



ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ  
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ  
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗ  
ΒΙΟΙΑΤΡΙΚΗ



ΕΛΛΗΝΙΚΟ ΙΝΣΤΙΤΟΥΤΟ ΠΑΣΤΕΡ  
INSTITUT PASTEUR HELLENIQUE

**«ΒΙΟΠΛΗΡΟΦΟΡΙΚΗ  
ΑΝΑΛΥΣΗ ΤΗΣ ΠΡΩΤΕΪΝΗΣ  
ΠΟΥ ΕΚΦΡΑΖΕΤΑΙ ΑΠΟ ΤΟ  
ΓΟΝΙΔΙΟ LdBK\_352470.1 ΤΟΥ  
ΠΡΩΤΟΖΩΟΥ *Leishmania  
donovani*»**

Πετσανά Μαρίνα

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Υπεύθυνοι:

κ. Μπράλιου Γεωργία  
Επίκουρος Καθηγήτρια

κ. Μπολέτη Χαραλαμπία  
Κύρια Ερευνήτρια

**Λομιά, 2020-2021**



ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ  
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ  
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΒΙΟΙΑΤΡΙΚΗ



ΕΛΛΗΝΙΚΟ ΙΝΣΤΙΤΟΥΤΟ ΠΑΣΤΕΡ  
INSTITUT PASTEUR HELLENIQUE

**«ΒΙΟΠΛΗΡΟΦΟΡΙΚΗ  
ΑΝΑΛΥΣΗ ΤΗΣ ΠΡΩΤΕΪΝΗΣ  
ΠΟΥ ΕΚΦΡΑΖΕΤΑΙ ΑΠΟ ΤΟ  
ΓΟΝΙΔΙΟ LdBK\_352470.1 ΤΟΥ  
ΠΡΩΤΟΖΩΟΥ *Leishmania  
donovani*»**

**Πετσανά Μαρίνα**

**ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ**

**Επιβλέπουσες:**

**κ. Μπράλιου Γεωργία  
Επίκουρος Καθηγήτρια**

**κ. Μπολέτη Χαραλαμπία  
Κύρια Ερευνήτρια**

**Λαμία, 2020-2021**

Με ατομική μου ευθύνη και γνωρίζοντας τις κυρώσεις <sup>(1)</sup>, που προβλέπονται από της διατάξεις της παρ. 6 του άρθρου 22 του Ν. 1599/1986, δηλώνω ότι:

1. Δεν παραθέτω κομμάτια βιβλίων ή άρθρων ή εργασιών άλλων αυτολεξεί **χωρίς να τα περικλείω σε εισαγωγικά** και χωρίς να αναφέρω το συγγραφέα, τη χρονολογία, τη σελίδα. Η αυτολεξεί παράθεση χωρίς εισαγωγικά χωρίς αναφορά στην πηγή, είναι λογοκλοπή. Πέραν της αυτολεξεί παράθεσης, λογοκλοπή θεωρείται και η παράφραση εδαφίων από έργα άλλων, συμπεριλαμβανομένων και έργων συμφοιτητών μου, καθώς και η παράθεση στοιχείων που άλλοι συνέλεξαν ή επεξεργάστηκαν, χωρίς αναφορά στην πηγή. Αναφέρω πάντοτε με πληρότητα την πηγή κάτω από τον πίνακα ή σχέδιο, όπως στα παραθέματα.
2. Δέχομαι ότι η αυτολεξεί **παράθεση χωρίς εισαγωγικά**, ακόμα κι αν συνοδεύεται από αναφορά στην πηγή σε κάποιο άλλο σημείο του κειμένου ή στο τέλος του, είναι αντιγραφή. Η αναφορά στην πηγή στο τέλος π.χ. μιας παραγράφου ή μιας σελίδας, δεν δικαιολογεί συρραφή εδαφίων έργου άλλου συγγραφέα, έστω και παραφρασμένων, και παρουσίασή τους ως δική μου εργασία.
3. Δέχομαι ότι υπάρχει επίσης περιορισμός στο μέγεθος και στη συχνότητα των παραθεμάτων που μπορώ να εντάξω στην εργασία μου εντός εισαγωγικών. Κάθε μεγάλο παράθεμα (π.χ. σε πίνακα ή πλαίσιο, κλπ), προϋποθέτει ειδικές ρυθμίσεις, και όταν δημοσιεύεται προϋποθέτει την άδεια του συγγραφέα ή του εκδότη. Το ίδιο και οι πίνακες και τα σχέδια
4. Δέχομαι όλες τις συνέπειες σε περίπτωση λογοκλοπής ή αντιγραφής.

Ημερομηνία: 29/03/2021

Ο – Η Δηλ.

(Υπογραφή)

(1) «Όποιος εν γνώσει του δηλώνει ψευδή γεγονότα ή αρνείται ή αποκρύπτει τα αληθινά με έγγραφη υπεύθυνη δήλωση του άρθρου 8 παρ. 4 Ν. 1599/1986 τιμωρείται με φυλάκιση τουλάχιστον τριών μηνών. Εάν ο υπαίτιος αυτών των πράξεων σκόπευε να προσπορίσει στον εαυτόν του ή σε άλλον περιουσιακό όφελος βλάπτοντας τρίτον ή σκόπευε να βλάψει άλλον, τιμωρείται με κάθειρξη μέχρι 10 ετών.

**«ΒΙΟΠΛΗΡΟΦΟΡΙΚΗ  
ΑΝΑΛΥΣΗ ΤΗΣ ΠΡΩΤΕΪΝΗΣ  
ΠΟΥ ΕΚΦΡΑΖΕΤΑΙ ΑΠΟ ΤΟ  
ΓΟΝΙΔΙΟ LdBK\_352470.1 ΤΟΥ  
ΠΡΩΤΟΖΩΟΥ *Leishmania  
donovani*»**

**Πετσανά Μαρίνα**

**Τριμελής Επιτροπή:**

κ. Μπράλιου Γεωργία, Επίκουρος καθηγήτρια

κ. Μπάγκος Παντελεήμων, Καθηγητής-Κοσμήτορας

κ. Χατζηγεωργίου Άρτεμις, Καθηγήτρια

## ΠΕΡΙΕΧΟΜΕΝΑ

ΠΕΡΙΛΙΨΗ.....	7
ΠΡΟΛΟΓΟΣ.....	7
ΕΥΧΑΡΙΣΤΙΕΣ.....	8
ΕΙΣΑΓΩΓΗ.....	8
ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΙ.....	17
<b>ΑΠΟΤΕΛΕΣΜΑΤΑ.....</b>	<b>21</b>
ΣΥΖΗΤΗΣΗ.....	44
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	47

## ΠΕΡΙΛΨΗ

Η παρούσα εργασία, αποτελεί Βιοπληροφορική ανάλυση της πρωτεΐνης που εκφράζεται από το γονίδιο *LdBPK\_352470.1* του πρωτόζωου *Leishmania donovani*. Η εργασία ως κύριο σκοπό φέρει την πρόβλεψη της δομής και της λειτουργίας της πρωτεΐνης αυτής, καθώς και το χαρακτηρισμό της, από την σκοπιά της Βιοπληροφορικής, ως Νεξίνη Διαλογής (Sorting Nexin). Αυτό, θα συνεισφέρει στον σχεδιασμό των αντίστοιχων εργαστηριακών πειραμάτων. Επίσης, επόμενο στόχο της εργασίας, αποτελεί η εμβάθυνση της ανάλυσης, και η επέκταση που έγκειται στον εντοπισμό και τη δομική κατάταξη όλων των πρωτεϊνών από τις Κινετοπλαστίδες, που περιέχουν λειτουργική περιοχή πρόσδεσης σε φωσφοϊνοσιτίδια PX.

Το τελευταίο, μπορεί να οδηγήσει σε μελλοντική αναθεώρηση των χαρακτηρισμών των πρωτεϊνών αυτών, μέσω των αποτελεσμάτων που ευρέθησαν και την επεξεργασία τους.

## ΠΡΟΛΟΓΟΣ

Η λείσμανίαση είναι ομάδα νόσων που προκαλούνται από τα είδη ενδοκυτταρικών παρασιτικών πρωτοζώων, του γένους *Leishmania* (Λείσμάνια). Οι πιο κοινές μορφές της ασθένειας είναι η σπλαχνική (καλά-αζάρ), η δερματική και η βλεννογονοδερματική. Η σπλαχνική λείσμανίαση είναι η σοβαρότερη μορφή της νόσου, προσβάλλει πολλά εσωτερικά όργανα (συνήθως το σπλήνα, το ήπαρ και το μυελό των οστών) και πιθανώς είναι θανατηφόρος αν δεν θεραπευτεί με χορήγηση χημειοθεραπείας. Υπάρχουν και ασυμπτωματικές μορφές.

Η παρούσα εργασία αποτελεί βιοπληροφορική ανάλυση πρωτεΐνης που μελετάται στο Εργαστήριο Ενδοκυτταρικού Παρασιτισμού, στο Ελληνικό Ινστιτούτο Παστέρ υπό την επίβλεψη της Δρ. Χαραλαμπίας Μπολέτη. Η πρωτεΐνη αυτή εκκρίνεται από το πρωτόζωο *Leishmania donovani* (Silverman et al 2008), και αποτελεί υποψήφιο μολυσματικό παράγοντα. Το προϊόν του γονιδίου *LdBPK\_352470.1*, υπάρχει ως καταχώρηση στη βάση δεδομένων UNIPROT με το όνομα E9BSB7. Στο εργαστήριο επιβεβαιώθηκε η έκφραση της πρωτεΐνης σε προμαστιγωτά παράσιτα *L. donovani* και ο μεταπτυχιακός φοιτητής κ. Δρόσος Κουρούνης υπό την επίβλεψη της ερευνήτριας κ. Χαραλαμπίας Μπολέτη σύμφωνα με τη διατριβή του πραγματοποίησε:

- α) την κλωνοποίηση του γονιδίου *LdBPK\_352470.1* σε διαφορετικούς πλασμιδιακούς φορείς
- β) την έκφραση της ανασυνδυασμένης πρωτεΐνης σε βακτήρια και κύτταρα θηλαστικών
- γ) τον εντοπισμό της ενδογενούς πρωτεΐνης σε κύτταρα *Leishmania* και σε μολυσμένα μακροφάγα, χρησιμοποιώντας το ειδικό πολυκλωνικό αντίσωμα anti-*LdPIBpNex* που παράχθηκε στο εργαστήριο,
- δ) τον εμπλουτισμό της πρωτεΐνης σε διαλυτά πρωτεϊνικά κλάσματα από αξενικά (καλλιεργημένα) προμαστιγωτά παράσιτα *L. donovani* και

ε) την έκκριση της πρωτεΐνης από καλλιέργεια προμαστιγωτών, στο εξωκυττάριο θρεπτικό μέσο.<sup>1</sup>

Στη συνέχεια, τέθηκε το ζήτημα του πειραματικού προσδιορισμού της λειτουργίας της πρωτεΐνης και των πιθανών της δράσεων. Για τον σκοπό αυτό, με τη χρήση των εργαλείων Βιοπληροφορικής, έγινε πρόβλεψη της δομής και της λειτουργίας της πρωτεΐνης, έτσι ώστε να υπάρξει δυνατότητα στοχευμένου σχεδιασμού των επόμενων εργαστηριακών πειραμάτων, που αφορούν τον προσδιορισμό της λειτουργίας της πρωτεΐνης.

## ΕΥΧΑΡΙΣΤΙΕΣ

Η παρούσα Πτυχιακή Εργασία εκπονήθηκε στα πλαίσια των προπτυχιακών σπουδών του τμήματος Πληροφορικής με Εφαρμογές στη Βιοϊατρική, του Πανεπιστημίου Θεσσαλίας κατά την διάρκεια του ακαδημαϊκού έτους 2020-21 σε συνεργασία με το Ελληνικό Ινστιτούτο Παστέρ.

Θα ήθελα να ευχαριστήσω θερμά τις επιβλέπουσες μου, κ. Μπράλιου και κ. Μπολέτη, για την πολύτιμη βοήθεια και τις συμβουλές τους που με συντρόφευσαν καθ' όλη την πορεία εκπόνησης αυτής της εργασίας. Είμαι ευγνώμων, για το γεγονός ότι αποτέλεσαν πηγή έμπνευσης για εμένα με τα ερεθίσματα που μου προσέφεραν και τη συμβολή τους. Με τη βοήθεια τους, άρχισα να ανακαλύπτω τον κόσμο της έρευνας και να αγαπώ το αντικείμενο των σπουδών μου ακόμη περισσότερο, καθώς τα ενδιαφέροντα μου σταδιακά βρήκαν προσανατολισμό και απέκτησαν νόημα.

Θα ήθελα να ευχαριστήσω επίσης, τον κ. Μπάγκο, διευθυντή του εργαστηρίου Μοριακής Βιολογίας, Γενετικής και Βιοπληροφορικής του τμήματος μας, που ήταν δίπλα μας, πρόθυμος να μοιραστεί τις γνώσεις και τις συμβουλές του, τον κ. Ahmed, F. Roumia, διδακτορικό υπότροφο στο Πανεπιστήμιο Θεσσαλίας για την πολύτιμη βοήθεια του και την κ. Ολύβια Τζιουβάρα, διδακτορική υπότροφο στο Εργαστήριο Ενδοκυτταρικού Παρασιτισμού του Ελληνικού Ινστιτούτου Παστέρ, η οποία μου δίδαξε πολύ σημαντικά πράγματα, προσφέροντας απλόχερα τις συμβουλές της κατά τη διάρκεια της πρακτικής μου άσκησης.

Τέλος, θα ήθελα να ευχαριστήσω όλους τους καθηγητές του τμήματος, για τις γνώσεις που μεταλαμπάδευαν σε μένα και τους συν σπουδαστές μου, όλα τα χρόνια της φοίτησης μας στο τμήμα, καθώς και για τα εφόδια που μας βοήθησαν να αποκομίσουμε, τα οποία θα μας συνοδεύουν στην υπόλοιπη ζωή μας.

## ΕΙΣΑΓΩΓΗ

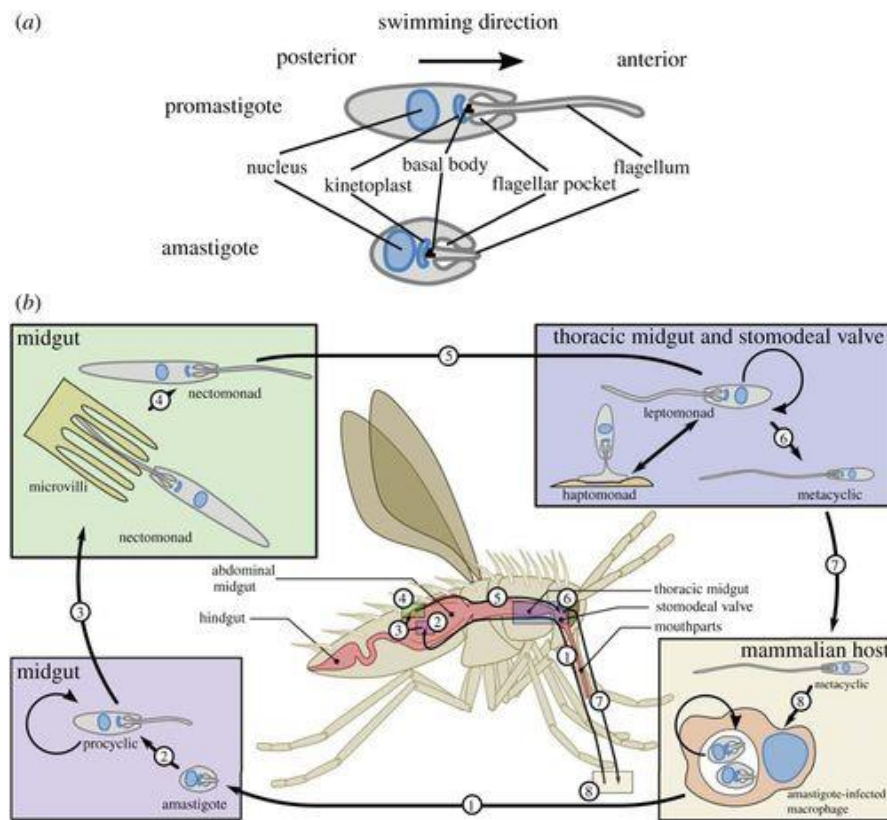
### 1.1 Ο κύκλος ζωής του πρωτοζώου *Leishmania*

Ο παρασιτικός κύκλος ζωής του πρωτοζώου *Leishmania* εκτυλίσσεται εντός δύο διαφορετικών ξενιστών, ενός ενδιάμεσου αρθρόποδου (σκνίπα) και ενός τελικού σπονδυλωτού.

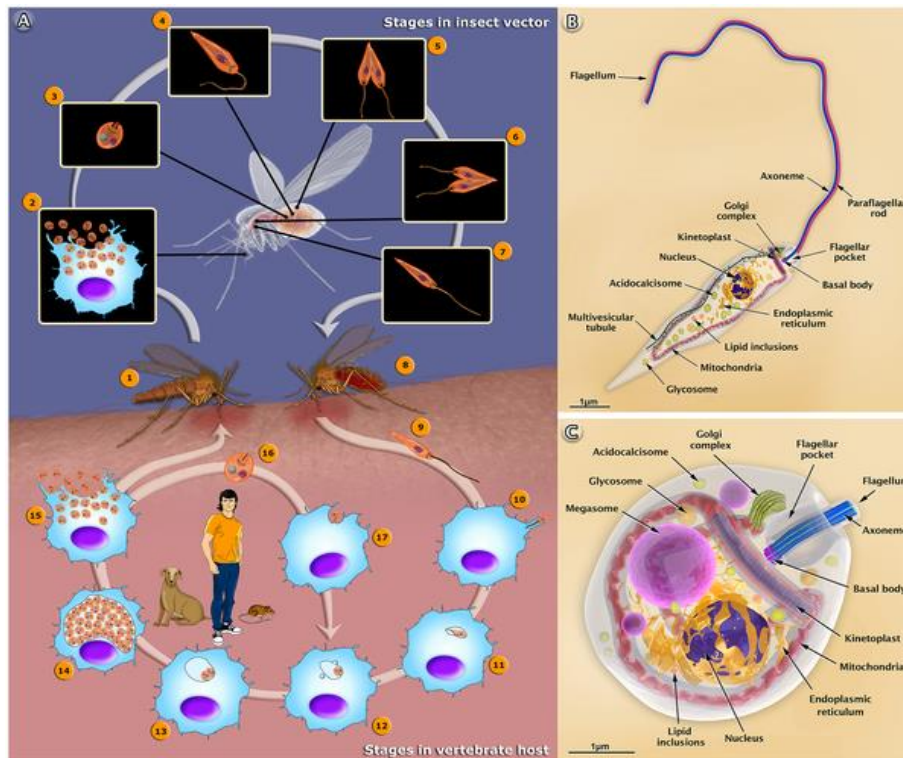
Ο κύκλος ζωής ξεκινά εντός της σκνίπας (αρθρόποδος ξενιστής), μέσω ενός γέυματος αίματος μιας θηλυκής σκνίπας από ένα μολυσμένο θηλαστικό. Τα παράσιτα βρίσκονται στην αμαστιγωτή μορφή τους εντός του αίματος και του χυμού των ιστών. Στο έντερο της σκνίπας, μετατρέπονται στην προμαστιγωτή τους μορφή, προσκολλώνται στο επιθήλιο (ώστε να μην απεκκριθούν μέσω της απεκκριτικής ροής) και διαιρούνται (Εικόνα 1). Στην πρόσδεση του παρασίτου, συμβάλλει το μόριο LPG



(Λιποπολισακχαρίτης) της εξωκυτταρικής μεμβράνης του, το οποίο εξασφαλίζει και την προστασία από την πέψη του αίματος. Έχει αποδειχθεί ότι τα αναπτυξιακά στάδια του παρασίτου εντός της σκνίπας διαρκούν οχτώ με είκοσι ημέρες μετά την μόλυνση. Μετά τη διαδικασία πολλαπλασιασμού του στο έντερο, το παράσιτο *Leishmania* διαφοροποιείται στην μετακυκλική μορφή και προχωρά στον φάρυγγα της σκνίπας. Εκεί, μπλοκάρει το σύστημα αναρρόφησης (sucking apparatus), δημιουργώντας την ανάγκη ενός νέου γεύματος αίματος. Κατά τη διάρκεια ενός νέου γεύματος αίματος μεταδίδεται το παράσιτο στον θηλαστικό ξενιστή.<sup>2 3</sup> (Εικόνα 1)



**Εικόνα 1.** Η διμορφική ζωή του παρασίτου και ο κύκλος ζωής του στον αρθρόποδο ξενιστή<sup>3</sup>



**Εικόνα 2.** Η διμορφική ζωή του παρασίτου, ως εξωκυττάρια και ενδοκυττάρια<sup>4</sup>

Οι μηχανισμοί που το παράσιτο *Leishmania* χρησιμοποιεί για να επιβιώσει και να προκαλέσει τη μόλυνση ξενιστών αποτελούν αντικείμενο έντονης ερευνητικής μελέτης. Πρωτεομική ανάλυση των βιομορίων που εκκρίνονται σε καλλιέργειες από παράσιτα του είδους *Leishmania donovani*, στην προμαστιγωτή τους μορφή, υπέδειξε την ύπαρξη πολλών πρωτεϊνών, που εκκρίνονται στο εξωκυττάριο περιβάλλον, ίσως μέσω εξωσωμάτων<sup>5</sup> Ανάμεσα στις πρωτεΐνες που ταυτοποιήθηκαν, αναδείχθηκε η πρωτεΐνη που κωδικοποιείται από το γονίδιο LdBPK\_352470.1 και εμφανίζει υψηλά ποσοστά συντήρησης στα είδη *Leishmania*. Η τελευταία, παρουσιάζει μια PX λειτουργική περιοχή πρόσδεσης σε φωσφοϊνοσιτίδια, που σε συνδυασμό με άλλα χαρακτηριστικά της δομής που εμφανίζει, κατηγοριοποιείται στην πρωτεϊνική οικογένεια των Νεξινών Διαλογής (Sorting Nexins). Οι πρωτεΐνες που ανήκουν στην οικογένεια αυτή, χαρακτηρίζονται από την πρόσδεση τους σε μεμβράνες, μέσω της αλληλεπίδρασης τους με λιπίδια ή μεμβρανικά συμπλέγματα πρωτεϊνών. Έρευνες έχουν δείξει ότι αρκετές Νεξίνες, εμφανίζουν συμμετοχή σε μονοπάτια που σχετίζονται με τη διαλογή πρωτεϊνών και τους ενδοκυττάριους μηχανισμούς μετακίνησης<sup>6 7 8</sup>

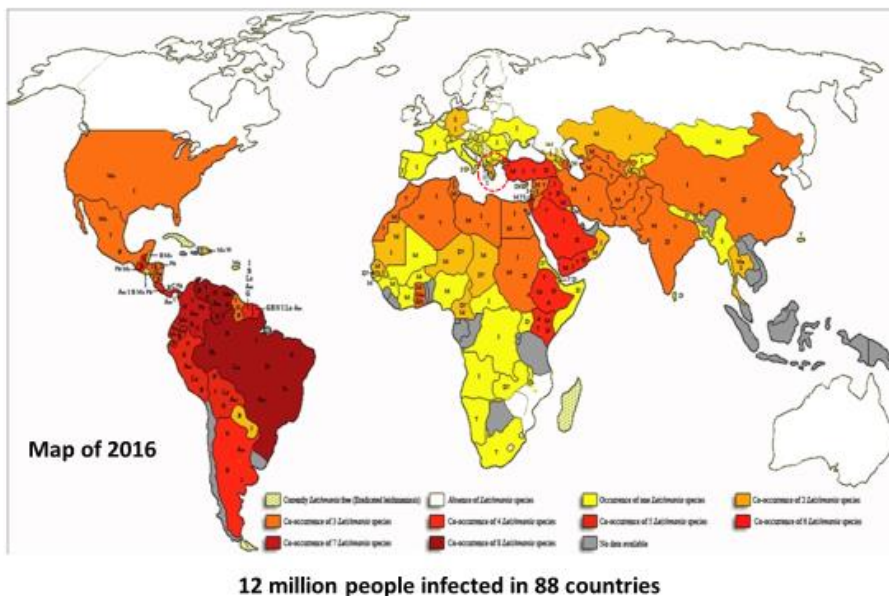
## 1.2 Γενικές Πληροφορίες για τη Λεισμανίαση

Η Λεισμανίαση είναι μια τροπική νόσος που προκύπτει από τη μόλυνση θηλαστικών από πρωτόζωα του γένους *Leishmania*. Χαρακτηρίζεται ως «νόσος των φτωχών», διότι συναντάται κυρίως στις αναπτυσσόμενες χώρες της Αφρικής, Ασίας και Ν. Αμερικής. Σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας (WHO) η Λεισμανίαση, ως ανθρώπινη νόσος, εμφανίζεται σε 98 κράτη με 2 εκατομμύρια περιστατικά παγκοσμίως (ανά έτος) και 20-30.000 θανάτους. (WHO, <https://www.who.int/news-room/fact-sheets/detail/leishmaniasis>). Το παράσιτο συνήθως ενδημεί σε περιοχές με τροπικό ή υποτροπικό κλίμα (Εικόνα 3), όπου συναντώνται και οι αρθρόποδοι ξενιστές (σκνίπες). (WHO, [https://apps.who.int/neglected\\_diseases/ntddata/leishmaniasis/leishmaniasis.ht](https://apps.who.int/neglected_diseases/ntddata/leishmaniasis/leishmaniasis.ht)

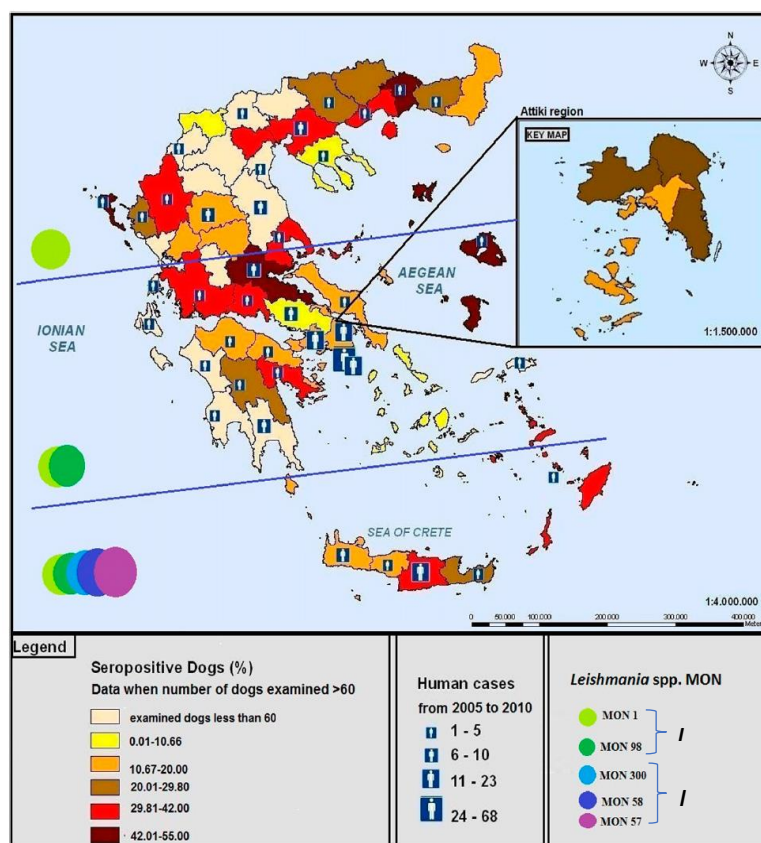
ml). Ωστόσο, έχει παρατηρηθεί η παρουσία ορισμένων ειδών και κατ' επέκταση περιπτώσεων Λεϊσμανίασης σε ανθρώπους και σκύλους σε περιοχές της Ελλάδας. (ΕΟΔΥ, <https://eody.gov.gr/disease/leismaniasi/>)

### 1.3. Κλινικές μορφές της νόσου

Ανάλογα με το είδος του παρασίτου *Leishmania*, που σχετίζεται με τη μόλυνση, μπορεί να προκληθούν διαφορετικοί τύποι Λεϊσμανίασης. Το φάσμα των κλινικών εκδηλώσεων της νόσου είναι αρκετά ευρύ και εξαρτάται από την περιπλοκότητα της αλληλεπίδρασης του μολυσματικού παρασίτου και του ξενιστή, από το ανοσοποιητικό σύστημα, τη διατροφική κατάσταση, την ηλικία και το γενετικό υπόβαθρο του υπό μόλυνση οργανισμού. Επίσης, ως σημαντικοί παράγοντες, εμφανίζονται ο πιθανός εμβολιασμός του ξενιστή και το μέγεθος του μικροβιακού φορτίου. Η ασθένεια, παρουσιάζεται συχνότερα με τις εξής κλινικές μορφές: τη σπλαχνική (VL, Visceral Leishmaniasis), τη δερματική (CL, Cutaneous Leishmaniasis) και τη βλεννογοδερματική λεϊσμανίωση (MCL, Mucocutaneous Leishmaniasis). Οι διαφορετικές εκδηλώσεις της ασθένειας, σχετίζονται με τον πολλαπλασιασμό της *Leishmania* στα φαγολυσοσώματα των μακροφάγων, στο χόριο, και στο ρινοστοματοφαρυγγικό βλεννογόνο<sup>9</sup>. Η Λεϊσμανίαση προκύπτει από περίπου 21 διαφορετικά είδη του πρωτοζώου *Leishmania*, που μεταδίδονται από περίπου 30 είδη σκνίπας των γενών *Phlebotomus* και *Lutzomyia*.<sup>10</sup>



**Εικόνα 3.** Γεωγραφική κατανομή και ενδημικότητα των παθογόνων για τον άνθρωπο στελεχών *Leishmania*<sup>11</sup>



**Εικόνα 4.** Γεωγραφική κατανομή και ενδημικότητα στελεχών *Leishmania* (*L. tropica*, *L. infantum*) σε ανθρώπους και ζώα στην Ελλάδα.<sup>12</sup>

### 1.3.1 Σπλαχνική Λεισμανίαση (VL)

Η σπλαχνική Λεισμανίαση (VL-Visceral Leishmaniasis), αν δεν αντιμετωπιστεί με κατάλληλη θεραπευτική προσέγγιση, μπορεί να εξελιχθεί σε θανατηφόρα. Προκαλείται από το παράσιτο *L. donovani*, στην Ανατολική Αφρική και την Ινδική υποήπειρο και από το *L. infantum* στην Ευρώπη, τη Βόρεια Αφρική και τη Λατινική Αμερική. Το *L. infantum* μολύνει κυρίως παιδιά και ανοσοκατασταλμένα άτομα, ενώ το *L. donovani* μολύνει όλες τις ηλικιακές ομάδες. Υπάρχουν δύο είδη VL, που ποικίλουν ανάλογα με τον τρόπο με τον οποίο μεταδίδονται, η ζωνοσός VL που μεταδίδεται από μολυσμένο ζώο φορέα στον άνθρωπο, καθώς και η ανθρωποσός VL που αφορά μολύνσεις μεταξύ ανθρώπων. Συνήθως, η ζωνοσός VL βρίσκεται σε περιοχές μετάδοσης *L. infantum* ενώ η ανθρωποσός VL βρίσκεται σε περιοχές μετάδοσης *L. donovani*.<sup>13</sup> Ως συμπτώματα, έχουν καταγραφεί πυρετός, νυκτερινές εφιδρώσεις, απώλεια βάρους, ρίγος, ανορεξία, σπληνομεγαλία καθώς και πανκυτταροπενία λόγω καταστολής του μυελού των οστών (αναιμία, θρομβοπενία, λευκοπενία). Στην περιοχή της Ινδίας, σπάνια συναντώνται περιστατικά ασθενών που εμφανίζουν υπέρχρωση του δέρματος, γεγονός που χαρακτήρισε την ασθένεια με τον όρο «kala-azar» (μαύρος πυρετός-γλώσσα Hindi). Οι νοσούντες μπορεί να εμφανίσουν επίσης, βλάβη στο δέρμα ή τους βλεννογόνους. Επιπλέον, έχουν καταγραφεί περιστατικά ασθενών που εμφανίζουν αιμοφαγοκυτταρικό σύνδρομο. Σε περίπτωση

συν-λοίμωξης με τον ιό HIV, μπορεί να παρατηρηθούν άτυπες εκδηλώσεις και σε άλλα συστήματα (όπως στο αναπνευστικό). (ΚΕΕΛΠΝΟ, 2016) Σε κάποιες περιπτώσεις ασθενών, παρουσιάζεται ως επιπλοκή της σπλαχνικής Λεισμανίασης, ένα σύνδρομο δερματικής Λεισμανίασης (post kala-azar dermal leishmaniasis), το οποίο περιέχει δερματικές βλάβες, κατά κύριο λόγο στην περιοχή του προσώπου, που εμφανίζονται συνήθως έξι μήνες έως ένα χρόνο μετά τη φαινομενική θεραπεία της VL. Η μορφή αυτή περιγράφεται κυρίως σε περιπτώσεις λοίμωξης από *L. donovani* και παρατηρείται συνήθως στην ανατολική Αφρική (WHO, <https://www.who.int/news-room/fact-sheets/detail/leishmaniasis>).

### 1.3.2 Δερματική Λεισμανίαση

Η δερματική μορφή της Λεισμανίασης, αποτελεί την συχνότερη εμφάνιση της ασθένειας και κατά κύριο λόγο δεν προκαλεί τον θάνατο του νοσούντα. Οφείλεται στα πρωτόζωα *L. major*, *L. tropica*, *L. guyanensis*, *L. mexicana*, *L. braziliensis*, *L. amazonensis*, *L. peruviana*, *L. panamensis*. Η εκδήλωση των συμπτωμάτων παρουσιάζεται στην επιφάνεια του δέρματος, στο σημείο που παρατηρείται το τσίμπημα από τον αρθρόποδο ξενιστή. Συνήθως, παρατηρείται εξέλιξη των δερματικών βλαβών από βλατίδες σε οζώδεις πλάκες και σε ελκωτικές βλάβες, με επηρμένα άκρα και κεντρικό κρατήρα. Τις περισσότερες φορές είναι ανώδυνες και θεραπεύονται εντός μερικών μηνών, αφήνοντας ως σημάδια δερματικές αλλοιώσεις (ουλές).<sup>14</sup>

### 1.3.3 Βλεννογονοδερματική Λεισμανίωση (Mucocutaneous Leishmaniasis MCL)

Συνήθως προκύπτει από το πρωτόζωο *L. braziliensis*. Συνήθως, συναντάται μετά από την εμφάνιση της δερματικής έκφρασης της νόσου. Οφείλεται στη διασπορά των παρασίτων *Leishmania* από το δέρμα στο βλεννογόνο του στόματος και του φάρυγγα. Τις περισσότερες φορές, παρουσιάζεται μετά από μεγάλο χρονικό διάστημα (σε σχέση με την αρχική δερματική μόλυνση και την ανεπαρκή ίαση της). Εντούτοις, βλάβες σε περιοχές του βλεννογόνου και του δέρματος μπορεί να εμφανιστούν παράλληλα, ενώ κάποιοι από τους νοσούντες έχουν υποκλινικές δερματικές λοιμώξεις. Τα αρχικά συμπτώματα που εκδηλώνονται είναι κατά κύριο λόγο ρινικά (ευαισθησία, αιμορραγία), μολονότι κάποιες φορές παρουσιάζονται πρώτα τα συμπτώματα στη στοματική κοιλότητα ή τον φάρυγγα. Σε περίπτωση μη αποτελεσματικής αντιμετώπισης, παρατηρείται εξέλιξη της ασθένειας, με καταστροφή του ρινοστοματοφαρυγγικού βλεννογόνου καθώς και γειτονικών ιστών. Ακόμη, υπάρχει η πιθανότητα πρόκλησης δευτεροπαθών λοιμώξεων, που επιφέρουν τον θάνατο του ασθενούς, παρά τη χορήγηση φαρμακευτικής αγωγής.<sup>14</sup>

## 1.4 Πρόληψη της νόσου

Για την επιτυχία της πρόληψης της ασθένειας της Λεισμανίασης, απαιτείται η συνδυαστική εφαρμογή στρατηγικών πρόληψης (WHO, <https://www.who.int/news-room/fact-sheets/detail/leishmaniasis>)

1. Η εγκυρότητα στη διάγνωση και τη θεραπεία, συντελεί στη μείωση του επιπολασμού της νόσου και αποτρέπει τη θνητότητα.
2. Ο έλεγχος του αριθμού των φορέων, δηλαδή της σκνίπας (αρθρόποδος ξενιστής), βοηθάει στη μείωση ή τη διακοπή της μετάδοσης της νόσου. Οι μέθοδοι ελέγχου

- περιλαμβάνουν ψεκάσμο εντομοκτόνου, χρήση διχτυών επεξεργασμένων με εντομοκτόνα, περιβαλλοντική διαχείριση και προσωπική προστασία.
3. Η αποτελεσματική παρακολούθηση της ασθένειας, συνδράμει στη δράση έναντι επιδημιών και καταστάσεων με υψηλά ποσοστά θνητότητας.
  4. Ο έλεγχος των ξενιστών, που πρέπει να πραγματοποιείται σύμφωνα με την κατάσταση της κάθε περιοχής.
  5. Η ευαισθητοποίηση των πολιτών σχετικά με το θέμα, η κινητοποίηση και η εκπαίδευση των κοινοτήτων με παρεμβάσεις αλλαγής συμπεριφορών που εξυπηρετούν την προστασία της τοπικής κοινωνίας.

### 1.5 Διάγνωση και θεραπεία της Λεισμανίασης

Στην σπλαχνική λεισμανίαση, η διάγνωση γίνεται με συνδυασμό κλινικών ενδείξεων και εξετάσεων (όπως ταχείες διαγνωστικές εξετάσεις). Σε δερματικές και βλενωδείς εκφάνσεις της νόσου, οι κλινικές ενδείξεις συνδυαστικά με παρασιτολογικές εξετάσεις επιβεβαιώνουν τη διάγνωση. Η θεραπεία της λεισμανίασης εξαρτάται από διάφορους παράγοντες, όπως ο τύπος της νόσου, οι ταυτόχρονες παθολογίες, τα είδη των παρασίτων και η γεωγραφική θέση. Η λεισμανίαση είναι μια ιάσιμη ασθένεια, η οποία απαιτεί την ύπαρξη ενός ισχυρού ανοσολογικού συστήματος, επειδή τα φάρμακα δεν απαλλάσσουν το σώμα από το παράσιτο, συνεπώς υπάρχει ο κίνδυνος υποτροπής σε περίπτωση ανοσοκαταστολής. Όλοι οι ασθενείς που έχουν διαγνωστεί με σπλαχνική λεισμανίαση χρειάζονται άμεση και πλήρη θεραπεία. (WHO, <https://www.who.int/news-room/fact-sheets/detail/leishmaniasis>).

### 1.6 Τα φωσφοϊνοσιτίδια (PIs)

Τα φωσφοϊνοσιτίδια αποτελούν φωσφολιπίδια που έχουν ως δράση τον άμεσο ή έμμεσο έλεγχο πολλαπλών κυτταρικών διεργασιών (π.χ. μετάδοση σημάτων, έλεγχος κυτταρικής ανάπτυξης, δυναμική του κυτταροσκελετού, ανταλλαγή λιπιδίων μεταξύ των κυτταρικών οργανιδίων και ρύθμιση της μεταφοράς φορτίου μέσω των διαμεμβρανικών πρωτεϊνών).<sup>15</sup>

Η παρουσία τους στην μεμβράνη συγκεκριμένων δομών, εμπλέκεται σε διεργασίες όπως βιογένεση φαγοσωμάτων και ενδοσωμάτων, εξωκύττωση, ενδοκύττωση, καθώς και ανταλλαγή κυστιδίων με το σύστημα Golgi.<sup>16</sup>

Εκκρινόμενες πρωτεΐνες που συμβάλλουν με ποικίλους τρόπους στη διαλογή (sorting) των πρωτεϊνών, προσδένονται στις μεμβράνες διαφορετικών οργανιδίων, βασιζόμενες στην ταυτοποίηση συγκεκριμένων δομικών περιοχών πρόσδεσης φωσφοϊνοσιτιδίων, οι οποίες συναντώνται στη δομή τους.

Ο κώδικας με τον οποίο πραγματοποιείται η αναγνώριση της στερεοειδικής οργάνωσης της φωσφορικής ομάδας γύρω από το δακτύλιο της ινοσιτόλης, εξαρτάται από τον τρόπο που κατανέμονται τα κατάλοιπα των πρωτεϊνών στην τριτοταγή δομή των περιοχών πρόσδεσης σε φωσφοϊνοσιτίδια (PIs, phosphoinositides)<sup>17</sup>. Παρατηρούνται διαφορετικά είδη δομικών περιοχών πρόσδεσης σε φωσφοϊνοσιτίδια. Με τη βοήθεια των ορθών πρωτεϊνικών αλληλεπιδράσεων με την πολική κεφαλή των PIs, τα τελευταία συμβάλλουν σημαντικά στη ρύθμιση της επικοινωνίας των κυτταρικών μεμβρανών. Επιπλέον, λειτουργούν ως σηματοδοτικά μόρια και συντελούν στον καθορισμό της ταυτότητας των κυτταρικών οργανιδίων. Η δράση τους επεκτείνεται και στον έλεγχο της διαπερατότητας των μεμβρανών και της οργάνωσης του κυτταροσκελετού, εκτός από τη μεταγωγή σημάτων και τον έλεγχο της διαμεμβρανικής επικοινωνίας.<sup>1</sup>

## 1.7 Οι Νεξίνες Διαλογής (Sorting Nexins)

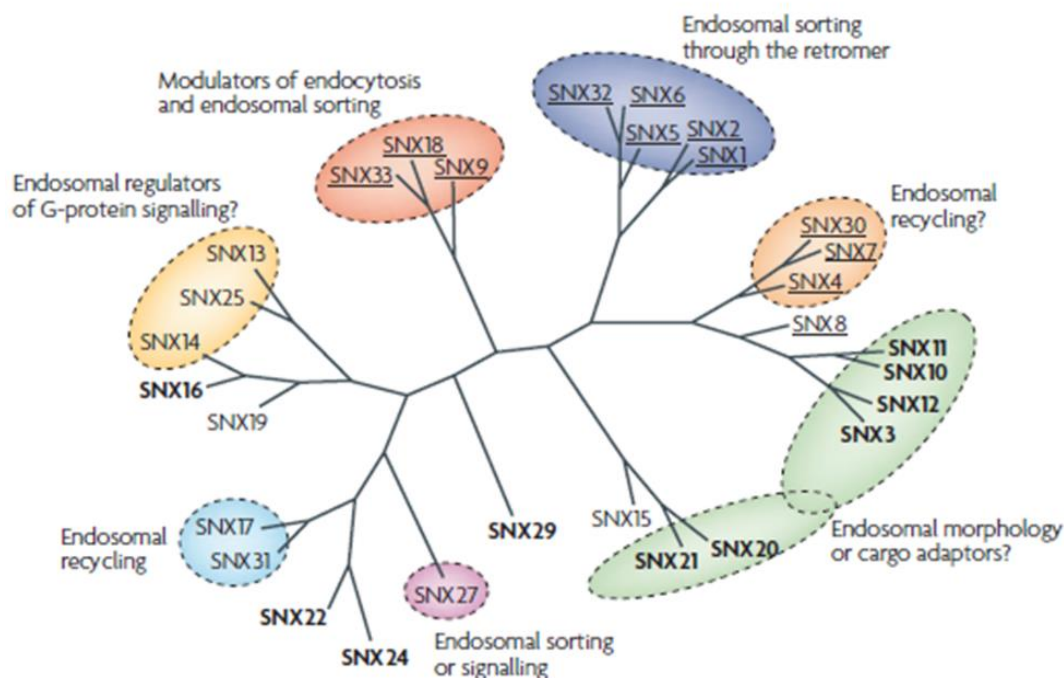
Οι Νεξίνες Διαλογής (Sorting Nexins, SNXs), αποτελούν μια οικογένεια πρωτεϊνών που κατηγοριοποιείται με βάση την ύπαρξη μιας λειτουργικής περιοχής πρόσδεσης σε φωσφοϊνοσιτίδια PX (Lemmon 2003). Σύμφωνα με έρευνες, όλες οι Νεξίνες προσελκύονται στην επιφάνεια μεμβρανών όπου απαντώνται τα Pls, μέσω μοτίβων πρόσδεσης σε φωσφοϊνοσιτίδια που υπάρχουν στη δομή τους.<sup>8</sup>

Οι Νεξίνες Διαλογής εντοπίζονται στο κυτταρόπλασμα καθώς και σε ενδοκυττάριας μεμβράνες και δρουν ως παράγοντες πρόσδεσης (μέσω της λειτουργικής περιοχής PX που διαθέτουν) σε φωσφοϊνοσιτίδια ή άλλες πρωτεΐνες.

Έχουν ταυτοποιηθεί αρκετές Νεξίνες σε θηλαστικά και μήκυτες, ωστόσο δεν έχει προσδιοριστεί με ακρίβεια η δράση των περισσοτέρων<sup>18</sup>. Οι Νεξίνες θεωρούνται ότι συμμετέχουν στη διαλογή των πρωτεϊνών και σε μονοπάτια σηματοδότησης.<sup>6</sup>

Εντός της οικογένειας των SNXs, εμφανίζονται πρωτεΐνες, που εκτός από τη λειτουργική περιοχή PX που αναφέρθηκε, περιέχουν μια επιπλέον λειτουργική περιοχή πρόσδεσης σε πρωτεϊνικές δομές, την περιοχή «BAR». Στις Βιοπληροφορικές βάσεις δεδομένων υπάρχει μια σύγκριση σχετικά με την ονοματολογία αυτής της δομικής περιοχής, καθώς συναντώνται επικαλύψεις. Πολλές φορές αναγράφεται και ως Vps5 (Vacuolar protein sorting-associated protein 5) πιθανώς λόγω του εντοπισμού της περιοχής για πρώτη φορά στον οργανισμό-μοντέλο *saccharomyces cerevisiae*, που αποτέλεσε σημείο αναφοράς για πολλές μελέτες (η περιοχή κωδικοποιείται από το γονίδιο Vps5 του *S.cerevisiae*).

Οι πρωτεΐνες που ανήκουν σε αυτή την υποοικογένεια (και περιλαμβάνουν τις περιοχές πρόσδεσης PX και BAR) είναι μεμβρανικές και παρουσιάζουν δομή διπλής Α-έλικας. Πιστεύεται πως έχουν την ιδιότητα να αναγνωρίζουν σημεία που οι μεμβράνες παρουσιάζουν κύρτωση και να προκαλούν αναδόμηση τους έτσι ώστε να δημιουργηθούν σωληνοειδείς μεμβρανικές προεκβολές.<sup>7</sup> Οι σωληνοειδείς δομές που ανήκουν στο πρώιμο σύστημα της ενδοκύττωσης, συντελούν στη μεταφορά και τη διαλογή πρωτεϊνών, γεγονός που καθιστά τις πρωτεΐνες αυτές σημαντικό αντικείμενο μελέτης.



**Εικόνα 5.** Εξελικτικό Φυλογενετικό δένδρο των ομάδων των διαφορετικών Νεξινών Διαλογής στα θηλαστικά και πιθανές ή αποδεδειγμένες λειτουργίες τους.<sup>8</sup>

### 1.8 Η κωδικοποιούμενη από το γονίδιο LdBPK\_352470.1 πρωτεΐνη E9BSB7

Το γονίδιο LdBPK\_352470.1 κωδικοποιεί την έκφραση μιας πρωτεΐνης αποτελούμενης από 417 αμινοξικά κατάλοιπα<sup>5</sup>. Ανάλυση της αμινοξικής αλληλουχίας της έδειξε πως η εν λόγω πρωτεΐνη περιλαμβάνει την λειτουργική περιοχή PX και την περιοχή πρόσδεσης BAR (Εικόνα 6), όπως συναντάται στην υποοικογένεια των SNXs που αναφέρθηκε στο κεφάλαιο 1.7. Αυτό, καθιστά πιθανή την ομολογία της δράσης της με πρωτεΐνες που χαρακτηρίζονται από αυτή τη δομή, στα ενδοκυτταρικά μονοπάτια που σχετίζονται με τη διαλογή των πρωτεϊνών στα φαγοκύτταρα του θηλαστικού ξενιστή.

Αν ισχύει αυτή η υπόθεση, η παραπάνω πρωτεΐνη θα μπορούσε να χαρακτηριστεί ως μολυσματικός παράγοντας, λόγω της πιθανής απορρύθμισης των αμυντικών μηχανισμών του θηλαστικού ξενιστή που σχετίζονται με την διαλογή πρωτεϊνών.

Δομικές περιοχές	Θέσεις
PX Domain	8-139
Non structured Region	164-185
Helix	214-234
Helix	318-338
Helix	340-367

**Εικόνα 6.** Δομικές περιοχές της πρωτεΐνης E9BSB7 σύμφωνα με τη βάση δεδομένων Uniprot και την πρόβλεψη της δευτεροταγούς δομής της με την χρήση του αλγορίθμου Psipred



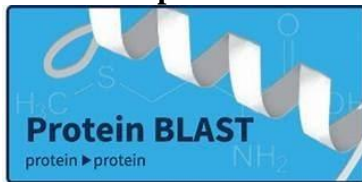
## ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΙ

### 2.1 MICROSOFT EXCEL



Το Microsoft Excel είναι ένα πρόγραμμα λογιστικών φύλλων που υλοποιήθηκε από τη Microsoft και χρησιμοποιείται για τη διαχείριση δεδομένων. Δίνεται η δυνατότητα με τη χρήση εντολών να εκτελεστούν μαθηματικές πράξεις και να παρουσιαστούν τα αποτελέσματα με τη μορφή γραφημάτων και πινάκων.

### 2.2 BLASTp



Το BLASTp, αποτελεί ένα βιοπληροφορικό εργαλείο, που χρησιμοποιείται για τη σύγκριση ακολουθιών, όπως ακολουθίες πρωτεϊνικών αμινοξέων ή νουκλεοτιδίων που αφορούν είτε DNA είτε RNA.<sup>19</sup>

### 2.3 HMMER



Το HMMER αποτελεί ένα πακέτο λογισμικού που χρησιμεύει στην ανάλυση ακολουθιών, ταυτοποιεί ομόλογες πρωτεϊνικές ή νουκλεοτιδικές ακολουθίες και πραγματοποιεί στοιχίσεις ακολουθιών. Προσδιορίζει την ομολογία συγκρίνοντας Hidden Markov Model Profiles.

Το πακέτο αυτό, περιέχει αλγορίθμους που επιτελούν διαφορετικές λειτουργίες όπως:

- **Δημιουργία HMM προφίλ**

hmmbuild – δημιουργία προφίλ HMMs από πολλαπλές στοιχίσεις

- **Εύρεση ομολογίας**

hmmsearch – αναζήτηση πρωτεϊνικών ακολουθιών εντός βάσης δεδομένων HMM

hmmsearch – αναζήτηση profile HMMs σε βάσεις δεδομένων

nhmmer – αναζήτηση DNA/RNA ερωτημάτων (queries) σε DNA/RNA βάσεις δεδομένων

nhmmer – αναζήτηση νουκλεοτιδικών ακολουθιών σε βάσεις δεδομένων.

phmmer – αναζήτηση πρωτεϊνικών ακολουθιών σε βάσεις δεδομένων.

- **Άλλες λειτουργίες**

hmmalign – στοιχίση ακολουθιών από HMM

hmmlgo – παραγωγή δεδομένων από HMM logo από HMM file<sup>20</sup>

## 2.4 CLUSTAL OMEGA



Το Clustal Omega, είναι ένα πρόγραμμα πολλαπλής στοίχισης που χρησιμοποιεί δέντρα-οδηγούς και Hidden Markov Model profile-profile τεχνικές για να αναπαράγει στοίχισεις μεταξύ τριών ή περισσότερων ακολουθιών.<sup>21</sup>

## 2.5 ITOL EMBL ALGORITHM



Ο αλγόριθμος ITOL, είναι ένα εργαλείο για την επεξεργασία φυλογενετικών δέντρων.<sup>22</sup>

## 2.6 JALVIEW



Το JALVIEW, αποτελεί ένα πρόγραμμα για πολλαπλή στοίχιση ακολουθιών, επεξεργασία και οπτικοποίηση.<sup>23</sup>

## 2.7 PSIPRED



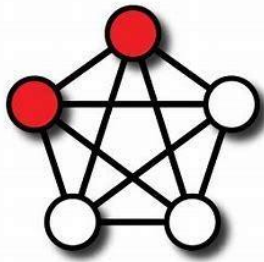
Είναι ένας αλγόριθμος που χρησιμοποιείται για την διερεύνηση της δομής πρωτεϊνών. Χρησιμοποιεί μηχανική μάθηση που βασίζεται σε Τεχνητά Νευρωνικά Δίκτυα. Χρησιμοποιήθηκε για την πρόβλεψη της δευτεροταγούς δομής της πρωτεΐνης.<sup>24</sup>

## 2.8 PHYRE2



Αποτελεί σημαντικό εργαλείο πρόβλεψης τριτοταγούς δομής πρωτεϊνών.<sup>25</sup>

## 2.9 STRING DATABASE



Η βάση δεδομένων STRING (The Search Tool for the Retrieval of Interacting Genes) είναι μια διαδικτυακή βάση που παρέχει πληροφορίες σχετικά με αλληλεπιδράσεις πρωτεϊνών: είτε άμεσες (φυσικές) είτε έμμεσες (λειτουργικές), συνοψίζοντας τόσο πειραματικά δεδομένα όσο και υπολογιστικές προβλέψεις.<sup>26</sup>

## 2.10 KEGG



Η KEGG (Kyoto Encyclopedia of Genes and Genomes) είναι μια συλλογή βάσεων δεδομένων, η οποία περιέχει πληροφορίες που βοηθούν στην κατανόηση λειτουργιών σε μοριακό επίπεδο, πληροφορίες για τα γονιδιώματα, βιολογικά μονοπάτια, ασθένειες, χημικές ουσίες και φάρμακα, αλληλεπιδράσεις μορίων, γονιδίων και πρωτεϊνών. Η KEGG είναι βασικό εργαλείο για την επιστήμη της βιοπληροφορικής, χρησιμεύει στη γονιδιωματική, στην προσομοίωση βιολογικών συστημάτων και στην ανάπτυξη φαρμάκων.<sup>27</sup>

## 2.11 SKYLIGN ALGORITHM

Αποτελεί αλγόριθμο που δημιουργεί LOGO βασισμένο σε HMM Profiles.<sup>28</sup>

## 2.12 HEATMAPPER



Είναι ένα εύχρηστο εργαλείο ελεύθερου λογισμικού το οποίο δίνει τη δυνατότητα στους χρήστες να οπτικοποιήσουν διαφορετικούς τύπους δεδομένων. Οι χρήστες μπορούν να οπτικοποιήσουν χάρτες με βάση την έκφραση, χάρτες συσχετίσεων, χάρτες επικάλυψης εικόνας. Στο αποτέλεσμα αποτυπώνονται οι διαβαθμίσεις την έντασης δύο χρωμάτων τα οποία χαρακτηρίζουν το φαινόμενο που μελετάται. Υπάρχει η δυνατότητα να μπορεί να τροποποιηθεί η εικόνα και να προσαρμοστεί χρησιμοποιώντας τις επιλογές και τις παραμέτρους σχεδίασης.<sup>29</sup>

## 2.13 UNIPROT DATABASE



Βιοπληροφορική βάση δεδομένων που αποτελεί πηγή δεδομένων για πρωτεϊνικές ακολουθίες και λειτουργίες.<sup>30</sup>

## 2.14 SWISS MODEL



Αλγόριθμος που χρησιμοποιείται για την πρόβλεψη της στερεοδιάταξης των πρωτεϊνών.<sup>31</sup>

### ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ ΒΗΜΑΤΩΝ ΜΕΘΟΔΟΛΟΓΙΑΣ

- ❖ Χρήση των αλγορίθμων BLAST και RHHMER για την εύρεση ορθόλογων πρωτεϊνών.
- ❖ Χρήση των ορθόλογων ακολουθιών που ευρέθηκαν, ως είσοδο στον αλγόριθμο Clustal Omega, για τη δημιουργία φυλογενετικών δέντρων και πολλαπλών στοιχίσεων σε συνδυασμό με τον αλγόριθμο ITOL EMBL και το πρόγραμμα JALVIEW.
- ❖ Πραγματοποίηση δευτεροταγούς και τριτοταγούς πρόβλεψης της υπό μελέτη πρωτεΐνης, με τη χρήση των αλγορίθμων PSIPRED και PHYRE2.
- ❖ Πρόβλεψη της λειτουργικότητας της πρωτεΐνης με τη βοήθεια των βάσεων δεδομένων String και KEGG.
- ❖ Αναζήτηση των πρωτεομάτων των Κινητοπλαστίδων στη βάση δεδομένων UNIPROT. Χρήση του LINUX, και του πακέτου HMM για τον προσδιορισμό και την οργάνωση των δομικών περιοχών (domain architecture) που συναντώνται στις ευρεθείσες πρωτεΐνες. Χρήση του προγράμματος Excel και της γλώσσας προγραμματισμού Bash Script για την επεξεργασία των δεδομένων και την τελική μορφοποίηση. Για την κατασκευή του HEATMAP, τα δεδομένα υπέστησαν ειδική μορφοποίηση ώστε να μπορέσουν να λειτουργήσουν ως είσοδος στον αλγόριθμο HEATMAPPER.

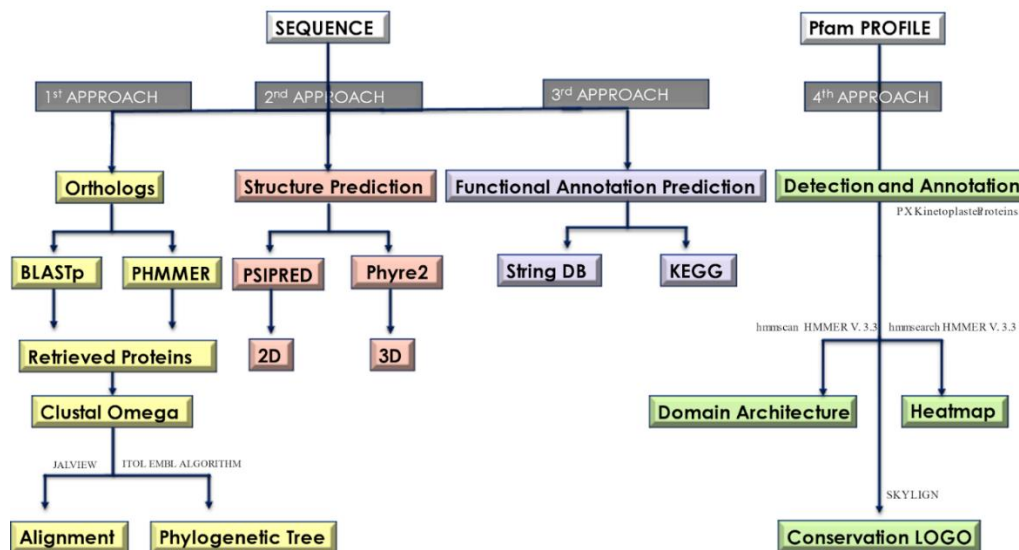
(για εκτενέστερη αναφορά βλ. κεφάλαιο **ΑΠΟΤΕΛΕΣΜΑΤΑ**)

## ΑΠΟΤΕΛΕΣΜΑΤΑ

Η Βιοπληροφορική ανάλυση ξεκίνησε, με την αναζήτηση της υπό μελέτη πρωτεΐνης, σε βάσεις δεδομένων. Μια από τις πιο διαδεδομένες βάσεις, αποτελεί η UNIPROT DB. Η βάση αυτή χαρακτηρίζει την πρωτεΐνη ως «Phosphoinositide-Binding-Protein» ενώ εμείς την κατατάσσουμε στην οικογένεια των Νεξινών Διαλογής (Sorting Nexins). Επιπλέον, δεν υπάρχει κανένας πιο σαφής χαρακτηρισμός σε καμία βάση δεδομένων ή στη βιβλιογραφία. Εμφανίζεται συνεπώς, η ανάγκη μιας καλής πρόβλεψης της λειτουργίας της. Σκοπός της παρούσας εργασίας, λοιπόν είναι ο επαναπροσδιορισμός του χαρακτηρισμού της πρωτεΐνης σε «Sorting Nexin» από τη σκοπιά της Βιοπληροφορικής, η πρόβλεψη της δομής της και τέλος η πρόβλεψη της λειτουργίας της, που θα συνδράμει στον σχεδιασμό των αντίστοιχων πειραμάτων στο εργαστήριο. Μετά την πρόβλεψη της δομής και λειτουργίας της πρωτεΐνης E9BSB7 από το παράσιτο *Leishmania donovani*, επεκτείναμε την ανάλυση στον εντοπισμό και τη δομική κατάταξη όλων των πρωτεϊνών από τις Κινετοπλαστίδες, που περιέχουν λειτουργική περιοχή PX, μεγάλο μέρος των οποίων πιθανώς να ανήκουν στην οικογένεια των Νεξινών Διαλογής (Sorting Nexins).

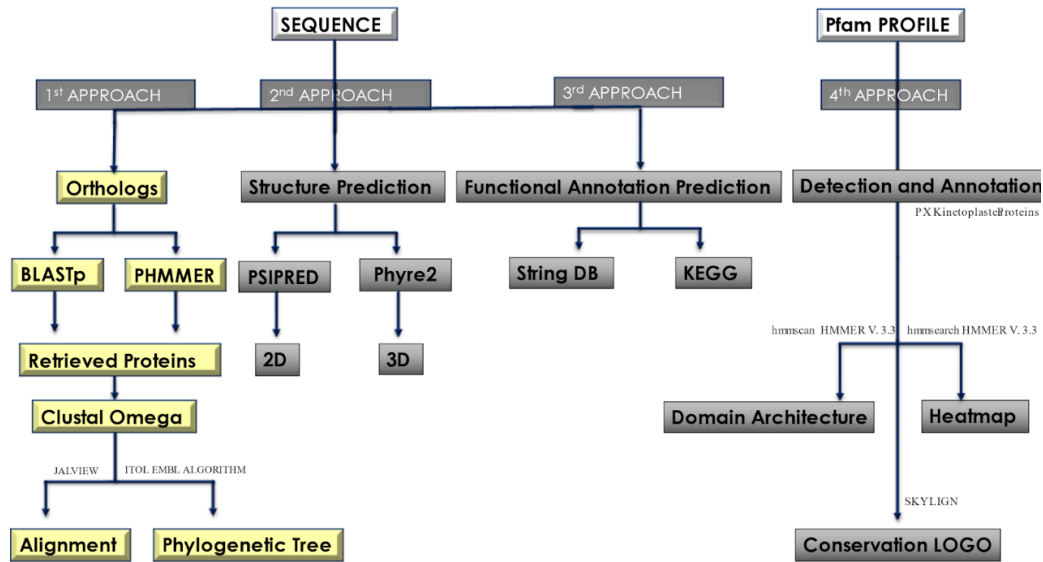
Για τον σκοπό αυτό ακολουθήθηκε η παρακάτω προσέγγιση που απαρτίζεται από πέντε βήματα που συνοψίζονται σχηματικά στην Εικόνα 7:

1. Εύρεση ορθόλογων πρωτεϊνών, πολλαπλή στοίχιση και κατασκευή φυλογενετικών δέντρων
2. Πρόβλεψη της Δευτεροταγούς και Τριτοταγούς δομής της E9BSB7
3. Πρόβλεψη της λειτουργίας της E9BSB7 με βάση αλληλεπιδράσεις με άλλες πρωτεΐνες
4. Εντοπισμός και μελέτη όλων των ομόλογων πρωτεϊνών που συναντώνται στις Κινετοπλαστίδες
5. Προσδιορισμός και οργάνωση των δομικών περιοχών (domain architecture) που συναντώνται σε όλες τις πρωτεΐνες των Κινετοπλαστιδών που διαθέτουν λειτουργική περιοχή PX, κατασκευή Heatmap και δημιουργία Logo συντηρημένων αμινοξέων των λειτουργικών περιοχών PX και VPS5.



**Εικόνα 7.** Σχηματική αναπαράσταση των βημάτων που ακολουθήθηκαν κατά τη Βιοπληροφορική ανάλυση της πρωτεΐνης E9BSB7.

### 3.1 Εύρεση ορθόλογων πρωτεϊνών, πολλαπλή στοίχιση και κατασκευή φυλογενετικών δέντρων



**Εικόνα 8.** Σχηματική αναπαράσταση του πρώτου βήματος της Βιοπληροφορικής ανάλυσης για την πρωτεΐνη E9BSB7

Πρώτο βήμα της ανάλυσης αποτέλεσε η αναζήτηση με βάση την ακολουθία, πρωτεϊνών που ανήκουν σε άλλους οργανισμούς και εμφανίζουν ομοιότητα στην αμινοξική αλληλουχία με την υπό μελέτη πρωτεΐνη. Για τον σκοπό αυτό, έγινε χρήση των αλγορίθμων BLAST και PHMMER. Ως κατώφλι, τέθηκε το  $e$ -value της τάξης του  $e^{-10}$ . Η αναζήτηση πραγματοποιήθηκε στις καταχωρήσεις της βάσης δεδομένων UNIPROT στην περίπτωση του πρώτου, ενώ στη δεύτερη περίπτωση έγινε αναζήτηση σε όλα τα πρωτεόματα που είναι καταχωρημένα σε όλες τις βάσεις δεδομένων. Ο σκοπός αυτής της προσέγγισης έγκειται στο ενδεχόμενο εύρεσης πρωτεϊνών με εξαιρετικά υψηλά ποσοστά ομοιότητας με την πρωτεΐνη που μας ενδιαφέρει, γεγονός που θα μπορούσε να προσανατολίσει την έρευνα μας καλύτερα. Υψηλά ποσοστά ομοιότητας σε επίπεδο ακολουθίας είναι πιθανό να σημαίνουν και συνάφεια ως προς την λειτουργία. Και οι δύο αλγόριθμοι έδωσαν αρκετά μεγάλη ομοιότητα στα αποτελέσματα. Επιλέχθηκε να χρησιμοποιηθούν αυτά που προέκυψαν από τον αλγόριθμο BLASTp, διότι τα στατιστικά σημαντικά αποτελέσματα ήταν περισσότερα.

BA002	BA002_LEM1 - Actin filament binding protein - Leishmania major - E-value: 1.1	WAF28	WAF28_784 - Phosphoinositide 3-kinase - Homo sapiens - E-value: 4.7e-17	AK113754	AK113754_1 - Phosphoinositide 3-kinase - Homo sapiens - E-value: 3.7e-24	Q68705	Q68705_1 - Phosphoinositide 3-kinase - Homo sapiens - E-value: 1.3e-22
BA003	BA003_LEM1 - Actin filament binding protein - Leishmania major - E-value: 1.1	WAF29	WAF29_784 - Phosphoinositide 3-kinase - Homo sapiens - E-value: 1.4e-16	AK113755	AK113755_1 - Phosphoinositide 3-kinase - Homo sapiens - E-value: 3.9e-24	Q68706	Q68706_1 - Phosphoinositide 3-kinase - Homo sapiens - E-value: 1.3e-22
BA004	BA004_LEM1 - Actin filament binding protein - Leishmania major - E-value: 1.1	WAF30	WAF30_784 - Phosphoinositide 3-kinase - Homo sapiens - E-value: 1.4e-16	AK113756	AK113756_1 - Phosphoinositide 3-kinase - Homo sapiens - E-value: 3.9e-24	Q68707	Q68707_1 - Phosphoinositide 3-kinase - Homo sapiens - E-value: 1.3e-22
BA005	BA005_LEM1 - Actin filament binding protein - Leishmania major - E-value: 1.1	WAF31	WAF31_784 - Phosphoinositide 3-kinase - Homo sapiens - E-value: 1.4e-16	AK113757	AK113757_1 - Phosphoinositide 3-kinase - Homo sapiens - E-value: 3.9e-24	Q68708	Q68708_1 - Phosphoinositide 3-kinase - Homo sapiens - E-value: 1.3e-22
BA006	BA006_LEM1 - Actin filament binding protein - Leishmania major - E-value: 1.1	WAF32	WAF32_784 - Phosphoinositide 3-kinase - Homo sapiens - E-value: 1.4e-16	AK113758	AK113758_1 - Phosphoinositide 3-kinase - Homo sapiens - E-value: 3.9e-24	Q68709	Q68709_1 - Phosphoinositide 3-kinase - Homo sapiens - E-value: 1.3e-22

Εικόνα 9. Μορφή αποτελεσμάτων BLASTp (130/250)

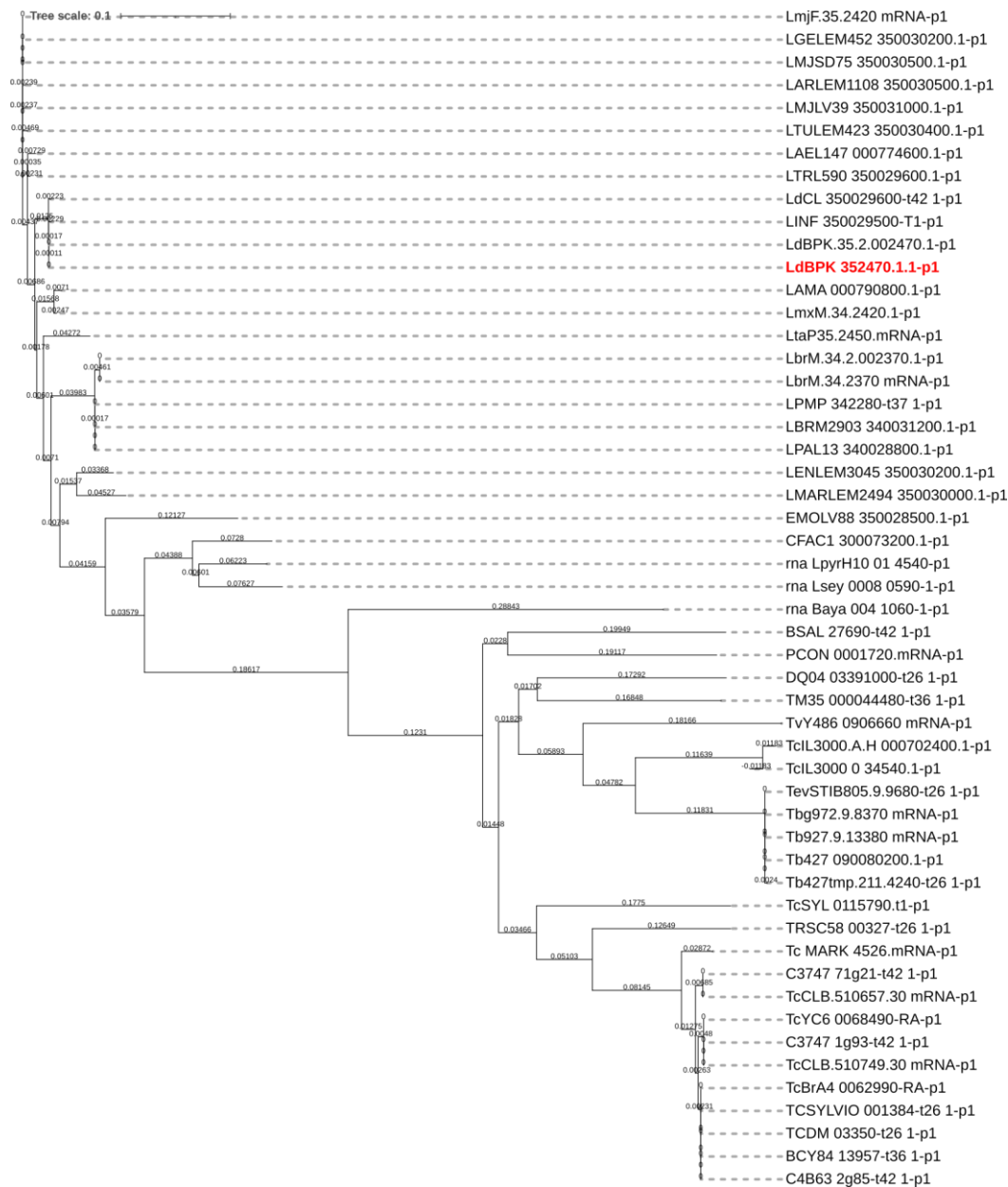
Στη συνέχεια, έγινε συμπληρωματική αναζήτηση ορθολογων πρωτεϊνών και στην βάση δεδομένων TRITRYP, που αφορά μόνο πρωτόζωα που ανήκουν στην οικογένεια των Τρυπανοσπιτιδών. Επειδή δεν αναγράφεται εντός της βάσης κάποιο στατιστικό μέτρο αξιοπιστίας, δεν είμαστε σε θέση να γνωρίζουμε την ακρίβεια αυτών των αποτελεσμάτων. Ωστόσο, στη συγκεκριμένη βάση συναντάμε πρωτεΐνες από αρκετά είδη *Leishmania*, που οι αλγόριθμοι BLAST και PHHMER δεν εντόπισαν.

Gene	Organism	Product	is syntenic	has comments
<a href="#">Baya_004_1060</a>	Blechnomonas ayalai B08-376	phosphoinositide-binding protein, putative	yes	<a href="#">no</a>
<a href="#">BSAL_27690</a>	Bodo saltans strain Lake Konstanz	phosphoinositide-binding protein, putative	no	<a href="#">no</a>
<a href="#">CFAC1_300073200</a>	Crithidia fasciculata strain Cf-CI	phosphoinositide-binding protein, putative	yes	<a href="#">no</a>
<a href="#">EMOLV88_350028500</a>	Endotrypanum monterogei strain LV88	phosphoinositide-binding protein, putative	yes	<a href="#">no</a>
<a href="#">LAEL147_000774600</a>	Leishmania aethiopia L147	phosphoinositide-binding protein, putative	yes	<a href="#">no</a>
<a href="#">LAMA_000790800</a>	Leishmania amazonensis MHOM/BR/71973/M2269	PX domain/Vps5 C terminal like, putative	yes	<a href="#">no</a>
<a href="#">LARLEM1108_350030500</a>	Leishmania arabica strain LEM1108	phosphoinositide-binding protein, putative	yes	<a href="#">no</a>
<a href="#">LBRM2903_340031200</a>	Leishmania braziliensis MHOM/BR/75/M2903	phosphoinositide-binding protein, putative	yes	<a href="#">no</a>
<a href="#">LbrM.34.2370</a>	Leishmania braziliensis MHOM/BR/75/M2904	Autophagy-related protein 24	yes	<a href="#">no</a>
<a href="#">LbrM.34.2.002370</a>	Leishmania braziliensis MHOM/BR/75/M2904 2019	PX domain/Vps5 C terminal like, putative	yes	<a href="#">no</a>
<a href="#">LdCL_350029600</a>	Leishmania donovani CL-SL	Autophagy-related protein 24	yes	<a href="#">no</a>
<a href="#">LdBPK.35.2.002470</a>	Leishmania donovani strain LV9	PX domain/Vps5 C terminal like, putative	yes	<a href="#">no</a>
<a href="#">LENLEM3045_350030200</a>	Leishmania enriettii strain LEM3045	phosphoinositide-binding protein, putative	yes	<a href="#">no</a>
<a href="#">LGELEM452_350030200</a>	Leishmania gerbilli strain LEM452	phosphoinositide-binding protein, putative	yes	<a href="#">no</a>
<a href="#">LINF_350029500</a>	Leishmania infantum JPCM5	Autophagy-related protein 24	yes	<a href="#">no</a>
<a href="#">LmjF.35.2420</a>	Leishmania major strain Friedlin	Autophagy-related protein 24	yes	<a href="#">no</a>

Εικόνα 10. Μορφή αποτελεσμάτων, που προέκυψε από την αναζήτηση ορθολογων πρωτεϊνών της E9BSB7 εντός της βάσης δεδομένων Try TripDB

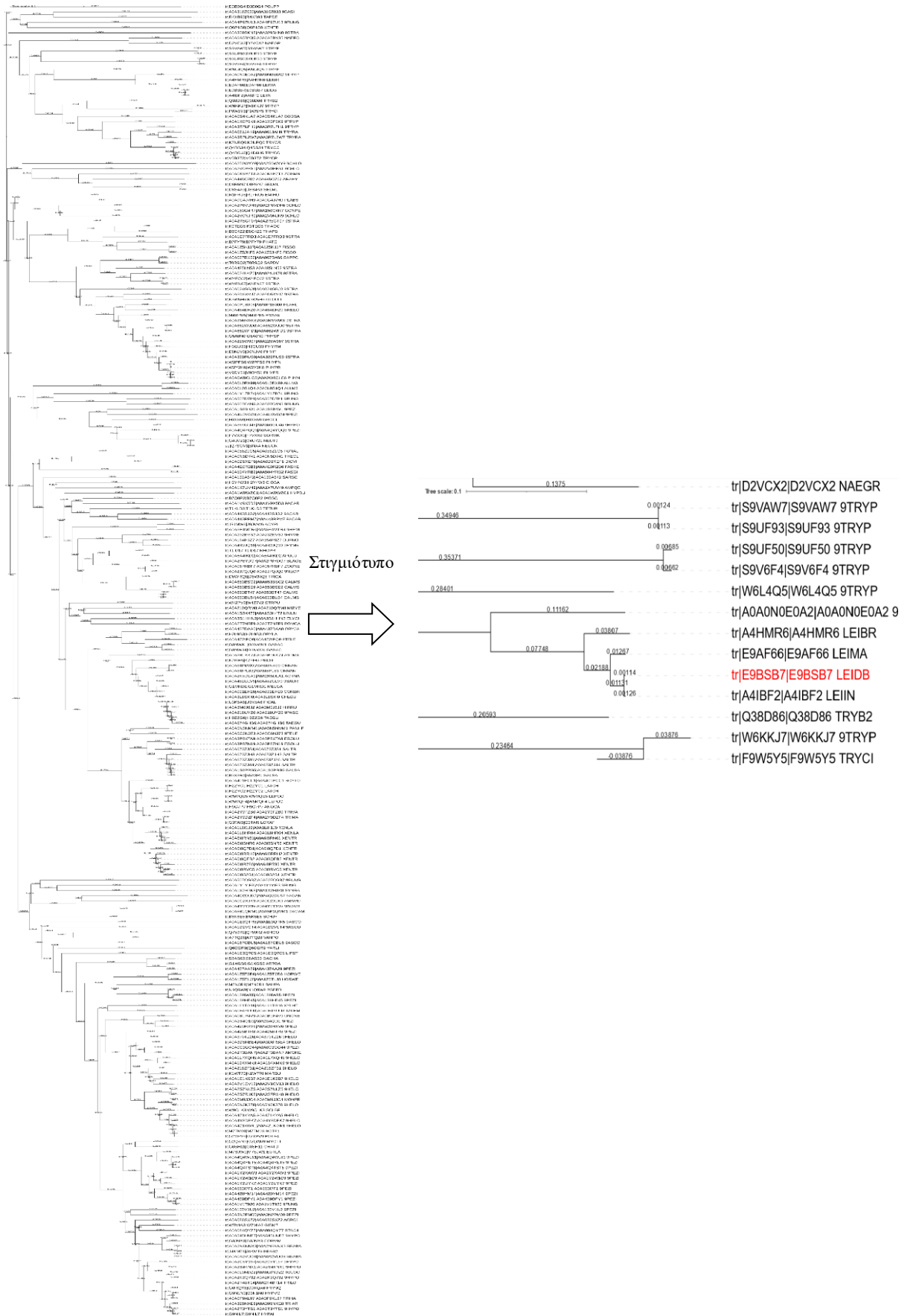
Μετά την εύρεση των ορθόλογων πρωτεϊνών, ακολούθησε η χρήση του αλγορίθμου Clustal Omega προκειμένου να γίνει κατασκευή φυλογενετικών δέντρων σε συνδυασμό με τον αλγόριθμο ITOL EMBL, καθώς και πολλαπλή στοίχιση η οποία οπτικοποιήθηκε με το εργαλείο Βιοπληροφορικής JALVIEW. Τρία φυλογενετικά δέντρα δημιουργήθηκαν με τις ορθολογες πρωτεΐνες που εντοπίστηκαν :

1. από τη βάση δεδομένων TryTriP και αφορά μόνο τις Κινητοπλαστίδες.
2. μέσω του αλγορίθμου BLAST.
3. μέσω της φόρμας BLAST MODEL NCBI και αφορούν μόνο οργανισμούς-μοντέλα.

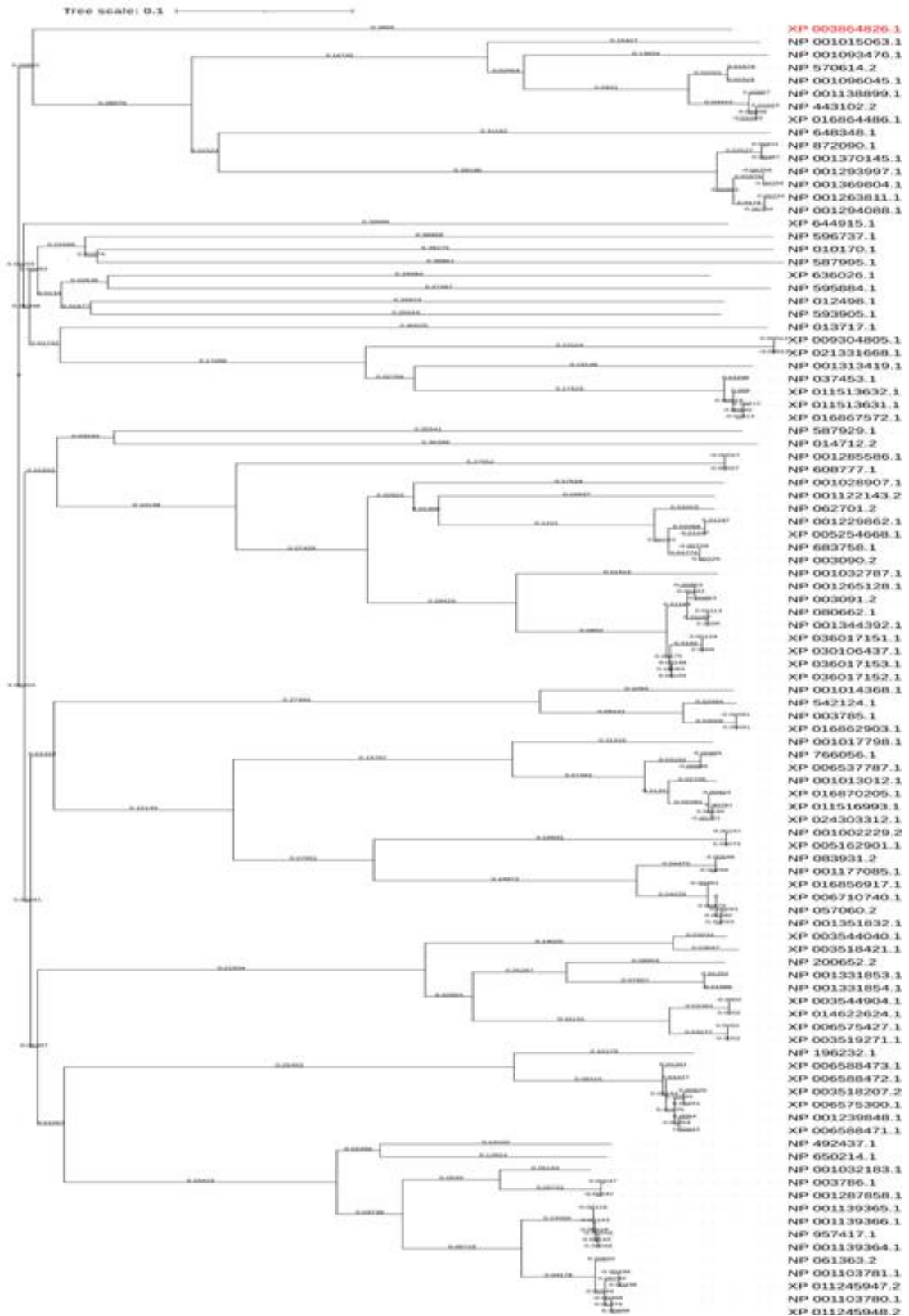


**Εικόνα 11.** Φυλογενετικό δέντρο που προέκυψε από τα αποτελέσματα της βάσης TryTriPDB. Με κόκκινο χρώμα υποδεικνύεται η υπό μελέτη πρωτεΐνη

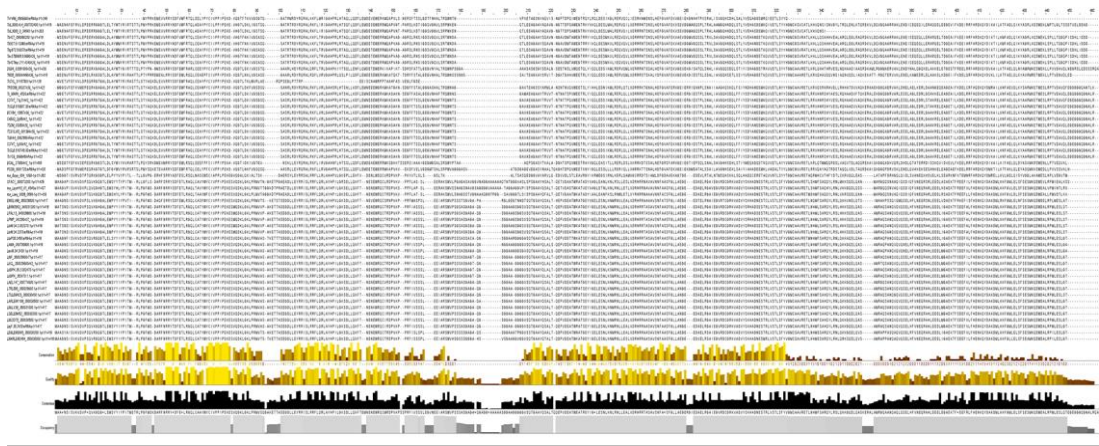




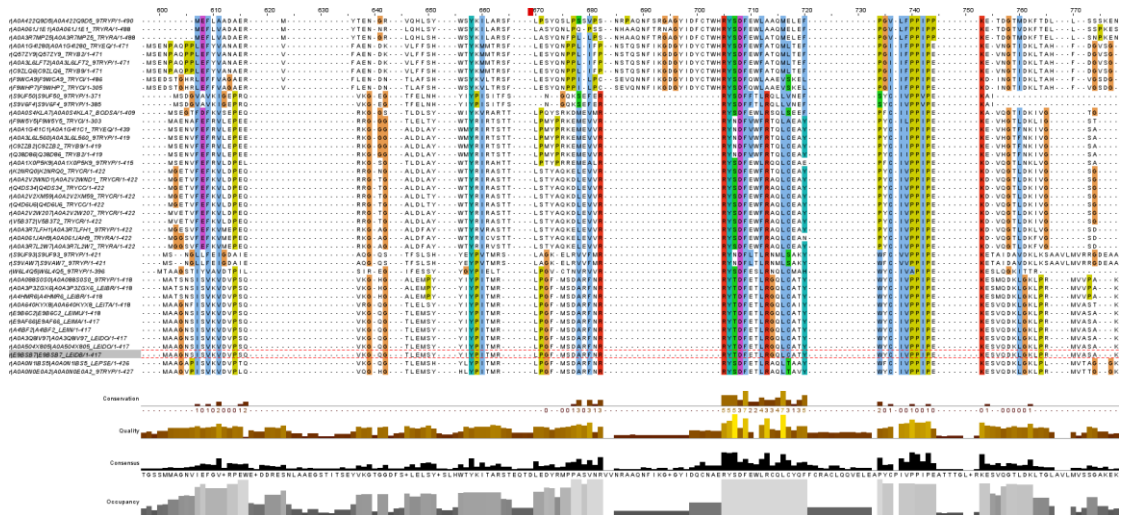
**Εικόνα 12.** Φυλογενετικό δέντρο που προέκυψε από τα αποτελέσματα του αλγορίθμου BLASTp. Με κόκκινο χρώμα υποδεικνύεται η υπό μελέτη πρωτεΐνη.



**Εικόνα 13.** Φυλογενετικό δέντρο που προέκυψε από τα αποτελέσματα του αλγορίθμου BLAST κατά την εφαρμογή του μόνο σε πρωτεΐνες που ανήκουν σε οργανισμούς-μοντέλα. Με κόκκινο χρώμα αναγράφεται η υπό μελέτη πρωτεΐνη.



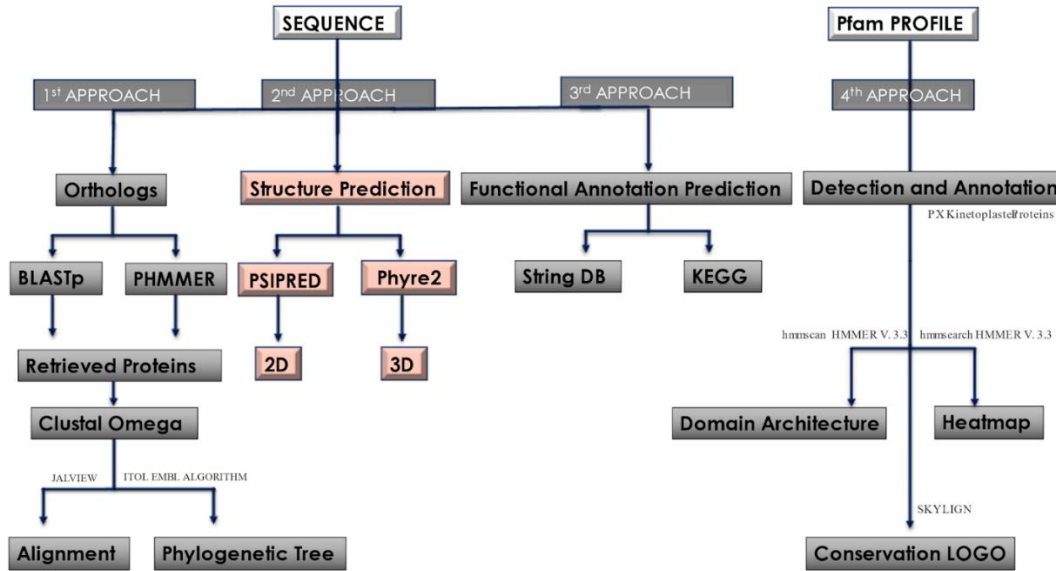
**Εικόνα 14.** Μορφή της απεικόνισης από την πολλαπλή στοίχιση των πρωτεϊνών που προέκυψε από τη χρήση του αλγορίθμου Clustal Omega και αφορά τις πρωτεΐνες που επιλέχθηκαν μέσω της βάσης δεδομένων TriTryp. Για την απεικόνιση, χρησιμοποιήθηκε το πρόγραμμα Jalview



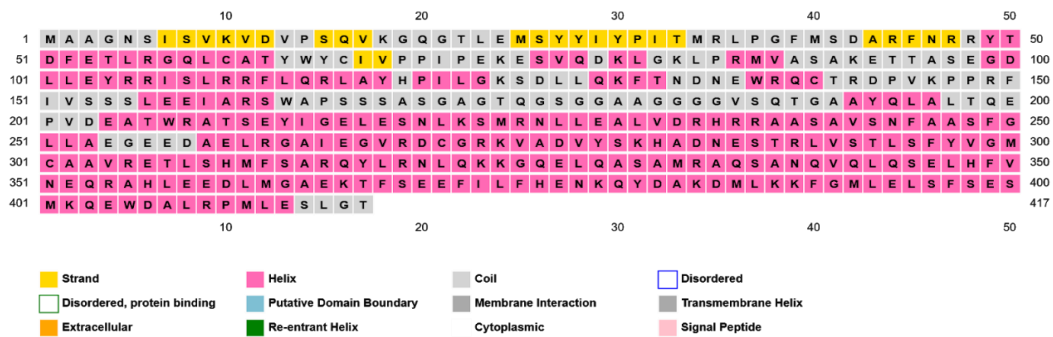
**Εικόνα 15.** Μορφή της απεικόνισης από την πολλαπλή στοίχιση των πρωτεϊνών που προέκυψε από τη χρήση του αλγορίθμου Clustal Omega και αφορά τις πρωτεΐνες που επιλέχθηκαν μέσω του αλγορίθμου BLAST. Για την απεικόνιση, χρησιμοποιήθηκε το πρόγραμμα Jalview

### 3.2 Πρόβλεψη της Δευτεροταγούς και Τριτοταγούς δομής της E9BSB7

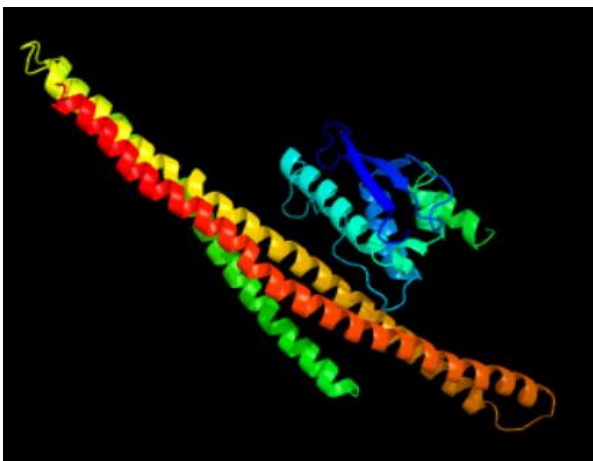
Επόμενο βήμα στην μελέτη μας ήταν η πρόβλεψη της δομής της πρωτεΐνης E9BSB7. Αρχικά έγινε πρόβλεψη της δευτεροταγούς δομής, με χρήση του αλγορίθμου PSIPRED και στη συνέχεια, προχωρήσαμε στη δημιουργία 3D μοντέλου της πρωτεΐνης χρησιμοποιώντας τον αλγόριθμο Phyre2, με στόχο την απεικόνιση της στερεοδιάταξής της. Αυτό το βήμα συντέλεσε στην εύρεση πρωτεϊνών με όμοια στερεοδιάταξη. Ως γνωστόν, η λειτουργία μιας πρωτεΐνης εξαρτάται από την τριτοταγή δομή της. Επομένως, η εύρεση δομικά όμοιων πρωτεϊνών, που πιθανώς να είναι πλήρως μελετημένες, μας φέρνει ένα βήμα πιο κοντά στην πρόβλεψη της λειτουργίας της υπό μελέτη πρωτεΐνης, κάτι που αποτελεί και τον τελικό στόχο αυτής της ανάλυσης.



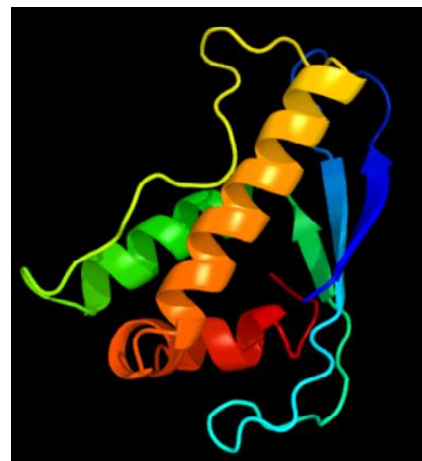
**Εικόνα 16.** Σχηματική αναπαράσταση του δεύτερου βήματος της Βιοπληροφορικής ανάλυσης που αφορά την πρωτεΐνη E9BSB7



**Εικόνα 17.** Πρόβλεψη δευτεροταγούς δομής της υπό μελέτη πρωτεΐνης E9BSB7 με χρήση του αλγορίθμου PSIPRED



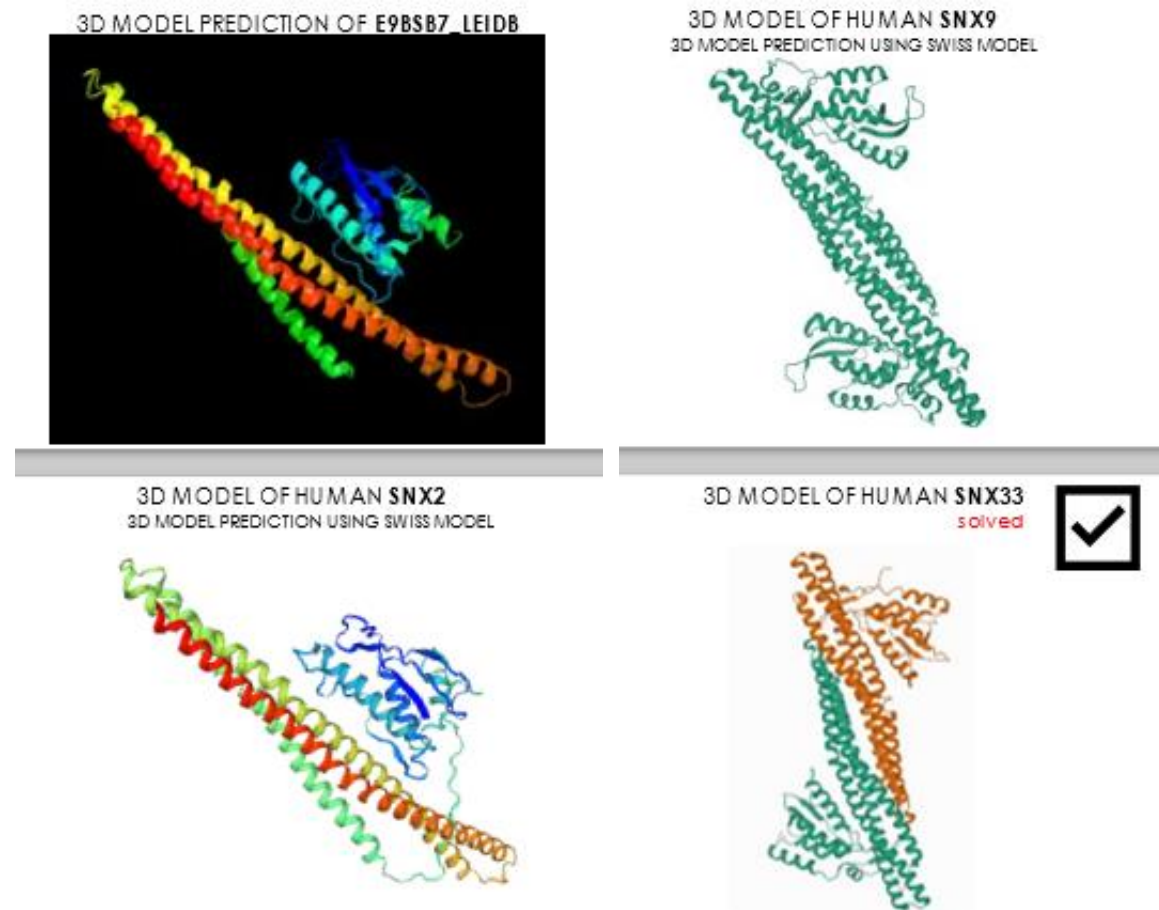
**Εικόνα 18.** Πρόβλεψη τριτοταγούς δομής της υπό μελέτη πρωτεΐνης E9BSB7 με τη χρήση του αλγορίθμου Phyre2



**Εικόνα 19.** Πρόβλεψη τριτοταγούς δομής της λειτουργικής περιοχής PX της πρωτεΐνης E9BSB8 με τη χρήση του αλγορίθμου Phyre2

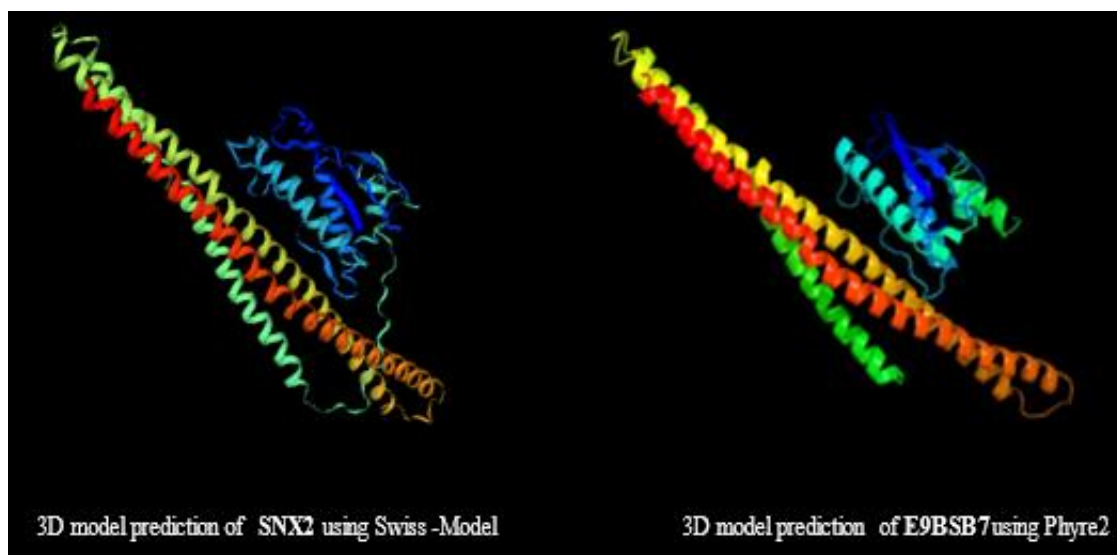
Η μοντελοποίηση της τριτοταγούς δομής της πρωτεΐνης, συντέλεσε στον προσδιορισμό των δομικών χαρακτηριστικών της. Η λειτουργική περιοχή PX αποτελείται από τρία Β-φύλλα και τρεις αντι-παράλληλες μικρού μήκους έλικες, ενώ η περιοχή του BAR domain, αποτελείται από τρεις αντι-παράλληλες έλικες. Η πρωτεΐνη αναδιπλώνεται με τέτοιο τρόπο, ώστε η διάταξη των στοιχείων του PX domain εμφανίζεται να είναι αντικριστή με εκείνη του BAR domain ως προς έναν κεντρικό άξονα παράλληλο με τις έλικες του BAR domain (Ει).

Είναι πολύ ενδιαφέρον ότι η 3D δομή της E9BSB7, εμφανίζει ομοιότητες με τις τρισδιάστατες δομές πρωτεϊνών, που κατατάσσονται στην οικογένεια των Νεξινών.

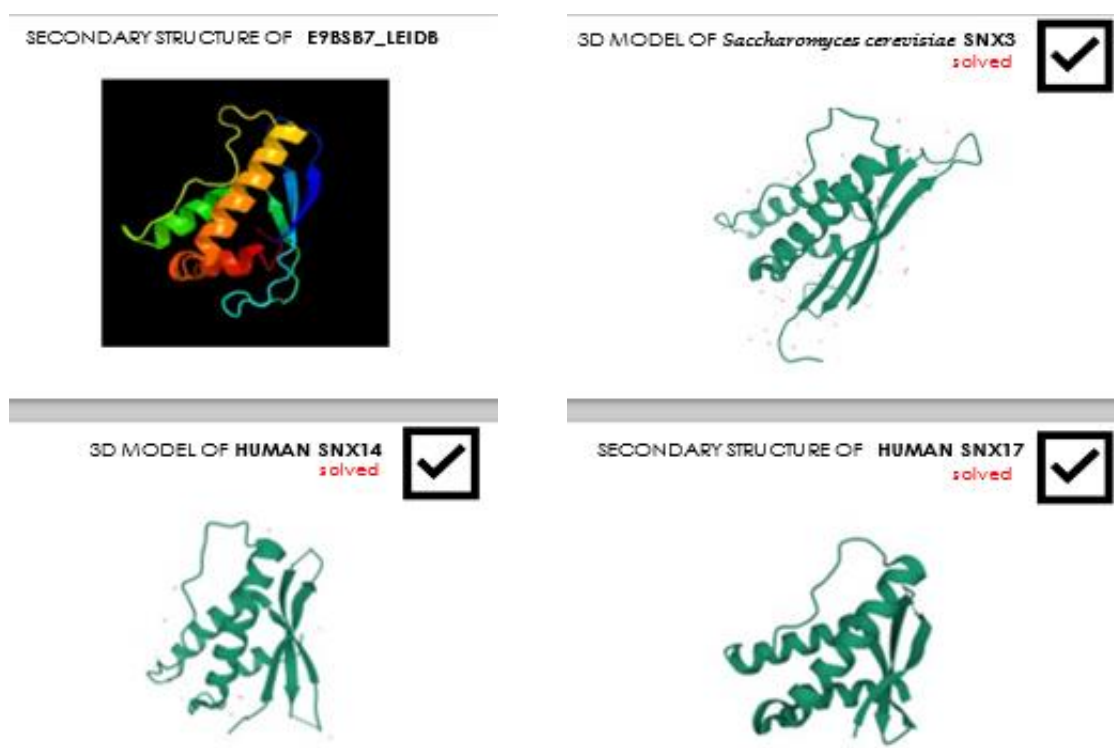


**Εικόνα 20.** Τρισδιάστατες δομές των πρωτεϊνών E9BSB7, ανθρώπινης Νεξίνης 2 (πρόβλεψη μέσω SWISS MODEL), ανθρώπινης Νεξίνης 9 (πρόβλεψη μέσω SWISS MODEL) και ανθρώπινης Νεξίνης 33 (λυμένη δομή).

Εξαιρετικά μεγάλη ομοιότητα εμφανίζεται σε επίπεδο 3D δομής με την ανθρώπινη Νεξίνη 2 (Sorting Nexin 2). Οι όμοιες τρισδιάστατες δομές των πρωτεϊνών αυτών, θα μπορούσαν να αποτελέσουν δείκτη ομοιότητας και σε επίπεδο λειτουργίας.



**Εικόνα 21.** Προβλέψεις τρισδιάστατων δομών για τις πρωτεΐνες E9BSB7 (Phyre2) και ανθρώπινη Νεξίνη 2/SNX2 (SWISS MODEL).



**Εικόνα 22.** Τρισδιάστατης δομές των λειτουργικών περιοχών PX των πρωτεϊνών E9BSB7, ανθρώπινης Νεξίνης 14 (λυμένη δομή), ανθρώπινης Νεξίνης 17 (λυμένη δομή) και Νεξίνης 3 (λυμένη δομή) του οργανισμού *Saccharomyces cerevisiae*.

### 3.3 Πρόβλεψη της λειτουργίας του γονιδιακού προϊόντος του LDBPK\_352470 με βάση τις αλληλεπιδράσεις

Σε αυτό το βήμα, πραγματοποιήθηκε πρόβλεψη της λειτουργίας της υπό μελέτη πρωτεΐνης E9BSB7, με τη βοήθεια των βάσεων String και KEGG.

Αρχικά, έγινε αναζήτηση στην βάση String, μέσω της ακολουθίας της E9BSB7, για να γίνει έλεγχος του ενδεχόμενου να υπάρχει ήδη κάποια καταχώρηση για την πρωτεΐνη, το οποίο απορρίφθηκε. Στη συνέχεια, προχωρήσαμε στην αναζήτηση κάποιας ομόλογης πρωτεΐνης της E9BSB7.

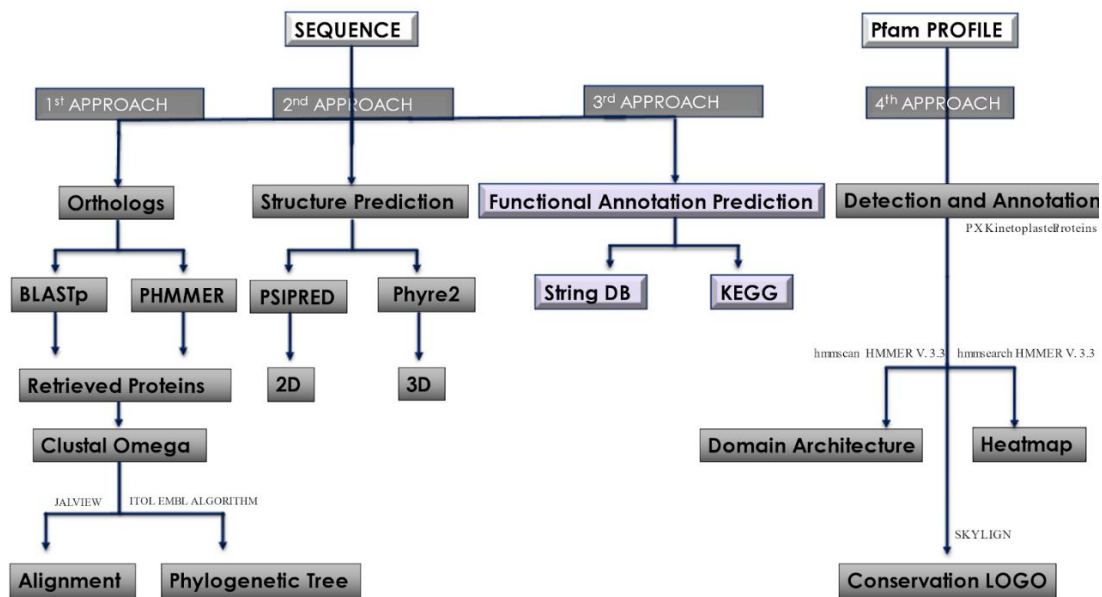
UNIPROT ID	STRING ID	BLAST % IDENTITY
Q38D86	EAN77234	30,30%
A4IBF2	XP_001469071.1	99,8%
Q4DS34	XP_817174.1	34,40%
Q4D6U6	XP_810095.1	34,40%
E9AF66	LmjF.35.2420	97,60%
A0A088S0S0	XP_010702654.1	NO BLAST RESULT

**Πίνακας Α.** Ομόλογες πρωτεΐνες της E9BSB7, που εμφανίζουν ομοιότητα σε επίπεδο ακολουθίας, οργάνωσης των δομικών περιοχών, 2D, 3D και διαθέτουν καταχώριση εντός της βάσης String.

Προκειμένου να επιλεγεί η κατάλληλη, ανατρέξαμε στα αποτελέσματα του BLAST (1<sup>st</sup> APPROACH) και σημειώθηκαν οι πρωτεΐνες που διαθέτουν (Πίνακας Α):

- ομοιότητα σε επίπεδο ακολουθίας
- ίδια οργάνωση των δομικών περιοχών (domain architecture, 5<sup>th</sup> APPROACH)
- ομοιότητα σε δομικό επίπεδο (2D, 3D)
- καταχώριση εντός της βάσης String

Ακολουθήθηκαν τα βήματα που περιγράφονται σχηματικά στην Εικόνα 23.

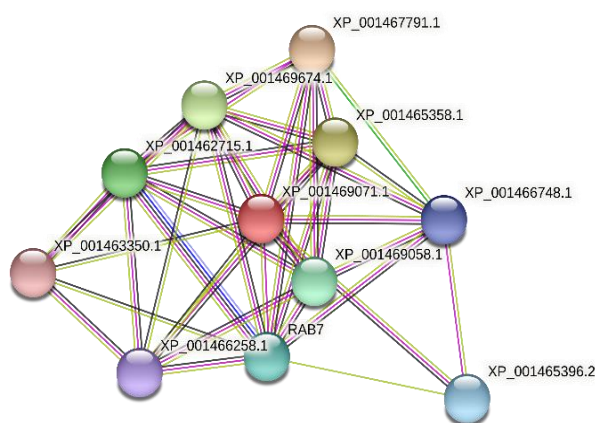


**Εικόνα 23.** Σχηματική αναπαράσταση του τρίτου βήματος της Βιοπληροφορικής ανάλυσης που αφορά την πρωτεΐνη E9BSB7

Από αυτές, επιλέχθηκε, η πρωτεΐνη με NCBI ID *XP\_001469071.1*, η οποία εκκρίνεται από το παράσιτο *Leishmania infantum*, το κοντινότερο συγγενικό είδος *Leishmania* στο είδος *Leishmania donovani*. Η πρωτεΐνη αυτή εμφανίζει εξαιρετικά υψηλή ομοιότητα (99,8 %) με την υπό μελέτη πρωτεΐνη και επομένως η εύρεση της λειτουργίας της θα μπορούσε να αποτελέσει σημαντικό δείκτη αναφοράς στην πρόβλεψη μας.

UNIPROT ID	NCBI ID	BLAST % IDENTITY	ORGANISM
A4IBF2	XP_001469071.1	99.8%	<i>Leishmania infantum</i>

Από τη μελέτη του δικτύου πρωτεϊνικών αλληλεπιδράσεων της πρωτεΐνης XP\_001469071.1 (Εικόνα 24) προέκυψαν τα δεδομένα που παρουσιάζονται στον Πίνακα Β και Εικόνα 25.



**Εικόνα 24.** Δίκτυο πρωτεϊνικών αλληλεπιδράσεων της ορθόλογης πρωτεΐνης XP\_001469071.1 του πρωτοζώου *Leishmania infantum*

Edges:

Edges represent protein-protein associations  
associations are meant to be specific and meaningful, i.e. proteins jointly contribute to a shared function; this does not necessarily mean they are physically binding each other.

Known Interactions  
 from curated databases  
 experimentally determined

Predicted Interactions  
 gene neighborhood  
 gene fusions  
 gene co-occurrence

Others  
 textmining  
 co-expression  
 protein homology

Your Input:

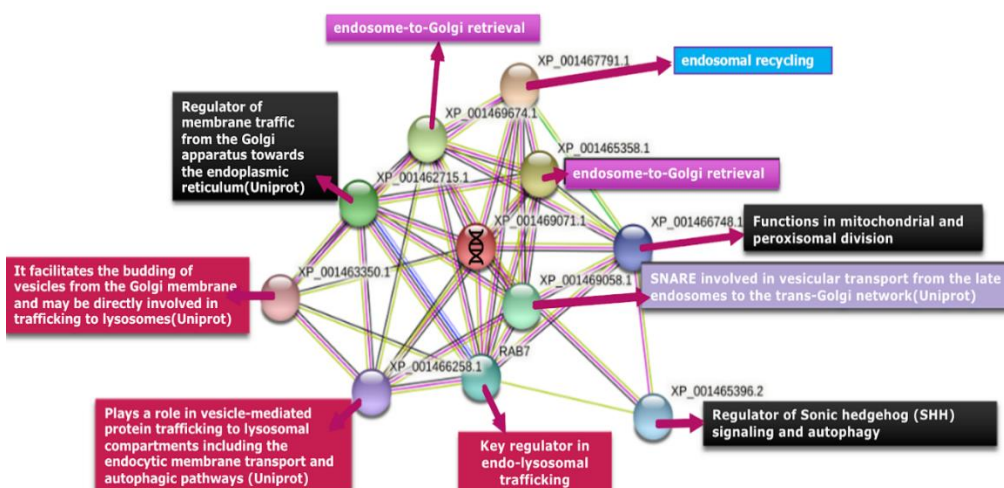
Predicted Functional Partners:

Protein ID	Description	Neighborhood	Gene Fusion	Cooccurrence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
XP_001467791.1	Vacuolar sorting-like protein; Belongs to the VPS29 family	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	0.977
XP_001465358.1	Vacuolar protein sorting-associated protein 35; Plays a role in vesicular protein sorting	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	0.975
XP_001469674.1	Vacuolar protein sorting-associated protein-like protein	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	0.975
XP_001462715.1	Small GTP binding protein rab6-like protein	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	0.765
XP_001469058.1	QA-SNARE protein putative	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	0.743
RAB7	Putative rab7 GTP binding protein	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	0.722
XP_001465396.2	Uncharacterized protein	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	0.707
XP_001466748.1	Putative GTP-binding protein; Belongs to the TRAFAC class dynamin-like GTPase superfamily. Dynamin/Fzo/YdjA fa...	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	0.701
XP_001466258.1	Vacuolar protein sorting-like protein; Belongs to the STXBP/unc-18/SEC1 family	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	0.699
XP_001463350.1	Putative adaptor complex protein (AP) 3 delta subunit 1; Belongs to the adaptor complexes large subunit family	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	0.697

Your Current Organism:  
 Leishmania infantum  
 NCBI taxonomy id: 5671  
 Other names: L. infantum, Leishmania (Leishmania) infantum, Leishmania donovani infantum, Leishmania infantum

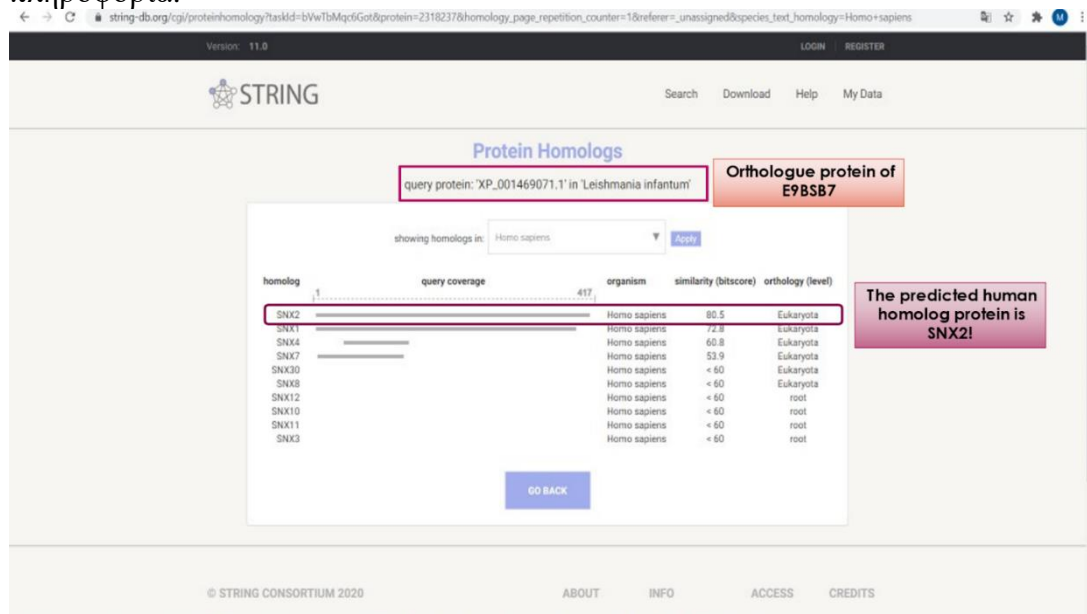
**Πίνακας Β.** Οι πρωτεΐνες που ανήκουν στο δίκτυο αλληλεπιδράσεων της XP\_001469071.1 σύμφωνα με την String DB





**Εικόνα 25.** Δίκτυο πρωτεϊνικών αλληλεπιδράσεων της πρωτεΐνης XP\_001469071.1. Στα ορθογώνια πλαίσια αναγράφονται οι χαρακτηρισμοί των ορθόλογων ανθρώπινων πρωτεϊνών που μέσω της βάσης String db ήρθαν σε αντιστοιχία με όλες τις πρωτεΐνες που εμπεριέχονται στο δίκτυο, έχοντας υποστεί χρωματική κατάταξη ανάλογα με τη λειτουργία τους.

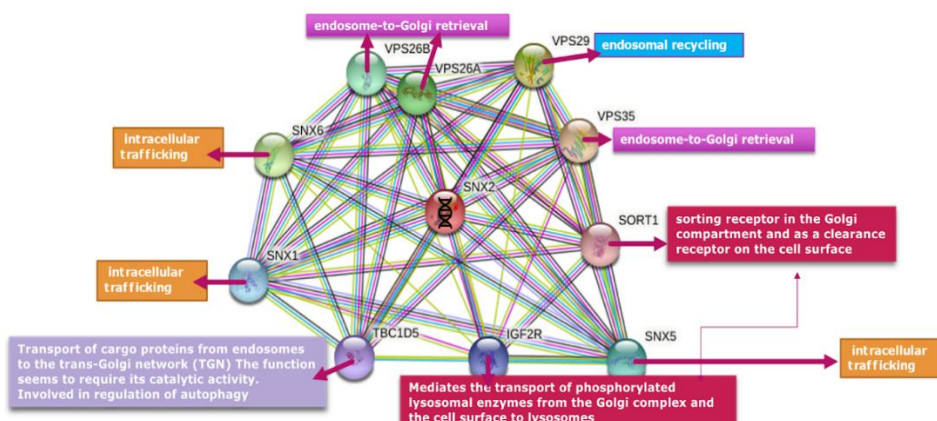
Αφού βρέθηκε το δίκτυο αλληλεπιδράσεων της ορθόλογης πρωτεΐνης της E9BSB7 από το παράσιτο *Leishmania infantum*, επόμενο βήμα αποτέλεσε η αναζήτηση της ομόλογης της στον άνθρωπο, μέσω της βάσης String. Το αποτέλεσμα της αναζήτησης αυτής (Εικόνα 26) έδωσε μια εξαιρετικά σημαντική πληροφορία.



**Εικόνα 26.** Αποτέλεσμα της αναζήτησης ανθρώπινων ορθόλογων πρωτεϊνών της XP\_001469071.1, με βάση τις αλληλεπιδράσεις εντός της βάσης δεδομένων STRING. Η πρωτεΐνη που αναδείχθηκε είναι η ανθρώπινη Νεξίνη Διαλογής 2 (Human SNX2).

Η πρωτεΐνη σημείο αναφοράς που ανήκει στο παράσιτο *Leishmania infantum* βρίσκεται σε λειτουργική αντιστοιχία με την ανθρώπινη Νεξίνη Διαλογής 2 (SNX2. Sorting Nexin 2). Η υπό μελέτη πρωτεΐνη, βρίσκεται σε αντιστοιχία (ορθόλογο

μόριο από το είδος *Leishmania donovani*) με την πρωτεΐνη αναφοράς του παρασίτου *Leishmania infantum* και κατ' επέκτασιν με την **SNX2**, γεγονός που επιβεβαιώνει την κατάταξη της E9BSB7 στην οικογένεια των Νεξινών Διαλογής. Συνεπώς, η παραδοχή αυτή μας οδηγεί στην αναζήτηση της αντιστοιχίας όλων των πρωτεϊνών που εμφανίζονται στο δίκτυο αλληλεπιδράσεων της XP\_001469071.1 με ανθρώπινες πρωτεΐνες μέσω της βάσης String, έτσι ώστε το δίκτυο να καταστεί συγκρίσιμο με αυτό της ανθρώπινης Νεξίνης Διαλογής 2 όπως φαίνεται στην Εικόνα 27.



**Εικόνα 27.** Δίκτυο πρωτεϊνικών αλληλεπιδράσεων της ανθρώπινης πρωτεΐνης SNX2. Στα ορθογώνια πλαίσια αναγράφονται οι χαρακτηρισμοί των ανθρώπινων πρωτεϊνών που αλληλεπιδρά, έχοντας υποστεί χρωματική κατάταξη ανάλογα με τη λειτουργία τους.

Επόμενο βήμα, αποτέλεσε η αναζήτηση στη βάση δεδομένων KEGG Pathways, στην οποία βρέθηκε καταχώρηση για την πρωτεΐνη που εκκρίνεται από το γονίδιο **LDBPK\_352470**. Τα στοιχεία της καταχώρησης, βρίσκονται στον Πίνακα Γ:

<b>ENTRY</b>	<b>LDBPK_352470</b>	<b>CDS</b>	<a href="#">T02289</a>
<b>DEFINITION</b>	(RefSeq) phosphoinositide-binding protein, putative		
<b>SSDB</b>	ORTHOLOGUE, PARALOGUE, CLUSTER, GFIT		
<b>MOTIF</b>	Pfam: Vps5 PX BAR_3_WASP_bdg P4Ha_N DUF4164 G0-G1_switch_2 DUF5405		
<b>OTHER DBS</b>	NCBI-GeneID: 13387704 NCBI-ProteinID: XP_003864826 UniProt: <b>A0A504X805 ,E9BSB7</b>		
<b>POSITION</b>	35:1010816..1012069		

**Πίνακας Γ.** Καταχώρηση για την πρωτεΐνη E9BSB7 εντός της βάσης δεδομένων KEGG

Αφού βρέθηκε η καταχώρηση, έγινε αναζήτηση λειτουργικά ομόλογων πρωτεϊνών στον άνθρωπο, μέσω της βάσης KEGG, με σκοπό την επαλήθευση της προηγούμενης πρόβλεψης. Τα αποτελέσματα βρίσκονται στην Εικόνα 28:

DBHit(Length)	Overlap	Ident. %	Score	Paralog	KO	Orth	Current Annotation
> ldo:LDBPK_352470(	417	99.8	2717		8		phosphoinositide-binding protein, putative
> lma:LMJF_35_2420	417	97.4	2667		10		putative phosphoinositide-binding protein
> lmi:LMXM_34_2420	417	96.9	2654		5		putative phosphoinositide-binding protein
> lpan:LPMP_342280	418	93.1	2553		4		phosphoinositide-binding protein, putative
> lbz:LBRM_34_2370	418	92.6	2545		2		putative phosphoinositide-binding protein
> tcr:510657.30(422)	420	34.8	756		6		phosphoinositide-binding protein
> tbg:TbgDa1_IX8370	429	31.5	692		3		phosphoinositide-binding protein, putative
> tbr:Tb09.211.4240	429	31.5	692		3		phosphoinositide-binding protein, putative
> ngr:NAEGRDRAFT_7	414	24.6	403	5	K17917	SNX1_2	phosphoinositide-binding protein; K17917 sorting nexin-1/2
> ehx:EMIHUDDRAFT_	439	26.2	350	170	K17917	SNX1_2	SNX1; sorting nexin 1/Sorting nexin 2/Vacuolar protein sorting pro
> pif:PITG_12915(527	422	26.1	349	9	K17917	SNX1_2	sorting nexin, putative; K17917 sorting nexin-1/2
> dgt:114522691(528	414	25.6	348	59	K17917	SNX1_2	sorting nexin-2-like; K17917 sorting nexin-1/2
> spis:111325367(492	417	24.9	346	22	K17917	SNX1_2	sorting nexin-2-like; K17917 sorting nexin-1/2
> pdam:113686695(4	417	24.5	343	20	K17917	SNX1_2	sorting nexin-2-like; K17917 sorting nexin-1/2
> dpce:113799498(56	427	23.0	341	7	K17917	SNX1_2	sorting nexin-2-like isoform X1; K17917 sorting nexin-1/2
> psj:PHYSODRAFT_	422	26.1	341	10	K17917	SNX1_2	hypothetical protein; K17917 sorting nexin-1/2
> pvm:113810186(79	417	23.5	335	474	K17917	SNX1_2	sorting nexin-2-like isoform X1; K17917 sorting nexin-1/2
> mju:108395693(521	424	25.7	329	29	K17917	SNX1_2	SNX2; sorting nexin-2 isoform X1; K17917 sorting nexin-1/2
> ola:101163736(743	417	25.7	329	42	K17917	SNX1_2	sorting nexin-1 isoform X1; K17917 sorting nexin-1/2
> spu:581074(511)	418	24.4	327	30	K17917	SNX1_2	sorting nexin-2; K17917 sorting nexin-1/2
> tmu:101358219(52	417	24.5	326	28	K17917	SNX1_2	sorting nexin-1 isoform X3; K17917 sorting nexin-1/2
> npr:108798661(504	416	24.3	324	13	K17917	SNX1_2	SNX1; sorting nexin-1; K17917 sorting nexin-1/2
> lak:106178690(528	419	23.6	322	39	K17917	SNX1_2	sorting nexin-2; K17917 sorting nexin-1/2
> amil:114954956(48	408	23.8	321	16	K17917	SNX1_2	sorting nexin-2-like; K17917 sorting nexin-1/2
> elk:111153150(521	416	24.5	321	27	K17917	SNX1_2	sorting nexin-1 isoform X1; K17917 sorting nexin-1/2
> lav:100658068(518	421	25.2	321	28	K17917	SNX1_2	SNX2; sorting nexin-2; K17917 sorting nexin-1/2
> oca:100338927(522	417	25.7	320	19	K17917	SNX1_2	SNX1; sorting nexin 1; K17917 sorting nexin-1/2
> pss:102458370(423	421	24.9	320	20	K17917	SNX1_2	SNX2; sorting nexin-2; K17917 sorting nexin-1/2
> ssc:100626970(521	421	24.9	320	34	K17917	SNX1_2	SNX2; sorting nexin-2 isoform X2; K17917 sorting nexin-1/2

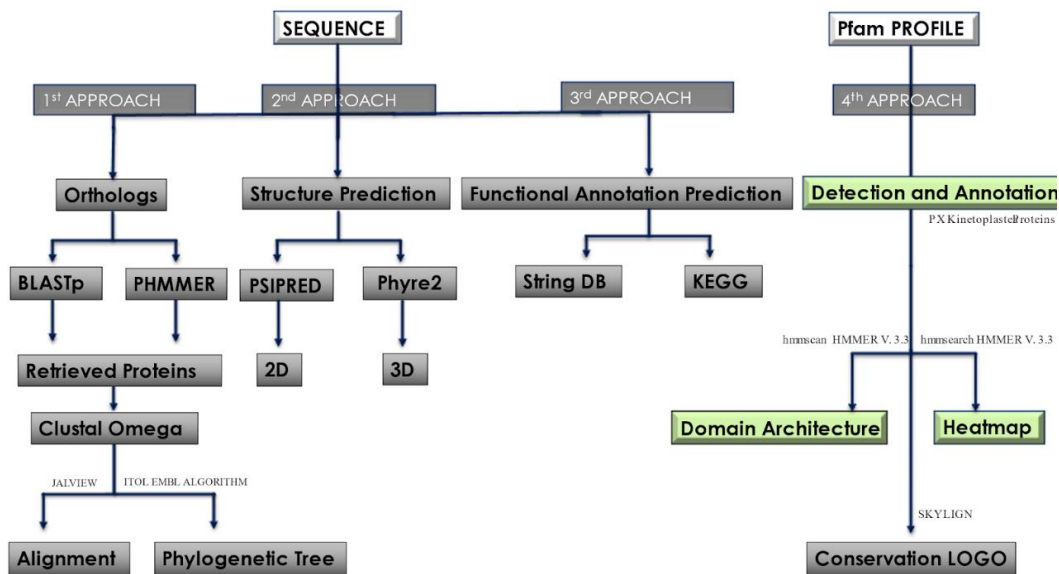
> aml:100480770(519	421	24.9	319	34	K17917	SNX1_2	SNX2; sorting nexin-2 isoform X1; K17917 sorting nexin-1/2
> btab:109042140(47	418	24.4	319	5	K17917	SNX1_2	sorting nexin-2 isoform X1; K17917 sorting nexin-1/2
> ccan:109694404(51	421	24.9	319	34	K17917	SNX1_2	Snx2; sorting nexin-2 isoform X1; K17917 sorting nexin-1/2
> cdk:105105061(519	421	24.9	319	38	K17917	SNX1_2	SNX2; sorting nexin-2 isoform X1; K17917 sorting nexin-1/2
> cfr:102511410(519	421	24.9	319	40	K17917	SNX1_2	SNX2; sorting nexin-2 isoform X1; K17917 sorting nexin-1/2
> cge:100774009(519	421	24.9	319	33	K17917	SNX1_2	Snx2; sorting nexin-2 isoform X3; K17917 sorting nexin-1/2
> dle:111182361(519	421	24.9	319	22	K17917	SNX1_2	SNX2; sorting nexin-2; K17917 sorting nexin-1/2
> eai:106842832(518	421	24.9	319	26	K17917	SNX1_2	SNX2; sorting nexin 2; K17917 sorting nexin-1/2
> ecb:100073099(518	421	24.9	319	32	K17917	SNX1_2	SNX2; sorting nexin-2; K17917 sorting nexin-1/2
> epa:110253515(479	413	23.5	319	32	K17917	SNX1_2	sorting nexin-2; K17917 sorting nexin-1/2
> epz:103557900(516	421	24.9	319	14	K17917	SNX1_2	SNX2; sorting nexin 2; K17917 sorting nexin-1/2
> fca:101083737(519	421	24.9	319	42	K17917	SNX1_2	SNX2; sorting nexin-2 isoform X2; K17917 sorting nexin-1/2
> lve:103089391(519	421	24.9	319	26	K17917	SNX1_2	SNX2; sorting nexin 2; K17917 sorting nexin-1/2
> oor:101290116(519	421	24.9	319	26	K17917	SNX1_2	SNX2; sorting nexin-2; K17917 sorting nexin-1/2
> pale:102889484(51	421	24.9	319	28	K17917	SNX1_2	SNX2; sorting nexin-2; K17917 sorting nexin-1/2
> pcd:102992715(51	421	24.9	319	25	K17917	SNX1_2	SNX2; sorting nexin-2; K17917 sorting nexin-1/2
> pcan:112553625(52	415	24.6	319	16	K17917	SNX1_2	sorting nexin-2-like; K17917 sorting nexin-1/2
> ray:107511143(519	421	24.9	319	24	K17917	SNX1_2	SNX2; sorting nexin-2 isoform X1; K17917 sorting nexin-1/2
> biu:109561859(402	421	24.9	318	14	K17917	SNX1_2	SNX2; sorting nexin-2; K17917 sorting nexin-1/2
> bom:102278052(51	421	24.9	318	22	K17917	SNX1_2	SNX2; sorting nexin 2; K17917 sorting nexin-1/2
> chx:102191692(519	421	24.9	318	28	K17917	SNX1_2	SNX2; sorting nexin 2; K17917 sorting nexin-1/2
> hai:109377441(547	417	24.9	318	32	K17917	SNX1_2	SNX1; sorting nexin 1; K17917 sorting nexin-1/2
> mun:110565942(51	421	24.9	318	24	K17917	SNX1_2	Snx2; sorting nexin-2; K17917 sorting nexin-1/2
> cjc:100413511(519	421	24.9	317	32	K17917	SNX1_2	SNX2; sorting nexin 2; K17917 sorting nexin-1/2
> mcal:110284501(51	421	24.9	317	28	K17917	SNX1_2	Snx2; sorting nexin-2; K17917 sorting nexin-1/2
> mcc:700507(519)	421	24.9	317	35	K17917	SNX1_2	SNX2; sorting nexin-2; K17917 sorting nexin-1/2
> mmu:67804(519)	421	24.9	317	23	K17917	SNX1_2	[sp:SNX2_MOUSE] Snx2, 0610030A03Rik; sorting nexin 2; K17917
> nle:100607779(519	421	24.9	317	41	K17917	SNX1_2	SNX2; sorting nexin-2 isoform X2; K17917 sorting nexin-1/2
> oaa:100091063(520	411	24.3	317	51	K17917	SNX1_2	SNX1; sorting nexin-1; K17917 sorting nexin-1/2
> rbb:108520383(518	421	24.9	317	14	K17917	SNX1_2	SNX2; sorting nexin 2; K17917 sorting nexin-1/2
> mo:291464(519)	421	24.9	317	26	K17917	SNX1_2	Snx2; sorting nexin 2; K17917 sorting nexin-1/2

**Εικόνα 28.** Εντοπισμός ανθρώπινων λειτουργικά ομόλογων πρωτεϊνών της πρωτεΐνης που κωδικοποιείται από το γονίδιο LDBPK\_352470 μέσω της βάσης KEGG

Παρατηρούμε, ότι και στην περίπτωση αυτής της βάσης δεδομένων εμφανίζεται η ανθρώπινη Νεξίνη Διαλογής 2 (SNX2). Οπότε η προηγούμενη πρόβλεψη επιβεβαιώνεται και η ανθρώπινη SNX2 προβλέπεται λειτουργικά ομόλογη με την πρωτεΐνη E9BSB7 του πρωτοζώου *Leishmania donovani*.

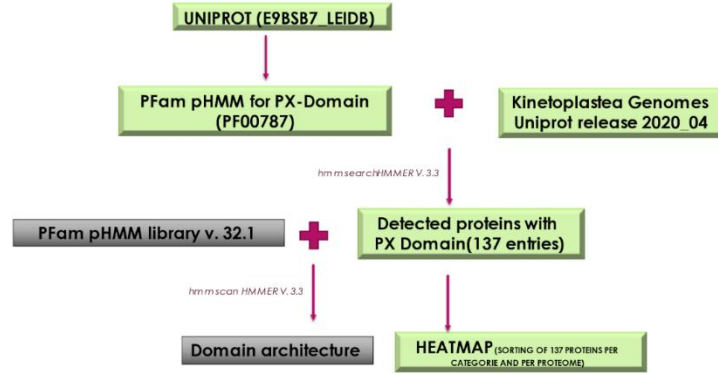
### 3.4 Εντοπισμός και μελέτη όλων των πρωτεϊνών με λειτουργική περιοχή PX που συναντώνται στις Κινετοπλαστίδες-Προσδιορισμός και οργάνωση των δομικών περιοχών (domain architecture) για τις ευρεθείσες πρωτεΐνες, Κατασκευή Heatmap και δημιουργία Logo συντηρημένων αμινοξέων των λειτουργικών περιοχών PX και VPS5.

Σε αυτό το στάδιο της ανάλυσης, προχωρήσαμε σε μια πιο γενικευμένη προσέγγιση, που αφορά όλες τις πρωτεΐνες των Κινετοπλαστιδών που περιέχουν περιοχή PX. Αρχικά, διαπιστώθηκε η ύπαρξη profile στην βάση PFAM, που αφορά την περιοχή PX. Κατεβάσαμε τα πρωτόμαχα των Κινετοπλαστιδών από τη βάση δεδομένων UNIPROT και με τη βοήθεια της εντολής hmmsearch, σε περιβάλλον LINUX, λάβαμε ως αποτέλεσμα 137 πρωτεΐνες των Κινετοπλαστιδών που εμπεριέχονται στο PX Domain Profile. Στη συνέχεια, πραγματοποιήθηκε επεξεργασία των δεδομένων που προσκομίστηκαν, ώστε να αποκτήσουν την κατάλληλη μορφή εισόδου στον αλγόριθμο Heatmapper. Έτσι κατασκευάστηκε ένα Heatmap, δηλαδή μια απεικόνιση στην οποία ταξινομήθηκαν οι πρωτεΐνες σε κατηγορίες ανάλογα με τον χαρακτηρισμό τους από την UNIPROT DB, ανά πρωτόμαχο.



**Εικόνα 29.** Σχηματική αναπαράσταση της επέκτασης της Βιοπληροφορικής ανάλυσης σε όλες τις πρωτεΐνες των Κινετοπλαστιδών που περιέχουν λειτουργική περιοχή PX.

4<sup>th</sup> APPROACH



**Εικόνα 30.** Αναλυτική περιγραφή της μεθοδολογίας που χρησιμοποιήθηκε στο βήμα που περιεγράφηκε στην Εικόνα 29.

Στη συνέχεια, με τη βοήθεια της εντολής *hmmscan*, στο ίδιο περιβάλλον έγινε αναζήτηση στη βιβλιοθήκη της PFAM όλων των δομικών περιοχών (domain architecture) που συναντώνται στις 137 πρωτεΐνες που μελετούμε. Για κάποιες από αυτές δεν ευρέθη χαρακτηρισμός και δεν συμπεριλήφθηκαν στην κατάταξη που πραγματοποιήθηκε.

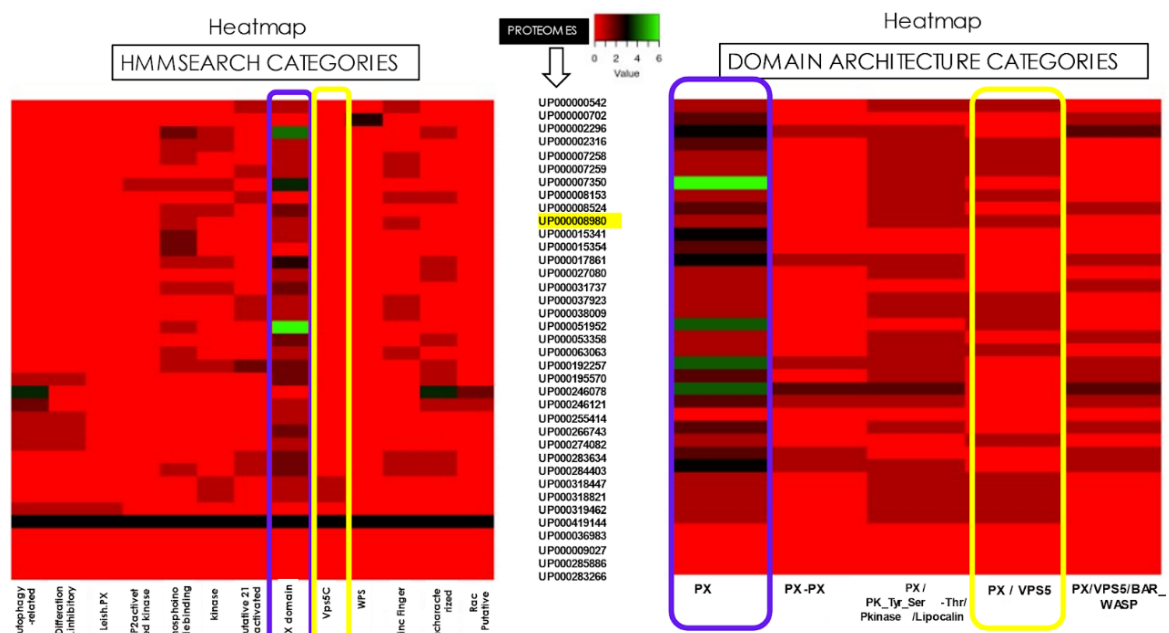
PF00787_25	E-value	Domain ID	Domain Name	Protein Name	
PF00787_25	4.1e-11	42.3	0.0	1e-10 41.0 0.0 1.6 1 0 0 1 1 1 1	1 Zinc finger protein kinase-like OS=Leishmania major OX=5664 GN=LJMF
PF00787_25	6.2e-11	41.7	0.0	9.7e-10 37.9 0.0 2.8 2 0 0 2 2 1 1	1 PK domain-containing protein OS=Leishmania major OX=5664 GN=LJMF 14
PF00787_25	2e-17	62.1	0.1	3.9e-17 61.2 0.1 1.5 1 0 0 1 1 1 1	1 WGS project CAE000000000 data, annotated contig 1462 (Fragment) OS=
PF00787_25	2e-10	39.6	0.0	9.2e-10 37.5 0.0 2.1 2 1 0 2 2 2 1	1 WGS project CAE000000000 data, annotated contig 680 (Fragment) OS=T
PF00787_25	5.3e-10	38.3	0.0	2.1e-09 36.4 0.0 2.1 2 0 0 2 2 1 1	1 Phosphoinositide-binding protein, annotated contig 2308 OS=Trypanosoma
PF00787_25	4.3e-17	62.7	0.0	1.1e-16 61.4 0.0 1.7 1 0 0 1 1 1 1	1 Phosphoinositide-binding protein, putative OS=Trypanosoma cruzi (st
PF00787_25	4.5e-17	62.6	0.0	1.1e-16 61.4 0.0 1.7 1 0 0 1 1 1 1	1 Phosphoinositide-binding protein, putative OS=Trypanosoma cruzi (st
PF00787_25	1.5e-16	61.0	0.0	2e-07 31.7 0.0 3.3 3 0 0 3 3 2 2	2 Uncharacterized protein OS=Trypanosoma cruzi (strain CL Brener) OX=
PF00787_25	1.9e-10	41.4	1.4	1.1e-09 38.9 0.0 3.1 3 1 0 3 3 1 1	1 PK domain-containing protein OS=Trypanosoma cruzi (strain CL Brener
PF00787_25	2.2e-10	41.2	0.1	8.2e-10 39.3 0.1 2.0 2 0 0 2 2 1 1	1 PK domain-containing protein OS=Trypanosoma cruzi (strain CL Brener
PF00787_25	2.2e-10	41.2	0.1	8.2e-10 39.3 0.1 2.0 2 0 0 2 2 1 1	1 PK domain-containing protein OS=Trypanosoma cruzi (strain CL Brener
PF00787_25	1.6e-07	32.0	0.1	4e-07 30.7 0.1 1.7 1 0 0 1 1 1 1	1 Protein kinase, putative OS=Trypanosoma cruzi (strain CL Brener) OX
PF00787_25	0.00018	22.2	0.0	0.00043 21.0 0.0 1.6 1 0 0 1 1 1 0	0 PK domain-containing protein (Fragment) OS=Trypanosoma cruzi (strai
PF00787_25	0.00021	22.0	0.0	0.0005 20.8 0.0 1.5 1 0 0 1 1 1 0	0 PK domain-containing protein OS=Trypanosoma cruzi (strain CL Brener
PF00787_25	3.8e-18	65.1	0.1	8.1e-18 64.0 0.1 1.6 1 0 0 1 1 1 1	1 Phosphoinositide-binding protein, putative OS=Trypanosoma brucei, ga
PF00787_25	3.1e-12	46.1	0.0	4.3e-11 42.5 0.0 2.5 2 0 0 2 2 1 1	1 T. brucei spp.-specific protein OS=Trypanosoma brucei gambiense (st
PF00787_25	2.1e-10	40.2	0.3	6e-10 38.8 0.3 1.8 1 0 0 1 1 1 1	1 PK domain-containing protein OS=Trypanosoma brucei gambiense (stai
PF00787_25	3.5e-08	33.1	0.0	1.1e-07 31.6 0.0 1.7 1 0 0 1 1 1 1	1 Protein kinase OS=Trypanosoma brucei gambiense (strain MHOM/CI/86/D
PF00787_25	6.2e-15	54.5	0.0	1.5e-14 53.3 0.0 1.7 1 0 0 1 1 1 1	1 Putative phosphoinositide-binding protein OS=Leishmania braziliensi
PF00787_25	1.5e-11	43.7	0.1	4e-11 42.3 0.1 1.6 1 0 0 1 1 1 1	1 Zinc finger protein kinase-like OS=Leishmania braziliensis OX=5660
PF00787_25	1.4e-10	40.6	0.3	1.3e-09 37.4 0.0 2.7 3 0 0 3 3 1 1	1 PK domain-containing protein OS=Leishmania braziliensis OX=5660 GN=
PF00787_25	4.7e-15	54.9	0.0	1.3e-14 53.5 0.0 1.7 1 0 0 1 1 1 1	1 Putative phosphoinositide-binding protein OS=Leishmania mexicana (s
PF00787_25	1e-11	44.2	0.0	2.4e-11 43.0 0.0 1.6 1 0 0 1 1 1 1	1 Zinc finger protein kinase-like OS=Leishmania mexicana (strain MHOM
PF00787_25	5.5e-11	41.9	0.0	9.4e-10 37.9 0.0 2.8 2 0 0 2 2 1 1	1 PK domain-containing protein OS=Leishmania mexicana (strain MHOM/GT
PF00787_25	5.5e-18	64.6	0.0	1.2e-17 63.5 0.0 1.6 1 0 0 1 1 1 1	1 Phosphoinositide-binding protein, putative OS=Trypanosoma brucei, mar
PF00787_25	4.5e-17	61.7	0.0	1e-08 34.9 0.0 3.6 3 0 0 3 3 1 1	1 PK domain-containing protein OS=Trypanosoma cruzi marinkellei OX=85
PF00787_25	5.6e-12	45.4	0.2	1.0e-11 43.7 0.2 1.8 1 0 0 1 1 1 1	1 p21-activated kinase 3, putative OS=Trypanosoma cruzi marinkellei O
PF00787_25	2.3e-11	43.4	3.2	4.4e-11 42.5 0.1 3.1 3 1 1 4 4 1 1	1 PK domain-containing protein OS=Trypanosoma cruzi marinkellei OX=85
PF00787_25	1.4e-10	40.9	1.9	2.7e-10 40.0 0.1 2.4 3 0 0 3 3 1 1	1 PK domain-containing protein OS=Trypanosoma cruzi marinkellei OX=85
PF00787_25	1.2e-07	31.5	0.1	2.9e-07 30.3 0.1 1.7 1 0 0 1 1 1 1	1 Protein kinase, putative OS=Trypanosoma cruzi marinkellei OX=85
PF00787_25	4.9e-05	23.1	0.0	0.00011 22.0 0.0 1.6 1 0 0 1 1 1 1	1 PK domain-containing protein OS=Trypanosoma cruzi marinkellei OX=85
PF00787_25	5.7e-15	54.7	0.0	1.3e-14 53.5 0.0 1.6 1 0 0 1 1 1 1	1 Putative phosphoinositide-binding protein OS=Leishmania infantum OX
PF00787_25	4.3e-11	42.2	0.0	1e-10 41.0 0.0 1.5 1 0 0 1 1 1 1	1 Zinc finger protein kinase-like OS=Leishmania infantum OX=5673 GN=
PF00787_25	1.1e-10	40.9	0.0	9.7e-10 37.9 0.0 2.6 2 0 0 2 2 1 1	1 Uncharacterized protein OS=Leishmania infantum OX=5671 GN=LJNF 14_0
PF00787_25	3.4e-18	65.1	0.1	7.2e-18 64.0 0.1 1.6 1 0 0 1 1 1 1	1 Phosphoinositide-binding protein, putative OS=Trypanosoma brucei, br
PF00787_25	2.6e-12	46.2	0.0	3.9e-11 42.5 0.0 2.5 2 0 0 2 2 1 1	1 PK domain-containing protein OS=Trypanosoma brucei brucei (strain 9
PF00787_25	1.9e-10	40.2	0.3	5.4e-10 38.8 0.3 1.8 1 0 0 1 1 1 1	1 PK domain-containing protein OS=Trypanosoma brucei brucei (strain 9
PF00787_25	3.1e-08	33.1	0.0	9.4e-08 31.6 0.0 1.7 1 0 0 1 1 1 1	1 Protein kinase OS=Trypanosoma brucei brucei (strain 927/4 GUT10.1
PF00787_25	5.6e-15	54.7	0.0	1.3e-14 53.5 0.0 1.6 1 0 0 1 1 1 1	1 Phosphoinositide-binding protein, putative OS=Leishmania donovani (
PF00787_25	3.3e-11	42.6	0.0	2.9e-10 39.6 0.0 2.6 2 0 0 2 2 1 1	1 PK domain-containing protein OS=Leishmania donovani (strain BPK282A
PF00787_25	4.2e-11	42.2	0.0	9.9e-11 41.0 0.0 1.5 1 0 0 1 1 1 1	1 Zinc finger protein kinase-like OS=Leishmania donovani (strain BPK2
PF00787_25	8.6e-12	45.3	0.2	8.6e-12 45.3 0.2 2.0 2 1 0 2 2 1 1	1 Phosphoinositide-binding protein OS=Angomonas deanei OX=59799 GN=AG
PF00787_25	9.1e-12	45.2	0.2	9.1e-12 45.2 0.2 2.0 2 1 0 2 2 1 1	1 Phosphoinositide-binding protein OS=Angomonas deanei OX=59799 GN=AG
PF00787_25	1.5e-06	58.5	0.1	4.5e-06 27.0 0.1 1.9 1 0 0 1 1 1 1	1 PK domain-containing protein OS=Angomonas deanei OX=59799 GN=AGDE 1
PF00787_25	2.7e-15	56.0	0.3	7.1e-15 54.6 0.3 1.7 1 0 0 1 1 1 1	1 Phosphoinositide-binding protein OS=Strigomonas culicis OX=28085 GN=
PF00787_25	2.7e-15	56.0	0.3	7.1e-15 54.6 0.3 1.7 1 0 0 1 1 1 1	1 Phosphoinositide-binding protein OS=Strigomonas culicis OX=28085 GN=
PF00787_25	2.6e-17	62.6	0.0	6.4e-17 61.4 0.0 1.7 1 0 0 1 1 1 1	1 Phosphoinositide-binding protein OS=Trypanosoma cruzi Dm28c OX=1416
PF00787_25	3.1e-11	43.1	1.5	1.8e-10 40.7 0.0 3.3 3 1 4 4 1 1	1 PK domain-containing protein OS=Trypanosoma cruzi Dm28c OX=1416333
PF00787_25	1.3e-10	41.2	0.1	4.8e-10 39.3 0.1 2.0 2 0 0 2 2 1 1	1 PK domain-containing protein OS=Trypanosoma cruzi Dm28c OX=1416333
PF00787_25	5.4e-08	32.8	0.2	1.4e-07 31.5 0.2 1.7 1 0 0 1 1 1 1	1 Protein kinase OS=Trypanosoma cruzi Dm28c OX=1416333 GN=TCDF 07107
PF00787_25	3.8e-05	23.7	0.0	7.0e-05 22.6 0.0 1.4 1 0 0 1 1 1 1	1 PK domain-containing protein OS=Trypanosoma cruzi Dm28c OX=1416333
PF00787_25	2.1e-11	42.9	0.5	7.8e-11 41.1 0.2 2.1 2 0 0 2 2 1 1	1 PK domain-containing protein OS=Phytomonas sp. isolate EMI OX=13400
PF00787_25	6.3e-10	38.2	0.1	1.4e-09 37.0 0.1 1.6 1 0 0 1 1 1 1	1 Uncharacterized protein OS=Phytomonas sp. isolate EMI OX=13400 GN=
PF00787_25	9.7e-17	60.2	0.1	2.2e-16 59.0 0.1 1.7 1 0 0 1 1 1 1	1 Phosphoinositide-binding protein OS=Trypanosoma rangeli SC58 OX
PF00787_25	1.4e-14	53.3	0.0	1.2e-06 27.0 0.0 3.1 3 0 0 3 3 1 1	1 PK domain-containing protein OS=Trypanosoma rangeli SC58 OX
PF00787_25	5.1e-12	45.1	0.0	2e-11 43.2 0.0 2.1 1 0 0 1 1 1 1	1 PK domain-containing protein OS=Trypanosoma rangeli SC58 OX
PF00787_25	9.6e-08	31.4	0.1	2.3e-07 30.1 0.1 1.6 1 0 0 1 1 1 1	1 Protein kinase OS=Trypanosoma rangeli SC58 OX=429131 GN=TRS
PF00787_25	7.9e-16	57.6	0.0	1.8e-15 56.5 0.0 1.6 1 0 0 1 1 1 1	1 Putative phosphoinositide-binding protein OS=Leptomonas pyrrochoris OX=1
PF00787_25	1.3e-11	44.1	0.0	6e-11 41.9 0.0 2.1 2 0 0 2 2 1 1	1 PK domain-containing protein OS=Leptomonas pyrrochoris OX=1
PF00787_25	1.8e-10	40.4	0.0	4e-10 39.3 0.0 1.5 1 0 0 1 1 1 1	1 Putative zinc finger protein kinase OS=Leptomonas pyrrochoris OX=1
PF00787_25	4.3e-16	58.3	0.0	1.3e-15 56.8 0.0 1.8 1 0 0 1 1 1 1	1 Putative phosphoinositide-binding protein OS=Leptomonas sey

**Εικόνα 31** Η έξοδος με τις πρωτεΐνες που εντοπίστηκαν μετά τη χρήση της εντολής *hmmscan* σε περιβάλλον Linux

#	target name	accession	tlen	query name	accession	qlen	full sequence	E-value	score	bias	# of	c-value	i-Evalue	score	bias	hmm	coord	al
							---	---	---	---	---	---	---	---	---	from	to	fr
Vps5	PF09325.11	236	tr	E9AF66 E9AF66_LEIMA	417	4.9e-17	62.3	4.0	1	1	7.2e-21	7.1e-17	81.8	4.8	30	231	21	
PX	PF00787.25	84	tr	E9AF66 E9AF66_LEIMA	417	1.2e-14	54.7	0.0	1	1	3.2e-18	2.9e-14	53.5	0.0	10	80	4	
Kinase	PF00669.26	264	tr	Q40B98 Q40B98_LEIMA	1323	2.1e-62	210.9	0.0	1	1	6.9e-66	3.2e-62	210.3	0.0	2	264	64	
Lipocalin_5	PF13924.7	142	tr	Q40B98 Q40B98_LEIMA	1323	1.6e-34	119.7	0.0	1	1	6.5e-38	3e-34	117.9	0.0	1	136	106	
PK_Tyr_Ser-Thr	PF07714.18	259	tr	Q40B98 Q40B98_LEIMA	1323	2.3e-32	112.3	0.0	1	1	8.1e-36	3.7e-32	111.7	0.0	2	220	64	
Vps5	PF00787.25	84	tr	Q40B98 Q40B98_LEIMA	1323	9.4e-11	42.3	0.0	1	1	5e-14	2.3e-10	41.0	0.0	6	82	14	
PX	PF00787.25	84	tr	Q40B98 Q40B98_LEIMA	1323	1.4e-10	41.7	0.0	1	1	1.2e-13	2.2e-09	37.9	0.0	4	80	4	
Vps5	PF09325.11	236	tr	F9W5Y5 F9W5Y5_TRYCI	303	6.2e-17	62.1	0.1	1	1	2e-20	1.2e-16	61.2	0.1	5	83	4	
PX	PF00787.25	84	tr	F9W5Y5 F9W5Y5_TRYCI	303	5.9e-09	35.8	0.0	1	1	1.7e-12	1.1e-08	34.9	0.0	43	150	15	
Vps5	PF09325.11	236	tr	F9W5Y5 F9W5Y5_TRYCI	303	1.5e-07	31.3	0.0	1	1	3.7e-11	2.2e-07	30.7	0.0	29	128	15	
PX	PF00787.25	84	tr	F9W5Y5 F9W5Y5_TRYCI	303	6.3e-10	39.6	0.0	1	1	1.6e-13	2.8e-09	37.5	0.0	8	80	6	
PX	PF00787.25	84	tr	F9W5Y5 F9W5Y5_TRYCI	303	1.6e-09	38.3	0.0	1	1	3.6e-13	6.6e-09	36.4	0.0	8	80	6	
Vps5	PF09325.11	236	tr	Q4D534 Q4D534_TRYCC	422	3.6e-23	82.4	3.1	1	1	7.8e-27	4.7e-23	82.0	3.1	8	234	17	
PX	PF00787.25	84	tr	Q4D534 Q4D534_TRYCC	422	4.1e-17	62.7	0.0	1	1	1.7e-20	1e-16	61.4	0.0	5	83	4	
PX	PF00787.25	84	tr	Q4D534 Q4D534_TRYCC	422	2.6e-07	30.4	0.0	1	1	2.2e-10	1.3e-06	28.1	0.0	43	226	15	
Vps5	PF09325.11	236	tr	Q4D6U6 Q4D6U6_TRYCC	422	9.2e-23	81.1	2.1	1	1	2e-26	1.2e-22	80.7	2.1	9	234	17	
PX	PF00787.25	84	tr	Q4D6U6 Q4D6U6_TRYCC	422	4.3e-17	62.6	0.0	1	1	1.7e-20	1e-16	61.4	0.0	5	83	4	
PX	PF00787.25	84	tr	Q4D6U6 Q4D6U6_TRYCC	422	5.9e-07	29.2	1.0	1	1	7.2e-10	4.4e-06	26.4	0.9	43	226	15	
BAR_3_WASP_bdg	PF10456.10	236	tr	Q4DW43 Q4DW43_TRYCC	590	1.4e-16	61.0	0.0	2	2	1.5e-09	2.7e-05	24.8	0.0	8	68	2	
PX	PF00787.25	84	tr	Q4DW43 Q4DW43_TRYCC	590	1.4e-16	61.0	0.0	2	2	1e-11	1.9e-07	31.7	0.0	7	66	25	
PX	PF00787.25	84	tr	Q4DT19 Q4DT19_TRYCC	489	1.8e-10	41.4	1.4	1	1	5.9e-14	1.1e-09	39.9	0.0	7	82	5	
PX	PF00787.25	84	tr	Q4DXU3 Q4DXU3_TRYCC	498	2.1e-10	41.2	0.1	1	1	4.3e-14	7.8e-10	39.3	0.1	6	71	16	
PX	PF00787.25	84	tr	Q4DXU3 Q4DXU3_TRYCC	498	1.2e-10	42.1	0.1	1	1	4.3e-14	7.8e-10	39.3	0.1	6	71	16	
Kinase	PF00669.26	264	tr	Q4DXB5 Q4DXB5_TRYCC	928	2.7e-63	213.8	0.0	1	1	1.1e-66	4.8e-63	213.0	0.0	2	264	42	
Lipocalin_5	PF13924.7	142	tr	Q4DXB5 Q4DXB5_TRYCC	928	1.3e-36	125.5	0.0	1	1	6.3e-40	2.9e-36	124.4	0.0	2	142	76	
PK_Tyr_Ser-Thr	PF07714.18	259	tr	Q4DXB5 Q4DXB5_TRYCC	928	1.5e-32	116.5	0.0	1	1	3.5e-37	1.6e-33	116.1	0.0	2	250	42	
Vps5	PF00787.25	84	tr	Q4DXB5 Q4DXB5_TRYCC	928	1.5e-07	32.0	0.1	1	1	8.3e-11	3.8e-07	30.7	0.1	6	82	13	
PX	PF00787.25	84	tr	C9Z2B2 C9Z2B2_TRYB9	419	1.7e-21	77.0	7.0	1	1	4.2e-25	2.5e-21	76.4	7.0	29	234	15	
Vps5	PF09325.11	236	tr	C9Z2B2 C9Z2B2_TRYB9	419	7.2e-18	65.1	0.1	1	1	2.5e-21	1.5e-17	64.0	0.1	5	83	4	
BAR_3_WASP_bdg	PF10456.10	236	tr	C9Z2B2 C9Z2B2_TRYB9	419	1.6e-10	40.9	5.6	1	1	2e-11	1.2e-07	31.5	5.3	42	226	15	
PX	PF00787.25	84	tr	C9Z2B2 C9Z2B2_TRYB9	419	5.8e-12	46.9	0.0	1	1	4.5e-15	8.2e-11	42.5	0.0	9	80	7	
PX	PF00787.25	84	tr	C9Z2B2 C9Z2B2_TRYB9	419	1.6e-10	40.2	0.3	1	1	6.2e-14	1.1e-09	38.8	0.3	2	74	12	
Kinase	PF00669.26	264	tr	D0A8B7 D0A8B7_TRYB9	858	1.4e-62	211.5	0.0	1	1	4.1e-66	1.9e-62	211.1	0.0	2	264	36	
Lipocalin_5	PF13924.7	142	tr	D0A8B7 D0A8B7_TRYB9	858	1.7e-39	133.4	0.2	1	1	2e-42	1.2e-38	132.8	0.2	2	141	71	
PK_Tyr_Ser-Thr	PF07714.18	259	tr	D0A8B7 D0A8B7_TRYB9	858	1.7e-35	122.6	0.0	1	1	5.6e-39	2.6e-35	122.0	0.0	2	250	36	
Vps5	PF09325.11	236	tr	D0A8B7 D0A8B7_TRYB9	858	6.6e-08	33.1	0.0	1	1	4.4e-11	2e-07	31.6	0.0	5	82	5	
PX	PF00787.25	84	tr	D0A8B7 D0A8B7_TRYB9	858	1.6e-16	60.6	0.0	1	1	2.6e-19	2.3e-15	60.1	0.0	31	231	23	
Vps5	PF09325.11	236	tr	A4HMRE A4HMRE_LEIBR	418	1.4e-14	54.5	0.0	1	1	3.7e-18	3.4e-14	53.3	0.0	10	80	4	
Kinase	PF00669.26	264	tr	A4HGMI A4HGMI_LEIBR	1313	3.8e-64	216.8	0.0	1	1	4.8e-67	4.6e-64	216.4	0.0	2	264	63	
PK_Tyr_Ser-Thr	PF07714.18	259	tr	A4HGMI A4HGMI_LEIBR	1313	1.9e-34	119.1	0.0	1	1	6.4e-38	2.9e-34	118.5	0.0	2	220	62	
Lipocalin_5	PF13924.7	142	tr	A4HGMI A4HGMI_LEIBR	1313	9.4e-34	116.2	0.0	1	1	3.5e-37	1.6e-33	115.5	0.0	1	135	105	
PX	PF00787.25	84	tr	A4HGMI A4HGMI_LEIBR	1313	3.4e-11	43.7	0.1	1	1	2e-14	9.1e-11	42.3	0.1	6	82	14	
PX	PF00787.25	84	tr	A4H7I2 A4H7I2_LEIBR	454	3.1e-10	40.6	0.3	1	1	1.7e-13	3e-09	37.4	0.0	3	74	4	
Vps5	PF09325.11	236	tr	E9B6C2 E9B6C2_LEIMU	418	6.8e-17	61.9	4.0	1	1	1.1e-20	1e-16	61.3	4.0	29	231	21	
PX	PF00787.25	84	tr	E9B6C2 E9B6C2_LEIMU	418	1.9e-14	54.9	0.0	1	1	3.2e-18	2.9e-14	53.5	0.0	10	80	4	
Kinase	PF00669.26	264	tr	E9AZ77 E9AZ77_LEIMU	1322	5.1e-63	212.9	0.0	1	1	1.6e-66	7.5e-63	212.4	0.0	2	264	64	
PK_Tyr_Ser-Thr	PF07714.18	259	tr	E9AZ77 E9AZ77_LEIMU	1322	1.7e-33	116.0	0.0	1	1	5.8e-37	2.7e-33	115.4	0.0	2	220	64	
Lipocalin_5	PF13924.7	142	tr	E9AZ77 E9AZ77_LEIMU	1322	6.8e-33	113.4	0.0	1	1	3.2e-36	1.5e-32	112.4	0.0	1	116	106	
PX	PF00787.25	84	tr	E9AZ77 E9AZ77_LEIMU	1322	2.3e-11	44.2	0.0	1	1	1.2e-14	5.5e-11	43.0	0.0	6	82	14	
PX	PF00787.25	84	tr	E9APL5 E9APL5_LEIMU	454	1.2e-10	41.9	0.0	1	1	1.2e-13	2.1e-09	37.9	0.0	3	80	4	
Vps5	PF09325.11	236	tr	K2NR09 K2NR09_TRYCR	422	1.2e-22	80.7	1.4	1	1	2.6e-26	1.6e-22	80.3	1.4	5	234	17	
PX	PF00787.25	84	tr	K2NR09 K2NR09_TRYCR	422	1e-17	64.6	0.0	1	1	3.7e-21	2.3e-17	63.5	0.0	3	83	2	
PX	PF00787.25	84	tr	K2NR09 K2NR09_TRYCR	422	8.2e-17	63.7	0.0	1	1	1e-12	1.9e-08	34.9	0.0	7	66	25	
PX	PF00787.25	84	tr	K2NBW3 K2NBW3_TRYCR	576	1e-11	45.4	0.2	1	1	1.8e-15	3.3e-11	43.7	0.2	11	73	3	
PX	PF00787.25	84	tr	K2NBW3 K2NBW3_TRYCR	576	1e-11	45.4	0.2	1	1	1.8e-15	3.3e-11	43.7	0.2	11	73	3	
PX	PF00787.25	84	tr	K2N6W1 K2N6W1_TRYCR	489	4.2e-11	43.4	3.2	1	1	4.4e-15	8e-11	42.5	0.1	7	82	2	
PX	PF00787.25	84	tr	K2N1F1 K2N1F1_TRYCR	498	6.8e-10	40.9	1.9	1	1	2.7e-14	5e-10	41.0	0.1	4	71	16	
Kinase	PF00669.26	264	tr	K2NTY1 K2NTY1_TRYCR	852	1.7e-63	214.5	0.0	1	1	6.2e-67	2.8e-63	213.8	0.0	2	264	34	
Lipocalin_5	PF13924.7	142	tr	K2NTY1 K2NTY1_TRYCR	852	3.8e-36	124.0	0.1	1	1	2e-39	9.3e-36	122.7	0.0	2	142	76	
PK_Tyr_Ser-Thr	PF07714.18	259	tr	K2NTY1 K2NTY1_TRYCR	852	3.2e-34	118.4	0.0	1	1	1.2e-37	5.6e-34	117.6	0.0	2	250	34	

Εικόνα 32. Εξοδος της εντολής hmmscan σε περιβάλλον Linux

Επιπλέον, κατασκευάστηκε ένα νέο Heatmap, μοιώς με παραπάνω, στο οποίο η κατηγοριοποίηση των πρωτεϊνών έγινε με βάση τη δομική αρχιτεκτονική τους (που ευρέθη παραπάνω) και όχι τον χαρακτηρισμό από τη UNIPROT.



Εικόνα 33. Heatmaps με βάση τους χαρακτηρισμούς της UNIPROT Database (αριστερά) και με βάση τη δομική αρχιτεκτονική που ευρέθη με τη χρήση της bash script σε LINUX (δεξιά)

PROTEIN	DOMAIN	DESCRIPTION	ORGANISM
A0A061ITV6	PX PF00787.25 380 442	PX domain-containing protein	Trypanosoma rangeli SC58
A0A061J1E1	PX PF00787.25 60 153	PX domain-containing protein	Trypanosoma rangeli SC58
A0A088RMZ2	PX PF00787.25 46 123	PX domain-containing protein	Leishmania panamensis
A0A0N0DTB2	PX PF00787.25 176 259	PX domain-containing protein	Leptomonas pyrrocoris
A0A0N1HW83	PX PF00787.25 175 258	PX domain-containing protein	Leptomonas seymouri
A0A0S4ITC8	PX PF00787.25 212 294	PX domain-containing protein	Bodo saltans (Flagellated pr
A0A0S4IZJ1	PX PF00787.25 160 242	PX domain-containing protein	Bodo saltans (Flagellated pr
A0A0S4JU41	PX PF00787.25 54 114	PX domain-containing protein	Bodo saltans (Flagellated pr
A0A0S4KMG1	PX PF00787.25 34 123	PX domain-containing protein	Bodo saltans (Flagellated pr
A0A1G4I280	PX PF00787.25 70 157	PX domain containing protein,	Trypanosoma equiperdum
A0A1G4IHR1	PX PF00787.25 127 205	PX domain containing protein,	Trypanosoma equiperdum
A0A1X0NNT5	PX PF00787.25 126 197	PX domain-containing protein	Trypanosoma theileri
A0A1X0P2S5	PX PF00787.25 47 119	Putative p21-activated kinase	Trypanosoma theileri
A0A1X0P5L8	PX PF00787.25 39 110	Putative p21-activated kinase	Trypanosoma theileri
A0A1X0PA65	PX PF00787.25 60 152	PX domain-containing protein	Trypanosoma theileri
A0A2V2VA52	PX PF00787.25 108 178	Uncharacterized protein	Trypanosoma cruzi
A0A2V2VGR3	PX PF00787.25 108 178	PX domain-containing protein	Trypanosoma cruzi
A0A2V2VNB6	PX PF00787.25 59 152	Autophagy-related protein 24	Trypanosoma cruzi
A0A2V2VUF8	PX PF00787.25 59 152	Autophagy-related protein 24	Trypanosoma cruzi
A0A2V2X1B9	PX PF00787.25 108 178	Uncharacterized protein	Trypanosoma cruzi
A0A2V2X5G7	PX PF00787.25 59 152	Autophagy-related protein 24	Trypanosoma cruzi
A0A3L6L4U9	PX PF00787.25 127 205	PX domain containing protein	Trypanosoma brucei equipe
A0A3L6LFT2	PX PF00787.25 70 157	PX domain containing protein	Trypanosoma brucei equipe
A0A3P3Z1A4	PX PF00787.25 46 123	PX_domain_containing_protei	Leishmania braziliensis MHC
A0A3R7LA20	PX PF00787.25 100 182	PX domain-containing protein	Trypanosoma rangeli
A0A3R7MPZ5	PX PF00787.25 60 150	PX domain-containing protein	Trypanosoma rangeli
A0A3S7WSS9	PX PF00787.25 47 129	PX domain containing protein,	Leishmania donovani
A0A422PZG9	PX PF00787.25 41 104	Putative p21-activated kinase	Trypanosoma conorhini
A0A422Q9D5	PX PF00787.25 61 154	PX domain-containing protein	Trypanosoma conorhini
A0A422QAE2	PX PF00787.25 107 179	PX domain-containing protein	Trypanosoma conorhini
A0A504XCY2	PX PF00787.25 164 247	PX domain family protein	Leishmania donovani
A0A504Y5F1	PX PF00787.25 164 246	PX domain family protein	Leishmania donovani

**Εικόνα 34.** Πρωτεΐνες των Κινητοπλαστιδών με δομή PX έπειτα από επεξεργασία στο πρόγραμμα EXCEL. Τα δεδομένα για τις στήλες «Description, Organism», συλλέχθηκαν από τη βάση δεδομένων Uniprot.

PROTEIN	DOMAIN	DESCRIPTION	ORGANISM
A0A640KD33	PX PF00787.25 47 129	PX domain-containing protein	Leishmania tarentolae (Saur
A4H7I2	PX PF00787.25 46 123	PX domain-containing protein	Leishmania braziliensis
A4HVV4	PX PF00787.25 47 129	PX_domain_containing_protei	Leishmania infantum
C9ZLQ6	PX PF00787.25 70 157	T. brucei spp.-specific protein	Trypanosoma brucei gambie
C9ZT62	PX PF00787.25 127 205	PX domain-containing protein	Trypanosoma brucei gambie
E9APL5	PX PF00787.25 46 129	PX domain-containing protein	Leishmania mexicana (strair
E9BBQ7	PX PF00787.25 47 130	PX domain-containing protein	Leishmania donovani (strair
F9WCA9	PX PF00787.25 69 157	WGS project CAEQ00000000 da	Trypanosoma congolense (s
F9WHP7	PX PF00787.25 69 157	WGS project CAEQ00000000 da	Trypanosoma congolense (s
K2MV78	PX PF00787.25 124 197	PX domain-containing protein	Trypanosoma cruzi marinke
K2N1F1	PX PF00787.25 108 178	PX domain-containing protein	Trypanosoma cruzi marinke
K2N6W1	PX PF00787.25 59 152	PX domain-containing protein	Trypanosoma cruzi marinke
K2NBW3	PX PF00787.25 34 96	p21-activated kinase 3, putativ	Trypanosoma cruzi marinke
K2NRG5	PX PF00787.25 296 357	PX domain-containing protein	Trypanosoma cruzi marinke
Q4DJ38	PX PF00787.25 108 178	PX domain-containing protein	Trypanosoma cruzi (strain Cl
Q4DTI9	PX PF00787.25 59 152	PX domain-containing protein	Trypanosoma cruzi (strain Cl
Q4DUX3	PX PF00787.25 108 178	PX domain-containing protein	Trypanosoma cruzi (strain Cl
Q4QFW3	PX PF00787.25 47 129	PX domain-containing protein	Leishmania major
Q57X66	PX PF00787.25 127 205	PX domain-containing protein	Trypanosoma brucei brucei (
Q57ZY9	PX PF00787.25 70 157	PX domain-containing protein	Trypanosoma brucei brucei (
S9TTG7	PX PF00787.25 118 200	PX domain-containing protein	Angomonas deanei
S9UF50	PX PF00787.25 35 107	Phosphoinositide-binding pro	Angomonas deanei
S9UF93	PX PF00787.25 43 139	Phosphoinositide-binding pro	Strigomonas culicis
S9V6F4	PX PF00787.25 35 107	Phosphoinositide-binding pro	Angomonas deanei
S9VAW7	PX PF00787.25 43 139	Phosphoinositide-binding pro	Strigomonas culicis
V5B677	PX PF00787.25 49 142	PX domain-containing protein	Trypanosoma cruzi Dm28c
V5BCZ4	PX PF00787.25 178 264	PX domain-containing protein	Trypanosoma cruzi Dm28c
V5BMA1	PX PF00787.25 108 178	PX domain-containing protein	Trypanosoma cruzi Dm28c
W6KPD0	PX PF00787.25 131 203	PX domain-containing protein	Phytomonas sp. isolate EM1
W6L5R0	PX PF00787.25 135 203	PX domain-containing protein	Phytomonas sp. isolate Hart

**Εικόνα 35.** Πρωτεΐνες των Κινητοπλαστιδών με δομή PX έπειτα από επεξεργασία στο πρόγραμμα EXCEL. Τα δεδομένα για τις στήλες «Description, «Organism», συλλέχθηκαν από τη βάση δεδομένων Uniprot.



PROTEIN	1 <sup>ST</sup> DOMAIN	2 <sup>ND</sup> DOMAIN	3 <sup>RD</sup> DOMAIN	4 <sup>TH</sup> DOMAIN	DESCRIPTION	ORGANISM
K2NTY1	PX PF00787.25 60 148	PK_Tyr_Ser-Thr PF07714.18 34	Pkinase PF00069.26 349 602	Lipocalin_5 PF13924.7 705 849	Protein kinase, putative	Trypanosoma cruzi marinkellei
A0A2V2XFU5	PX PF00787.25 60 148	PK_Tyr_Ser-Thr PF07714.18 35	Pkinase PF00069.26 354 607	Lipocalin_5 PF13924.7 710 854	Putative rac serine-threonin	Trypanosoma cruzi
A0A2V2WM36	PX PF00787.25 60 148	PK_Tyr_Ser-Thr PF07714.18 35	Pkinase PF00069.26 355 608	Lipocalin_5 PF13924.7 710 855	Putative rac serine-threonin	Trypanosoma cruzi
A0A2V2W234	PX PF00787.25 60 148	PK_Tyr_Ser-Thr PF07714.18 35	Pkinase PF00069.26 355 608	Lipocalin_5 PF13924.7 711 855	Putative rac serine-threonin	Trypanosoma cruzi
V5BFB5	PX PF00787.25 60 148	PK_Tyr_Ser-Thr PF07714.18 35	Pkinase PF00069.26 355 608	Lipocalin_5 PF13924.7 711 855	Protein kinase	Trypanosoma cruzi Dm28c
A0A3R7N046	PX PF00787.25 59 148	PK_Tyr_Ser-Thr PF07714.18 36	Pkinase PF00069.26 362 615	Lipocalin_5 PF13924.7 721 865	Zinc finger protein kinase	Trypanosoma rangeli
A0A061JY78	PX PF00787.25 59 148	PK_Tyr_Ser-Thr PF07714.18 36	Pkinase PF00069.26 363 616	Lipocalin_5 PF13924.7 722 866	Protein kinase	Trypanosoma rangeli SC58
A0A3R7LMC9	PX PF00787.25 61 148	PK_Tyr_Ser-Thr PF07714.18 36	Pkinase PF00069.26 363 616	Lipocalin_5 PF13924.7 725 868	Zinc finger protein kinase	Trypanosoma conorhini
A0A1G4I7W9	PX PF00787.25 59 148	PK_Tyr_Ser-Thr PF07714.18 36	Pkinase PF00069.26 364 617	Lipocalin_5 PF13924.7 712 854	Differentiation inhibitory ki	Trypanosoma equiperdum
A0A3L6K5S2	PX PF00787.25 59 148	PK_Tyr_Ser-Thr PF07714.18 36	Pkinase PF00069.26 364 617	Lipocalin_5 PF13924.7 712 854	Differentiation inhibitory ki	Trypanosoma brucei equiperdum
D0A8B7	PX PF00787.25 59 148	PK_Tyr_Ser-Thr PF07714.18 36	Pkinase PF00069.26 364 617	Lipocalin_5 PF13924.7 712 854	Protein kinase	Trypanosoma brucei gambiense (strain M
Q384A2	PX PF00787.25 59 148	PK_Tyr_Ser-Thr PF07714.18 36	Pkinase PF00069.26 364 617	Lipocalin_5 PF13924.7 712 854	Protein kinase	Trypanosoma brucei brucei (strain 927/4 c
A0A1X0NSM0	PX PF00787.25 59 147	PK_Tyr_Ser-Thr PF07714.18 41	Pkinase PF00069.26 415 666	Lipocalin_5 PF13924.7 783 929	Protein kinase	Trypanosoma theileri
Q4DXB5	PX PF00787.25 130 218	PK_Tyr_Ser-Thr PF07714.18 42	Pkinase PF00069.26 425 678	Lipocalin_5 PF13924.7 781 925	Protein kinase, putative	Trypanosoma cruzi (strain CL Brener)
W6L5D5	PX PF00787.25 94 167	PK_Tyr_Ser-Thr PF07714.18 59	Pkinase PF00069.26 598 852	Lipocalin_5 PF13924.7 973 1086	Uncharacterized protein	Phytomonas sp. isolate Hart1
W6KMR7	PX PF00787.25 94 167	PK_Tyr_Ser-Thr PF07714.18 60	Pkinase PF00069.26 600 854	Lipocalin_5 PF13924.7 975 1084	Uncharacterized protein	Phytomonas sp. isolate EM1
A0A0N0VHJ3	PX PF00787.25 147 234	PK_Tyr_Ser-Thr PF07714.18 62	Pkinase PF00069.26 627 928	Lipocalin_5 PF13924.7 1070 1189	Putative zinc finger protein	Leptomonas pyrrocoris
A0A088RU16	PX PF00787.25 141 230	PK_Tyr_Ser-Thr PF07714.18 62	Pkinase PF00069.26 629 908	Lipocalin_5 PF13924.7 1052 1196	Zinc finger protein kinase-li	Leishmania panamensis
A0A3P3ZB12	PX PF00787.25 141 230	PK_Tyr_Ser-Thr PF07714.18 62	Pkinase PF00069.26 629 908	Lipocalin_5 PF13924.7 1052 1197	Differentiation_inhibitory ki	Leishmania braziliensis MHOM/BR/75/M2
A4HGMI	PX PF00787.25 148 237	PK_Tyr_Ser-Thr PF07714.18 63	Pkinase PF00069.26 636 915	Lipocalin_5 PF13924.7 1059 1204	Zinc finger protein kinase-li	Leishmania braziliensis
A0A0N11MC0	PX PF00787.25 146 234	PK_Tyr_Ser-Thr PF07714.18 63	Pkinase PF00069.26 637 932	Lipocalin_5 PF13924.7 1076 1192	Putative zinc finger protein	Leptomonas seymouri
A0A640KLF8	PX PF00787.25 141 230	PK_Tyr_Ser-Thr PF07714.18 64	Pkinase PF00069.26 643 922	Lipocalin_5 PF13924.7 1066 1217	Zinc finger protein kinase-li	Leishmania tarentolae (Saurleishmania 1
E9AZY7	PX PF00787.25 141 230	PK_Tyr_Ser-Thr PF07714.18 64	Pkinase PF00069.26 643 922	Lipocalin_5 PF13924.7 1066 1187	Zinc finger protein kinase-li	Leishmania mexicana (strain MHOM/GT/2
Q4Q898	PX PF00787.25 141 230	PK_Tyr_Ser-Thr PF07714.18 64	Pkinase PF00069.26 643 922	Lipocalin_5 PF13924.7 1066 1213	Zinc finger protein kinase-li	Leishmania major
A0A3Q8IQN3	PX PF00787.25 141 230	PK_Tyr_Ser-Thr PF07714.18 64	Pkinase PF00069.26 643 922	Lipocalin_5 PF13924.7 1066 1218	Differentiation_inhibitory ki	Leishmania donovani
A0A504WX74	PX PF00787.25 141 230	PK_Tyr_Ser-Thr PF07714.18 64	Pkinase PF00069.26 643 922	Lipocalin_5 PF13924.7 1066 1218	Differentiation_inhibitory ki	Leishmania donovani
A0A504Y059	PX PF00787.25 141 230	PK_Tyr_Ser-Thr PF07714.18 64	Pkinase PF00069.26 643 922	Lipocalin_5 PF13924.7 1066 1218	Protein kinase domain fami	Leishmania donovani
A4I3P6	PX PF00787.25 141 230	PK_Tyr_Ser-Thr PF07714.18 64	Pkinase PF00069.26 643 922	Lipocalin_5 PF13924.7 1066 1218	Differentiation_inhibitory ki	Leishmania infantum
E9BJQ1	PX PF00787.25 141 230	PK_Tyr_Ser-Thr PF07714.18 64	Pkinase PF00069.26 643 922	Lipocalin_5 PF13924.7 1066 1218	Zinc finger protein kinase-li	Leishmania donovani (strain BPK282A1)

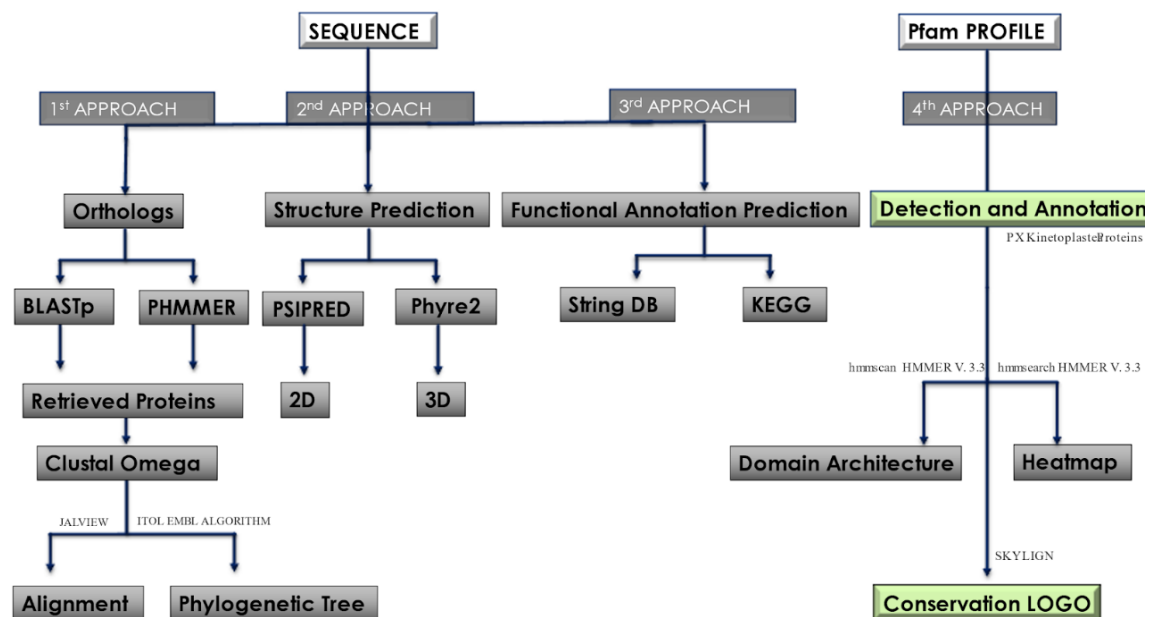
**Εικόνα 36.** Πρωτεΐνες των Κινητοπλαστιδών με δομή PX-PK\_Tyr\_Ser\_Thr-Pkinase-Lipocalin\_5 έπειτα από επεξεργασία στο πρόγραμμα EXCEL. Τα δεδομένα για τις στήλες «Description», «Organism», συλλέχθηκαν από τη βάση δεδομένων Uniprot.

PROTEIN	1 <sup>ST</sup> DOMAIN	2 <sup>ND</sup> DOMAIN	DESCRIPTION	ORGANISM
A0A2V2V2E1	PX PF00787.25 29 94	PX PF00787.25 291 352	Uncharacterized protein	Trypanosoma cruzi
V5BFS0	PX PF00787.25 29 94	PX PF00787.25 291 352	Uncharacterized protein	Trypanosoma cruzi Dm28c
A0A2V2VFT3	PX PF00787.25 31 94	PX PF00787.25 296 357	Uncharacterized protein	Trypanosoma cruzi
A0A2V2WEX7	PX PF00787.25 31 92	PX PF00787.25 297 358	Uncharacterized protein	Trypanosoma cruzi
Q4DW43	PX PF00787.25 31 92	PX PF00787.25 297 358	Uncharacterized protein	Trypanosoma cruzi (strain
A0A422NLT1	PX PF00787.25 305 366	PX PF00787.25 29 90	Uncharacterized protein	Trypanosoma rangeli
A0A3R7L2S8	PX PF00787.25 31 95	PX PF00787.25 300 362	Uncharacterized protein	Trypanosoma conorhini
A0A1X0NKZ7	PX PF00787.25 38 99	PX PF00787.25 308 368	Uncharacterized protein	Trypanosoma theileri

**Εικόνα 37.** Πρωτεΐνες των Κινητοπλαστιδών με δομή PX-PX έπειτα από επεξεργασία στο πρόγραμμα EXCEL. Τα δεδομένα για τις στήλες «Description», «Organism», συλλέχθηκαν από τη βάση δεδομένων Uniprot.

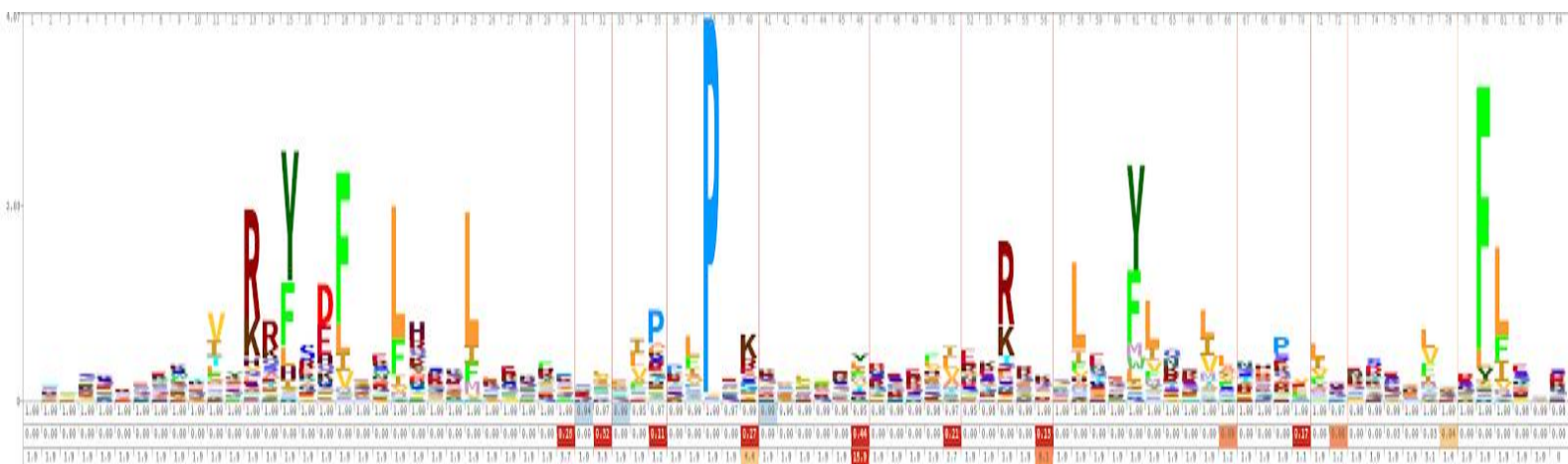
PROTEIN	1 <sup>ST</sup> DOMAIN	2 <sup>ND</sup> DOMAIN	3 <sup>RD</sup> DOMAIN	DESCRIPTION	ORGANISM
A0A2V2XM59	PX PF00787.25 41 129	Vps5 PF09325.11 175 402	BAR_3_WASP_bdg PF10456	Autophagy-related protei	Trypanosoma cruzi
A0A2V2WND1	PX PF00787.25 41 129	Vps5 PF09325.11 174 402	BAR_3_WASP_bdg PF10456	Autophagy-related protei	Trypanosoma cruzi
Q4DS34	PX PF00787.25 41 129	Vps5 PF09325.11 174 402	BAR_3_WASP_bdg PF10456	Phosphoinositide-binding	Trypanosoma cruzi (strain CL Brener)
A0A2V2W207	PX PF00787.25 41 129	Vps5 PF09325.11 175 402	BAR_3_WASP_bdg PF10456	Autophagy-related protei	Trypanosoma cruzi
Q4D6U6	PX PF00787.25 41 129	Vps5 PF09325.11 175 402	BAR_3_WASP_bdg PF10456	Phosphoinositide-binding	Trypanosoma cruzi (strain CL Brener)
V5B3T2	PX PF00787.25 41 129	Vps5 PF09325.11 175 402	BAR_3_WASP_bdg PF10456	Phosphoinositide-binding	Trypanosoma cruzi Dm28c
A0A061JAH9	PX PF00787.25 39 129	Vps5 PF09325.11 190 402	BAR_3_WASP_bdg PF10456	Phosphoinositide-binding	Trypanosoma rangeli SC58
W6L4Q5	PX PF00787.25 44 122	Vps5 PF09325.11 191 388	BAR_3_WASP_bdg PF10456	PX domain-containing prc	Phytomonas sp. isolate Hart1
A0A3R7LFH1	PX PF00787.25 39 129	Vps5 PF09325.11 193 402	BAR_3_WASP_bdg PF10456	Putative phosphoinositid	Trypanosoma conorhini
A0A3R7L2W7	PX PF00787.25 39 129	Vps5 PF09325.11 194 402	BAR_3_WASP_bdg PF10456	Putative phosphoinositid	Trypanosoma rangeli
F9W5Y5	PX PF00787.25 41 129	Vps5 PF09325.11 195 297	BAR_3_WASP_bdg PF10456	WGS project CAEQ000000	Trypanosoma congolense (strain IL300
A0A1G41C1	PX PF00787.25 41 129	Vps5 PF09325.11 195 402	BAR_3_WASP_bdg PF10456	Autophagy-related protei	Trypanosoma equiperdum
A0A3L6L560	PX PF00787.25 41 129	Vps5 PF09325.11 195 402	BAR_3_WASP_bdg PF10456	Autophagy-related protei	Trypanosoma brucei equiperdum
C9ZZB2	PX PF00787.25 41 129	Vps5 PF09325.11 195 402	BAR_3_WASP_bdg PF10456	Phosphoinositide-binding	Trypanosoma brucei gambiense (strain
Q38D86	PX PF00787.25 41 129	Vps5 PF09325.11 195 402	BAR_3_WASP_bdg PF10456	Phosphoinositide-binding	Trypanosoma brucei brucei (strain 927,
A0A1X0P5K9	PX PF00787.25 41 129	Vps5 PF09325.11 196 406	BAR_3_WASP_bdg PF10456	Putative phosphoinositid	Trypanosoma theileri
A0A3Q8IV97	PX PF00787.25 44 131	Vps5 PF09325.11 211 406		Autophagy-related protei	Leishmania donovani
A0A504X805	PX PF00787.25 44 131	Vps5 PF09325.11 211 406		Autophagy-related prote	Leishmania donovani
A4IBF2	PX PF00787.25 44 131	Vps5 PF09325.11 211 406		Autophagy-related prote	Leishmania infantum
E9AF66	PX PF00787.25 44 131	Vps5 PF09325.11 211 406		Putative phosphoinositid	Leishmania major
E9BSB7	PX PF00787.25 44 131	Vps5 PF09325.11 211 406		Phosphoinositide-binding	Leishmania donovani (strain BPK282A1
A0A640KYX8	PX PF00787.25 44 131	Vps5 PF09325.11 211 407		Phosphoinositide-binding	Leishmania tarentolae (Sauroleishman
E9B6C2	PX PF00787.25 44 131	Vps5 PF09325.11 211 407		Putative phosphoinositid	Leishmania mexicana (strain MHOM/G
A0A088S0S0	PX PF00787.25 44 131	Vps5 PF09325.11 213 407		Phosphoinositide-binding	Leishmania panamensis
A0A3P3ZGX6	PX PF00787.25 44 131	Vps5 PF09325.11 213 407		Autophagy-related prote	Leishmania braziliensis MHOM/BR/75/
A4HMR6	PX PF00787.25 44 131	Vps5 PF09325.11 213 407		Putative phosphoinositid	Leishmania braziliensis
A0A0N1IB55	PX PF00787.25 44 132	Vps5 PF09325.11 219 419		Putative phosphoinositid	Leptomonas seymouri
A0A0N0E0A2	PX PF00787.25 44 132	Vps5 PF09325.11 221 419		Putative phosphoinositid	Leptomonas pyrrocoris
K2NRQ0	PX PF00787.25 39 129	Vps5 PF09325.11 171 402		Phosphoinositide-binding	Trypanosoma cruzi marinellei
A0A054KLA7	PX PF00787.25 42 127	Vps5 PF09325.11 191 402		Phosphoinositide-binding	Bodo saltans (Flagellated protozoan)

**Εικόνα 38** Πρωτεΐνες των Κινητοπλαστιδών με δομή PX-Vps5, PX-Vps5-BAR\_3\_WASP) έπειτα από επεξεργασία στο πρόγραμμα EXCEL. Τα δεδομένα για τις στήλες «Description, «Organism», συλλέχθηκαν από τη βάση δεδομένων Uniprot. Σε αυτή την κατηγορία, συναντάται και η πρωταΐνη E9BSB7 που μελετούμε.

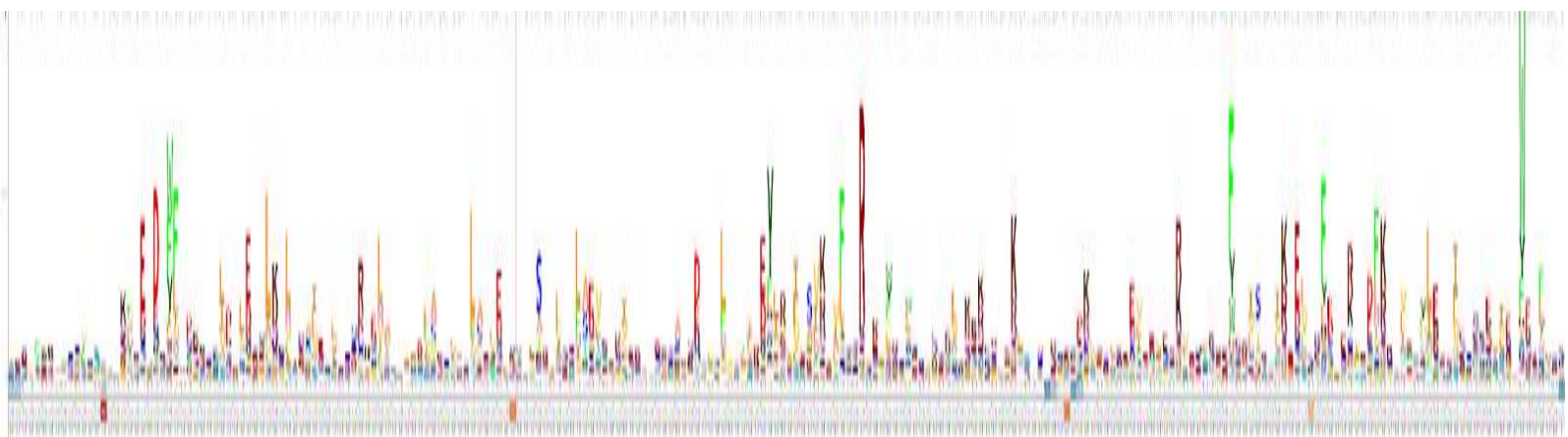


**Εικόνα 39.** Σχηματική αναπαράσταση του βήματος της ανάλυσης που σχετίζεται με την κατασκευή LOGO συντηρημένων αμινοξέων.

Σε αυτό το βήμα, έγινε κατασκευή LOGO συντηρημένων αμινοξέων για τις περιοχές PX, VPS5. Χρησιμοποιήθηκαν τα PFAM Profiles των περιοχών αυτών, τα οποία αποτέλεσαν είσοδο στον αλγόριθμο SKYLIGN που δημιούργησε τα LOGO.



**Εικόνα 40.** LOGO συντηρημένων αμινοξέων του HMM προφίλ της περιοχής PX



**Εικόνα 41.** LOGO συντηρημένων αμινοξέων του HMM προφίλ της περιοχής Vps5

Παρατηρείται πως και οι δύο περιοχές εμφανίζουν αρκετά υψηλή συντήρηση αμινοξέων. Τα γράμματα αναπαριστούν τα 20 αμινοξέα. Όσο μεγαλύτερο μέγεθος εμφανίζουν, τόσο μεγαλύτερη η συντήρηση του εκάστοτε αμινοξέος στη συγκεκριμένη θέση.

## ΣΥΖΗΤΗΣΗ

Η νόσος της Λεισμανίασης προκαλείται από τα πρωτόζωα του γένους *Leishmania*. Εμφανίζει ετερογένεια στα συμπτώματα που προκαλεί και σε πολλές περιπτώσεις μπορεί να καταστεί θανατηφόρος. Παρά τις πολύχρονες μέχρι τώρα ερευνητικές μελέτες, δεν έχουν προσδιοριστεί με ακρίβεια οι μηχανισμοί που συντελούν σε αυτό, με αποτέλεσμα η θεραπευτική προσέγγιση που χρησιμοποιείται να περιλαμβάνει μόνο χημειοθεραπείες, που λόγω της φύσης τους πρέπει να πραγματοποιούνται εντός του νοσοκομείου ενώ έχουν σημαντικές παρενέργειες.

Η παρούσα πτυχιακή διατριβή είχε ως στόχο την Βιοπληροφορική ανάλυση μιας εκκρινόμενης πρωτεΐνης από το παράσιτο *Leishmania donovani* που εμφανίζεται με κωδικό E9BSB7 στην βάση δεδομένων UniProt. Αυτή, συμπεριέλαβε και πρόβλεψη της δομής της καθώς και υποθετικής λειτουργικής δράσης με χρήση διαφορετικών εργαλείων Βιοπληροφορικής.

Στη συνέχεια, για λόγους συγκριτικής ανάλυσης, πραγματοποιήθηκε επέκταση της μελέτης σε πρωτεΐνες όλων των Κινητοπλαστιδίων που διαθέτουν τη λειτουργική περιοχή PX, έτσι ώστε να εντοπιστεί το σύνολο των πρωτεϊνών με αυτή τη δομική περιοχή και να διερευνηθούν και τα διαφορετικά δομικά στοιχεία που φέρουν αυτές οι πρωτεΐνες ώστε να κατηγοριοποιηθούν ανάλογα με αυτές τις προβλέψεις δομών σε ομάδες με συγγενή λειτουργική δράση.

Για την επίτευξη αυτών των στόχων, η μελέτη ξεκίνησε από την εύρεση των ορθόλογων πρωτεϊνών του γονιδιακού προϊόντος που μας ενδιαφέρει, με χρήση του αλγορίθμου BLASTp, Έγινε επίσης κατασκευή φυλογενετικών δέντρων με χρήση του Clustal Omega σε συνδυασμό με τον αλγόριθμο ITOL EMBL και το πρόγραμμα Jalview. Η προσέγγιση αυτή οδήγησε στην ταυτοποίηση όλων οι πρωτεϊνικών ακολουθιών που εμφανίζουν υψηλά ποσοστά ομοιότητας σε επίπεδο πρωτοταγούς δομής με την E9BSB7 και στην ταξινόμηση τους σε φυλογενετικά δέντρα, ώστε να είναι διακριτή η εξελικτική τους σχέση. Όταν τα ποσοστά ομοιότητας των πρωτεϊνικών ακολουθιών είναι υψηλά, πολλά από τα τμήματα τους είναι πολύ καλά συντηρημένα, οπότε πιθανότατα είναι σημαντικά ή και απαραίτητα για την επιβίωση των οργανισμών αυτών κατά τη διάρκεια της εξέλιξης των ειδών. Η αύξηση της ομοιότητας σε επίπεδο ακολουθίας αυξάνει την πιθανότητα να έχουν συνάφεια και ως προς τη λειτουργία τους οι πρωτεΐνες αυτές. Οι στοιχίσεις αυτές μπορούν να δώσουν σημαντικά συμπεράσματα, ιδιαίτερα αν κάποιες από τις ορθόλογες πρωτεΐνες σε άλλους οργανισμούς έχουν ήδη μελετηθεί.

Στα αποτελέσματα του αλγορίθμου BLASTp, εμφανίστηκαν να έχουν στατιστικά σημαντική ομοιότητα, κυρίως πρωτεΐνες που είναι χαρακτηρισμένες ως «Phosphoinositide-binding proteins», «PX domain-containing proteins», «Autophagy-related proteins», πρωτεΐνες που είναι μη χαρακτηρισμένες και πρωτεΐνες της οικογένειας των «Sorting nexins» σε Κινητοπλαστίδες, μύκητες, παράσιτα ζώων και φυτών αλλά και ανώτερους ευκαρυωτικούς οργανισμούς. Αξιοσημείωτη ομοιότητα της πρωταγούς δομής της E9BSB7, που χαρακτηρίζεται από τιμή e-value 0.0, παρατηρείται στις ορθόλογες πρωτεΐνες άλλων ειδών *Leishmania* A4IBF2 (*L. infantum*), E9AF66 (*L. major*), A4HMR6 (*L. braziliensis*), καθώς και στην πρωτεΐνη A0A0N0E0A2 του παρασίτου *Leptomonas pyrrhocoris* που ανήκει στις Τρυπανοσωματίδες.

Για την πρόβλεψη της δευτεροταγούς δομής στην μελέτη μας, χρησιμοποιήθηκε ο αλγόριθμος PSIPRED. Με τη βοήθεια της απεικόνισης αυτής, προσδιορίστηκε η ακριβής θέση για τις  $\alpha$ -έλικες, τις  $\beta$ -πτυχώσεις και τις στροφές (turns) της πρωτεΐνης, ευρήματα που συνετέλεσαν στην πρόβλεψη της τρισδιάστατης δομής της. Επίσης, διαπιστώθηκε μεγάλη ομοιότητα με την δευτεροταγή δομή λυμένων πρωτεϊνών, όπως η SNX33. Η πρόβλεψη της τριτοταγούς δομής της πρωτεΐνης, υλοποιήθηκε με χρήση του αλγορίθμου Phyre2 και συνετέλεσε στην εξαγωγή συμπερασμάτων με σημαντικότερο την εξαιρετική ομοιότητα της στερεοδιάταξης της πρωτεΐνης E9BSB7 του παρασίτου *L. donovani* με εκείνη της Νεξίνης Διαλογής 2 (SNX2) του ανθρώπου. Η συγκεκριμένη υψηλή ομοιότητα στις δομές αυτών των πρωτεϊνών, αποτελεί σημαντικό δείκτη πρόβλεψης ομοιότητας στη λειτουργική τους δράση.

Ενίσχυση της πρόβλεψης της λειτουργικής δράσης μιας πρωτεΐνης αποτελεί και η ανεύρεση αλληλεπιδράσεων με άλλα βιομόρια με γνωστή λειτουργία. Έτσι στη μελέτη μας προχωρήσαμε σε σχετική πρόβλεψη πρωτεϊνικών αλληλεπιδράσεων για την πρωτεΐνη E9BSB7.

Αρχικά χρησιμοποιήθηκε η String DB. Έγινε αναζήτηση της ύπαρξης δικτύου πρωτεϊνικών αλληλεπιδράσεων, με βάση την ακολουθία της E9BSB7, ωστόσο δεν υπήρχε καμία καταχώρηση. Στη συνέχεια ερευνήθηκαν ορθόλογες πρωτεΐνες που είχαν εντοπιστεί με τη χρήση του BLASTp, έτσι ώστε να επιλεγεί κάποια ορθόλογη, εξαιρετικά όμοια πρωτεΐνη που να υπάρχει ως καταχώρηση στη βάση δεδομένων String. Επιλέχθηκε η πρωτεΐνη με NCBI ID XP\_001469071.1, η οποία ανήκει στο πρωτόζωο *Leishmania infantum* (οργανισμός που ανήκει στο σύμπλεγμα *Leishmania donovani*). Η XP\_001469071.1 εμφανίζει ίδια δομική οργάνωση (domain architecture) και ομοιότητα σε επίπεδο αμινοξικής ακολουθίας 99,8% με την E9BSB7. Μελέτη του δικτύου πρωτεϊνικών αλληλεπιδράσεων της XP\_001469071.1, οδήγησε στην εξαγωγή εξαιρετικά σημαντικών συμπερασμάτων. Η βάση δεδομένων String, παρέχει τη δυνατότητα εύρεσης των ανθρώπινων πρωτεϊνών που σύμφωνα με τις πρωτεϊνικές αλληλεπιδράσεις τους είναι ομόλογες με την XP\_001469071.1 και κατ' επέκταση με την E9BSB7. Η στατιστικά σημαντικότερη λειτουργικά ομόλογη πρωτεΐνη που εντοπίστηκε ήταν η ανθρώπινη Νεξίνη Διαλογής 2 (SNX2). Βρέθηκε ότι οι δύο αυτές πρωτεΐνες, εμφανίζουν παρόμοιες πρωτεϊνικές αλληλεπιδράσεις και πιθανώς εμπλέκονται στα ίδια μονοπάτια που σχετίζονται με την σηματοδότηση στην ενδοκύττωση.

Επαλήθευση του αποτελέσματος αυτού, έγινε μέσω της βάσης δεδομένων KEGG. Σε αυτή την περίπτωση, υπήρχε καταχώρηση για την υπό μελέτη πρωτεΐνη E9BSB7. Πραγματοποιήθηκε, η ίδια αναζήτηση λειτουργικά ομόλογων ανθρώπινων πρωτεϊνών και το στατιστικά επικρατέστερο αποτέλεσμα και σε αυτή την περίπτωση ήταν η ανθρώπινη SNX2. **Τα ευρήματα αυτά, σε συνδυασμό με την εξαιρετικά μεγάλη δομική ομοιότητα σε επίπεδο τριτοταγούς δομής, οδήγησαν στην υπόδειξη της ανθρώπινης SNX2, ως λειτουργικά και δομικά ομόλογη με την πρωτεΐνη που εκφράζεται από το γονίδιο LdBK\_352470 και ανήκει στο παράσιτο *Leishmania donovani*.** Αυτή η πληροφορία, είναι εξαιρετικά σημαντική για τον σχεδιασμό πειραμάτων για τον προσδιορισμό της λειτουργίας της πρωτεΐνης στο εργαστήριο.

Μετά την πρόβλεψη της δομής και λειτουργίας της πρωτεΐνης E9BSB7 από το παράσιτο *Leishmania donovani*, επεκτείναμε την ανάλυση στον εντοπισμό και τη δομική κατάταξη όλων των πρωτεϊνών από τις Κινετοπλαστίδες, που περιέχουν

λειτουργική περιοχή PX, μεγάλο μέρος των οποίων πιθανώς να ανήκουν στην οικογένεια των Νεξινών Διαλογής.

Για τον σκοπό αυτό, έγινε χρήση του περιβάλλοντος LINUX και της γλώσσας προγραμματισμού bash script. Με τη βοήθεια του πακέτου HMMER, και της εντολής hmmssearch εντοπίστηκαν όλες οι πρωτεΐνες που έχουν λειτουργική περιοχή PX και ανήκουν στις Κινητοπλαστίδες (συνολικά 137 πρωτεΐνες). Στη συνέχεια, με τη βοήθεια της εντολής hmmscan έγινε προσδιορισμός και οργάνωση των δομικών περιοχών (domain architecture) που συναντώνται στις πρωτεΐνες αυτές. Για την απεικόνιση των αποτελεσμάτων, κατασκευάστηκαν δύο Heatmaps, δηλαδή απεικονίσεις στις οποίες ταξινομήθηκαν οι πρωτεΐνες σε κατηγορίες ανάλογα με τον χαρακτηρισμό τους από την UniProt, ανά πρωτόμα στην πρώτη περίπτωση, ενώ στη δεύτερη ανάλογα με τη δομική τους οργάνωση (domain architecture), ανά πρωτόμα. Η σύγκριση των δύο Heatmaps, οδήγησε στο συμπέρασμα ότι η κατηγοριοποίηση με βάση τη δομική οργάνωση ήταν αποδοτικότερη, διότι στην πρώτη περίπτωση υπήρχε επικάλυψη των κατηγοριών.

Τέλος, δημιουργήθηκαν με βάση τα HMM profiles τους, τα LOGO που εκφράζουν τη συντήρηση των αμινοξέων που εμφανίζεται στις λειτουργικές περιοχές PX και BAR. Αυτό, θα μπορούσε να συμβάλει στον σχεδιασμό εργαστηριακών πειραμάτων που αφορούν τη δημιουργία μεταλλαγμάτων της πρωτεΐνης για να μελετηθεί ο ρόλος της δράσης της κάθε περιοχής στην αλληλεπίδραση με μοριακούς partners αλλά και στους μηχανισμούς δράσης.

Η παρούσα εργασία, παρέχει σημαντική πληροφορία για τον πλήρη χαρακτηρισμό της δομής και της λειτουργικής δράσης της πρωτεΐνης E9BSB7 του παρασίτου *Leishmania donovani*. Τα πειράματα που θα βασιστούν στα αποτελέσματα που παρουσιάστηκαν, θα υποδείξουν αν τελικά η πρωτεΐνη E9BSB7 θα μπορούσε να αποτελέσει σημείο αναφοράς στη δημιουργία μιας αναθεωρημένης θεραπευτικής προσέγγισης ενάντια στη νόσο της Λείσμανιάσης.

## ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Koyrounis D. MSc 2017 «Cloning and Molecular Characterization of the Secreted Nexin LdPIBPnex with Phosphoinositide Binding Domain from the Protist *Leishmania Donovanii*».
2. Bates PA. Leishmania sand fly interaction: progress and challenges. *Curr Opin Microbiol.* 2008;11(4):340-344. doi:10.1016/j.mib.2008.06.003
3. Sunter J, Gull K. Shape, form, function and *Leishmania* pathogenicity: from textbook descriptions to biological understanding. *Open Biol.* 2017;7(9):170165. doi:10.1098/rsob.170165
4. Teixeira DE, Benchimol M, Rodrigues JCF, Crepaldi PH, Pimenta PFP, de Souza W. The Cell Biology of *Leishmania*: How to Teach Using Animations. Knoll LJ, ed. *PLoS Pathog.* 2013;9(10):e1003594. doi:10.1371/journal.ppat.1003594
5. Maxwell MJ, Chan SK, Robinson DP, et al. Proteomic analysis of the secretome of *Leishmania donovani*. *Genome Biol.* 2008;9(2):R35. doi:10.1186/gb-2008-9-2-r35
6. Cullen J. An approximate diatomics in molecules formulation of generalized valence bond theory. *J Comput Chem.* 2008;29(4):497-504. doi:10.1002/jcc.20808
7. van Weering JRT, Verkade P, Cullen PJ. SNX-BAR proteins in phosphoinositide-mediated, tubular-based endosomal sorting. *Semin Cell Dev Biol.* 2010;21(4):371-380. doi:10.1016/j.semcdb.2009.11.009
8. Carlton J, Bujny M, Rutherford A, Cullen P. Sorting nexins - Unifying trends and new perspectives. *Traffic.* 2005;6(2):75-82. doi:10.1111/j.1600-0854.2005.00260.x
9. Herwaldt BL. Leishmaniasis. *Lancet.* 1999;354(9185):1191-1199. doi:10.1016/S0140-6736(98)10178-2
10. Ashford RW. The leishmaniasis as emerging and reemerging zoonoses. *Int J Parasitol.* 2000;30(12-13):1269-1281. doi:10.1016/S0020-7519(00)00136-3
11. Akhoundi M, Downing T, Votýpka J, et al. *Leishmania* infections: Molecular targets and diagnosis. *Mol Aspects Med.* 2017;57:1-29. doi:10.1016/j.mam.2016.11.012
12. Ntais P, Sifaki-Pistola D, Christodoulou V, et al. Leishmaniasis in Greece. *Am J Trop Med Hyg.* 2013;89(5):906-915. doi:10.4269/ajtmh.13-0070
13. Chappuis F, Sundar S, Hailu A, et al. Visceral leishmaniasis: What are the needs for diagnosis, treatment and control? *Nat Rev Microbiol.* 2007;5(11):873-882. doi:10.1038/nrmicro1748
14. Pearson RD, Hargus JL, Roberts D, Donowitz GR. Differential survival of *Leishmania donovani* amastigotes in human monocytes. *J Immunol.* 1983;131(4).

15. Di Paolo G, De Camilli P. Phosphoinositides in cell regulation and membrane dynamics. *Nature*. 2006;443(7112):651-657. doi:10.1038/nature05185
16. De Matteis MA, Godi A. PI-loting membrane traffic. *Nat Cell Biol*. 2004;6(6):487-492. doi:10.1038/ncb0604-487
17. Kutateladze TG. Translation of the phosphoinositide code by PI effectors. *Nat Chem Biol*. 2010;6(7):507-513. doi:10.1038/nchembio.390
18. Pourcher M, Santambrogio M, Thazar N, et al. Analyses of SORTING NEXINs Reveal Distinct Retromer-Subcomplex Functions in Development and Protein Sorting in Arabidopsis thaliana. *Plant Cell*. 2010;22(12):3980-3991. doi:10.1105/tpc.110.078451
19. BLAST QuickStart - Comparative Genomics - NCBI Bookshelf. (n.d.). Retrieved February 9, 2021 from <https://www.ncbi.nlm.nih.gov/books/NBK1734>. BLAST QuickStart - Comparative Genomics - NCBI Bookshelf. Accessed February 9, 2021. <https://www.ncbi.nlm.nih.gov/books/NBK1734/>
20. HMMER. Accessed February 9, 2021. <http://hmmer.org/documentation.html>
21. Clustal Omega Help and Documentation - Job Dispatcher Sequence Analysis Tools - EMBL-EBI. Accessed February 9, 2021. <https://www.ebi.ac.uk/seqdb/confluence/display/JDSAT/Clustal+Omega+Help+and+Documentation>
22. Letunic I, Bork P. Interactive Tree Of Life (iTOL): An online tool for phylogenetic tree display and annotation. *Bioinformatics*. 2007;23(1):127-128. doi:10.1093/bioinformatics/btl529
23. Procter JB, Carstairs GM, Soares B, et al. Alignment of Biological Sequences with Jalview. In: *Methods in Molecular Biology*. Vol 2231. Humana Press Inc.; 2021:203-224. doi:10.1007/978-1-0716-1036-7\_13
24. McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. *Bioinformatics*. 2000;16(4):404-405. doi:10.1093/bioinformatics/16.4.404
25. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*. 2015;10(6):845-858. doi:10.1038/nprot.2015.053
26. Szklarczyk D, Franceschini A, Kuhn M, et al. The STRING database in 2011: Functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res*. 2011;39(SUPPL. 1). doi:10.1093/nar/gkq973
27. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res*. 2000;28(1):27-30. doi:10.1093/nar/28.1.27
28. Wheeler TJ, Clements J, Finn RD. Skylign: A tool for creating informative, interactive logos representing sequence alignments and profile hidden Markov models. *BMC Bioinformatics*. 2014;15(1). doi:10.1186/1471-2105-15-7
29. Babicki S, Arndt D, Marcu A, et al. Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res*. 2016;44(W1):W147-W153. doi:10.1093/nar/gkw419
30. Bateman A, Martin MJ, O'Donovan C, et al. UniProt: The universal protein knowledgebase. *Nucleic Acids Res*. 2017;45(D1):D158-D169. doi:10.1093/nar/gkw1099



31. Waterhouse A, Bertoni M, Bienert S, et al. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* 2018;46(W1):W296-W303. doi:10.1093/nar/gky427

