιτρέπει εξελιγμένες αναζητήσεις σε βάση δεδομένων

```
while( $x != count($chars3) ){  
  
    $query = "SELECT DISTINCT diana_transcripts.internal_tid, diana_transcripts.internal_gid,  
    diana_transcripts.enstid, diana_transcripts.species, diana_transcripts.strand,  
    diana_transcripts.chromosome, diana_transcripts.start, diana_transcripts.stop  
        FROM diana_transcription_factors, diana_tfbs, diana_transcripts  
        WHERE diana_transcription_factors.internal_tf_id = diana_tfbs.internal_tf_id  
        AND diana_tfbs.internal_tid = diana_transcripts.internal_tid  
        AND diana_transcription_factors.name = '$chars3[$x]' ";  
  
    $tempRow3[$x] = Yii::app()->getDb()->getCommandBuilder()->createSqlCommand($query)->query();  
    $x=$x+1;  
}  
  
$this->render('results1',array('tempRow1'=>$tempRow1, 'tempRow2'=>$tempRow2, 'tempRow3'=>$tempRow3, 'i'=>$i, 'j'=>$j, 'x'=>$x ));  
}
```

Εικόνα 17: Κώδικας για το query για την επιλογή Gene

Όμοια λοιπόν με το υποκεφάλαιο 6.3.1. αποθηκεύουμε τα αποτελέσματα της αναζήτησης στη βάση δεδομένων στους πίνακες temprow1,temprow2,temprow3 και κάνουμε ανακατεύθυνση στη σελίδα results1.php.

Στη results1.php τρέχουμε με τη βοήθεια των εντολών:

```
while(($row1=$tempRow1[$z]->read())!==false)  
while(($row1=$tempRow2[$z]->read())!==false)  
while(($row1=$tempRow3[$z]->read())!==false)
```

τους πίνακες και εκτυπώνουμε όλα τα πεδία του πίνακα diana_transcripts όπως φαίνεται στον παρακάτω κώδικα:

```
<?php  
  
echo '<br/>';  
echo '<br/>';  
  
?>  
  
<table border="2" >  
  
  <tr>  
    <td><u><b> <?php echo 'internal_tid';?></b></u></td>  
    <td><u><b> <?php echo 'internal_gid';?></b></u> </td>  
    <td><u><b> <?php echo 'enstid';?></b></u> </td>  
    <td><u><b> <?php echo 'species';?></b></u> </td>  
    <td><u><b> <?php echo 'strand';?></b></u> </td>  
    <td><u><b> <?php echo 'chromosome';?></b></u> </td>  
    <td><u><b> <?php echo 'start';?></b></u> </td>  
    <td><u><b> <?php echo 'stop';?></b></u> </td>  
  </tr>
```

```
<?php  
$z=0;  
while($z < $i) {  
  while (($row1=$tempRow1[$z]->read())!==false) {  
    ?><tr>  
      <td><?php echo $row1['internal_tid']; ?> </td>  
      <td><?php echo $row1['internal_gid']; ?> </td>  
      <td><?php echo $row1['enstid']; ?> </td>  
      <td><?php echo $row1['species']; ?> </td>  
      <td><?php echo $row1['strand']; ?> </td>  
      <td><?php echo $row1['chromosome']; ?> </td>  
      <td><?php echo $row1['start']; ?> </td>  
      <td><?php echo $row1['stop']; ?> </td>  
    <?php  
  }  
  $z=$z+1;  
}
```

```
$z=0;
while($z < $j){
    while(($row1=$tempRow2[$z]->read())!==false) {
        ?><tr>
        <td><?php echo $row1['internal_tid']; ?> </td>
        <td><?php echo $row1['internal_gid']; ?> </td>
        <td><?php echo $row1['enstid']; ?> </td>
        <td><?php echo $row1['species']; ?> </td>
        <td><?php echo $row1['strand']; ?> </td>
        <td><?php echo $row1['chromosome']; ?> </td>
        <td><?php echo $row1['start']; ?> </td>
        <td><?php echo $row1['stop']; ?> </td>
        <?php
    }
    $z=$z+1;
}

$z=0;
while($z < $x){
    while(($row1=$tempRow3[$z]->read())!==false) {
        ?><tr>
        <td><?php echo $row1['internal_tid']; ?> </td>
        <td><?php echo $row1['internal_gid']; ?> </td>
        <td><?php echo $row1['enstid']; ?> </td>
        <td><?php echo $row1['species']; ?> </td>
        <td><?php echo $row1['strand']; ?> </td>
        <td><?php echo $row1['chromosome']; ?> </td>
        <td><?php echo $row1['start']; ?> </td>
        <td><?php echo $row1['stop']; ?> </td>
        <?php
    }
    $z=$z+1;
}
?>
</table>
```

Εικόνα 18: Κώδικας για τη σελίδα results1.php

6.7.3. Περίπτωση MatureMirna

Στο υποκεφάλαιο αυτό εξετάζουμε την περίπτωση που ο χρήστης θα επιλέξει στο 1ο επίπεδο MatureMirna και έτσι θα ενεργοποιηθεί το 3ο if της συνάρτησης public function actionSearch2() του site controller και θα γίνει ανακατεύθυνση στη σελίδα search4.php της οποίας ο κώδικας εμφανίζεται παρακάτω.

```
<?php
//Create a list that performs the "actionsearch", of the "SiteController".

echo CHtml::beginForm(array('site/result2'));
echo '<br/>';
echo '<br/>';
echo '<b>Here you can fill with one of the options bellow or a combination of them : </b>';
echo '<br/>';
echo '(diana_transcripts$enstid or diana_protein_genes$ensgid or diana_protein_genes$names or diana_protein_gene_refseqs$refseq_id or
echo '<br/>';
echo '<u>Please distinct the different options by using</u>: <b></b> ' ;
echo '<br/>';
echo '<br/>';
echo CHtml::textArea('ch1','');
echo "<div class=\"row\">";
echo '<br/>';
echo '<br/>';
echo CHtml::submitButton('Search');
echo "</div>";
echo CHtml::endForm();

?>
```

Εικόνα 19: Κώδικας της σελίδας search3.php

Όπως φαίνεται στον κώδικα δημιουργούμε μια φόρμα η οποία περιέχει ένα πεδίο textArea. Στο συγκεκριμένο πεδίο ο χρήστης μπορεί να εισάγει δεδομένα του τύπου:

diana_transcripts\$enstid or
diana_protein_genes\$ensgid or
diana_protein_genes\$names or
diana_protein_gene_refseqs\$refseq_id or
diana_protein_gene_names\$common_name

Μόλις εισάγει ο χρήστης τα δεδομένα που θέλει και πατήσει το submitButton τα δεδομένα στέλνονται στη συνάρτηση

public function actionResult2()

του site controller και εκτελείται ο παρακάτω κώδικας.

```
public function actionResult2()
{
    $ch1=$_POST['ch1'];

    $chars1 = preg_split('/',/, $ch1, -1);

    $i=0;

    while( $i != count($chars1) ){

        $query = "SELECT DISTINCT diana_mature_mirnas.internal_mimat_id, diana_mature_mirnas.name, diana_mature_mirnas.mimat,
        diana_mature_mirnas.sequence, diana_mature_mirnas.species
        FROM diana_mature_mirnas, diana_microT_5_interactions, diana_transcripts, diana_protein_genes,
        diana_protein_gene_refseqs, diana_protein_gene_names
        WHERE diana_mature_mirnas.internal_mimat_id = diana_microT_5_interactions.internal_mimat_id
        AND diana_microT_5_interactions.internal_tid = diana_transcripts.internal_tid
        AND diana_transcripts.internal_gid = diana_protein_genes.internal_gid
        AND diana_protein_genes.internal_gid = diana_protein_gene_refseqs.internal_gid
        AND diana_protein_genes.internal_gid = diana_protein_gene_names.internal_gid
        AND (diana_transcripts.enstid = '$chars1[$i]' OR diana_protein_genes.ensgid = '$chars1[$i]'
        OR diana_protein_genes.names = '$chars1[$i]' OR diana_protein_gene_refseqs.refseq_id = '$chars1[$i]'
        OR diana_protein_gene_names.common_name = '$chars1[$i]') ";

        $tempRow1[$i] = Yii::app()->getDb()->getCommandBuilder()->createSqlCommand($query)->query();
        $i=$i+1;
    }
}
```

Εικόνα 20: Κώδικας της συνάρτησης actionResult2 του Site Controller

Σύμφωνα λοιπόν με όσα περιγράψαμε για τις περιπτώσεις των υποκεφαλαίων 6.3.1 και 6.3.2 η συνάρτηση

public function actionResult2()

παίρνει τα δεδομένα από τη σελίδα search4.php με τη μέθοδο POST και με τη βοήθεια της συνάρτησης preg_split εισάγει τα δεδομένα που έβαλε ο χρήστης το καθένα σε μια θέση του πίνακα \$chars1. Στη συνέχεια κάνει την επερώτηση στη βάση δεδομένων με τους κατάλληλους περιορισμούς, κρατάει στον πίνακα temprow1 τα αποτελέσματα της επερώτησης στη βάση και στέλνει τον πίνακα στη σελίδα results2.php

Στον κώδικα της σελίδας results2.php όμοια με πριν με τη βοήθεια της εντολής `while(($row1=$tempRow1[$z]->read())!==false)` τρέχουμε τον πίνακα `temprow1` και τυπώνουμε στη συνέχεια τα περιεχόμενα του. Δηλαδή τυπώνουμε όλα τα πεδία του πίνακα `diana_mature_mirnas` με τις εγγραφές που επιστράφηκαν από την επερώτηση στη βάση δεδομένων βάσει των περιορισμών στη συνθήκη `WHERE` της επερώτησης.

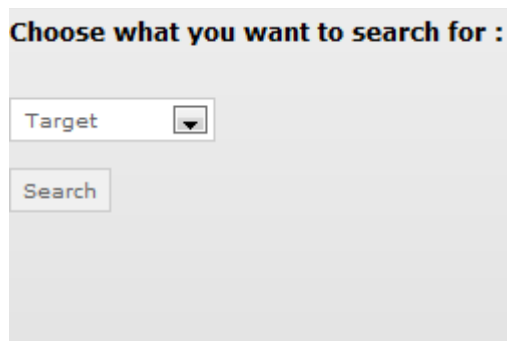
Για να κατανοηθούν πιο εύκολα όσα περιγράψαμε σε αυτό το κεφάλαιο ας διαβάσουμε τα παραδείγματα τρεξίματος του κώδικα που ακολουθούν.

7. Παραδείγματα τρεξίματος του κώδικα...

Στο κεφάλαιο αυτό θα δούμε τα αποτελέσματα που επιστρέφονται στην οθόνη του χρήστη για ενδεικτικά τρεξίματα του κώδικα που γράφτηκε στα πλαίσια της παρούσας διπλωματικής εργασίας.

7.1. Επιλογή Target

Έστω λοιπόν πως στο 1ο επίπεδο αναζήτησης ο χρήστης επιλέγει Target όπως φαίνεται στην εικόνα παρακάτω:



Εικόνα 21: Επιλογή Target

Μόλις πατήσει το κουμπί της αναζήτησης εμφανίζεται στο χρήστη μια εικόνα με 3 πεδία. Έστω λοιπόν πως ο χρήστης γεμίζει τα κενά πεδία όπως φαίνεται στην εικόνα παρακάτω:

Υλοποίηση online εφαρμογής που θα επιτρέπει εξελιγμένες αναζητήσεις σε βάση δεδομένων

Please give the threshold of the score that you want :

Here you can fill with one of the options bellow or a combination of them :
(diana_transcripts\$enstid or diana_protein_genes\$ensgid or diana_protein_genes\$names or diana_protein_gene_refseqs\$refseq_id or diana_protein_gene_names\$common_name)
Please distinct the different options by using: ,

Here you can fill with one of the options bellow or a combination of them :
(diana_mature_mirnas\$name or diana_mature_mirnas\$mimat)
Please distinct the different options by using: ,

Εικόνα 22: Εισαγωγή δεδομένων για επιλογή Target

Δηλαδή επιλέγει να βάλει στο 1ο πεδίο ένα score, στο 2ο πεδίο δύο διαφορετικά \$enstid από τον πίνακα diana_transcripts και στο 3ο πεδίο ένα \$name και ένα mimat από τον πίνακα diana_mature_mirnas.

Μόλις πατήσει το κουμπί της αναζήτησης θα του επιστραφούν στην οθόνη του τα αποτελέσματα που ζήτησε από τη βάση δεδομένων όπως φαίνεται παρακάτω:

Υλοποίηση online εφαρμογής που θα επιτρέπει εξελιγμένες αναζητήσεις σε βάση δεδομένων

internal tid	internal mimat id	score
6	7678	0.379
6	141	0.367
6	1751	0.396
6	8626	0.397
6	4676	0.399
6	4701	0.392
6	8858	0.389
6	9824	0.359
6	289	0.362
6	6555	0.389
6	5121	0.395
6	4965	0.362
6	8139	0.399
6	3565	0.389
6	9558	0.367
6	3586	0.368
6	8181	0.389
6	7399	0.397
6	6564	0.374
6	3468	0.354
6	9034	0.378
6	3365	0.401
6	4199	0.397
6	9906	0.383
6	5329	0.368
6	958	0.404
6	2085	0.365
6	8001	0.372
6	2742	0.403
6	9207	0.380

Εικόνα 23: Αποτελέσματα για επιλογή Target

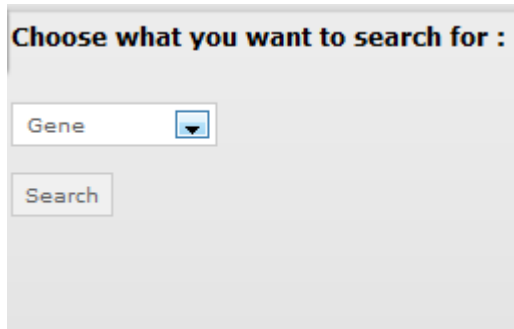
127207	1	0.371
84917	1	0.401
31800	1	0.389
10456	1	0.394
69001	1	0.381
163898	1	0.388
37163	1	0.399
69002	1	0.385
69006	1	0.376
111169	1	0.378
5184	1	0.400
47775	1	0.387
84929	1	0.405
116496	1	0.399
5194	1	0.406
21244	1	0.388
21246	1	0.372
58429	1	0.401
121841	1	0.383
137868	1	0.395
105838	1	0.373
132523	1	0.399
121846	1	0.373
100699	1	0.398
132526	1	0.399
58441	1	0.376
63789	1	0.372
37209	1	0.374
142971	1	0.376
100710	1	0.364
111206	1	0.374

Εικόνα 24: Αποτελέσματα για επιλογή Target

Τα αποτελέσματα που εμφανίζονται στην οθόνη του χρήστη μετά από την παραπάνω αναζήτηση που έγινε είναι πολλά και ενδεικτικά δείχνουμε στις παραπάνω εικόνες κάποια από αυτά. Όπως φαίνεται ο χρήστης επιλέγοντας στο 1ο επίπεδο της αναζήτησης, Target, περιμένει να του επιστραφούν από τον πίνακα `diana_microT5_interactions` τα πεδία: `internal_tid`, `internal_mimat_id` και `score`. Όπως φαίνεται στους πίνακες το `score` κυμαίνεται για όλες τις εγγραφές που επιστράφηκαν κάτω από 0.407 που εισήγαγε ο χρήστης σαν κατώφλι. Συνεπώς παρατηρούμε πως μας επιστράφηκαν τα σωστά αποτελέσματα από την αναζήτηση. Ακόμα για να το ελέγξουμε μπορούμε να μπούμε στη βάση δεδομένων και να ελέγξουμε τις εγγραφές που υπάρχουν στους πίνακες των αποτελεσμάτων με τα αντίστοιχα `score`.

7.2. Επιλογή Gene

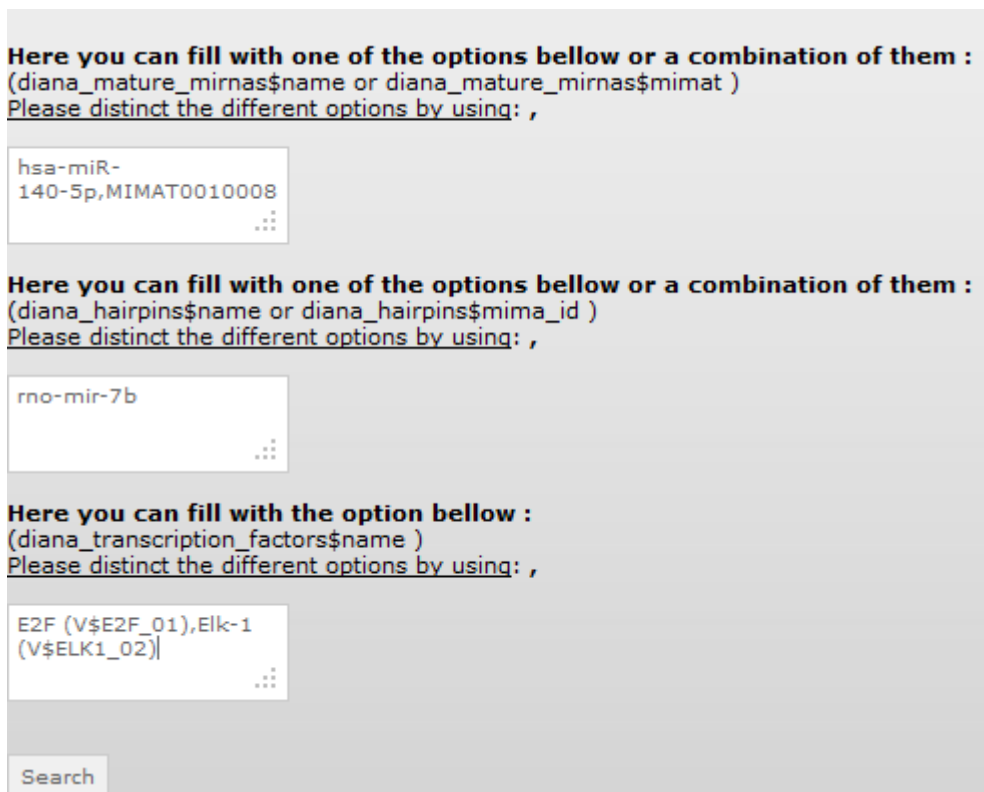
Έστω λοιπόν πως στο 1ο επίπεδο αναζήτησης ο χρήστης επιλέγει Gene όπως φαίνεται στην εικόνα παρακάτω:



The screenshot shows a web interface with the heading "Choose what you want to search for :". Below the heading is a dropdown menu with "Gene" selected. To the right of the dropdown is a small blue downward arrow. Below the dropdown is a "Search" button.

Εικόνα 25: Επιλογή Gene

Μόλις πατήσει το κουμπί της αναζήτησης εμφανίζεται στο χρήστη μια εικόνα με 3 πεδία. Έστω λοιπόν πως χρηστής γεμίζει τα κενά πεδία όπως φαίνεται στην εικόνα παρακάτω:



The screenshot shows a web interface with three input fields for search criteria. Each field has a heading and instructions. The first field is labeled "Here you can fill with one of the options bellow or a combination of them : (diana_mature_mirnas\$name or diana_mature_mirnas\$mimat)" and "Please distinct the different options by using: ,". The input field contains "hsa-miR-140-5p,MIMAT0010008". The second field is labeled "Here you can fill with one of the options bellow or a combination of them : (diana_hairpins\$name or diana_hairpins\$mima_id)" and "Please distinct the different options by using: ,". The input field contains "rno-mir-7b". The third field is labeled "Here you can fill with the option bellow : (diana_transcription_factors\$name)" and "Please distinct the different options by using: ,". The input field contains "E2F (V\$E2F_01),Elk-1 (V\$ELK1_02)". At the bottom of the form is a "Search" button.

Εικόνα 26: Εισαγωγή δεδομένων για επιλογή Gene

Υλοποίηση online εφαρμογής που θα επιτρέπει εξελιγμένες αναζητήσεις σε βάση δεδομένων

Δηλαδή επιλέγει να βάλει στο 1ο πεδίο ένα \$name και ένα \$mimat από τον πίνακα diana_mature_mirnas, στο 2ο πεδίο ένα \$name από τον πίνακα diana_hairpins και στο 3ο πεδίο 2 διαφορετικά \$name από τον πίνακα diana_transcriptions_factors.

Μόλις πατήσει το κουμπί της αναζήτησης θα του επιστραφούν στην οθόνη του τα αποτελέσματα που ζήτησε από τη βάση δεδομένων όπως φαίνεται παρακάτω:

internal tid	internal gid	enstid	species	strand	chromosome	start	stop
15879	28632	ENST00000371937	HOMO SAPIENS	-1	1	46871000	46906686
79466	104622	ENST00000313565	HOMO SAPIENS	-1	16	70451090	70475551
127207	28962	ENST00000290881	HOMO SAPIENS	-1	16	65767010	65775384
63747	74159	ENST00000324537	HOMO SAPIENS	1	15	81907287	82078495
84917	42634	ENST00000394530	HOMO SAPIENS	1	1	92187720	92252571
37159	107908	ENST00000376403	HOMO SAPIENS	-1	20	25702654	25791837
31800	21920	ENST00000377009	HOMO SAPIENS	-1	20	23379041	23381482
116481	47168	ENST00000293829	HOMO SAPIENS	1	17	7283413	7288971
53055	105328	ENST00000355797	HOMO SAPIENS	-1	17	72183404	72218651
10456	95856	ENST00000390656	HOMO SAPIENS	1	5	169864155	170096212
6	67334	ENST00000394810	HOMO SAPIENS	-1	10	75074653	75085838
105810	22541	ENST00000398783	HOMO SAPIENS	-1	3	51408338	51492884
69001	22404	ENST00000273283	HOMO SAPIENS	1	3	52786648	52801117
163898	29187	ENST00000297056	HOMO SAPIENS	-1	7	6415288	6454110
37163	52508	ENST00000310574	HOMO SAPIENS	-1	2	27275963	27288575
84919	25015	ENST00000286800	HOMO SAPIENS	1	21	29593091	29640340
69002	45427	ENST00000348993	HOMO SAPIENS	1	15	82975695	82985249
79473	52740	ENST00000303521	HOMO SAPIENS	-1	15	74459119	74984799
95335	56965	ENST00000296736	HOMO SAPIENS	-1	5	149352878	149360410
9	91570	ENST00000397019	HOMO SAPIENS	-1	1	43164115	43190639
142932	38819	ENST00000358481	HOMO SAPIENS	1	1	17571328	17600782
21224	13032	ENST00000369454	HOMO SAPIENS	-1	X	154140721	154147046
153453	66560	ENST00000383662	HOMO SAPIENS	1	3	123288120	123322672
47768	9078	ENST00000242375	HOMO SAPIENS	1	7	137411736	137453590
69006	44721	ENST00000382743	HOMO SAPIENS	-1	11	205031	226362
100678	49973	ENST00000274118	HOMO SAPIENS	1	4	187349972	187371605
127220	24030	ENST00000220325	HOMO SAPIENS	-1	15	39978933	40052038
37168	11454	ENST00000301608	HOMO SAPIENS	-1	17	71449191	71487039
26543	79562	ENST00000328195	HOMO SAPIENS	1	8	37738833	37756443
37171	96770	ENST00000338785	HOMO SAPIENS	1	21	37661749	37809550

Εικόνα 27: Αποτελέσματα για την επιλογή Gene

Υλοποίηση online εφαρμογής που θα επιτρέπει εξελιγμένες αναζητήσεις σε βάση δεδομένων

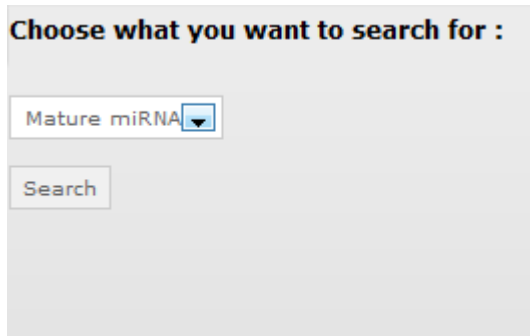
111169	63075	ENST00000218104	HOMO SAPIENS	1	X	152643517	152663410
5184	7971	ENST00000269503	HOMO SAPIENS	-1	18	68354895	68362542
95345	30990	ENST00000345080	HOMO SAPIENS	1	6	105511616	105637899
47775	45532	ENST00000330793	HOMO SAPIENS	-1	17	70048842	70053877
116490	21249	ENST00000252034	HOMO SAPIENS	1	7	73080367	73122173
15901	37398	ENST00000078652	HOMO SAPIENS	1	11	32871368	32958392
84929	4906	ENST00000320892	HOMO SAPIENS	1	2	6974974	7101761
53067	5636	ENST00000392693	HOMO SAPIENS	-1	11	125331927	125438397
121830	13174	ENST00000355377	HOMO SAPIENS	1	16	19087139	19186055
105829	56857	ENST00000309086	HOMO SAPIENS	1	3	113113169	113177784
63762	59530	ENST00000261520	HOMO SAPIENS	-1	15	58499100	58558636
163916	54119	ENST00000372927	HOMO SAPIENS	1	X	100242082	100305326
116496	34060	ENST00000392335	HOMO SAPIENS	1	19	19005471	19029978
132512	36064	ENST00000382859	HOMO SAPIENS	1	8	172382	187342
5194	19029	ENST00000377334	HOMO SAPIENS	1	5	171656129	171657985
153474	28853	ENST00000381059	HOMO SAPIENS	1	X	7770303	7772183
21244	49527	ENST00000380649	HOMO SAPIENS	-1	2	26267009	26321098
21246	9070	ENST00000356347	HOMO SAPIENS	1	4	196418	239769
163922	67486	ENST00000316081	HOMO SAPIENS	1	6	46763825	46780013
84936	37571	ENST00000367677	HOMO SAPIENS	-1	1	173249751	173259184
69025	47370	ENST00000264646	HOMO SAPIENS	-1	17	38105820	38150574
58429	39943	ENST00000376875	HOMO SAPIENS	-1	X	48201866	48210824
31824	95622	ENST00000330446	HOMO SAPIENS	1	19	46122010	46148403
121841	100297	ENST00000327490	HOMO SAPIENS	1	17	71773043	71778968
90266	12767	ENST00000288943	HOMO SAPIENS	-1	2	96172644	96174906
47792	96337	ENST00000307394	HOMO SAPIENS	1	16	1763230	1770629
137868	52694	ENST00000382927	HOMO SAPIENS	1	18	46659433	46728256
105838	78213	ENST00000225519	HOMO SAPIENS	-1	17	3458305	3486365
132523	51182	ENST00000375943	HOMO SAPIENS	1	1	15637525	15645740
42466	63457	ENST00000301463	HOMO SAPIENS	-1	12	51744656	51759437
95372	90405	ENST00000361470	HOMO SAPIENS	1	X	75564521	75568148

Εικόνα 28: Αποτελέσματα για την επιλογή Gene

Τα αποτελέσματα που εμφανίζονται στην οθόνη του χρήστη μετά από την παραπάνω αναζήτηση που έγινε είναι πολλά και ενδεικτικά δείχνουμε στις παραπάνω εικόνες κάποια από αυτά. Όπως φαίνεται ο χρήστης επιλέγοντας στο 1ο επίπεδο της αναζήτησης, Gene, περιμένει να του επιστραφούν από τον πίνακα `diana_transcripts` τα πεδία `internal_tid`, `internal_gid`, `enstid`, `species`, `strand`, `chromosome`, `start` και `stop`. Προφανώς θα μπορούσαμε να προσθέσουμε και τα υπόλοιπα πεδία του πίνακα, ωστόσο στα πλαίσια της παρούσας διπλωματικής εμφανίζουμε τα πρώτα 8 πεδία του πίνακα αυτού.

7.3.Επιλογή Mature Mirna

Έστω πως ο χρήστης επιλέγει στο 1ο επίπεδο Mature Mirna όπως φαίνεται παρακάτω:



The image shows a search interface with the heading "Choose what you want to search for :". Below the heading is a dropdown menu with "Mature miRNA" selected. To the right of the dropdown is a "Search" button.

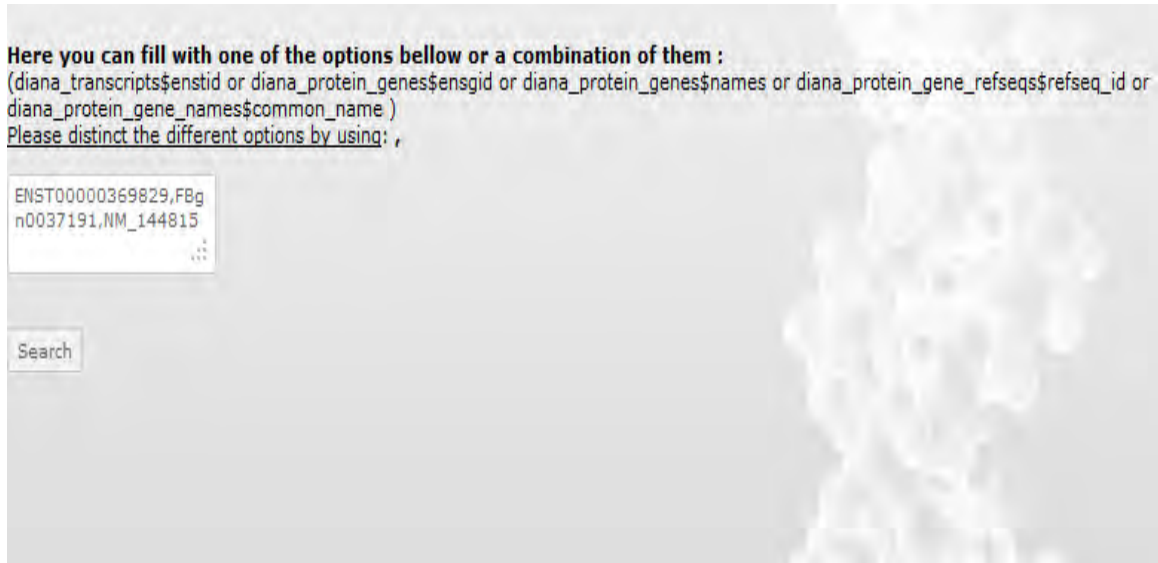
Εικόνα 29: Επιλογή Mature miRNA

Μόλις πατήσει το κουμπί της αναζήτησης θα του εμφανιστεί ένα πεδίο στο οποίο πρέπει να εισάγει όσες από τις παρακάτω μεταβλητές θέλει όπως περιγράψαμε σε προηγούμενο κεφάλαιο.

diana_transcripts\$enstid or
diana_protein_genes\$ensgid or
diana_protein_genes\$names or
diana_protein_gene_refseqs\$refseq_id or
diana_protein_gene_names\$common_name

Έστω λοιπόν πως ο χρήστης εισάγει τα παρακάτω δεδομένα:

Υλοποίηση online εφαρμογής που θα επιτρέπει εξελιγμένες αναζητήσεις σε βάση δεδομένων



Εικόνα 30: Εισαγωγή δεδομένων για επιλογή Mature miRNA

Δηλαδή στη συγκεκριμένη περίπτωση ο χρήστης έχει εισάγει μια εγγραφή \$enstid από τον πίνακα diana_transcripts, μια εγγραφή \$ensgid από τον πίνακα diana_protein_genes και μια εγγραφή \$refseq_id από τον πίνακα diana_protein_gene_refseqs.

Εφόσον έχουν εισαχθεί τα δεδομένα από το χρήστη, με το πάτημα της αναζήτησης περιμένει να του εμφανισθούν τα δεδομένα στην οθόνη όμοια με το προηγούμενο παράδειγμα.

Ωστόσο στη συγκεκριμένη περίπτωση αντιμετωπίζουμε θέμα μνήμης και αναμένοντας τα αποτελέσματα της αναζήτησης στη βάση δεδομένων μας εμφανίζεται το συγκεκριμένο μήνυμα:

```
Fatal error: Maximum execution time of 60 seconds exceeded in C:\xampp\htdocs\yii-1.0.8.r1317\framework\db\CDbCommand.php on line 302
```

Αυτό σημαίνει ότι δεν αρκεί η συγκεκριμένη μνήμη που μας παρέχει ο εικονικός server xampp αλλά χρειάζεται επιπλέον μνήμη για να τρέξει το πρόγραμμα.

Το πρόβλημα εντοπίζεται στο σημείο του κώδικα που προσπαθούμε να κρατήσουμε στη μεταβλητή query το αποτέλεσμα της αναζήτησης στη βάση όπως φαίνεται στην εικόνα παρακάτω:


```
while( $i != count($chars1) ){  
  
$query = "SELECT DISTINCT diana_mature_mirnas.internal_mimat_id, diana_mature_mirnas.name,  
diana_mature_mirnas.mimat, diana_mature_mirnas.sequence, diana_mature_mirnas.species  
FROM diana_mature_mirnas, diana_microT_5_interactions, diana_transcripts, diana_protein_genes,  
diana_protein_gene_refseqs, diana_protein_gene_names  
WHERE diana_mature_mirnas.internal_mimat_id = diana_microT_5_interactions.internal_mimat_id  
AND diana_microT_5_interactions.internal_tid = diana_transcripts.internal_tid  
AND diana_transcripts.internal_gid = diana_protein_genes.internal_gid  
AND diana_protein_genes.internal_gid = diana_protein_gene_refseqs.internal_gid  
AND diana_protein_genes.internal_gid = diana_protein_gene_names.internal_gid  
AND (diana_transcripts.enstid = '$chars1[$i]' OR diana_protein_genes.ensgid = '$chars1[$i]' OR  
diana_protein_genes.names = '$chars1[$i]' OR diana_protein_gene_refseqs.refseq_id = '$chars1[$i]'  
OR diana_protein_gene_names.common_name = '$chars1[$i]') ";  
  
$tempRow1[$i] = Yii::app()->getDb()->getCommandBuilder()->createSqlCommand($query)->query();  
$i=$i+1;  
}
```

Εικόνα 31: Κώδικας query για επιλογή Mature miRNA

Στο σημείο αυτό προσπαθεί να κάνει πολλές συνενώσεις πινάκων και ο απαιτούμενος χρόνος που χρειάζεται για αυτό το σκοπό ξεπερνάει τα ανώτερο όριο χρόνου.

Για την επίλυση του προβλήματος προσπαθήσαμε να κάνουμε τον κώδικα όσο πιο απλοποιημένο γίνεται. Σκεφτήκαμε λοιπόν να αλλάξουμε τον παραπάνω κώδικα και να τοποθετήσουμε το εξωτερικό while που τρέχει για όλο τον πίνακα \$chars1 στο εσωτερικό του query, δηλαδή κάτω από τις συνθήκες WHERE.

Ωστόσο παρατηρήσαμε ότι δεν άλλαξε τίποτα καθώς το πρόβλημα εντοπίζεται καθαρά στη συνένωση πολλών πινάκων οι οποίοι περιέχουν πολλές εγγραφές έκαστος. Συνεπώς επειδή ακολουθήσαμε ίδια λογική στη συγγραφή του κώδικα με την επιλογή Gene γνωρίζουμε ότι η λογική μας είναι σωστή και πως τα αποτελέσματα θα εμφανιστούν σωστά στην οθόνη του χρήστη σε κάποιο μηχάνημα που χρησιμοποιεί πραγματικό server και όχι εικονικό.

Ωστόσο για να επαληθεύσουμε τον κώδικα μπορούμε να βγάλουμε κάποιους περιορισμούς. Έστω λοιπόν ότι τρέχουμε τον κώδικά μας μονάχα για τον έλεγχο: diana_transcripts.enstid = '\$chars1[\$i]'

και όχι για όλες τις συνθήκες OR όπως φαίνεται στον πίνακα

Αν τρέξουμε ένα νέο παράδειγμα :

Here you can fill with one of the options bellow or a combination of them :
(diana_transcripts\$enstid or diana_protein_genes\$ensgid or diana_protein_genes\$names or diana_protein_gene_refseqs\$refseq_id or
diana_protein_gene_names\$common_name)
Please distinct the different options by using: ,

Εικόνα 32: Εισαγωγή δεδομένων για επιλογή Mature miRNA

Όπου ο χρήστης εισάγει δύο διαφορετικά \$enstid από τον πίνακα diana_transcripts, επιστρέφονται τα παρακάτω αποτελέσματα στο χρήστη:

Υλοποίηση online εφαρμογής που θα επιτρέπει εξελιγμένες αναζητήσεις σε βάση δεδομένων

internal mimat id	name	mimat	sequence	species
8169	mmu-let-7g-star	MIMAT0004519	ACUGUACAGGCCACUGCCUUGC	MUS MUSCULUS
2562	mmu-miR-103	MIMAT0000546	AGCAGCAUUGUACAGGGCUAUGA	MUS MUSCULUS
564	mmu-miR-105	MIMAT0004856	CCAAGUGCUCAGAUGCUUGUGGU	MUS MUSCULUS
3576	mmu-miR-106a	MIMAT0000385	CAAAGUGCUAACAGUGCAGGUAG	MUS MUSCULUS
8151	mmu-miR-106b	MIMAT0000386	UAAAGUGCUGACAGUGCAGAU	MUS MUSCULUS
9159	mmu-miR-107	MIMAT0000647	AGCAGCAUUGUACAGGGCUAUCA	MUS MUSCULUS
4393	mmu-miR-1199	MIMAT0005860	UCUGAGUCCCGUCGCGCGG	MUS MUSCULUS
4567	mmu-miR-122	MIMAT0000246	UGGAGUGUGACAAUGGUGUUUG	MUS MUSCULUS
2739	mmu-miR-125a-3p	MIMAT0004528	ACAGGUGAGGUUCUUGGGAGCC	MUS MUSCULUS
3323	mmu-miR-125a-5p	MIMAT0000135	UCCUGAGACCCUUUAACCUGUGA	MUS MUSCULUS
7882	mmu-miR-125b-5p	MIMAT0000136	UCCUGAGACCCUAACUUGUGA	MUS MUSCULUS
2447	mmu-miR-128	MIMAT0000140	UCACAGUGAACCGGUCUCUUU	MUS MUSCULUS
6739	mmu-miR-130b-star	MIMAT0004583	ACUCUUUCCUGUUGCACUACU	MUS MUSCULUS
4012	mmu-miR-133a	MIMAT0000145	UUUGGUCCCCUUAACCAGCUG	MUS MUSCULUS
8778	mmu-miR-133b	MIMAT0000769	UUUGGUCCCCUUAACCAGCUA	MUS MUSCULUS
216	mmu-miR-134	MIMAT0000146	UGUGACUGGUUGACCAGAGGGG	MUS MUSCULUS
3920	mmu-miR-135a	MIMAT0000147	UAUGGCUUUUUAUCCUAUGUGA	MUS MUSCULUS
2672	mmu-miR-135b	MIMAT0000612	UAUGGCUUUUCAUCCUAUGUGA	MUS MUSCULUS
3017	mmu-miR-137	MIMAT0000149	UUUUUGCUUAAGAAUACGCGUAG	MUS MUSCULUS
3403	mmu-miR-140	MIMAT0000151	CAGUGGUUUUACCCUAUGGUAG	MUS MUSCULUS
5233	mmu-miR-141-star	MIMAT0004533	CAUCUCCAGUGCAGUGUUGGA	MUS MUSCULUS
5803	mmu-miR-143	MIMAT0000247	UGAGAUGAAGCACUGUAGCUC	MUS MUSCULUS
9641	mmu-miR-145	MIMAT0000157	GUCCAGUUUUCCAGGAUCCCU	MUS MUSCULUS
7328	mmu-miR-145-star	MIMAT0004534	AUCCUGGAAAUACUGUUCUUG	MUS MUSCULUS
9661	mmu-miR-148a	MIMAT0000516	UCAGUGCACUACAGAACUUUGU	MUS MUSCULUS
7645	mmu-miR-148a-star	MIMAT0004617	AAAGUUCUGAGACACUCCGACU	MUS MUSCULUS
5705	mmu-miR-154	MIMAT0000164	UAGGUUAUCCGUGUUGCCUUCG	MUS MUSCULUS
1802	mmu-miR-15a	MIMAT0000526	UAGCAGCACAUAAUGGUUUUGUG	MUS MUSCULUS
4813	mmu-miR-15b	MIMAT0000124	UAGCAGCACAUCAUGGUUUACA	MUS MUSCULUS
4663	mmu-miR-15b-star	MIMAT0004521	CGAAUCAUUUUUGCUGCUCUA	MUS MUSCULUS

Εικόνα 33: Αποτελέσματα για την επιλογή Mature miRNA

Ενδεικτικά εμφανίζουμε τις πρώτες εγγραφές που δόθηκαν στο χρήστη ως αποτελέσματα της αναζήτησης στη βάση δεδομένων για τα στοιχεία που εισήγαγε στην επιλογή Mature Mirnas.

Παρατηρούμε λοιπόν ότι ο κώδικας μας τρέχει σωστά και απλά είναι θέμα μνήμης η αδυναμία να εμφανίσει αποτελέσματα για όλες τις συνθήκες OR. Ο χρήστης από την παραπάνω αναζήτηση περιμένει να του επιστραφούν από τον πίνακα `diana_mature_mirnas` τα πεδία: `internal_mimat_id`, `name`, `mimat`, `sequence` και `species`.

8. Επίλογος

Στο κεφάλαιο αυτό συνοψίζεται η παρούσα διπλωματική εργασία. Παρουσιάζονται οι βασικοί στόχοι της που επετεύχθησαν και επισημαίνονται οι αρχικές απαιτήσεις της. Ακόμα, τονίζονται τα διάφορα συμπεράσματα που προέκυψαν καθώς επίσης και οι μελλοντικές επεκτάσεις που σημειώθηκαν για την ανάπτυξη και βελτίωση της συγκεκριμένης εργασίας.

8.1. Σύνοψη και συμπεράσματα

Όπως έχει αναφερθεί και στην εισαγωγή, η μελέτη των microRNAs διαδραματίζει καθοριστικό ρόλο στην ανάπτυξη των οργανισμών και στην πρόβλεψη πολλών ασθενειών.

Στη συγκεκριμένη εργασία αναπτύχθηκε ένα σύστημα αναζήτησης βιολογικών πληροφοριών. Το κοινό στο οποίο απευθύνεται η εργασία αποτελείται από βιολόγους ερευνητές, αποδεικνύοντας με αυτόν τρόπο πόσο σημαντική είναι η συνεργασία του κλάδου της βιολογίας με τον κλάδο της πληροφορικής.

Η εφαρμογή που αναπτύχθηκε καλείται να συνδράμει στην έρευνα των microRNA παρέχοντας πρόσβαση σε μια βάση δεδομένων. Η συγκεκριμένη βάση δεδομένων περιλαμβάνει ένα πλήθος από σημεία πρόσδεσης των μορίων microRNA πάνω στα μόρια mRNA. Στη συγκεκριμένη διπλωματική αναπτύχθηκε αλγόριθμος αναζήτησης στα δεδομένα ώστε ο χρήστης να μπορεί να εφαρμόσει μια απλή ή και μια σύνθετη αναζήτηση στην βάση. Η αναζήτηση αυτή μπορεί να αφορά βιολογικούς στόχους ενός γονιδίου, ενός microRNA ή ακόμα και ενός συνδυασμού αυτών.

Κατά τη διάρκεια της συγγραφής παρουσιάστηκαν διάφορα ζητήματα σχετικά με την απόδοση του συστήματος. Συγκεκριμένα στην προσπάθειά μας να τρέξουμε δοκιμαστικά παραδείγματα παρατηρήσαμε ότι δημιουργούνταν προβλήματα μνήμης καθώς η βάση δεδομένων mirna2 περιέχει πάρα πολλές εγγραφές δεδομένων και η αναζήτηση στους πίνακες καθυστερούσε για μεγάλο χρονικό διάστημα. Ωστόσο για να μπορέσουμε να δούμε αποτελέσματα από τις αναζητήσεις μας αφαιρέσαμε κάποιες αναζητήσεις και κρατήσαμε ενδεικτικά κάποιες άλλες. Για παράδειγμα στο 1^ο επίπεδο επιλέγοντας ο χρήστης Target κρατήσαμε μονάχα την αναζήτηση για δεδομένα \$enstid του πίνακα diana_transcripts. Παρατηρήσαμε ότι το πρόβλημα μνήμης που εμφανιζόταν αφορούσε τα διάφορα OR που χρησιμοποιούσαμε στις συνθήκες WHERE του query στη

βάση δεδομένων, συνεπώς κρατήσαμε ενδεικτικά τα παραπάνω για να επαληθεύσουμε τον κώδικά μας.

8.2 Μελλοντικές Επεκτάσεις

Η συγκεκριμένη εφαρμογή που αναπτύχθηκε στα πλαίσια της παρούσας διπλωματικής εργασίας έχει στόχο όπως αναφέραμε την εξυπηρέτηση των αναγκών ερευνητών βιολόγων με τον καλύτερο δυνατό τρόπο. Συνεπώς εφαρμογές τέτοιου είδους απαιτούν τη συνεχή ενημέρωσή τους για βελτιστοποιήσεις. Έχουμε ήδη σκεφτεί κάποιες από αυτές όπως την αποθήκευση κάποιων αναζητήσεων στη βάση δεδομένων από χρήστες. Για παράδειγμα με την είσοδο του χρήστη στο site του DIANA LAB όταν αυτός επιλέξει την συγκεκριμένη εφαρμογή και του επιστραφούν κάποια αποτελέσματα να μπορεί να τα αποθηκεύει για να τα κάνει χρήση στο μέλλον. Ουσιαστικά δηλαδή να είναι σε θέση να έχει κάθε χρήστης ένα χώρο «αγαπημένων» αναζητήσεων και να μην αναγκάζεται να κάνει εκ νέου αναζητήσεις με τις οποίες έχει ασχοληθεί ξανά.

Ακόμα, όσον αφορά στη συγγραφή του κώδικα θα μπορούσε να χρησιμοποιηθεί ένα αρχείο XML για να αποφύγουμε τη συγγραφή πολλών διαφορετικών σελίδων. Με τη χρήση αυτού, διευκολύνουμε όποια μελλοντική επέκταση εφόσον θα είναι πιο εύκολα διαχειρίσιμο ένα τέτοιο αρχείο από επόμενους προγραμματιστές.

9. Βιβλιογραφία

- Alexiou et al. - Nucleic acids research - 2010 - miRGen 2.0 a database of microRNA genomic information and regulation
- Database_paper_v2.4
- Maragkakis et al. - Nucleic Acids Research - 2009 - DIANA-microT web server elucidating microRNA functions through target prediction
- Maragkakis M, Reczko M, Simossis VA, Alexiou P, Papadopoulos GL, Dalamagas T, Giannopoulos G, Goumas G, Koukis E, Kourtis K, Vergoulis T, Koziris N, Sellis T, Tsanakas P, Hatzigeorgiou AG. DIANA-microT web server: elucidating microRNA functions through target prediction. Nucleic Acids Res. 2009 Jul 1; 37(Web Server issue):W273-6. Epub 2009 Apr 30. PubMed PMID: 19406924; PubMed Central PMCID: PMC2703977.
- Sethupathy P, Corda B, Hatzigeorgiou AG. TarBase: A comprehensive database of experimentally supported animal microRNA targets. RNA. 2006 Feb; 12(2):192-7. Epub 2005 Dec 22. PubMed PMID: 16373484; PubMed Central PMCID: PMC1370898.
- Alexiou P, Vergoulis T, Gleditsch M, Prekas G, Dalamagas T, Megraw M, Grosse I, Sellis T, Hatzigeorgiou AG. miRGen 2.0: a database of microRNA genomic information and regulation. Nucleic Acids Res. 2010 Jan; 38(Database issue):D137-41. Epub 2009 Oct 22. PubMed PMID: 19850714; PubMed Central PMCID: PMC2808909.
- 17: Papadopoulos GL, Alexiou P, Maragkakis M, Reczko M, Hatzigeorgiou AG. DIANA-mirPath: Integrating human and mouse microRNAs in pathways. Bioinformatics. 2009 Aug 1; 25(15):1991-3. Epub 2009 May 12. PubMed PMID: 19435746.
- <http://www.yiiframework.com/doc/api/>
- <http://php.net/docs.php>
- <http://www.isth.gr/?news=3021>
- <http://www.news-medical.net/news/20110118/4/Greek.aspx%28mirnas%29>