



Σχολή Επιστημών Υγείας και Πρόνοιας
Τμήμα Βιοϊατρικών Επιστημών



ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Ανάλυση ρυθμιστικών μορίων miRNAs που εμπεριέχονται στα
εξωσώματα υγιών και μολυσμένων με λείσμάνια σκύλων



ΟΝΟΜΑΤΕΠΩΝΥΜΟ ΦΟΙΤΗΤΗ:

Κωνσταντίνος Ντούνης

ΑΜ: 20678228

ΕΠΙΒΛΕΠΩΝ:

Δρ. Αντωνία Ευσταθίου

ΑΘΗΝΑ 2025



Faculty of Health and Care Sciences
Department of Biomedical Sciences



UNDERGRADUATE THESIS

Analysis of regulatory miRNA molecules contained in the exosomes
of healthy and Leishmania infected dogs



Student

Konstantinos Dounis

AM: 20678228

Supervisor

Dr Antonia Efstathiou

ATHENS 2025



Σχολή Επιστημών Υγείας και Πρόνοιας
Τμήμα Βιοϊατρικών Επιστημών



ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Ανάλυση ρυθμιστικών μορίων miRNAs που εμπεριέχονται στα
εξωσώματα υγιών και μολυσμένων με λεισμάνια σκύλων

ΕΠΙΤΡΟΠΗ ΕΞΕΤΑΣΗΣ ΣΥΜΠΕΡΙΛΑΜΒΑΝΟΜΕΝΟΥ ΚΑΙ ΤΟΥ ΕΙΣΗΓΗΤΗ

	Ονοματεπώνυμο	Υπογραφή
1 ^{ος} ΕΞΕΤΑΣΤΗΣ	Ευσταθίου Αντωνία	
2 ^{ος} ΕΞΕΤΑΣΤΗΣ	Βογιατζάκη Χρυσάνθη	
3 ^{ος} ΕΞΕΤΑΣΤΗΣ	Μπελούκας Απόστολος	

*Συνεπιβλέποντες: Ευδοκία Καραγκούνη, PhD
Διευθύντρια Ερευνών, Αντιπρόεδρος ΔΣ,
Ελληνικό Ινστιτούτο Παστέρ

Τιμοκράτης Καραμήτρος, PhD
Κύριος Ερευνητής
Ελληνικό Ινστιτούτο Παστέρ

Δήλωση συγγραφέα προπτυχιακής διπλωματικής εργασίας

Ο κάτωθι υπογεγραμμένος Ντούνης Κωνσταντίνος του Παναγιώτη, με αριθμό μητρώου 20678228 φοιτητής του Τμήματος Βιοϊατρικών Επιστημών της Σχολής Επιστημών Υγείας και Πρόνοιας του Πανεπιστημίου Δυτικής Αττικής, δηλώνω ότι: «Είμαι συγγραφέας αυτής της πτυχιακής/διπλωματικής εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, οι όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε ακριβώς είτε παραφρασμένες, αναφέρονται στο σύνολό τους, με πλήρη αναφορά στους συγγραφείς, τον εκδοτικό οίκο ή το περιοδικό, συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο. Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από μένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου, όσο και του Ιδρύματος. Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του πτυχίου μου».

Επιθυμώ την απαγόρευση πρόσβασης στο πλήρες κείμενο της πτυχιακής μου μέχρι 6 μήνες και έπειτα από αίτησή μου στη βιβλιοθήκη και έγκριση του επιβλέποντα καθηγητή.

Υπογραφή Φοιτητή



Κωνσταντίνος Ντούνης

Υπογραφή Επιβλέποντα

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΕΥΧΑΡΙΣΤΙΕΣ	8
ΠΕΡΙΛΗΨΗ	9
ABSTRACT	10
1. ΛΕΙΣΜΑΝΙΑ	11
1.1 Παλιός και Νέος κόσμος	11
1.2 Φλεβοτόμος	12
1.3 Κλινικές περιπτώσεις	12
1.3.1 Δερματική λείσμανίαση (CL)	12
1.3.2 Βλεννογονοδερματική λείσμανίαση (ML)	13
1.3.3 Σπλαχνική λείσμανίαση (VL)	13
1.3.4 Λείσμανίαση στους σκύλους (Canine Leishmaniasis) (CanL)	14
1.4 Μορφές και κύκλος ζωής της λείσμανίας	15
1.5 Γονιδιακή οργάνωση της λείσμανίας	16
1.6 Παθογένεια της λείσμανίασης στους σκύλους	17
1.7 Ανοσολογική απόκριση στην λείσμανίαση	18
1.8 Θεραπεία της λείσμανίασης	18
1.8.1 Συμβατική χημειοθεραπεία	18
1.8.2 Συνδυαστική χημειοθεραπεία	19
1.8.3 Τοπικές θεραπείες της CL	19
1.8.4 Θεραπεία της λείσμανίασης στους σκύλους (CanL)	20
1.8.5 Εμβόλιο κατά της λείσμανίασης στους ανθρώπους	20
1.8.6 Εμβόλιο κατά της λείσμανίασης στους σκύλους	20
2. MIRNA	21
2.1 Κανονική οδός βιογένεσης των miRNAs	21
2.2 RNA induced silencing complex (RISC)	23
2.3 miRNA ονοματολογία	25
2.4 Μη κανονικές οδοί βιογένεσης miRNA	25
2.5 miRNAs και εξωσώματα	26
2.6 miRNAs και Παράσιτα	28

2.7 miRNA και λεισμάνια.....	28
2.8 Μέθοδοι ανίχνευσης και ποσοτικοποίησης των miRNA	28
2.8.1 Northern Blot	29
2.8.2 Real-time qPCR.....	29
2.8.3 Τεχνολογία μικροσυστοιχιών (Microarray)	30
2.8.4 Αλληλούχιση Επόμενης Γενιάς (NGS)	30
3. ΠΕΙΡΑΜΑΤΙΚΟ ΜΕΡΟΣ	32
3.1 Σκοπός.....	32
3.2 Μεθοδολογία ανάλυσης ακατέργαστων δεδομένων αλληλούχισης μορίων miRNA	32
3.2.1 Ανάλυση δειγμάτων και εξαγωγή δεδομένων - Ion GeneStudio S5.....	33
3.2.2 Αρχεία SAM/BAM.....	33
3.2.3 Αρχεία FASTA/FASTQ	34
3.2.4 Ποιοτικός έλεγχος των δεδομένων αλληλούχισης, FastQC	35
3.2.4.1 Μετατροπή αρχείου BAM σε FASTQ.....	35
3.2.4.2 Εισαγωγή αρχείου στο FastQC.....	36
3.2.4.3 FastQC περιεχόμενα αποτελεσμάτων	36
3.2.5 Αντάπτορες και Cutadapt.....	37
3.2.6 Ταυτοποίηση miRNA μέσω του MiRDeep2.....	40
3.2.7 Μεθοδολογία ανάλυσης των δεδομένων από το MiRDeep2 με την χρήση της R.....	46
3.2.7.1 Εγκατάσταση πακέτων και φόρτωση αρχείων στην R.....	46
3.2.7.2 Φιλτράρισμα και αφαίρεση του δείγματος B6	48
3.2.7.3 Συγχώνευση δεδομένων και εξαγωγή σε αρχείο Excel.....	50
3.2.7.4 Ανάλυση διαφορικής έκφρασης μέσω της R	54
3.2.7.5 Απεικόνιση των αποτελεσμάτων με την χρήση ggplot2	58
4 Αποτελέσματα ανάλυσης ακατέργαστων δεδομένων αλληλούχισης μορίων miRNA από εξωσώματα ορού υγιών και μολυσμένων με λεισμανίαση σκύλων	62
4.1 Δείγματα προς ανάλυση	62
4.2 Αποτελέσματα ποιοτικού ελέγχου FastQC.....	62
4.2.1 Βασικές στατιστικές πληροφορίες:	63
4.2.2 Διάγραμμα ποιότητας ανά βάση	65
4.2.3 Βαθμολογία ποιότητας ανά αλληλουχία	66
4.2.4 Περιεχόμενο αλληλουχίας ανά βάση	67
4.2.5 Περιεχόμενο GC ανά βάση.....	68
4.2.6 Περιεχόμενο N ανά θέση	72
4.2.7 Κατανομή μήκους ακολουθίας.....	73
4.2.8 Επίπεδο διπλασιασμού αλληλουχίας	75
4.2.9 Υπερεκπροσωπούμενες Αλληλουχίες.....	76
4.2.10 Περιεχόμενο σε Αντάπτορες	79
4.3 Σύγκριση FastQC πριν και μετά τη χρήση του Cutadapt στα αρχεία fastq.....	79
4.4 Αφαίρεση δείγματος B6.....	82
4.5 Αποτελέσματα ανάλυσης διαφορικής έκφρασης miRNA στα δείγματα.....	83

5. ΣΥΜΠΕΡΑΣΜΑΤΑ	86
ΒΙΒΛΙΟΓΡΑΦΙΑ	90

ΕΥΧΑΡΙΣΤΙΕΣ

Η εκπόνηση της πτυχιακής μου εργασίας ολοκληρώθηκε με την πολύτιμη βοήθεια και καθοδήγηση της Δρ. Αντωνίας Ευσταθίου, από το Εργαστήριο Μοριακής Μικροβιολογίας και Ανοσολογίας του Πανεπιστημίου Δυτικής Αττικής και του Ελληνικού Ινστιτούτου Παστέρ, της Δρ. Ευδοκίας Καραγκούνη, από το Εργαστήριο Ανοσολογίας Λοίμωξης του Ελληνικού Ινστιτούτου Παστέρ, καθώς και του Δρ. Τιμοκράτη Καραμήτρου και της κ. Μαρίας Μπούσαλη, υποψήφιας διδάκτορος, από τη Μονάδα Βιοπληροφορικής & Εφαρμοσμένης Γενωμικής του ίδιου Ινστιτούτου. Τους ευχαριστώ θερμά για την στήριξη, τις γνώσεις και την καθοδήγησή τους.

Θα ήθελα να ευχαριστήσω ιδιαίτερα τη Δρ. Δήμητρα Τουμπανάκη και τον κ. Οδυσσέα Τζωρτζάτο, υποψήφιο διδάκτορα στο Εργαστήριο Ανοσολογίας Λοίμωξης του Ελληνικού Ινστιτούτου Παστέρ, για τις πολύτιμες συμβουλές και τη διαρκή τους υποστήριξη.

Ένα μεγάλο ευχαριστώ στους γονείς και τα αδέρφια μου για την στήριξη και ενθάρρυνση, καθώς και στους φίλους μου, που με την κατανόηση και την υπομονή τους στάθηκαν δίπλα μου σε κάθε βήμα.

Τέλος, ένα ξεχωριστό ευχαριστώ στην Πέννυ, για τη σταθερή της παρουσία και την αμέριστη υποστήριξή της σε όλη τη διάρκεια των σπουδών μου και της εκπόνησης αυτής της εργασίας.

ΠΕΡΙΛΗΨΗ

Η πτυχιακή αυτή εργασία επικεντρώνεται στη μελέτη των miRNA ως βασικών ρυθμιστικών μορίων στις κυτταρικές λειτουργίες και την εμπλοκή τους στη λοίμωξη από το παράσιτο λεισμάνια. Οι διάφορες μορφές της λεισμανίασης επηρεάζουν ανθρώπους και ζώα, με την Canine Leishmaniasis (CanL) να αποτελεί σοβαρό κτηνιατρικό και ανθρώπινο ζήτημα στις χώρες της λεκάνης της Μεσογείου. Οι περίπλοκοι μοριακοί μηχανισμοί παθογένεσης και οι δυσκολίες στη θεραπεία, όπως για παράδειγμα η ανάπτυξη ανθεκτικών στελεχών και η έλλειψη αποτελεσματικών εμβολίων, υπογραμμίζουν την ανάγκη εκτεταμένης έρευνας για την αντιμετώπισή της.

Τα miRNA, μέσω της αλληλεπίδρασής τους με mRNA, ρυθμίζουν τη γονιδιακή έκφραση και εμπλέκονται στην ανοσολογική απόκριση, ενώ τα εξωσώματα αποτελούν φορείς που μεταφέρουν miRNA και άλλα βιοδραστικά μόρια, τα οποία εμπλέκονται σε διάφορες βιολογικές διαδικασίες, όπως η ανοσολογική απόκριση και η πορεία της λοίμωξης.

Η ανάλυση της διαφορικής έκφρασης των μορίων miRNA στην εργασία αυτή, βασίστηκε σε τεχνολογίες αλληλούχισης επόμενης γενιάς (NGS) για την ταυτοποίηση των miRNA στα εξωσώματα ορού υγιών και μολυσμένων σκύλων, με τη χρήση εργαλείων όπως το MiRDeep2 και τη στατιστική ανάλυση μέσω της γλώσσας R. Η μελέτη αποκάλυψε διαφορές στην έκφραση τους, που πιθανώς οφείλονται στη μόλυνση και ανέδειξε τη σημασία των miRNA ως βιοδεικτών για τη διάγνωση και την κατανόηση της παθογένειας της λεισμανίασης.

Μεταξύ των πιο ανωρρυθμισμένων miRNAs σε μολυσμένα σκυλιά ήταν τα cfa-miR-222, cfa-miR-92b, cfa-miR-29b, cfa-miR-129 και cfa-miR-486, τα οποία σχετίζονται με τη σηματοδότηση του ανοσοποιητικού και τις φλεγμονώδεις αποκρίσεις. Αντίθετα, βασικά κατωρρυθμισμένα miRNAs περιλάμβαναν τα cfa-miR-30b, cfa-miR-30c, cfa-miR-107, cfa-miR-374b και cfa-miR-301a, πολλά από τα οποία εμπλέκονται στην ενεργοποίηση των μακροφάγων και στις Th1 ανοσολογικές αποκρίσεις. Η εργασία ενισχύει την κατανόηση της μοριακής αλληλεπίδρασης παρασίτου-ξενιστή και προσφέρει στοιχεία που θα μπορούσαν να συμβάλουν στην ανάπτυξη νέων θεραπευτικών προσεγγίσεων ή ακόμα και στη βελτίωση των εμβολίων για την πρόληψη της λεισμανίασης, ως μέρος μιας συνολικής προσπάθειας για την αντιμετώπιση μιας παραμελημένης τροπικής ασθένειας.

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Λεισμανίαση, *Leishmania infantum*, εξωσώματα, miRNA, ανάλυση διαφορικής έκφρασης, DESeq2, αλληλούχιση νέας γενιάς, NGS, R, MiRdeep2.

ABSTRACT

Leishmaniasis, caused by *Leishmania* parasites, is a significant anthroponozoonotic disease affecting both humans and animals. Canine leishmaniasis (CanL) is of major importance in Mediterranean region, because of the adverse effects on the health of the affected animals and its role in parasite transmission.

miRNAs are post-transcriptional gene regulators that are involved in immune responses and disease progression. Exosomes, small extracellular vesicles, facilitate miRNA transport, influencing immune modulation and host-pathogen interactions.

This study investigates the differential expression of miRNAs in exosomes derived from the serum of healthy and *Leishmania*-infected dogs. Through next-generation sequencing (NGS) and differential expression analysis using DESeq2 in R, several miRNAs were found to be significantly dysregulated. Among the most upregulated miRNAs in infected dogs, were *cfa-miR-222*, *cfa-miR-92b*, *cfa-miR-29b*, *cfa-miR-129*, and *cfa-miR-486*, which are associated with immune signaling and inflammatory responses. Conversely, key downregulated miRNAs included *cfa-miR-30b*, *cfa-miR-30c*, *cfa-miR-107*, *cfa-miR-374b*, and *cfa-miR-301a*, many of which are implicated in macrophage activation and Th1 immune responses.

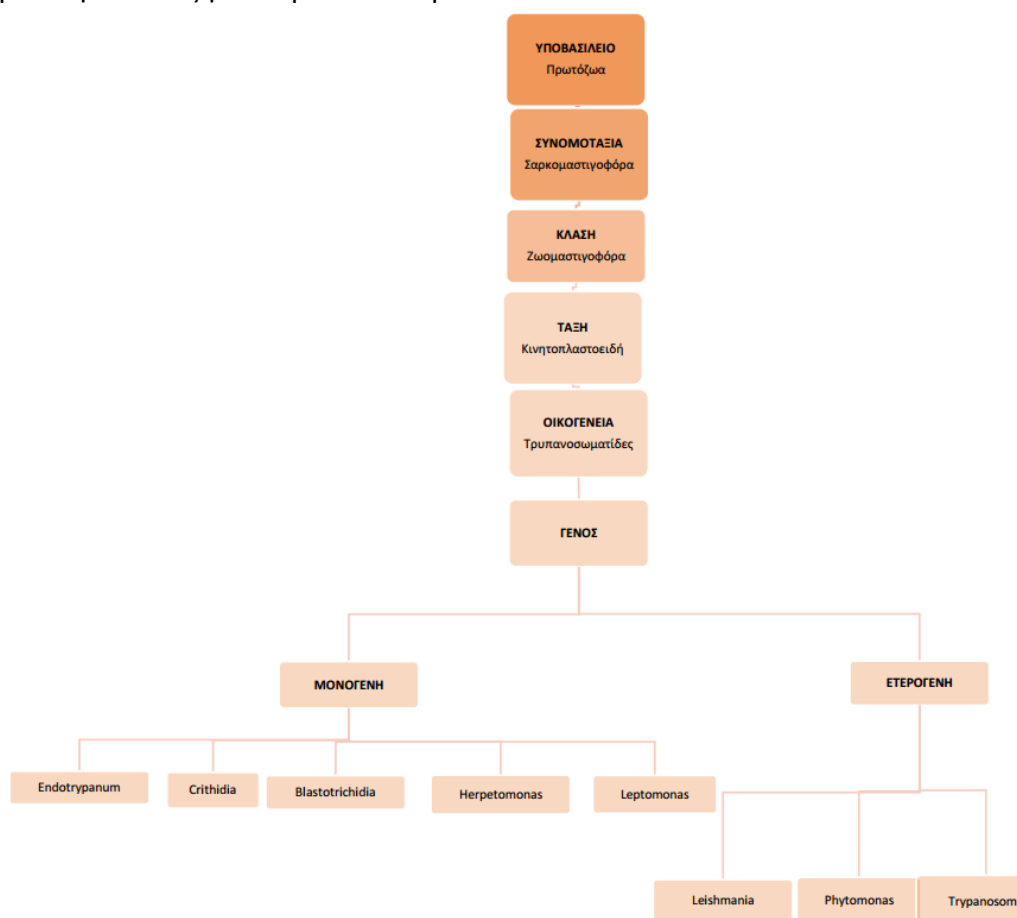
These results indicate that exosomal miRNAs may contribute to the *Leishmania* infection of the host through the modulation of immune responses, which may involve the suppression of inflammatory pathways to avoid immune recognition. The current study shows the possibility of using exosomal miRNAs as biomarkers for the diagnosis and monitoring of canine leishmaniasis. The functions of these exosomal miRNAs in host-pathogen interaction may be helpful for the development of new diagnostic and therapeutic methods of this disease.

KEYWORDS: Leishmaniasis, *Leishmania infantum*, exosomes, miRNA, differential expression analysis, DESeq2, next generation sequencing, NGS, R, MiRdeep2

1. Λεισμάνια

Η λεισμάνια είναι ένα πρωτοζωϊκό παράσιτο το οποίο ανήκει στην οικογένεια των Τρυπανοσωματίδων και είναι υπεύθυνη για την ζωοανθρωπονόσο λεισμανίαση. Η λεισμανίαση, μια παραμελημένη τροπική ασθένεια, εντοπίζεται με διαφορετικές κλινικές μορφές. Οι 3 κύριες μορφές είναι η δερματική λεισμανίαση (cutaneous leishmaniasis) (CL), η σπλαχνική λεισμανίαση (Visceral Leishmaniasis) (VL) και η βλεννογονοδερματική λεισμανίαση (Mucocutaneous Leishmaniasis) (ML). Η λεισμανίαση μπορεί επίσης να εμφανιστεί με μορφή διάχυτης δερματικής λεισμανίασης (Diffuse Cutaneous Leishmaniasis) (DCL) και με μορφή λεισμανίασης μετά το Kala-azar (Post Kala-azar Leishmaniasis) (PKDL).

Συστηματική κατάταξη των τρυπανοσωματίδων:



Εικόνα 1. Συστηματική κατάταξη των τρυπανοσωματίδων (πηγή: Ευσταθίου Αντωνία (2014). Μελέτη μορφών που παίζουν σημαντικό ρόλο στην ανάπτυξη των πρωτόζων *Leishmania* και *Trypanosoma*)

1.1 Παλιός και Νέος κόσμος

Λαμβάνοντας υπόψη την μεγάλη ποικιλομορφία των ειδών, η λεισμανίαση έχει γεωγραφικά χωριστεί σε Παλιό Κόσμο και Νέο Κόσμο. Ο Παλιός Κόσμος, που αναφέρεται στο ανατολικό ημισφαίριο, περιλαμβάνει την Ασία, την Αφρική και τη Νότια Ευρώπη. Ο Νέος Κόσμος αναφέρεται στο Δυτικό Ημισφαίριο και συγκεκριμένα

από το νότιο-κεντρικό Τέξας έως την κεντρική και Νότια Αμερική (εκτός από την Χιλή και την Ουρουγουάη) [1].

1.2 Φλεβοτόμος

Το μαστιγοφόρο πρωτόζωο *Leishmania* μεταδίδεται με το νύγμα ενός μολυσμένου θηλυκού φλεβοτόμου. Οι φλεβοτόμοι είναι μικρά δίπτερα με περιορισμένες πτητικές ικανότητες, τα οποία τρέφονται από χυμούς φυτών. Είναι ενεργοί τις νυχτερινές ώρες από την δύση του ηλίου μέχρι και τις πρώτες ώρες της ανατολής. Τα θηλυκά χρειάζονται αίμα για την ωρίμανση των αυγών τους και για αυτό τον λόγο, μέσω νύξης το προμηθεύονται από τα σπονδυλωτά [2] .

Είναι γνωστά πάνω από 900 είδη φλεβοτόμου, τα περισσότερα από τα οποία ανήκουν στα γένη *Phlebotomus* και *Sergentomyia* στον Παλαιό Κόσμο και *Lutzomyia* στον Νέο Κόσμο. Τουλάχιστον 20 παθογόνα είδη *Leishmania* μεταδίδονται στον άνθρωπο από περίπου 98 είδη sand fly σε 88 χώρες [3].

1.3 Κλινικές περιπτώσεις

1.3.1 Δερματική λεισμανίαση (CL)

Με 600.000 έως 1 εκατομμύριο νέες περιπτώσεις ετησίως παγκοσμίως, η CL είναι η πιο διαδεδομένη κλινική εκδήλωση λεισμανίασης. Η CL προκαλείται από συγκεκριμένα είδη του παρασίτου *Leishmania* και την χαρακτηρίζουν έντονες δερματικές αλλοιώσεις. Τα είδη του φλεβοτόμου τα οποία χρησιμοποιεί ως φορείς είναι κυρίως *Phlebotomus spp.* στον Παλαιό Κόσμο και *Lutzomyia* στον Νέο Κόσμο. Τα είδη λεισμανίας τα οποία την προκαλούν είναι κυρίως οι *L. major*, *L. tropica*, *L. aethiopica*, *L. killicki* και *L. donovani* στον Παλαιό Κόσμο και *L. braziliensis*, *L. guyanensis*, *L. panamensis*, *L. amazonensis*, *L. mexicana*, *L. venezuelensis*, *L. peruviana*, *L. lainsoni*, *L. colombiensis*, *L. garnhami* και *L. rifanoi* στο Νέο Κόσμο. Επίσης, έχουν αναφερθεί και λοιμώξεις από *L. infantum* στον Παλαιό και Νέο Κόσμο.

Το πρώτο σημάδι μόλυνσης συνήθως είναι ένα ερύθημα, το οποίο εμφανίζεται κάποιο χρονικό διάστημα μετά την νύξη του ξενιστή από μολυσμένο φλεβοτόμο. Το ερύθημα εξελίσσεται σε βλατίδα και στη συνέχεια σε ένα οζίδιο που εξελκώνεται προοδευτικά σε μια περίοδο 2 εβδομάδων με 6 μηνών ώστε να γίνει η χαρακτηριστική πληγή της τοπικής δερματικής λεισμανίασης LCL (Local Cutaneous Leishmaniasis). Η αυθόρμητη ανάρρωση συνήθως οδηγεί σε διαρκή ανοσία έναντι ασθενειών, η οποία δυνητικά επεκτείνεται πέρα από το αρχικό είδος *Leishmania*, από το οποίο έγινε η μόλυνση. Η διαδικασία επούλωσης συχνά αφήνει μια μόνιμη ουλή του δέρματος, που ανάλογα με το μέγεθος και την τοποθέτησή της μπορεί να προκαλέσει σημαντικό ψυχολογικό και κοινωνικό τραύμα στους πάσχοντες [4] .

1.3.2 Βλεννογονοδερματική λεισμανίαση (ML)

Η ML θεωρείται δευτερογενές αποτέλεσμα μια αρχικής λοίμωξης CL η οποία έχει εξαπλωθεί στον βλεννογόνο της ρινοστοματοφαρυγγικής/λαρυγγικής περιοχής. Η διάδοση αυτή του παρασίτου συμβαίνει σε λιγότερο από 5% των περιπτώσεων CL. Οι βλάβες του βλεννογόνου μπορεί να εκδηλωθούν χρόνια μετά την αρχική εμφάνιση των δερματικών πληγών. Η ML τυπικά εκδηλώνεται ως ελκώδεις βλάβες στην ανώτερη αεροφόρο και πεπτική οδό, που συνήθως επηρεάζει τον βλεννογόνο της ρινικής κοιλότητας, τη στοματική κοιλότητα, τον στοματοφάρυγγα και τον λάρυγγα, με το ρινικό διάφραγμα και το στόμα να είναι οι πιο συχνά προσβεβλημένες περιοχές. Των ευρημάτων αυτών προηγούνται συμπτώματα ρινικής απόφραξης, επίσταξης, χρόνιου ξηρού βήχα, βραχνάδα ή/και οδυνοφαγίας. Σημαντική ανησυχία αποτελεί και η πιθανότητα αλλοίωσης των χαρακτηριστικών του προσώπου ως επακόλουθο της διάβρωσης του ρινικού χόνδρου [5].

1.3.3 Σπλαχνική λεισμανίαση (VL)

Η σπλαχνική λεισμανίαση η οποία είναι γνωστή και ως kala-azar («μαύρος πυρετός» στα Χίντι) αποτελεί την πιο επικίνδυνη μορφή της λεισμανίασης φτάνοντας την 2^η και 7^η θέση μεταξύ των τροπικών ασθενειών σε θνησιμότητα και απώλειας ετών ζωής προσαρμοσμένης στην αναπηρία αντίστοιχα. Συμπεριλαμβάνεται στην ομάδα των παραμελημένων τροπικών ασθενειών (the neglected diseases) καθώς λόγω της επικράτεια της κυρίως σε αναπτυσσόμενες χώρες, οι φαρμακευτικές εταιρείες δεν δείχνουν μεγάλο ενδιαφέρον στο να επενδύσουν στην έρευνα για την αντιμετώπιση της ασθένειας. Η VL είναι ενδημική ασθένεια σε περισσότερες από 70 χώρες αλλά το μεγαλύτερο ποσοστό της είναι διαμοιρασμένο ανάμεσα σε 7 χώρες (Βραζιλία, Αιθιοπία, Ινδία, Κένυα, Σομαλία, Σουδάν και Νότιο Σουδάν). Η ανατολική μεσόγειος και άλλες περιοχές της Ευρώπης θεωρούνται επίσης ενδημικές για την ασθένεια, αλλά τα αναφερόμενα κρούσματα είναι λιγότερα εν συγκρίσει με τις περιοχές της νότιο-ανατολικής Ασίας και της Αφρικής. Το Ευρωπαϊκό Κέντρο Πρόληψης και Ελέγχου Νόσων (ECDC) παρακολούθησε περιπτώσεις σπλαχνικής λεισμανίασης (VL) σε όλη την Ευρώπη μεταξύ 2014 και 2019. Τα ευρήματα αποκάλυψαν αύξηση των περιστατικών VL στην Αρμενία, το Αζερμπαϊτζάν, τη Γαλλία, την Ελλάδα και τη Λιβύη, ενώ σημειώθηκαν αξιοσημείωτες μειώσεις στην Αλβανία, στην Αλγερία, στην Κροατία, στη Γεωργία, στο Μαρόκο και στην Τυνησία. Επιπλέον, το ECDC παρατήρησε αναδυόμενες περιπτώσεις VL σε υποεθνικό επίπεδο στα καθιερωμένα ενδημικά έθνη όπως η Ιταλία, η Ελλάδα και η Ισπανία, καθώς και σε πρώην μη ενδημικές χώρες όπως η Ρουμανία. Τα παράσιτα τα οποία προκαλούν κυρίως την ασθένεια της VL, είναι το *Leishmania donovani* και το *Leishmania infantum*, το οποίο ονομάζεται και *Leishmania chagasi* στην Νότια Αμερική.

Η σπλαχνική λεισμανίαση (VL) χωρίς την χορήγηση θεραπείας μπορεί να αποβεί μοιραία για τον ασθενή, είτε από την ίδια την ασθένεια, είτε από μολυσματικές ή αιμορραγικές επιπλοκές. Το παράσιτο μπορεί να παραμείνει σε λανθάνουσα κατάσταση και να εμφανιστεί αρκετά χρόνια μετά την μόλυνση του ασθενούς,

συνήθως σε περίπτωση ανάπτυξης ανοσοανεπάρκειας. Στα συμπτώματα της VL περιλαμβάνονται: ο πυρετός, η απώλεια σωματικού βάρους, η σπληνομεγαλία, η μεταβλητή ηπατομεγαλία, η πανκυτταροπενία (συχνότερα αναιμία και θρομβοπενία), τα αυξημένα ηπατικά ένζυμα και η υπολευκωματιναιμία με υπεργαμμασφαιριναιμία [6].

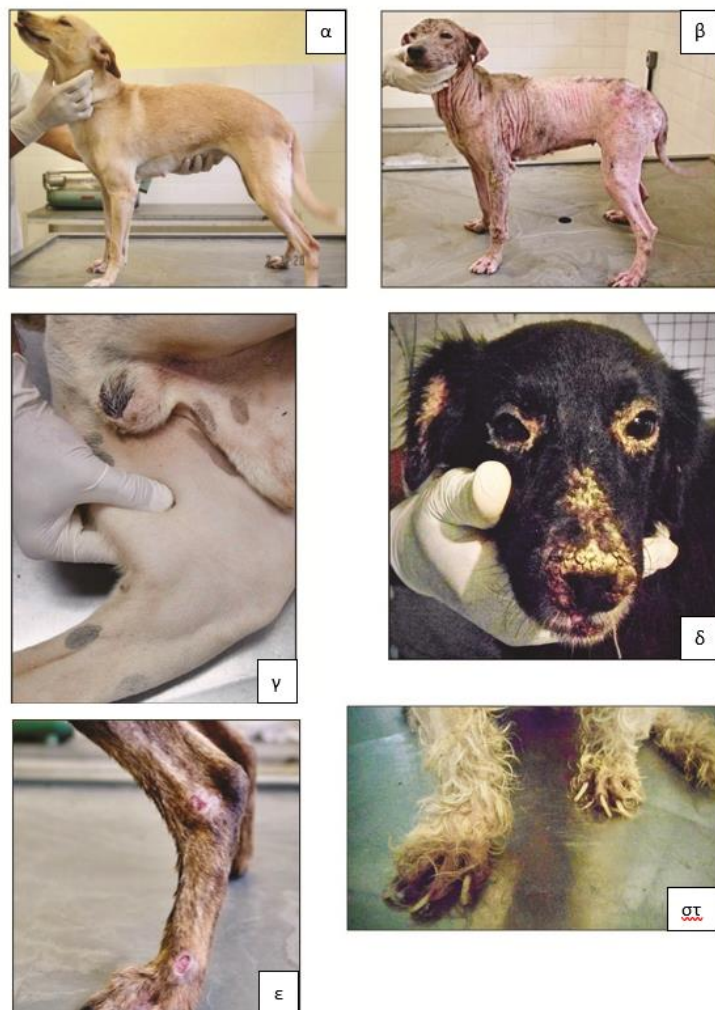
1.3.4 Λεισμανίαση στους σκύλους (Canine Leishmaniasis) (CanL)

Οι σκύλοι αποτελούν τη κύρια πηγή για το *L. Infantum*, το οποίο προκαλεί την CanL, η οποία είναι θανατηφόρα για τα σκυλιά που δεν τους παρέχεται θεραπεία. Η μεσόγειος αποτελεί ενδημική περιοχή για την CanL και αποτελεί ένα σημαντικό κτηνιατρικό πρόβλημα της περιοχής. Λόγω της κλιματικής αλλαγής και της εισαγωγής σκύλων από ενδημικές περιοχές η ασθένεια επεκτείνεται και σε προηγουμένως μη ενδημικές περιοχές της πιο βόρειας Ευρώπης.

Τα τυπικά κλινικά συμπτώματα της CanL είναι:

- Η καχεξία, όπου ο σκύλος είναι σε κακή σωματική κατάσταση και συνήθως τείνει προς την ανορεξία.
- Σημάδια στο δέρμα όπως, ξεφλούδισμα του δέρματος, δερματικές βλάβες και απολεπιστική δερματίτιδα
- Η αλωπεκία
- Η ονυχογρύπωση, ή η διεύρυνση των νυχιών, η οποία είναι πολύ συχνή σε σκύλους με CanL
- Η αδενοπάθεια, η διεύρυνση των λεμφαδένων ως συνέπεια της υπερτροφίας της δομής τους. Αυτή η διεύρυνση δεν σχετίζεται με το φορτίο του παρασίτου στον ιστό ή την κλινική κατάσταση του ζώου.
- Η σπληνομεγαλία, λόγω της διήθησης των μακροφάγων και της αλλαγής της μικροδομής στο όργανο.
- Η ηπατομεγαλία
- Επίσης σε σκύλους με CanL συνήθως παρουσιάζεται ρινορραγία, αιματοουρία, αιμορραγική διάρροια λόγω της διαταραχής της πήξης που προκαλεί η ασθένεια.

Τα όργανα στα οποία το παράσιτο της λεισμανία βρίσκεται σε μεγαλύτερες ποσότητες είναι ο σπλήνας, το ήπαρ και ο μυελός των οστών [2].



Εικόνα 2. Κλινικές εκδηλώσεις σκύλων που έχουν μολυνθεί φυσικά με *Leishmania (Leishmania) infantum*: (α) ασυμπτωματικός σκύλος (φαινομενικά υγιής αλλά μολυσμένος), (β) γενικευμένη μη κνηστώδης αλωπεκία και πολλαπλές άλλες δερματολογικές ανωμαλίες, (γ) λεμφαδενομεγαλία ιγνυακών λεμφαδένων, (δ) διμερής βλεφαρίτιδα και εκτεταμένη εμπλοκή της ρύγχους με έντονες αποφολιδωτικές ελκωτικές αλλοιώσεις, (ε) ελκωτικές αλλοιώσεις στις οστικές προεξοχές του οπίσθιου σκέλους, (στ) ονυχογρύπωση.

(Πηγή: Ribeiro, Raul Rio, Michalick, Marilene Suzan Marques, da Silva, Manoel Eduardo, dos Santos, Cristiano Cheim Peixoto, Frézard, Frédéric Jean Georges, da Silva, Sydnei Magno, *Canine Leishmaniasis: An Overview of the Current Status and Strategies for Control*, BioMed Research International, 2018, 3296893, 12 pages) [7].

1.4 Μορφές και κύκλος ζωής της λεισμάνιας.

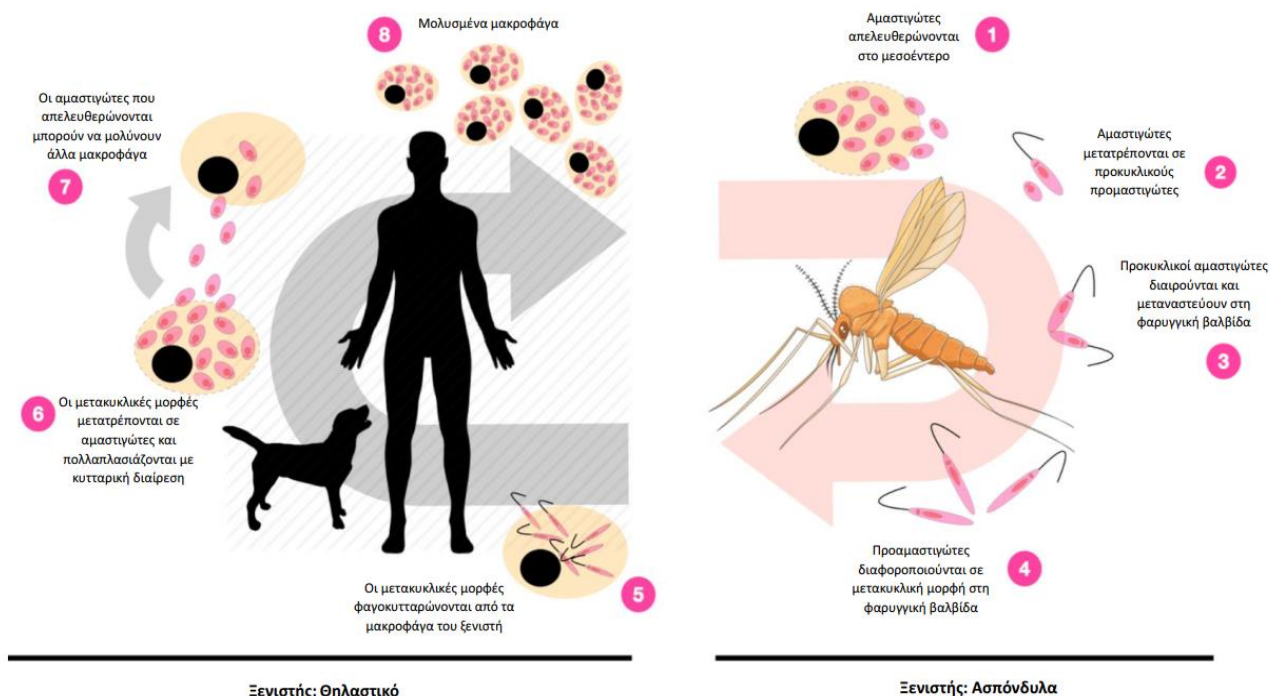
Όπως έχει ήδη αναφερθεί, η λεισμάνια μεταδίδεται με το νύγμα ενός θηλυκού φλεβοτόμου, ο οποίος δρα ως φορέας του παρασίτου. Το παράσιτο της λεισμάνια είναι διφασικό, δηλαδή αλλάζει την φάση του ανάλογα με το είδος του ξενιστή στον οποίο βρίσκεται. Στον ασπόνδυλο ξενιστή, όπως ο φλεβοτόμος, το παράσιτο βρίσκεται σε προμαστιγωτή φάση ενώ σε σπονδυλωτό ξενιστή, όπως ο άνθρωπος ή ο σκύλος, το παράσιτο βρίσκεται σε αμαστιγωτή φάση [2].

Στην προμαστιγωτή μορφή το παράσιτο φέρει μαστίγιο ώστε να μπορεί να κινείται στο έντερο του φλεβοτόμου. Όταν ο φλεβοτόμος απομυζεί αίμα από ένα

σπονδυλωτό, το παράσιτο εγχέεται στο δέρμα στην προμαστιγώτη του μορφή και φαγοκυτταρώνεται από μονοκύτταρα του αίματος. Μέσα σε αυτά τα μονοκύτταρα το παράσιτο θα διαφοροποιηθεί από προμαστιγώτο σε αμαστιγώτο.

Η αμαστιγώτη μορφή είναι η ενδοκυτταρική μορφή του παρασίτου και σε αυτή την μορφή το παράσιτο μπορεί και πολλαπλασιάζεται στο δικτυοενδοθηλιακό σύστημα σε μακροφάγα του σπλήνα, σε μονοκύτταρα, ουδετερόφιλα, στο ήπαρ, στον μυελό των οστών και στους επιχώριους λεμφαδένες [8].

Μετά την είσοδο στο μακροφάγο, ο προμαστιγώτης υφίσταται μεταμόρφωση από την κινητική του μορφή, που χαρακτηρίζεται από μακρύ μαστίγιο και επιμήκη σχήμα κυττάρου, σε μορφή αμαστιγώτη με βραχύ μαστίγιο, που διαθέτει μόνο ένα μικρό βολβώδες άκρο που προεξέχει από ένα πιο στρογγυλεμένο πλέον κυτταρικό σώμα. Αυτή η σημαντική αλλαγή στη μορφολογία των κυττάρων έχει ως αποτέλεσμα μια μειωμένη αναλογία επιφάνειας προς όγκο, ελαχιστοποιώντας έτσι την περιοχή που εκτίθεται στο εχθρικό περιβάλλον του μακροφάγου [9].



Εικόνα 3. Κύκλος ζωής της *Leishmania* (πηγή: Assis, L. H. C., Andrade-Silva, D., et al. (2021). *Cell Cycle, Telomeres, and Telomerase in Leishmania spp.: What Do We Know So Far?* *Cells*, 10(11), 3195. <https://doi.org/10.3390/cells10113195>) [80]

1.5 Γονιδιακή οργάνωση της λεισμάνιας

Το μέγεθος των γονιδιωμάτων της λεισμάνιας κυμαίνεται από 29 Mb (*L. amazonensis*) έως 33 Mb (*L. major*, *L. infantum* και *L. braziliensis*), και είναι οργανωμένα σε ποικίλο αριθμό χρωμοσωμάτων (π.χ. 34 στη *L. amazonensis* και τη *L. mexicana*, 35 στη *L. brasiliensis* και 36 στην *L. major*, την *L. donovani* και την *L. infantum*) [10]. Τα

Leishmania spp. παρουσιάζουν μια μοναδική γονιδιωματική δομή μεταξύ των ευκαρυωτών καθώς τα γονίδια τους δεν έχουν ιντρόνια και τα χρωμοσώματά τους είναι μικρά, περιέχοντας μεγαλύτερη πυκνότητα γονιδίων σε περιορισμένο πυρηνικό χώρο. Τα γονίδια της *Leishmania* βρίσκονται στον ένα ή και στους δύο κλώνους του DNA και μεταγράφονται ως πολυκιστρονικά mRNA από απροσδιόριστους υποκινητές, παρόμοια με την προκαρυωτική μεταγραφή. Ένα ιδιαίτερο χαρακτηριστικό της επεξεργασίας του mRNA στην λεισμανία είναι η χρήση του μηχανισμού trans-splicing [11].

1.6 Παθογένεια της λεισμανιάσης στους σκύλους

Στον σκύλο όπως και στον άνθρωπο, η ικανότητα του παράσιτου της λεισμανίας να προκαλέσει σπλαχνική (VL), δερματική (CL), βλεννογονοδερματική (ML) και μετά-το-Kala-azar (PKDL) λεισμανίαση, διαφέρει μεταξύ των ειδών και των στελεχών της λεισμανίας. Για παράδειγμα, το *L. braziliensis* προκαλεί δερματική ή βλεννογονοδερματική λεισμανίαση στους σκύλους ενώ το *L. infantum* μπορεί να προκαλέσει και δερματική και σπλαχνική λεισμανίαση.

Η σημερινή γνώση της παθογένειας της λεισμανιάσης σε σκύλους προκύπτει από έρευνες που έχουν πραγματοποιηθεί σε φυσικά ή πειραματικά μολυσμένους σκύλους, σε ανθρώπους με (VL) ή με την μόλυνση εργαστηριακών ζώων με κάποιο σπλαχνοτροπικό είδος λεισμανίας. Τα παράσιτα της λεισμανίας μπορούν να μολύνουν πολλούς διαφορετικούς τύπους κυττάρων, συμπεριλαμβανομένων αυτών του μονοπύρηνου φαγοκυτταρικού συστήματος, δενδριτικά κύτταρα, ινοβλάστες, ενδοθηλιακά κύτταρα, ηπατοκύτταρα, ουδετερόφιλα, ηωσινόφιλα και ακόμη και νεοπλασματικά κύτταρα. Με αυτόν τον τρόπο το παράσιτο μπορεί να προσβάλει σχεδόν όλους τους ιστούς και τα όργανα του ξενιστή συμπεριλαμβανομένου και του κεντρικού νευρικού συστήματος. Τα κύρια κύτταρα-ξενιστές του παρασίτου θεωρούνται τα φαγοκύτταρα, καθώς παρόλο που η μόλυνση μη φαγοκυτταρικών κυττάρων μπορεί να συμβάλει στην μεγαλύτερη χρονικά επιβίωση του παρασίτου, σπάνια είναι παραγωγική [12].

Μετά την είσοδο του παρασίτου σε ένα θηλαστικό, το προμαστιγωτό θα πρέπει να προσπαθήσει να αποφύγει την δράση του συμπληρώματος μέχρι να φαγοκυτταρωθεί από ένα μακροφάγο. Η προμαστιγωτή μορφή του παρασίτου είναι ευάλωτη στην δράση του συμπληρώματος σε αντίθεση με την αμαστιγωτή μορφή, η οποία ειδικεύεται στην μετάδοση του παρασίτου στον ξενιστή και μπορεί να αποφύγει τελείως την λύση από το συμπλήρωμα.

Τα προμαστιγωτά παράσιτα προσδένονται στα μακροφάγα και εσωτερικεύονται στο σχετικά μη εχθρικό περιβάλλον του ενδοσώματος, και εκεί ξεκινάνε να διαφοροποιούνται σε αμαστιγώτες. Οι προμαστιγώτες είναι ιδιαίτερα ευαίσθητοι στο όξινο υδρολυτικό περιβάλλον του φαγολυσσοσώματος, όμως με την διαφοροποίηση σε αμαστιγωτή μορφή το παράσιτο μπορεί να επιβιώσει στο

αφιλόξενο αυτό περιβάλλον του ενδοσώματος, καθώς αποτελεί μια πιο ανθεκτική μορφή του.

Οι αμαστιγώτες θα πολλαπλασιαστούν εντός του μακροφάγου έως ότου προκαλέσουν την λύση του κυττάρου και θα συνεχίσουν με την μόλυνση άλλων περιφερειακών μακροφάγων [13].

1.7 Ανοσολογική απόκριση στην λείσμανίαση

Οι κυτταρικές ανοσολογικές αποκρίσεις που ρυθμίζονται από την υποομάδα Th1, ιδιαίτερα μέσω των κυτταροκινών IFN-γ και TNF-α, κυριαρχούν σε ασυμπτωματικούς σκύλους που παρουσιάζουν εμφανή αντίσταση στη σπλαχνική λείσμανίαση. Αντιθέτως, ενώ ο ρόλος των Th2 κυτταροκινών, όπως η IL-4, η IL-10 και η IL-13 στα συμπτωματικά ζώα παραμένει αμφιλεγόμενος, υπάρχουν αυξανόμενες ενδείξεις που συσχετίζουν αυτές τις κυτταροκίνες με την προοδευτική νόσο [14]. Σε παλαιότερες μελέτες σε ποντίκια, υπήρχε ξεκάθαρος διαχωρισμός μεταξύ Th1 ως προ-φλεγμονώδη και Th2 ως αντιφλεγμονώδη απόκριση, αλλά σήμερα γνωρίζουμε πως οι ανοσολογικές αποκρίσεις σε ανθρώπους και σε σκύλους είναι αρκετά πιο περίπλοκες.

Οι ανοσολογικοί μηχανισμοί που προσδίδουν στον σκύλο αντίσταση ή ευαισθησία στη μόλυνση και την επακόλουθη νόσο της σπλαχνικής λείσμανιάσης παραμένουν άγνωστοι. Φαίνεται όμως ότι παράγοντες όπως η φυλή, εάν βρίσκεται ο σκύλος σε ανοσοκαταστολή, εάν νοσεί από κάποια συνυπάρχουσα ασθένεια και η θρεπτικότητα της τροφής που καταναλώνει μπορούν να επηρεάσουν το αποτέλεσμα της μόλυνσης [15].

1.8 Θεραπεία της λείσμανιάσης

Η λείσμανίαση θεωρείται ως μια από τις σημαντικότερες τροπικές παραμελημένες ασθένειες από τον Παγκόσμιο Οργανισμό Υγείας. Ακόμη και μετά από περισσότερο από έναν αιώνα, εξακολουθούμε να έχουμε λίγα φάρμακα στην διάθεση μας για την θεραπεία της νόσου.

Τόσο η συνεχής εμφάνιση καινούργιων ανθεκτικών στελεχών, όσο και η έλλειψη σημαντικών επενδύσεων για την έρευνα και την ανάπτυξη καινούργιων φαρμάκων, αποτελούν εμπόδια για την εξέλιξη της θεραπείας της ασθένειας [16].

1.8.1 Συμβατική χημειοθεραπεία

Τα πεντασθενή άλατα αντιμονίου (Sb^{V}) ήταν ιστορικά το φάρμακο πρώτης γραμμής για σχεδόν 7 δεκαετίες κατά της λείσμανιάσης, όμως έχουν συσχετισθεί με καρδιοτοξικότητα, κίρρωση, παγκρεατική τοξικότητα και κίνδυνο ανάπτυξης ανθεκτικότητας του παρασίτου.

Η αμφοτεरिकίνη Β και η λιποσωμική αμφοτεरिकίνη αποτελούν τα φάρμακα 2^{ης} γραμμής κατά της λεισμανίασης. Μετά τη μείωση της αποτελεσματικότητας των αντιμονικών φαρμάκων εξελέγη ως φάρμακο πρώτης γραμμής κατά της VL.

Η Μιλτεφοσίνη μέχρι σήμερα αποτελεί το μοναδικό φάρμακο κατά της λεισμανίασης το οποίο έχει στοματικό τρόπο χορήγησης. Παρόλο που έχει καλή αποτελεσματικότητα ενάντια σε συγκεκριμένα είδη λεισμανίας, η τερατογένεση και η ανάπτυξη ανθεκτικότητας από τα παράσιτα είναι πιθανές παρενέργειες.

Η Παρομομυκίνη είναι ένα αμινογλυκοσιδικό αντιβιοτικό που έχει αποδειχθεί ως ένας αποτελεσματικός αντιλεισμανικός παράγοντας. Η θεραπεία με παρομομυκίνη σε κλινική δοκιμή στην Ινδία βρέθηκε να είναι επιτυχής κατά της λεισμανίασης (94%) [16], αλλά ήταν λιγότερο αποτελεσματική στον Σουδανικό πληθυσμό [17].

Η Πενταμιδίνη από την δεκαετία του 40 όπου και έγινε η σύνθεσή της, έχει μελετηθεί για την αντιλεισμανιακή της δράση. Η δράση της όμως αυτή ποικίλει μεταξύ των διαφορετικών ειδών λεισμανίας [17] [16].

1.8.2 Συνδυαστική χημειοθεραπεία

Με στόχο την καταπολέμηση ανθεκτικών στελεχών της λεισμανίας, την μείωση του χρόνου θεραπείας, των παρενεργειών και του κόστους, αναπτύχθηκε η συνδυαστική χημειοθεραπεία. Συνήθως προτείνεται σε άτομα τα οποία δεν ανταποκρίνονται στην μονοθεραπεία με Sb^v.

1.8.3 Τοπικές θεραπείες της CL

Οι τοπικές θεραπείες κατά της CL περιλαμβάνουν:

- Την φωτοδυναμική θεραπεία (PDT), στην οποία αμινολαβουλινικό οξύ (ALA) ή μεθυλο-αμινολαβουλινικό οξύ εφαρμόζονται τοπικά στο δέρμα και ακολουθεί ακτινοβολία με λέιζερ ή έντονο παλμικό φως.
- Την κρυοθεραπεία, στην οποία χρησιμοποιείται υγρό άζωτο (στους -195 °C) σε δερματικές βλάβες από την λεισμανίαση μία ή δύο φορές εβδομαδιαίως. Η βλάβη των μολυσμένων κυττάρων και η καταστροφή των αμαστιγωτών επέρχεται με τον ενδοκυτταρικό και εξωκυτταρικό σχηματισμό παγοκρυστάλλων και με αλλαγές στην κυτταρική μεμβράνη.
- Τα παράσιτα της λεισμανία τα οποία προκαλούν δερματική νόσο είναι ευαίσθητα σε θερμοκρασίες άνω των 39°C. Με την χρήση ραδιοσυχνότητας (RF) η θερμότητα μπορεί να φθάσει έως και 4mm εντός του δέρματος σκοτώνοντας τους αμαστιγώτες χωρίς να βλάψει τον περιφερικό ιστό [17] [16].

1.8.4 Θεραπεία της λεισμανίασης στους σκύλους (CanL)

Η θεραπεία της CanL έχει ως στόχο την διαχείριση κλινικών συμπτωμάτων και αλλοιώσεων, την ενίσχυση της κυτταρικής ανοσίας του σκύλου, την μείωση του φορτίου των παρασίτων, την πρόληψη των υποτροπών και την μείωση του ρυθμού μετάδοσης στους φορείς. Το παρασιτικό φορτίο μπορεί να μειωθεί σε ασυμπτωματικό επίπεδο με την χρήση της χημειοθεραπείας. Οι αντιλεισμανιακές θεραπείες, οι οποίες χρησιμοποιούνται για την αντιμετώπιση της σπλαχνικής λεισμανίασης στους σκύλους, είναι παρόμοιες με εκείνες που χρησιμοποιούνται στον άνθρωπο.

Η επιλογή κατάλληλης θεραπείας εξαρτάται από το ποιο στάδιο της ασθένειας βρίσκεται ο σκύλος σύμφωνα με το σύστημα σταδιοποίησης Solano-Gallego *et al.* (2011) [18], το οποίο κατατάσσει 4 στάδια της ασθένειας. Στο πρώτο στάδιο, στο οποίο ο σκύλος είναι ασυμπτωματικός ή παρουσιάζει ήπια νόσο, δεν απαιτείται θεραπεία όμως ο σκύλος μπορεί να υποβληθεί σε θεραπεία μόνο με αλλοπουρινόλη. Στο δεύτερο στάδιο, ο σκύλος παρουσιάζει μέτρια νόσο και στο τρίτο στάδιο παρουσιάζει σοβαρή νόσο που σχετίζεται με χρόνια νεφρική νόσο. Συνιστάται η θεραπεία των σκύλων δεύτερου και τρίτου σταδίου με χορήγηση συνδυασμού αλλοπουρινόλης και αντιμονιακών/μυλτεφοσίνης. Οι σκύλοι που βρίσκονται στο τέταρτο στάδιο παρουσιάζουν πολύ σοβαρή νόσο, που περιλαμβάνει νεφρωσικό σύνδρομο και θα πρέπει να λάβουν θεραπεία μόνο με αλλοπουρινόλη ώστε να αποφεύγεται περαιτέρω νεφρική βλάβη [2].

1.8.5 Εμβόλιο κατά της λεισμανίασης στους ανθρώπους

Υπάρχει επιτακτική ανάγκη για ένα αποτελεσματικό εμβόλιο για την παροχή μακροχρόνιας ανοσίας για τον έλεγχο των παρασιτικών λοιμώξεων. Επι του παρόντος δεν υπάρχει εγκεκριμένο εμβόλιο για την ανθρώπινη λεισμανίαση, όμως εκτεταμένες προσπάθειες βρίσκονται σε εξέλιξη για την ανάπτυξη ποικίλων τρόπων εμβολιασμού με πολλά υποσχόμενα αποτελέσματα παγκοσμίως. Υποψήφια εμβόλια πρώτης, δεύτερης και ακόμη και τρίτης γενιάς βρίσκονται σε στάδιο κλινικών δοκιμών. Επίσης θεωρείται πως τα εμβόλια για τα ζώα θα διαδραματίσουν έναν ιδιαίτερα σημαντικό ρόλο στην πρόληψη της μετάδοσης της λεισμανίασης στον άνθρωπο [19].

1.8.6 Εμβόλιο κατά της λεισμανίασης στους σκύλους

Ένας από τους βασικούς τρόπους αποτροπής της λεισμανίασης στους σκύλους (CanL) είναι ο εμβολιασμός. Οι μολυσμένοι σκύλοι μπορούν να αποτελέσουν εστία μετάδοσης μέσω των φλεβοτόμων σε άλλους σκύλους και ανθρώπους. Από το 2004, έχουν εγκριθεί 4 εμβόλια κατά της λεισμανίασης, 2 εκ των οποίων στην Βραζιλία (Το Leishmune® και το Leish-Tec®) και 2 στην Ευρώπη (Canileish® και το LetiFend®). Το 2014, λόγω έλλειψης στοιχείων αποδοτικότητας στην τρίτη φάση των κλινικών δοκιμών, το εμβόλιο Leishmune® αποσύρθηκε από την αγορά από το υπουργείο γεωργίας της Βραζιλίας [20] [21].

Ένα ιδανικό εμβόλιο κατά της CanL θα πρέπει:

- να είναι ασφαλές και προσιτό σε τιμή
- να διεγείρει την ανοσολογική αντίδραση σε CD4⁺ και CD8⁺ T κύτταρα αλλά και την μακροχρόνια ανοσολογική μνήμη
- να είναι εύκολη η παραγωγή του και να μπορεί να αποθηκευτεί σε θερμοκρασία δωματίου ή στους 4°C (Χωρίς να χρειάζεται δηλαδή η αποθήκευση του στους -20 °C ή -80 °C) [21].

2. miRNA

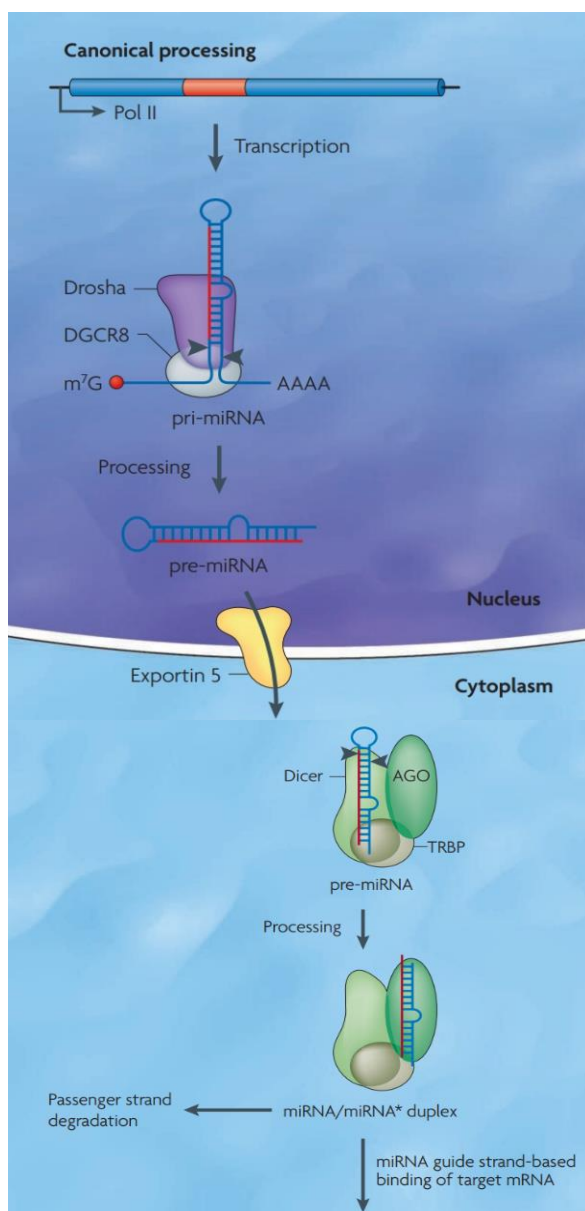
Τα microRNAs (miRNAs) είναι μόρια περίπου 22 nt σε μέγεθος και έχουν την ικανότητα ρύθμισης βασικών βιολογικών λειτουργιών όπως είναι ο κυτταρικός θάνατος, η μεταγωγή σημάτων, η ανάπτυξη, ο κυτταρικός πολλαπλασιασμός και η κυτταρική διαφοροποίηση. Τα miRNA ανήκουν στην τάξη των RNA που δεν κωδικοποιούν πρωτεΐνες και έχουν σημαντικό ρόλο στην μετα-μεταγραφική ρύθμιση της γονιδιακής έκφρασης. Μέσω της αλληλεπίδρασης τους με την 3' αμετάφραστη περιοχή των mRNA μπορούν να τα οδηγήσουν σε αλλοίωση και αποσταθεροποίηση. Υπάρχουν αναφορές για αλληλεπιδράσεις των miRNA και σε άλλες θέσεις των mRNA όπως η 5' αμετάφραστη περιοχή, η κωδική περιοχή και οι υποκινητές των γονιδίων [22] [23] [24].

2.1 Κανονική οδός βιογένεσης των miRNAs

Η βιογένεση των miRNA, όταν ακολουθείται η συμβατική οδός, χαρακτηρίζεται από δύο διασπάσεις του αρχικού μορίου RNA οι οποίες πραγματοποιούνται μία στον πυρήνα και η άλλη στο κυτταρόπλασμα του κυττάρου [25].

Τα γονίδια των miRNA κατά κύριο λόγο βρίσκονται στα ιντρόνια κωδικοποιητικών και μη κωδικοποιητικών μεταγράφων. Η μεταγραφή τους πραγματοποιείται κυρίως από την RNA πολυμεράση II αλλά και από την RNA πολυμεράση III. Το primary miRNA (pri – miRNA) το οποίο θα προκύψει λόγω της αναδίπλωσης του μεταγράφου στον χώρο, θα αποκτήσει μία δομή φουρκέτας ενώ στο 3' και 5' άκρο θα βρίσκονται προεξοχές με ελεύθερα νουκλεοτίδια [25] [26].

Η πρώτη επεξεργασία του μη ώριμου miRNA πραγματοποιείται στο εσωτερικό του πυρήνα του κυττάρου με την πρώτη διάσπαση του μορίου από το πρωτεϊνικό σύμπλεγμα της Drosha, η οποία ανήκει στην οικογένεια ριβονουκλεάσεων III και της πρωτεΐνης δέσμευσης δίκλωνου RNA (DiGeorge syndrome critical region gene 8) DGCR8. Το σύμπλεγμα αυτό είναι γνωστό και ως microprocessor, με την DGCR8 να προσδένεται στη διασταύρωση ssRNA-dsRNA και να κατευθύνει τη Drosha να διασπάσει ~11 bp μακριά από τη διασταύρωση. Έτσι απομακρύνονται οι προεξοχές των 3' και 5' άκρων σχηματίζοντας το precursor miRNA (pre-miRNA) [26].



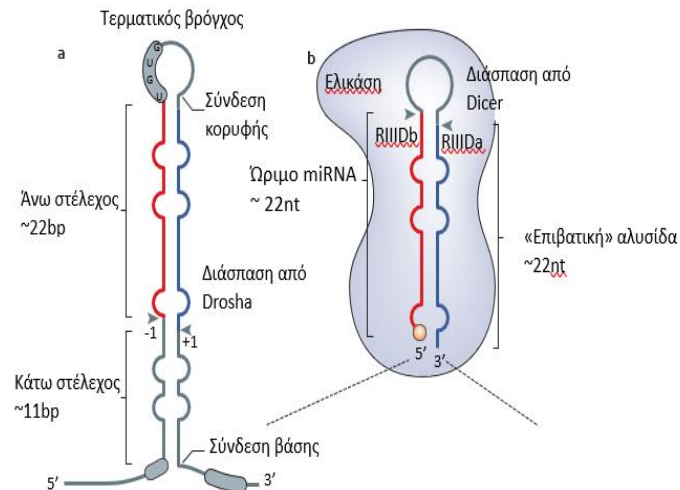
Εικόνα 4. Απεικόνιση κανονικής οδού βιογένεσης των miRNAs (Πηγή: Krol, J., Loedige, I., & Filipowicz, W. (2010). *The widespread regulation of microRNA biogenesis, function and decay.*

Nature reviews. Genetics, 11(9), 597–610.
<https://doi.org/10.1038/nrg2843>

Η επεξεργασία του pre-miRNA λαμβάνει χώρα στο κυτταρόπλασμα. Το pre-miRNA σχηματίζει σύμπλοκο με τον πυρηνοκυτταροπλασματικό μεταφορέα Exportin-5 και το RanGTP, το οποίο το προστατεύει από την αποικοδόμηση του στον πυρήνα και βοηθά στην μεταφορά του στο κυτταρόπλασμα [26].

Η δεύτερη διάσπαση του miRNA πραγματοποιείται στο κυτταρόπλασμα του κυττάρου από την ριβονουκλεάση III (RNase III) Dicer με την βοήθεια της TRBP (transactivation response RNA binding protein), η οποία είναι ένας κρίσιμος παράγοντας που βοηθά τη Dicer να βρίσκει και να διασπά αποτελεσματικά το pre-miRNA μεταξύ μιας μεγάλης ποσότητας κυτταρικών RNAs. Η Dicer προσδένεται στο pre-miRNA με προτίμηση στην θέση προεξοχής των δύο βάσεων στο 3' άκρο που είχε δημιουργηθεί κατά την επεξεργασία του μορίου από την Drosha [27].

Η δομή φουρκέτας του pre-miRNA ευθυγραμμίζεται με τη Dicer και επειδή οι καταλυτικές περιοχές της Dicer RNase III, RIIIIda και RIIIDb, βρίσκονται περίπου 22 νουκλεοτίδια (nt) μακριά από τα άκρα του μορίου, μετά την διάσπαση και απομάκρυνση της φουρκέτας θα προκύψει ένα δίκλωνο RNA μόριο περίπου 22nt σε μήκος [25] [28].



Εικόνα 5. Διάσπαση του miRNAs από Dicer και Drosha (Πηγή: Ha, M., Kim, V. Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol 15, 509–524 (2014). <https://doi.org/10.1038/nrm3838>) [25]

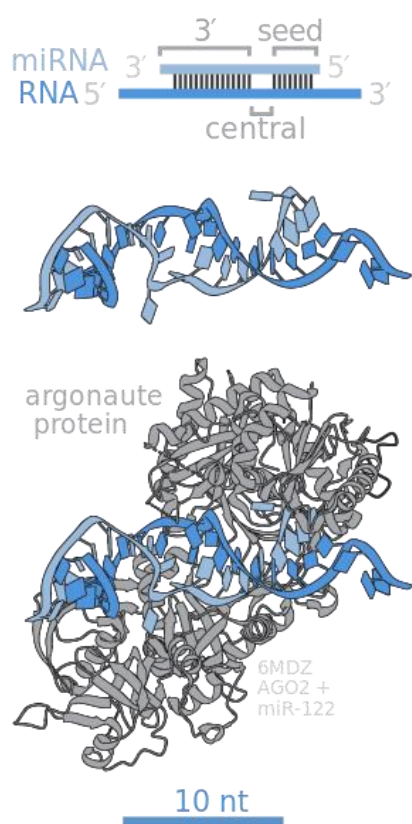
2.2 RNA induced silencing complex (RISC)

Το δίκλωνο μόριο RNA, το οποίο προκύπτει μετά την επεξεργασία του από την ριβονουκλεάση Dicer φορτώνεται σε μία πρωτεΐνη αργοναύτης AGO (Argonaut) και σχηματίζει το RISC (RNA induced silencing complex). Στον άνθρωπο, η Argonaute2 (Ago2) είναι η καταλυτική ενδονουκλεάση του RISC με δομικές μελέτες να ταυτοποιούν την περιοχή PIWI ως το καταλυτικό κέντρο της πρωτεΐνης.

Μία από τις δύο αλυσίδες του miRNA θα απομακρυνθεί από το RISC και θα αποτελεί την «επιβατική» αλυσίδα (passenger strand) ενώ εκείνη που θα παραμείνει αποτελεί την οδηγό αλυσίδα (guide strand). Μετά την φόρτωση του δίκλωνου miRNA στην Ago2 σχηματίζεται το pre-RNA induced silencing complex. Εάν το δίκλωνο μόριο είναι συμπληρωματικό στο κέντρο του, η AGO2 θα διασπάσει την επιβατική αλυσίδα (passenger strand) και με την βοήθεια της πρωτεΐνης C3PO, η αλυσίδα θα απομακρυνθεί από το RISC και θα σχηματιστεί το ώριμο (mature) RISC.

Αυτός ο μηχανισμός χρησιμοποιείται σπάνια στην οδό του miRNA, καθώς τα περισσότερα δίκλωνα miRNA έχουν κεντρικές αναντιστοιχίες που εμποδίζουν την διάσπαση από την AGO2, ενώ οι ανθρώπινες AGO1, AGO3 και AGO4 στερούνται της ικανότητας διάσπασης. Συνεπώς, το ξετύλιγμα των δίκλωνων miRNA χωρίς διάσπαση, είναι πιο συνηθισμένο από την διάσπαση της επιβατικής αλυσίδας. Οι αναντιστοιχίες στην οδηγό αλυσίδα στις θέσεις νουκλεοτιδίων 2–8 και 12–15, προάγουν το ξετύλιγμα των δίκλωνων miRNA [29] [25].

Το RISC χρησιμοποιεί την οδηγό αλυσίδα για να εντοπίσει το mRNA στόχο και να προκαλέσει τη μείωση της έκφρασης του γονιδίου που αυτό κωδικοποιεί. Το RISC

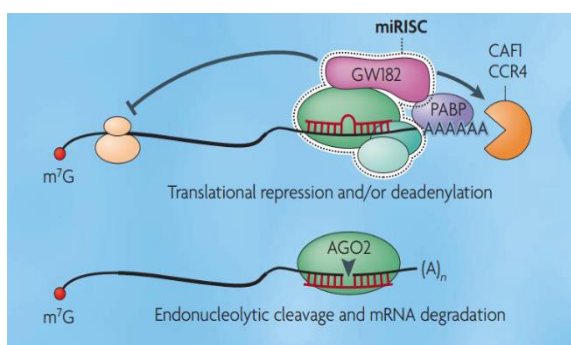


Εικόνα 6. Απεικόνιση RISC: AGO2 (γκρι) με ένα microRNA (ανοιχτό μπλε) και το mRNA στόχο του (σκούρο μπλε) (Πηγή: J.J.Froehlich, CC BY-SA 4.0, <https://commons.wikimedia.org/w/index.php?curid=122466245>)

μπορεί να επιτύχει αυτή την μείωση, είτε με την διάσπαση του mRNA στόχου, εφόσον το miRNA είναι επαρκώς συμπληρωματικό σε αυτό, είτε με την καταστολή της μετάφρασης του. Η αναγνώριση του mRNA στόχου πραγματοποιείται από την περιοχή του miRNA "Seed Region", στην οποία ανήκουν οι θέσεις 2-8 του 5' άκρου του miRNA και θα πρέπει να είναι απόλυτα συμπληρωματική με την 3' μη μεταφραζόμενη περιοχή (UTR) του mRNA [30] [23].

Ο βαθμός συμπληρωματικότητας του miRNA καθορίζει αν θα υπάρξει AGO2-εξαρτώμενη διάσπαση του mRNA-στόχου ή μεταφραστική καταστολή με τη μεσολάβηση του RISC. Στους ζωικούς οργανισμούς, τα περισσότερα miRNA είναι μόνο μερικώς συμπληρωματικά με την περιοχή UTR του mRNA στόχου τους και οι πρωτεΐνες AGO δεν αρκούν για την αποσιώπηση του.

Στην περίπτωση αυτή θα χρειαστεί το RISC να αλληλεπιδράσει με πρωτεΐνες της οικογένειας GW182 δημιουργώντας το σύμπλοκο AGO-GW182. Το σύμπλοκο αυτό έχει την ικανότητα αποσιώπησης του mRNA στόχου, καταστέλλοντας τη μετάφραση και επιταχύνοντας την αποδόμηση του [31].



Εικόνα 7. Διάσπαση ή αποσιώπηση του mRNA στόχου, από το miRISC (Πηγή: Krol, J., Loedige, I., & Filipowicz, W. (2010). The widespread regulation of microRNA biogenesis, function and decay. *Nature reviews. Genetics*, 11(9), 597–610. <https://doi.org/10.1038/nrg2843>)

2.3 miRNA ονοματολογία

Η συνεπής ονοματολογία των miRNAs είναι ιδιαίτερα σημαντική τόσο για τον διαχωρισμό των miRNAs από τα υπόλοιπα μικρά RNAs αλλά και για την αποφυγή περιπτώσεων πλεονασμών ή λανθασμένων επισημάνσεων. Επί του παρόντος, τα miRNAs ονομάζονται με το πρόθεμα «miR» το οποίο ακολουθεί ένα αριθμητικό αναγνωριστικό. Τα miRNAs ζώων και φυτών ξεχωρίζονται με την χρήση ή όχι της παύλας ανάμεσα στο πρόθεμα και το αριθμητικό αναγνωριστικό (π.χ. miR-21 για ζώο, miR21 για φυτό). Τα pre-miRNAs ονομάζονται με το πρόθεμα τους σε πεζά και σε πλάγιους (*italic*) χαρακτήρες (π.χ. *mir-21*). Ίδια miRNAs από διαφορετικά είδη ξεχωρίζουν μεταξύ τους με ένα πρόθεμα τριών γραμμάτων ακολουθούμενο από μία παύλα πριν από το «miR» πρόθεμα τους. Για παράδειγμα στον σκύλο (*Canis lupus familiaris*) πριν το «miR» τοποθετείται το πρόθεμα «cfa-» (π.χ. cfa-miR-21). Σε ώριμα miRNAs, τα οποία είναι στενά συγγενή μεταξύ τους αλλά εκφράζονται από διαφορετικούς γονιδιωματικούς τόπους προστίθεται μια κατάληξη ενός γράμματος (π.χ. cfa-miR-20a). Ώριμα miRNAs τα οποία είναι πανομοιότυπα μεταξύ τους, ανήκουν στο ίδιο είδος, αλλά εκφράζονται από διαφορετικούς γονιδιωματικούς τόπους διαχωρίζονται με την προσθήκη ενός αριθμού στην κατάληξή τους μετά από μία παύλα (π.χ. cfa-miR-16-1). Σε ορισμένες περιπτώσεις miRNAs με διαφορετικές ώριμες αλληλουχίες εκφράζονται από τον ίδιο precursor. Σε αυτή την περίπτωση στην κατάληξη του ονόματος τοποθετείται ένας αριθμός μετά από τελεία (π.χ. osa-miR159a.1). Οι καταλήξεις «-3p» και «-5p» υιοθετήθηκαν για να υποδεικνύουν δύο διαφορετικά ώριμα miRNAs που δημιουργήθηκαν από τα 3' και 5' άκρα της pre-miRNA φουρκέτας (π.χ. hsa-miR-21-5p). Ορισμένα γονίδια miRNA που βρέθηκαν σε πρώτες γενετικές μελέτες ονομάστηκαν από τους φαινοτύπους τους (για παράδειγμα, lin-4, let-7 και lsy-6) [32] [33] [25].

2.4 Μη κανονικές οδοί βιογένεσης miRNA

Πέρα από την κανονική οδό βιογένεσης υπάρχουν και μη κανονικές οδοί, οι οποίες μπορούν να συνθέσουν miRNAs ή μόρια, τα οποία μοιάζουν με miRNAs στην δομή και στην λειτουργία τους. Συνήθως αυτές οι οδοί συμπεριλαμβάνουν περισσότερα ή λιγότερα βήματα από το κανονικό μονοπάτι ωρίμανσης του miRNA και κατηγοριοποιούνται σε τρεις ομάδες: Ανεξάρτητες από Drosha/DGCR8, ανεξάρτητες από Dicer και εξαρτημένες από τελικές ουριδυλοτρανσφεράσες (TUT).

Στις ανεξάρτητες από Drosha/DGCR8 οδούς, σχηματίζονται μόρια φουρκέτας που μοιάζουν με pre-miRNA και λειτουργούν ως υπόστρωμα για την ριβονουκλεάση Dicer.

Σε αυτή την ομάδα συμπεριλαμβάνονται:

- Τα mirtrons, τα οποία δημιουργούνται από το συναρμόσωμα (spliceosome) και το ένζυμο DBR1.

- Τα miRNA με 5' καλύπτρα τα οποία παράγονται απευθείας από την RNA πολυμεράση II και στο 5' άκρο τους διαθέτουν μια καλύπτρα μεθυλογουανωσίνης που βοηθάει στην εξαγωγή του miRNA από τον πυρήνα.
- miRNAs τα οποία προέρχονται από άλλα μη κωδικοποιητικά RNAs, όπως tRNAs, μικρά nucleolar miRNAs (snoRNAs) ή ιϊκά RNAs που μοιάζουν με μικρά πυρηνικά RNAs. Για παράδειγμα, σε κύτταρα μολυσμένα με τον γ-ερπητοϊό 68 τα ιϊκά miRNAs συν-μεταγράφονται με τα tRNAs από την RNA πολυμεράση III και διασπώνται από την κυτταρική ριβονουκλεάση Z (RNaseZ) στο 3' άκρο του tRNA για να απελευθερώσει το pre-miRNA, το οποίο στην συνέχεια θα υποστεί επεξεργασία από την ριβονουκλεάση Dicer.

Παρόλο που οι περισσότερες οδοί βιογένεσης εξαρτώνται από τη Dicer, η βιογένεση ορισμένων miRNAs στηρίζεται στην καταλυτική δράση της AGO2 πρωτεΐνης. Σε αυτά τα miRNAs μετά την δράση της Drosha, το μόριο φουρκέτας το οποίο προκύπτει είναι πολύ μικρό για να επεξεργαστεί από τη Dicer και ενσωματώνεται απευθείας στην AGO2. Εκεί, το μόριο διασπάται στο καταλυτικό της κέντρο ώστε να προκύψει ένα ώριμο miRNA που όμως θα έχει διαφορετικά στοιχεία από τα άλλα miRNAs που παράχθηκαν μέσω της κανονικής οδού βιογένεσης.

Σε αντίθεση με τα pre-miRNA, τα οποία έχουν προέκταση δύο νουλεοτιδίων στο 3' άκρο τους και ανήκουν στην ομάδα I των pre-miRNAs, κάποια miRNAs έχουν προέκταση ενός μόνο νουκλεοτιδίου στο 3' άκρο τους και ανήκουν στην ομάδα II των pre-miRNAs. Στα pre-miRNA της ομάδας II προστίθεται ένα νουκλεοτίδιο στο 3' άκρο μέσω μονοουριδυλίωσης με τη μεσολάβηση των τερματικών ουριδυλυλικών τρανσφεράσων. Η προσθήκη αυτή του ενός νουκλεοτιδίου είναι απαραίτητη για την αποτελεσματικότητα της επεξεργασίας του μορίου από την ριβονουκλεάση Dicer [25] [24].

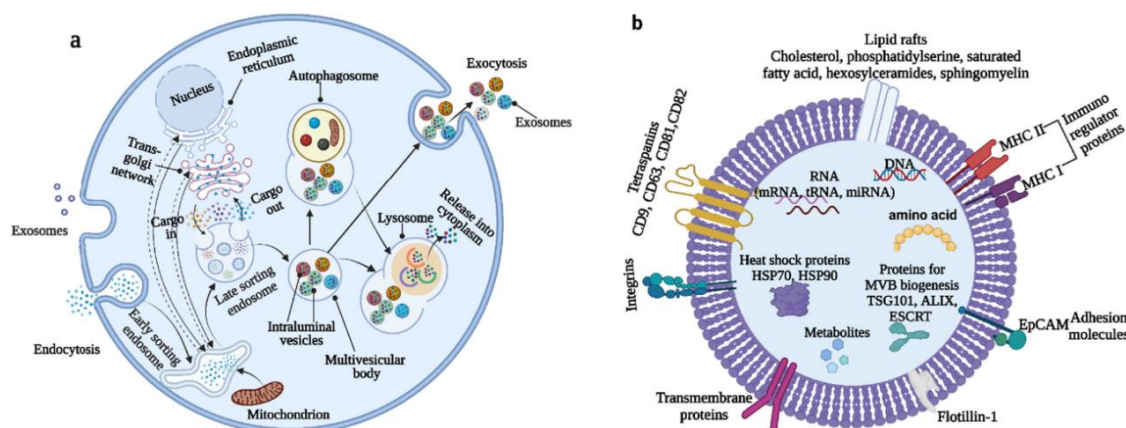
2.5 miRNAs και εξωσώματα

Τα εξωσώματα, με διάμετρο 30 – 150nm, θεωρούνται μία από τις σημαντικότερες και πολυπληθέστερες μορφές εξωκυττάρων κυστιδίων (Extracellular vesicles) (EVs). Διαδραματίζουν έναν καθοριστικό ρόλο στη διακυτταρική σηματοδότηση, στη μετάδοση βιοδραστικών μορίων και στην ανταπόκριση στο διαρκώς μεταβαλλόμενο κυτταρικό περιβάλλον.

Τα εξωκυττάρια κυστίδια (EVs) είναι κυστίδια ενδοσωματικής προέλευσης που περιβάλλονται από μεμβράνη λιπιδικής διπλοστιβάδας και μεταφέρουν διάφορους τύπους μορίων. Κάποια από αυτά τα μόρια είναι: Πρωτεΐνες, νουκλεϊκά οξέα (miRNAs, μακρά μη κωδικοποιητικά RNAs, άθικτα και μεταλλαγμένα mRNAs και θραύσματα DNA), λιπίδια αλλά και άλλους μεταβολίτες.

Η σύνθεση του εξωσωμικού φορτίου υπόκειται σε ρύθμιση από διάφορα σηματοδοτικά μονοπάτια και εξωκυτταρικές συνθήκες. Επιπλέον, η σύνθεση είναι ικανή να μεταβάλλεται σε απόκριση σε διάφορα ερεθίσματα και κυτταρικά σήματα.

Η βιογένεση των εξωσωμάτων πραγματοποιείται σε δύο φάσεις. Στην πρώτη φάση, λαμβάνει χώρα η αναδίπλωση της κυτταρικής μεμβράνης προς τα μέσα, σχηματίζοντας το πρώιμο ενδόσωμα διαλογής (Early Sorting Endosome) (ESE), στο οποίο ενσωματώνονται πρωτεΐνες της κυτταρικής επιφάνειας, εξωκυττάρια διαλυτές πρωτεΐνες και άλλα μικρομόρια. Στη συνέχεια, τα ESEs ωριμάζουν και μετατρέπονται σε LSEs (Late Sorting Endosomes). Η δεύτερη φάση περιλαμβάνει την εσωτερική αναδίπλωση της εξωτερικής μεμβράνης των ενδοσωμάτων και την ενσωμάτωση σε αυτά συστατικά από το ενδοπλασματικό δίκτυο, το δίκτυο trans Golgi (TGN), καθώς και μικρά βιομόρια όπως νουκλεϊκά οξέα. Αυτή η διαδικασία οδηγεί στο σχηματισμό πολυκυστιδιακών σωμάτων (Multivesicular Bodies, MVBs). Στη συνέχεια, τα ενδοκυττάρια αυτά κυστίδια απελευθερώνονται σταδιακά ως εξωσώματα στο μικροπεριβάλλον μέσω της σύντηξης των MVBs με την πλασματική μεμβράνη.



Εικόνα 8. α) Βιογένεση εξωσωμάτων β) Δομή εξωσώματος

Nail, H.M., Chiu, CC., Leung, CH. et al. Exosomal miRNA-mediated intercellular communications and immunomodulatory effects in tumor microenvironments. *J Biomed Sci* 30, 69 (2023). <https://doi.org/10.1186/s12929-023-00964-w> [34]

Πολλές είναι οι μελέτες που έχουν δείξει ότι τα εξωκυττάρια miRNAs μπορούν να ασκούν βιολογικές μεταβολές στα κύτταρα στόχους, ρυθμίζοντας τις λειτουργίες τους και δρώντας έτσι ως διακυτταρικά σηματοδοτικά μόρια.

Κάποιες από τις συνηθέστερες μεθόδους απομόνωσης εξωσωμάτων είναι η διαφορική υπερφυγοκέντρωση (Differential Ultracentrifugation), η απομόνωση βάσει μεγέθους (Size-Based Isolation) και η κατακρήμνιση με βάση τα πολυμερή (Polymer-Based Precipitation). Η επιλογή της μεθόδου εξαρτάται από τον όγκο και την καθαρότητα που απαιτείται [35] [23] [36] [37] [38] [39].

2.6 miRNAs και Παράσιτα

Τα παράσιτα, στη διαδικασία της συν-εξέλιξής τους με τους ξενιστές έχουν αναπτύξει διάφορους μηχανισμούς για την αποφυγή και ρύθμιση των μηχανισμών άμυνας του ξενιστή τους προκειμένου να εγκαθιδρύσουν χρόνιες λοιμώξεις. Τα miRNAs αποτελούν ιδιαίτερα χρήσιμα εργαλεία για τα παράσιτα καθώς τους επιτρέπουν να τροποποιούν την γονιδιακή έκφραση των κυττάρων του ξενιστή. Η τροποποίηση αυτή μπορεί να περιλαμβάνει τη μετα-μεταγραφική ρύθμιση σε γονίδια που εμπλέκονται στις φλεγμονώδεις και ανοσολογικές αποκρίσεις, τον κυτταρικό κύκλο, την απόπτωση, την αυτοφαγία και την αναδιοργάνωση του κυτταροσκελετού. Η απορρύθμιση της έκφρασης των miRNA του ξενιστή έχει συσχετισθεί με εξασθενημένη ανοσοποιητική απόκριση και αύξηση του αποικισμού του παρασίτου.

Ιδιαίτερο ενδιαφέρον παρουσιάζουν τα miRNAs που μεταφέρονται μέσω εξωκυττάρων κυστιδίων καθώς μπορούν να είναι διαμεσολαβητές της επικοινωνίας μεταξύ παρασίτου και ξενιστή. Αυτά μπορούν να επηρεάσουν την ανοσολογική απόκριση και τη συνολική πορεία της λοίμωξης επιτελώντας τη συμβατική τους λειτουργία ή δρώντας ως συνδέτης για τους υποδοχείς τύπου toll (TLRs) που υπάρχουν στην επιφάνεια των κυττάρων του ανοσοποιητικού συστήματος.

Η μελέτη των miRNAs στις παρασιτικές ασθένειες, πέρα από τη χρήση τους ως διαγνωστικά και προγνωστικά εργαλεία, ανοίγει προοπτικές για νέες θεραπευτικές στρατηγικές. Τα miRNAs μπορούν να χρησιμοποιηθούν ως στόχοι για καινοτόμες προσεγγίσεις στη χημειοθεραπεία και την ανοσοθεραπεία [24].

2.7 miRNA και λεισμάνια

Τα παράσιτα του γένους λεισμάνια έχουν αναπτυχθεί με ιδιαίτερη ικανότητα να επηρεάζουν τα συστήματα των ξενιστών τους μέσω της τροποποίησης της έκφρασης των miRNA. Η απορρύθμιση στην έκφραση των miRNA που προκαλείται από τα παράσιτα μπορεί να εξασθενήσουν τις ανοσολογικές άμυνες του ξενιστή, επιτρέποντας έτσι στο παράσιτο να αποφεύγει την ανοσοποιητική απόκριση του οργανισμού. Κατά τη διάρκεια των αλληλεπιδράσεων μεταξύ ξενιστή και παρασίτου, τόσο τα miRNA του παρασίτου, όσο και του ξενιστή, συμβάλλουν στον καθορισμό της πιθανότητας μόλυνσης και στην εξέλιξη της νόσου. Μέσω αυτών των αλλαγών, το παράσιτο δημιουργεί ένα ευνοϊκό μικροπεριβάλλον για την ανάπτυξή του, αποφεύγοντας το ανοσοποιητικό σύστημα και υποστηρίζοντας την παραμονή του μέσα στον ξενιστή [40] [24] [41].

2.8 Μέθοδοι ανίχνευσης και ποσοτικοποίησης των miRNA

Η ανίχνευση και ποσοτικοποίηση των miRNA είναι απαραίτητη για την μελέτη της έκφρασης τους σε διάφορα βιολογικά συστήματα και απαιτεί την χρήση αξιόπιστων και ευαίσθητων μεθόδων.

2.8.1 Northern Blot

Το Northern Blot αποτελεί την πιο καθιερωμένη μέθοδο ανίχνευσης miRNA καθώς είναι ικανό να ανιχνεύσει ώριμα αλλά και precursor miRNA. Μπορεί να προσδιορίσει το μέγεθος των μορίων miRNA καθώς και τα επίπεδα έκφρασης τους. Η βασική αρχή της μεθόδου είναι :

- 1) Η πέψη του δείγματος RNA με μία περιοριστική ενδονουκλεάση
- 2) Διαχωρισμός του δείγματος σε πηκτή αγαρόζης
- 3) Αποδιάταξη των μορίων RNA και μεταφορά τους σε μεμβράνη νιτροκυτταρίνης ή νάιλον, σύμφωνα με την θέση τους στο πήκτωμα, όπου και σταθεροποιούνται
- 4) Ακολουθεί αντίδραση με ανιχνευτές που είναι επισημασμένοι με ραδιοϊσότοπα ή άλλους δείκτες
- 5) Μετά την έκπλυση των μη δεσμευμένων ανιχνευτών, τα miRNAs μπορούν να ανιχνευθούν χρησιμοποιώντας τις κατάλληλες τεχνικές

Αποτελεί μια συγκριτικά φθηνή μέθοδο που δεν απαιτεί εξειδικευμένο εξοπλισμό αλλά παρόλα αυτά είναι χρονοβόρα με σχετικά χαμηλή απόδοση και ευαισθησία [42] [43].

2.8.2 Real-time qPCR

Η ανίχνευση των προϊόντων PCR σε πραγματικό χρόνο καθίσταται δυνατή με τη συμμετοχή ενός φθορίζοντος μορίου αναφοράς σε κάθε αντίδραση, το οποίο αποδίδει αυξημένο φθορισμό με την αύξηση της ποσότητας του προϊόντος DNA. Ο μετρούμενος φθορισμός είναι ανάλογος της συνολικής ποσότητας του αμπλικονίου, και η μεταβολή του φθορισμού με την πάροδο του χρόνου χρησιμοποιείται για τον υπολογισμό της ποσότητας του παραχθέντος αμπλικονίου σε κάθε κύκλο.

PCR μέθοδοι για τη μελέτη των miRNAs παρέχουν ποσοτικά δεδομένα με υψηλή απόδοση, εξειδίκευση και ευαισθησία, επιτρέποντας τη γρήγορη ανίχνευση διαφορών στην έκφραση RNA μεταξύ δειγμάτων. Η ανίχνευση των miRNAs μέσω PCR αποτελεί πρόκληση καθώς πρόκειται για μόρια με πολύ μικρό μήκος αλληλουχίας.

Το miRNA στόχος πρώτα μετατρέπεται σε συμπληρωματικό DNA (cDNA) με την μέθοδο της αντίστροφης μεταγραφής. Στην σύνθεση του cDNA με αντίστροφη μεταγραφή συνήθως χρησιμοποιείται εκκινητής stem-loop, poly(T) αντάπτορας ή εκκινητής ειδικού γονιδίου (GSP) ενώ χρησιμοποιούνται 2 μέθοδοι φθορισμού για τον ποσοτικό έλεγχο της PCR: χρήση ανιχνευτή υδρόλυσης (τύπου TaqMan) και η χρήση της χρώσης SYBR green.

Το SYBR Green αποτελεί ένα από την πιο κοινώς χρησιμοποιούμενη χρώση σε εφαρμογές RT-qPCR. Εμφανίζει χαμηλό φθορισμό στο διάλυμα αλλά μόλις προσδεθεί στην διπλή έλικα DNA, ο φθορισμός αυξάνεται καθιστώντας το ιδανικό για την ανίχνευση της συγκέντρωσης του DNA. Ωστόσο, μπορεί να παρουσιάσει ανακρίβειες από πιθανή επιμόλυνση ή δημιουργίας διμερών από τους εκκινητές.

Οι ειδικοί ανιχνευτές Taqman μπορούν να χρησιμοποιηθούν για τον εντοπισμό του ρυθμού αύξησης του DNA. Ο ανιχνευτής αυτός είναι ένα ολιγονουκλεοτιδικό μόριο, συμπληρωματικός με μια αλυσίδα DNA του μορίου στόχου και φέρει ένα φθοριόφορο και έναν αναστολέα στα δύο άκρα του. Κατά την ενίσχυση, ο ανιχνευτής απομακρύνεται από την αλυσίδα DNA, οδηγώντας σε διάσπαση του φθοριόφορου και αύξηση του φθορισμού.

Η real-time qPCR αποτελεί μια μέθοδο που δεν απαιτεί εξειδικευμένο εξοπλισμό, ωστόσο τόσο η μειωμένη ακρίβεια της στην αναγνώριση miRNA σε χαμηλότερη έκφραση, όσο και η δυσκολία σχεδιασμού των εκκινητών, περιορίζουν τη χρήση της [42] [43].

2.8.3 Τεχνολογία μικροσυστοιχιών (Microarray)

Η τεχνολογία μικροσυστοιχιών αποτελεί ένα ισχυρό εργαλείο για την μέτρηση των αλλαγών στο προφίλ έκφρασης των miRNA. Στη μέθοδο αυτή, κομμάτια γνωστών νουκλεϊκών οξέων είναι προσκολλημένα σε στερεή επιφάνεια η οποία ονομάζεται Chip. Κατά τη διάρκεια της αντίστροφης μεταγραφής, τα cDNA μόρια του δείγματος επισημαίνονται ταυτόχρονα με ετικέτες, όπως βιοτίνη ή φθοριοφόρα. Τα επισημασμένα cDNA μόρια υβριδοποιούνται με συμπληρωματικές αλληλουχίες που είναι σταθεροποιημένες στην επιφάνεια του chip, οι οποίες αντιστοιχούν σε συγκεκριμένα miRNA στόχους. Στη συνέχεια, η φθορίζουσα ένδειξη που παράγεται ανιχνεύεται από ειδικό εξοπλισμό, επιτρέποντας την ποσοτική ανάλυση των miRNA.

Παρόλης της μεγάλης αποδοτικότητας της τεχνικής αυτής, το κόστος της είναι αρκετά υψηλό. Η μέθοδος επίσης αντιμετωπίζει δυσκολίες στην ανάλυση των miRNA με παρόμοιες αλληλουχίες, καθώς και σε miRNA που είναι πολύ μικρά ή miRNA με μικρό αριθμό αντιγράφων [42] [43].

2.8.4 Αλληλούχιση Επόμενης Γενιάς (NGS)

Το NGS (Next Generation Sequencing) είναι μία τεχνική που επιτρέπει όχι μόνο τη δημιουργία προφίλ έκφρασης για γνωστά μόρια miRNA, αλλά και την ανίχνευση νέων μορίων miRNA. Οι χρήσεις του NGS στην αλληλούχιση των miRNA έχουν εξελιχθεί ραγδαία τα τελευταία χρόνια. Αρχικά, για την εφαρμογή της μεθόδου small RNA sequencing (small RNA-seq), δημιουργείται μία RNA βιβλιοθήκη στην οποία πραγματοποιείται εμπλουτισμός των μικρών RNA, ενώ ταυτόχρονα τα μόρια ενώνονται με αντάπτορες στα 5' και 3' άκρα τους. Το συνολικό τελικό μικρό cDNA δημιουργείται μέσω της αντίστροφης μεταγραφής και ακολουθούν τα κομμάτια της ενίσχυσης και της αλληλούχισης από το μηχάνημα NGS. Η μαζική παράλληλη αλληλούχιση πραγματοποιείται μετά την φόρτωση της cDNA βιβλιοθήκης σε μήτρες αλληλούχισης του εκάστοτε μηχανήματός NGS. Διαφορετικά μηχανήματα διαθέτουν και διαφορετικές μήτρες αλληλούχισης. Για παράδειγμα, το NGS sequencer της Illumina χρησιμοποιεί κυψέλες ροής, ενώ το NGS sequencer της Ion Torrent

χρησιμοποιεί chips αλληλούχισης. Ωστόσο, ο στόχος τους είναι ο ίδιος, δηλαδή να επιτρέπουν τη μαζική παράλληλη αλληλούχιση όλων των cDNA μορίων του δείγματος. Το βάθος της αλληλούχισης miRNA που παρέχει το NGS επιτρέπει την ταυτοποίηση νέων εξωκυτταρικών μικρών RNAs από το αίμα και το εγκεφαλονωτιαίο υγρό, ανοίγοντας νέους δρόμους στην έρευνα των miRNA.

Πέρα από την ιδιαίτερα προσεκτική προετοιμασία των δειγμάτων και του εξειδικευμένου εξοπλισμού που χρησιμοποιείται για την απόκτηση των NGS δεδομένων, λόγω του μεγάλου όγκου και της πολυπλοκότητάς τους, απαιτείται και η κατάλληλη υπολογιστική υποδομή για την ανάλυσή τους με εργαλεία βιοπληροφορικής.

Παρόλο που το NGS είναι πιο οικονομικό στις μέρες μας, παραμένει μια ακριβή μέθοδος που απαιτεί εκτεταμένη βιοπληροφορική ανάλυση, γεγονός που περιορίζει τις εφαρμογές του. Παρ' όλα αυτά, η αμερόληπτη απόκτηση δεδομένων και η εις βάθος αλληλούχιση, καθιστούν το NGS απaráμιλλη μέθοδο ανάλυσης miRNA και τη μοναδική που δίνει την δυνατότητα ταυτοποίησης νέων miRNA [42] [44].

3. ΠΕΙΡΑΜΑΤΙΚΟ ΜΕΡΟΣ

3.1 Σκοπός

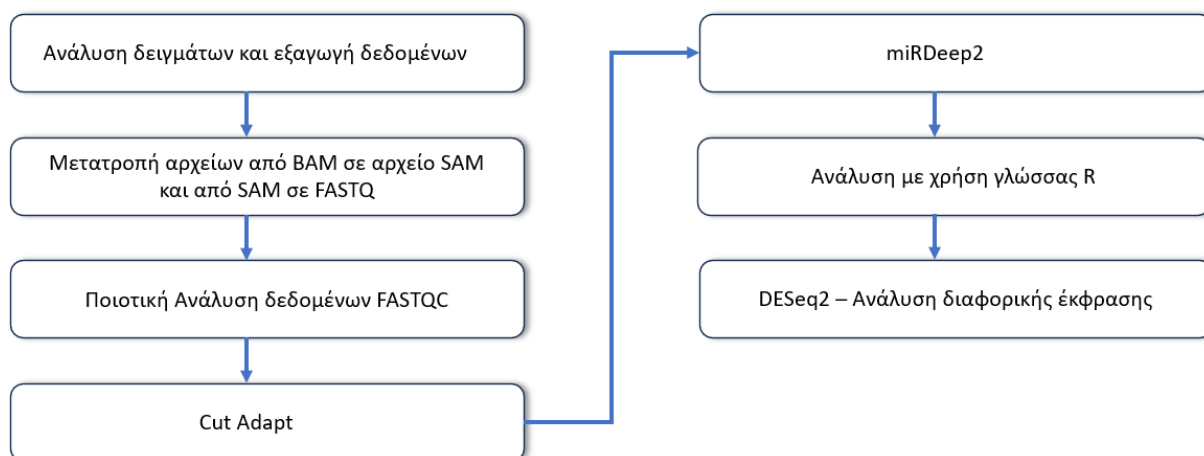
Ο σκοπός της παρούσας μελέτης είναι η διερεύνηση της διαφορικής έκφρασης των microRNA (miRNAs) που εντοπίζονται στα εξωσώματα του ορού σκύλων, με έμφαση στη σύγκριση μεταξύ μολυσμένων και μη μολυσμένων ζώων από *Leishmania infantum*.

Μέσω ανάλυσης ακατέργαστων NGS δεδομένων με τη χρήση βιοπληροφορικών εργαλείων και μεθόδων, επιδιώκεται η ταυτοποίηση miRNAs με διαφοροποιημένη έκφραση μεταξύ των δύο ομάδων, με ορισμένα εξ αυτών να παρουσιάζουν ανωρρύθμιση στους μολυσμένους σκύλους και άλλα κατωρρύθμιση, γεγονός που μπορεί να υποδηλώνει τον πιθανό ρόλο τους στην παθογένεια της λοίμωξης. Ιδιαίτερο ενδιαφέρον παρουσιάζουν miRNAs που ίσως σχετίζονται με ανοσολογικές αποκρίσεις και μονοπάτια φλεγμονής, καθώς θα μπορούσαν να αποτελέσουν πιθανούς βιοδείκτες της μόλυνσης ή ρυθμιστικούς παράγοντες της αλληλεπίδρασης ξενιστή-παθογόνου.

Τα αποτελέσματα της μελέτης ενδέχεται να συμβάλουν στην πληρέστερη κατανόηση του ρόλου των εξωσωμάτων στη μόλυνση από *Leishmania infantum*, προσφέροντας νέα δεδομένα για τη δυναμική των miRNAs ως μοριακών δεικτών ή πιθανών στόχων θεραπευτικής παρέμβασης.

3.2 Μεθοδολογία ανάλυσης ακατέργαστων δεδομένων αλληλούχισης μορίων miRNA

Σχηματικά η μεθοδολογία και τα βήματα της ανάλυσης αποτυπώνονται ως εξής:



Εικόνα 9. Διάγραμμα ροής μεθοδολογίας διαφορικής ανάλυσης miRNA

3.2.1 Ανάλυση δειγμάτων και εξαγωγή δεδομένων - Ion GeneStudio S5

Για την ανάλυση και εξαγωγή των ακατέργαστων δεδομένων χρησιμοποιείται το σύστημα αλληλούχισης νέας γενιάς (NGS). Στη συγκεκριμένη περίπτωση χρησιμοποιήθηκε το Ion GeneStudio S5 της εταιρίας Thermo Fisher Scientific, ένα ευέλικτο και προηγμένο σύστημα αλληλούχισης νέας γενιάς (NGS) που χρησιμοποιεί τεχνολογία ημιαγωγών για την ακριβή ανίχνευση και ανάλυση νουκλεϊκών οξέων. Η τεχνολογία αυτή χρησιμοποιεί την αρχή της ανίχνευσης της απελευθέρωσης ιόντων υδρογόνου κατά την ενσωμάτωση νέου νουκλεοτιδίου στο αναπτυσσόμενο πρότυπο DNA. Η ανίχνευση αυτή πραγματοποιείται με την χρήση ενός μικροτσίπ, το οποίο περιέχει εκατομμύρια πηγαδάκια (wells), που στο κάθε ένα από αυτά υπάρχει ένα κομμάτι DNA από την βιβλιοθήκη που έχει δημιουργηθεί. Όταν προστίθεται το κατάλληλο νουκλεοτίδιο, το σύστημα ανιχνεύει την αλλαγή στο ιοντικό φορτίο που προκαλείται από την απελευθέρωση του ιόντος υδρογόνου και αυτή η αλλαγή μεταφράζεται σε αλληλουχία DNA. Εάν υπάρχουν δύο ίδιες βάσεις σε σειρά στην αλυσίδα του DNA τότε είναι και διπλάσια η αλλαγή του φορτίου που θα πραγματοποιηθεί, οπότε το μικροτσίπ θα ανιχνεύσει δύο ίδιες βάσεις.

Η αλληλουχία (sequence) αναπαρίσταται επίσημα ως μια διάταξη από νουκλεοτίδια (βάσεις), που λαμβάνεται μέσω τεχνολογίας αλληλούχισης και είναι γνωστή ως read (ανάγνωση). Το read αποτελεί το θεμέλιο της αλληλούχισης νέας γενιάς, και η διαδικασία αλληλούχισης περιλαμβάνει συνήθως τη δημιουργία, συναρμολόγηση και ανάλυση των reads.

Η διαδικασία είναι ταχύτερη σε σύγκριση με την παραδοσιακή αλληλούχιση, καθώς παρακάμπτονται τα βήματα που περιλαμβάνουν σήμανση με φθορισμό και ανίχνευση μέσω λέιζερ, παρέχοντας ταυτόχρονα εξαιρετική ακρίβεια [45] [46].

3.2.2 Αρχεία SAM/BAM

Τα αρχεία SAM (Sequence Alignment Map) και BAM (Binary Alignment Map) είναι κρίσιμες μορφές στην βιοπληροφορική για την αποθήκευση ευθυγραμμισμένων δεδομένων από NGS. Τα αρχεία αυτά δίνουν την δυνατότητα χαρτογράφησης των ακατέργαστων δεδομένων αλληλούχισης σε ένα γονιδιωμα αναφοράς, διευκολύνοντας διάφορες αναλύσεις, όπως η σύγκριση γονιδιακής έκφρασης.

Αρχεία SAM: Τα αρχεία SAM (Sequence Alignment Map) είναι ένας τύπος αρχείου κειμένου που περιέχει πληροφορίες ευθυγράμμισης διαφόρων αλληλουχιών που αντιστοιχίζονται σε αλληλουχίες αναφοράς. Τα αρχεία αυτά μπορεί επίσης να περιέχουν μη αντιστοιχισμένες αλληλουχίες. Δεδομένου ότι τα αρχεία SAM είναι μορφής αρχείου κειμένου, είναι πιο ευανάγνωστα από τον άνθρωπο.

Αρχεία BAM: Τα αρχεία BAM (Binary Alignment Map) βρίσκονται σε δυαδική μορφή και δεν μπορούν να διαβαστούν από τον άνθρωπο. Περιέχει τις ίδιες πληροφορίες με τα αρχεία SAM αλλά είναι μικρότερα σε μέγεθος και πιο αποδοτικά για υπολογιστικές διαδικασίες. Τα αρχεία BAM υποστηρίζονται ευρέως από πολλά

εργαλεία βιοπληροφορικής, όπως τα SAMtools, IGV (Integrative Genomics Viewer) και GATK (Genome Analysis Toolkit), καθιστώντας εύκολη την οπτικοποίηση, ανάλυση και χειρισμό των δεδομένων αλληλουχίας [47].

3.2.3 Αρχεία FASTA/FASTQ

Αρχεία FASTA: Το αρχείο FASTA εφευρέθηκε από τον Bill Pearson και είναι μια μορφή κειμένου για την αναπαράσταση είτε νουκλεοτιδικών είτε πεπτιδικών ακολουθιών, όπου τα νουκλεοτίδια ή τα αμινοξέα αναπαρίστανται με κωδικούς ενός γράμματος. Συνήθως το αρχείο ξεκινάει με το σύμβολο ">". Η πρώτη γραμμή αποτελεί τη γραμμή περιγραφής, η οποία μπορεί να περιέχει περιγραφικές πληροφορίες για την ακολουθία. Η περιγραφή μπορεί να είναι το αναγνωριστικό ή το όνομα της αλληλουχίας όπως ονόματα γονιδίων.

```
>AB819629.1 Canis lupus familiaris TNFA gene for tumor necrosis factor-alpha, partial cds,
allele: TNFA-Beagle3
AGGCCGCAGACCTGAACCTCAGGAAAAAATGGAGGGAATGGGAATTGGGGGCAGAGGGAACCTGGGGTC
CAGTCTCCAGGGTCTACACAATGTACACAGCAGTGAAGTGGCCAGGAAGTCTCCCTCAGAATCCGAGGG
GGGCGAGTAGGAAGTATCCTTGATGCCTGGGTGTCCCAACTTTCCAAACCCCGCCCCACGATGGAGA
AGAAACCGAGACCGAAGGTGTAGGGCCCCACCGCTTCTCCAGATGAGCTCATGGTTTCTCCACCAAGG
AAGTTTTCCGCTGCTTGATTGAGAGCCTTTCCCGCCGCCCTCTCTGCCCCGCGGGTATATAAATGCAG
CTGTTTGACACCCAGCCAGCAGGAGCTCCCTCAGCGAGGACACAGGGGACCAGCGGGGAGAGACACG
CAACTCCAGAGGCCCTGGAAATAACTTCTCAGACGACGCCCCCAGACAAGCAGCCAGGCGATTCTCT
CCCTCTAATATACAAGCCCCAGGGCTCTACCGTCTCCCCACTGGACTTGAGCCCCCTCTCAGAAGCACAC
CATGAGCACTGAAGCATGATCCGGGACGTGGAGCTGGCCGAGGAGCGCTCCCAAGAAAGCGGGGGG
CCCCCGGGCTCCAGAAGGTGCTTGTGCTCAGCCTCTTCTCCTTCTCTCTGTCGAGGGGCCACACAC
TCTTGTGCTGCTGCACTTTGGAGTGATCGGGCCCCAGAGGGAAGAGGTGAGTGCCCCACGGAGGACTTG
GCTCATTTCTCCGCCAGGGAGGGTGGTGGTGTGGTGGGGACCATTGAGAGAGGAAATGTGGAGAAAGA
TGGGGAGGCAGAAGGAGATATGGGGAGAGATGGAAGAGGGAGAGCGAGGAGCACATGGAAGATGCTCATG
GATTGCTTCTGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAGCACGCAGACAGAC
AT
```

Εικόνα 10. Παράδειγμα αρχείου FASTA από το NCBI (National Center for Biotechnology Information)

Αρχεία FASTQ:

Το αρχείο FASTQ είναι μια μορφή αρχείου κειμένου που χρησιμοποιείται στη βιοπληροφορική για την αποθήκευση τόσο των αλληλουχιών DNA όσο και των αντίστοιχων βαθμών ποιότητας για κάθε βάση, που προκύπτουν από τεχνολογίες αλληλούχισης υψηλής απόδοσης.

Η μορφή του αρχείου FASTQ περιλαμβάνει τέσσερις τύπους γραμμών. Πρώτον, μια γραμμή τίτλου που ξεκινά με το σύμβολο '@' και συχνά περιέχει ένα μοναδικό αναγνωριστικό. Δεύτερον, η γραμμή της αλληλουχίας DNA ή RNA, η οποία, όπως και στη μορφή FASTA, περιέχει την αλληλουχία νουκλεοτιδίων. Τρίτον, ακολουθεί μια γραμμή που ξεκινά με το σύμβολο '+' και λειτουργεί ως διαχωριστικό ανάμεσα στην αλληλουχία και τη συμβολοσειρά ποιότητας. Τέλος, υπάρχει η γραμμή ποιότητας, η οποία περιέχει τιμές ποιότητας που αναπαρίστανται με χαρακτήρες ASCII [48].

```

@19TQL:07992:12925
AGCCCAGGA
+
<<<<7<<9>
@19TQL:07992:12927
TCCCCCGTCGCT
+
; ; ; ; 0<>==>>
@19TQL:07993:12901
CGCGACCTCAGATCAGACG
+
===@?9?==<<====>>==
@19TQL:07993:12920
TCCTGTAAGTGGTGGCCCGA
+
==9==>====<<<8888>8=
@19TQL:07994:12883
GGCTGGTCCGATGGTAGTGGTTATCAGAAC
+
=5<<<9==9<<<<8=<=<==7=6<<; ;=@9=

```

Εικόνα 11. Παράδειγμα αρχείου FASTQ που περιέχει δεδομένα αλληλούχισης μικρών RNA (small RNA-seq), το οποίο εμφανίζει τις ακολουθίες και τις αντίστοιχες τιμές ποιότητας.

3.2.4 Ποιοτικός έλεγχος των δεδομένων αλληλούχισης, FastQC

Ο έλεγχος ποιότητας στα αποτελέσματα της αλληλούχισης είναι απαραίτητος, καθώς βοηθά στον εντοπισμό προβλημάτων που μπορεί να επηρεάσουν την αξιοπιστία των δεδομένων. Προβλήματα όπως reads χαμηλής ποιότητας ή μόλυνση από αντάπτορες (adaptors), μπορούν να επηρεάσουν τα τελικά αποτελέσματα της ανάλυσης. Ο ποιοτικός έλεγχος επιτρέπει τη λήψη αποφάσεων σχετικά με το αν απαιτείται περικοπή ή φιλτράρισμα των δεδομένων, εξασφαλίζοντας έτσι την ακρίβεια και την αξιοπιστία των τελικών συμπερασμάτων.

Για την εφαρμογή του ελέγχου ποιότητας χρησιμοποιήθηκε το εργαλείο βιοπληροφορικής FastQC (έκδοση 0.12.1) μέσω του Galaxy (έκδοση 24.1.3). Το Galaxy είναι μια διαδικτυακή πλατφόρμα ανοιχτού κώδικα για βιοπληροφορικές αναλύσεις, παρέχοντας ένα ευρύ φάσμα εργαλείων στον χρήστη.

3.2.4.1 Μετατροπή αρχείου BAM σε FASTQ

Επειδή τα δεδομένα από το Ion GeneStudio S5 αποθηκεύονται σε μορφή αρχείου BAM, αυτά τα αρχεία πρέπει πρώτα να μετατραπούν σε αρχεία FastQ, τα οποία είναι κατάλληλα για εισαγωγή στο FastQC. Για αυτή τη μετατροπή χρησιμοποιήθηκαν μέσω του Galaxy τα εργαλεία: “BAM to SAM” (για τη μετατροπή του BAM αρχείου σε SAM) και “SamToFastq” (για τη μετατροπή του SAM αρχείου σε FASTQ).

Τα εργαλεία εκτελέστηκαν με τις προεπιλεγμένες ρυθμίσεις τους, εκτός από το εργαλείο “SamToFastq” όπου στην επιλογή “Output format” επιλέχθηκε το “Single-end” καθώς τα δεδομένα περιέχουν single end reads.

3.2.4.2 Εισαγωγή αρχείου στο FastQC

Τα αρχεία FASTQ εισήχθησαν στο περιβάλλον εργασίας του FastQC και ο ποιοτικός έλεγχος εκτελέστηκε με τις προεπιλεγμένες ρυθμίσεις του εργαλείου.

FastQC Read Quality reports (Galaxy Version 0.74+galaxy1)

Tool Parameters

Raw read data from your current history *

8: D16.fastq

accepted formats ▼

Contaminant list - optional

Nothing selected

accepted formats ▼

tab delimited file with 2 columns: name and sequence. For example: Illumina Small RNA RT Primer CAAGCAGAAGACGGCATACGA

Adapter list - optional

Nothing selected

accepted formats ▼

List of adapters adapter sequences which will be explicitly searched against the library. It should be a tab-delimited file with 2 columns: name and sequence. (--adapters)

Submodule and Limit specifying file - optional

Nothing selected

accepted formats ▼

a file that specifies which submodules are to be executed (default=all) and also specifies the thresholds for the each submodules warning parameter

Disable grouping of bases for reads >50bp

No

Using this option will cause fastqc to crash and burn if you use it on really long reads, and your plots may end up a ridiculous size. You have been warned! (--nogroup)

Lower limit on the length of the sequence to be shown in the report - optional

As long as you set this to a value greater or equal to your longest read length then this will be the sequence length used to create your read groups. This can be useful for making directly comparable statistics from datasets with somewhat variable read lengths. (--min_length)

Length of Kmer to look for *

7

Note: the Kmer test is disabled and needs to be enabled using a custom Submodule and limits file (--kmers)

Εικόνα 12. Εισαγωγή αρχείου FASTQ για το δείγμα D16 στο περιβάλλον εργασίας του FastQC

3.2.4.3 FastQC περιεχόμενα αποτελεσμάτων

Τα αποτελέσματα του FastQC παρουσιάζουν μια σειρά από διαγνωστικές πληροφορίες για την ποιότητα των reads. Αναλυτικά αναγράφονται τα διαγράμματα:

- Ποιότητας ανά βάση (Per base sequence quality), το οποίο είναι ένα διάγραμμα τύπου «boxplot» που δείχνει τη βαθμολογία ποιότητας για κάθε βάση.
- Βαθμολογίας ποιότητας ανά αλληλουχία (Per sequence quality scores), ένα διάγραμμα το οποίο απεικονίζει την κατανομή της βαθμολογίας ποιότητας σε όλες τις αλληλουχίες.
- Περιεχόμενου αλληλουχίας ανά βάση (Per base sequence content), ένα διάγραμμα το οποίο δείχνει τη συχνότητα εμφάνισης των βάσεων (A, T, C, G) σε κάθε θέση.
- Περιεχόμενου N ανά θέση (Per base N content), καταγράφει τη συχνότητα εμφάνισης της βάσης N (άγνωστη βάση) ανά θέση. Στην αλληλούχιση, όταν το μηχάνημα αλληλούχισης δεν μπορεί να αναγνωρίσει με βεβαιότητα μια βάση ως A, T, C ή G, αποδίδει ένα «N» στη θέση αυτή. Υψηλά επίπεδα περιεκτικότητας σε N μπορεί να υποδηλώνουν προβλήματα με την ποιότητα της αλληλούχισης, όπως κακή ποιότητα ανάγνωσης ή σφάλματα στην επιλογή βάσεων.

- Περιεχόμενου GC ανά βάση (Per Base GC Content), αυτή η ενότητα μετρά το περιεχόμενο GC σε όλο το μήκος κάθε αλληλουχίας σε ένα αρχείο και το συγκρίνει με μια μοντελοποιημένη κανονική κατανομή του περιεχομένου GC.
- Κατανομής μήκους ακολουθίας (Sequence length distribution), ένα διάγραμμα στο οποίο παρουσιάζει την κατανομή των μηκών των αλληλουχιών.
- Επιπέδων διπλασιασμού αλληλουχίας (Sequence duplication levels), ένα διάγραμμα το οποίο δείχνει πόσες φορές επαναλαμβάνονται αλληλουχίες στα δεδομένα.
- Υπερεκπροσωπούμενων Αλληλουχιών (Overrepresented Sequences), αναγράφονται αλληλουχίες που εμφανίζονται πολύ συχνά.
- Περιεχόμενου σε Αντάπτορες (Adapter content), ελέγχει το περιεχόμενο ανταπτόρων στα reads, εφόσον έχει προστεθεί στο εργαλείο το αρχείο που τους περιλαμβάνει.

Στην αριστερή πλευρά της αναφοράς του FastQC εμφανίζεται μια περίληψη των ενοτήτων που εκτελέστηκαν και μια γρήγορη αξιολόγηση των αποτελεσμάτων.

Εάν τα αποτελέσματα της ενότητας φαίνονται απολύτως φυσιολογικά, θα εμφανιστεί ένα πράσινο τικ, εάν τα αποτελέσματα είναι ελαφρώς ανώμαλα θα εμφανιστεί προειδοποίηση με πορτοκαλί θαυμαστικό και εάν τα αποτελέσματα είναι πολύ ασυνήθιστα θα εμφανιστεί σφάλμα με κόκκινο "X".

3.2.5 Αντάπτορες και Cutadapt

Οι αντάπτορες είναι σύντομες συνθετικές αλληλουχίες DNA που συνδέονται στα άκρα των τμημάτων DNA που πρόκειται να αλληλουχηθούν. Στην αλληλούχιση Ion Torrent, αφού συνδεθούν οι αντάπτορες, η βιβλιοθήκη φορτώνεται σε σφαιρικά σωματίδια ιόντων (Ion Sphere Particles) (ISPs). Οι αντάπτορες διευκολύνουν τη σύνδεση του DNA σε αυτά τα ISPs, όπου εκτελείται emulsion PCR (ePCR) για την κλωνική ενίσχυση των τμημάτων DNA.

επίσης, οι αντάπτορες συμβάλλουν στη διασφάλιση της σωστής προσκόλλησης των μορίων DNA στο τσιπ αλληλούχισης ημιαγωγών, επιτρέποντας στο σύστημα Ion S5 να ανιχνεύσει την ενσωμάτωση νουκλεοτιδίων κατά την αλληλούχιση [49] .

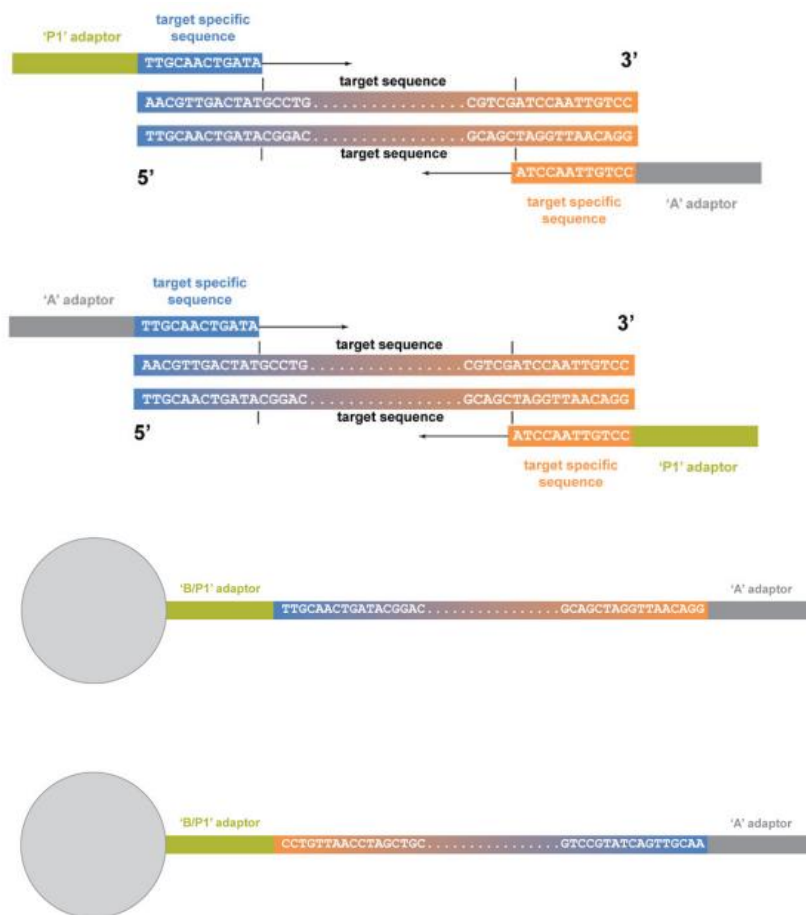
Οι αντάπτορες οι οποίοι χρησιμοποιήθηκαν για την αλληλούχιση των δειγμάτων της παρούσας εργασίας ήταν οι :

>Adapter_a

CCATCTCATCCCTGCGTGTCTCCGACTCAG

>Adapter_b

CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGGTGAT



Εικόνα 13. Σχηματική αναπαράσταση της σύνδεσης του cDNA στα ISPs με την βοήθεια των ανταπτόρων. Ion torrent, Application Note, Amplicon Sequencing, 2011, Life Technologies Corporation.

Καθώς στα δείγματα πραγματοποιείται ταυτόχρονη αλληλούχιση, ειδικοί barcode ανταπτόρες προστέθηκαν για την ταυτοποίηση τους. Οι 5' και 3' barcode ανταπτόρες προστέθηκαν κατά την παρασκευή των βιβλιοθηκών με την βοήθεια μίας RNA λιγάσης και αλληλουχήθηκαν μαζί με το cDNA σε ένα μόνο read.

Η πρώτη απομάκρυνση των ανταπτόρων και των barcode ανταπτόρων πραγματοποιείται από το μηχάνημα Ion GeneStudio S5 με την χρήση του πακέτου λογισμικού Torrent Suite που παρέχεται από την Thermo Fisher Scientific.

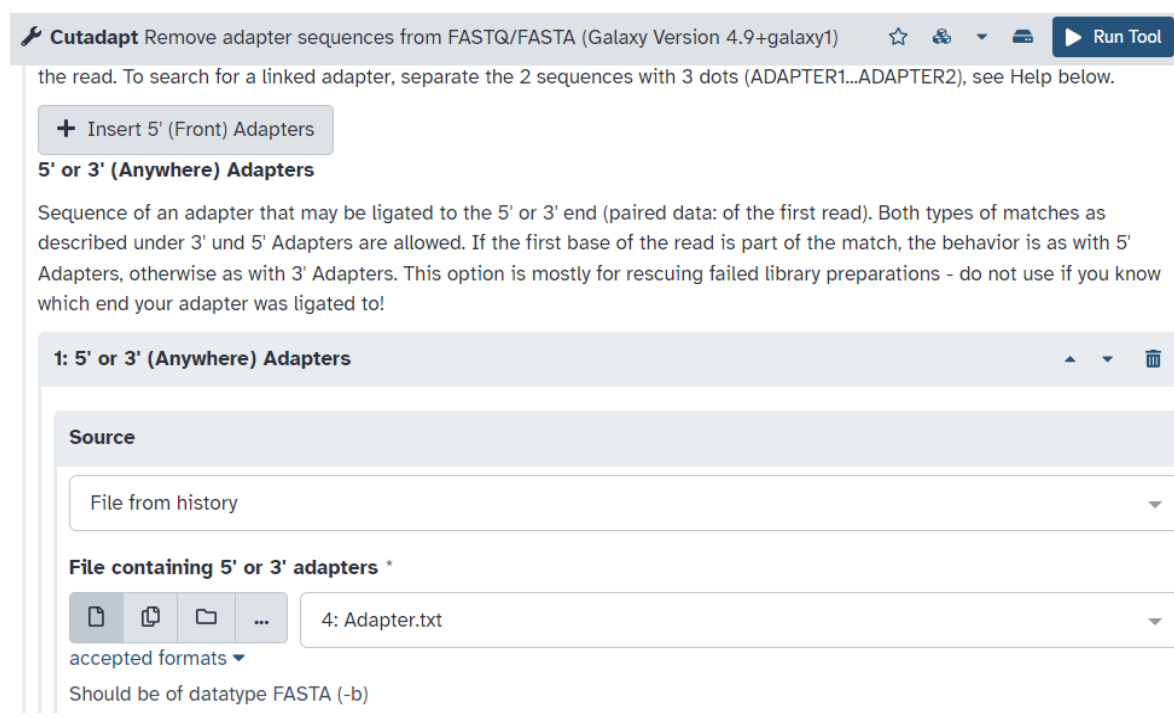
Για τον έλεγχο και την απομάκρυνση τυχών υπολειμμάτων των ανταπτόρων στα δεδομένα, χρησιμοποιήθηκε μέσω του Galaxy (έκδοση 24.1.3) το εργαλείο Cutadapt (έκδοση του Galaxy 4.9+galaxy1). Το Cutadapt βρίσκει και αφαιρεί αλληλουχίες ανταπτόρων, εκκινητές, poly-A ουρές και άλλους τύπους ανεπιθύμητων αλληλουχιών από reads αλληλούχισης υψηλής απόδοσης [50].

Το αρχείο fastq προστέθηκε για κάθε δείγμα ξεχωριστά στο περιβάλλον εργασίας του Cutadapt (Εικόνα 14).



Εικόνα 14. Εισαγωγή αρχείου Fastq στο περιβάλλον εργασίας του Cutadapt (Galaxy)

Οι αλληλουχίες των ανταπτόρων που θα απομακρυνθούν φορτώθηκαν στο ιστορικό του Galaxy σε μορφή αρχείου fasta και προστέθηκαν στο περιβάλλον εργασίας του Cutadapt στην επιλογή: 5' ή 3' Αντάπτορες (5' or 3' (Anywhere) Adapters) (Εικόνα 15).



Εικόνα 15. Προσθήκη ανταπτόρων στο περιβάλλον εργασίας του Cutadapt (Galaxy)

Στις υπόλοιπες επιλογές διατηρήθηκαν οι προεπιλεγμένες ρυθμίσεις του εργαλείου.

3.2.6 Ταυτοποίηση miRNA μέσω του MiRDeep2

Για την ταυτοποίηση των μορίων miRNA στα δείγματα, χρησιμοποιήθηκε μέσω του Galaxy το εργαλείο MiRDeep2. Το εργαλείο αυτό έχει σχεδιαστεί ώστε να μπορεί να επεξεργάζεται ακατέργαστα reads, να εντοπίζει γνωστά και νέα miRNAs και να παρέχει λεπτομερή προφίλ έκφρασης.

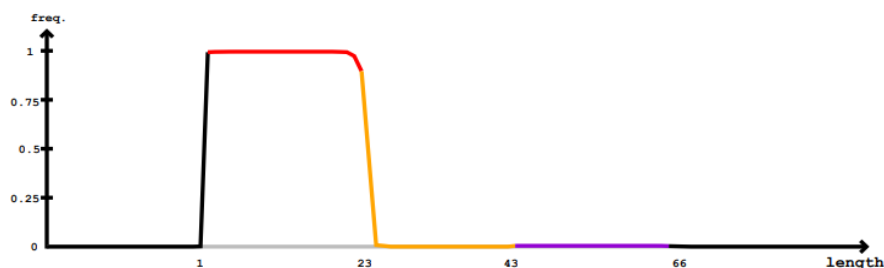
Πιο αναλυτικά κάποιες από τις λειτουργίες του εργαλείου είναι:

- Η ταυτοποίηση γνωστών miRNA, ευθυγραμμίζοντας τα reads με ένα γονιδίωμα αναφοράς και αντιστοιχίζοντας τα με γνωστές αλληλουχίες miRNA από βάσεις δεδομένων όπως η miRBase.
- Η ταυτοποίηση νέων miRNA, αναγνωρίζοντας την χαρακτηριστική δομή φουρκέτας του precursor miRNA και ελέγχοντας για μοτίβα επεξεργασίας του Dicer.
- Η ποσοτικοποίηση των επιπέδων έκφρασης των miRNA, υπολογίζοντας τον αριθμό των ώριμων reads για κάθε miRNA στο εκάστοτε δείγμα.
- Η επικύρωση των miRNAs μέσω σύγκρισης με γνωστές αλληλουχίες, το οποίο εξασφαλίζει την αξιοπιστία των αποτελεσμάτων όσον αφορά την ανίχνευσή τους. Το MiRDeep2 συγκρίνει τα προβλεπόμενα miRNAs με μια βάση δεδομένων αναφοράς (miRBase), για την επικύρωση των ανιχνευόμενων miRNAs.

Το MiRDeep2 βελτιώνει την αξιοπιστία της ανίχνευσης με πρόσθετα μέτρα που βασίζονται στην εντροπία. Η εντροπία είναι το μέτρο για τη μεταβλητότητα και την αξιοπιστία του αριθμού των reads. Έτσι, ο έλεγχος της κατανομής των reads και της ομοιομορφίας της, επιτρέπει στο MiRDeep2 να διακρίνει την πραγματική ένδειξη miRNA από τον θόρυβο. Αυτή η προσέγγιση που ευνοεί τα miRNA με χαμηλή εντροπία, μειώνει το ενδεχόμενο ψευδώς θετικών αποτελεσμάτων και δίνει ένα αξιόπιστο αποτέλεσμα, επιβεβαιώνοντας έτσι τη βιολογική σημασία των ανιχνευμένων miRNAs.

Κάποιοι από τους λόγους επιλογής του MiRDeep2 για την ανάλυση είναι:

- Το φιλικό προς τον χρήστη περιβάλλον εργασίας που παρέχει το Galaxy για το εργαλείο.
- Η απόδοση του εργαλείου όσον αφορά τον χρόνο και την κατανάλωση μνήμης.
- Η ικανότητα διάκρισης των miRNA από άλλα μικρά RNA τα οποία συνδέονται με Argonaute πρωτεΐνες [51].

[illegible]

Η ανάλυση πραγματοποιήθηκε με την χρήση του MiRDeep2 (έκδοση του Galaxy 2.0.1.2+galaxy0) και ακολουθήθηκαν τα εξής βήματα:

- 41

MiRDeep2 Mapper process and map reads to a reference genome (Galaxy Version 2.0.0.8.1)

Tool Parameters

Pool multiple read sets

No

Deep sequencing reads *

5: CA_B8.fastq

accepted formats ▼

Reads in fastq or FASTA format

Remove reads with non-standard nucleotides

☒ No

Remove all entries that have a sequence that contains letters other than a,c,g,t,u,n,A,C,G,T,U,N. (-j)

Convert RNA to DNA alphabet (to map against genome)

☒ No

(-i)

Clip 3' Adapter Sequence

Don't Clip

(-k)

Discard reads shorter than this length *

18

Set to 0 to keep all reads. (-l)

Collapse reads and/or Map

Collapse reads and Map

(-m) and/or (-p)

Will you select a reference genome from your history or use a built-in index?

Use one from the history

Map to genome. (-p)

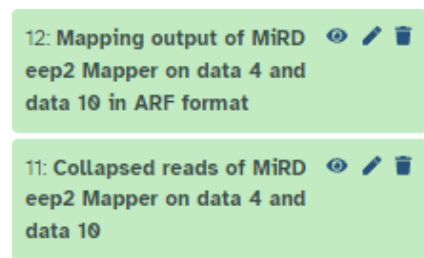
Select the reference genome *

12: canFam4.fa

accepted formats ▼

Εικόνα 17. Περιβάλλον εργασίας MiRdeep2 mapper

Το εργαλείο δημιουργεί 2 αρχεία, το αρχείο με τα collapsed reads που περιέχει μοναδικές αλληλουχίες από τα reads, οι οποίες συνοδεύονται από τον αριθμό εμφάνισης της κάθε ακολουθίας στα αρχικά δεδομένα και το mapping output αρχείο, το οποίο περιέχει πληροφορίες για τον τρόπο και την περιοχή ευθυγράμμισης της κάθε αλληλουχίας με το γονιδίωμα αναφοράς (canFam4).



Εικόνα 18. Απεικόνιση των αρχείων “Mapping output” και “Collapsed reads” που δημιουργούνται στο περιβάλλον εργασίας του Galaxy

Στη συνέχεια τα παραπάνω αρχεία εισάχθηκαν στο εργαλείο MiRDeep2 μαζί με τις αλληλουχίες των ώριμων και precursor miRNA του σκύλου, που αποκτήθηκαν από την βάση δεδομένων miRbase (Εικόνα 19).

- Στην επιλογή collapsed deep sequencing reads προστέθηκε το collapsed reads αρχείο από το MiRDeep2 Mapper,
- όπως και αντίστοιχα στην επιλογή Mappings, το mapping output αρχείο. Στην επιλογή “Genome” προστέθηκε το αρχείο canfam4,
- στην επιλογή “Mature miRNA sequences for this species” προστέθηκε το αρχείο των ώριμων miRNA από την miRBase,
- στην επιλογή “Precursor sequences” προστέθηκε το αρχείο των precursor miRNA από την miRBase
- και στην επιλογή “Search in species” επιλέχθηκε η επιλογή dog (σκύλος).

Το MiRDeep2 δημιουργεί μία html αναφορά η οποία παρέχει μια φιλική προς το χρήστη περίληψη των αποτελεσμάτων των προβλέψεων των miRNA. Συγκεκριμένα, η αναφορά περιέχει απεικονίσεις, όπως διαγράμματα δευτεροταγούς δομής για κάθε υποψήφιο miRNA, γραφήματα συχνότητας και λεπτομερή ευθυγράμμιση αναγνώσεων για κάθε miRNA.

MiRDeep2 identification of novel and known miRNAs (Galaxy Version 2.0.1.2+galaxy0)
☆
📄
▶ Run Tool

Tool Parameters

Collapsed deep sequencing reads *

📄
📄
📄
...
15: Collapsed reads of MiRDeep2 Mapper on data 12 and data 5

accepted formats ▼

Reads in fasta format. The identifier should contain a prefix, a runningnumber and a '_x' to indicate the number of reads that have this sequence. There should be no redundancy in the sequences.

Genome *

📄
📄
📄
...
12: canFam4.fa

accepted formats ▼

Genome contigs in fasta format. The identifiers should be unique.

Mappings *

📄
📄
📄
...
16: Mapping output of MiRDeep2 Mapper on data 12 and data 5 in ARF format

accepted formats ▼

Reads mapped against genome. Mappings should be in ARF format.

Mature miRNA sequences for this species - optional

📄
📄
📄
...
13: mature_miRNA_no_spaces.txt

accepted formats ▼

miRBase miRNA sequences in fasta format. These should be the known mature sequences for the species being analyzed.

Mature miRNA sequences for related species - optional

📄
📄
📄
...
Nothing selected

accepted formats ▼

miRBase miRNA sequences in fasta format. These should be the pooled known mature sequences for 1-5 species closely related to the species being analyzed.

Precursor sequences - optional

📄
📄
📄
...
14: precursor_miRNA_no_spaces.txt

accepted formats ▼

miRBase miRNA precursor sequences in fasta format. These should be the known precursor sequences for the species being analyzed.

Search in species *

dog

If not searching in a specific species all species in your files will be analyzed. (-t)

Star sequences - optional

📄
📄
📄
...
Nothing selected

accepted formats ▼

From miRBase in fasta format (optional) (-s)

Minimum read stack height - optional

minimum read stack height that triggers analysis. Using this option disables automatic estimation of the optimal value and all detected precursors are analyzed. (-a)

Maximum precursors *

50000

Maximum number of precursors to analyze when automatic excision gearing is used. If set to -1 all precursors will be analyzed. (-g).

Minimum miRNA score *

0

Minimum score cut-off for predicted novel miRNAs to be displayed in the overview table. (-b)

Disable randfold analysis

☒ No

(-c)

Detailed fasta output

☒ No

Εικόνα 19. Απεικόνιση περιβάλλοντος εργασίας MiRdeep2

Επίσης αναγράφει πίνακες για γνωστά αλλά και νέα miRNA που περιέχουν σημαντικές πληροφορίες όπως:

- **Provisional ID:** Στον πίνακα των νέων miRNA οι αλληλουχίες δεν ταιριάζουν με κανένα γνωστό miRNA, οπότε το MiRDeep2 παράγει ένα μοναδικό προσωρινό αναγνωριστικό, χαρακτηρίζοντάς το ως νέο ή υποθετικό miRNA για περαιτέρω επικύρωση. Αυτά τα προσωρινά αναγνωριστικά βοηθούν στη διάκριση των νέων υποψηφίων miRNA από τα ήδη γνωστά.
- **Mature miRBase miRNA:** Αυτή η στήλη εμφανίζεται στον πίνακα των γνωστών miRNAs και καταγράφει το miRNA από την miRBase που έχει αντιστοιχιστεί σε κάθε αλληλουχία.
- **MiRDeep2 score:** Το score του MiRDeep2 είναι ένα μέτρο αξιοπιστίας της πρόβλεψη του κάθε miRNA.
- **Significant randfold.p.value:** το randfold p-value είναι ένα στατιστικό μέτρο που χρησιμοποιείται για την αξιολόγηση της σταθερότητας της προβλεπόμενης δομής φουρκέτας ενός πρόδρομου miRNA και εμφανίζει “yes” ή “no”.
"yes": Το προβλεπόμενο precursor miRNA έχει μια στατιστικά σημαντική δομή σύμφωνα με το randfold test. Αυτό σημαίνει ότι το μοτίβο αναδίπλωσης της αλληλουχίας RNA είναι δομής φουρκέτας, τηρεί τα κριτήρια, όπως υψηλό ποσοστό συμπληρωματικότητας βάσεων στην ώριμη περιοχή, και δεν είναι πιθανό να είναι τυχαίο, γεγονός που υποδηλώνει ότι μπορεί να είναι ένα πραγματικό precursor miRNA.
"no": Η δομή δεν πέρασε το τεστ σημαντικότητας randfold, πράγμα που σημαίνει ότι θα μπορούσε να είναι μια τυχαία δομή και όχι ένας βιολογικά σημαντικός precursor του miRNA.
- **Mature Reads:** Αριθμός των reads για κάθε ώριμο miRNA
- **Precursor Coordinate:** Είναι η θέση του precursor ενός miRNA στο γονιδίωμα του σκύλου.

Tag.IDs.MirDeep2	Score.the.log.ods.score.for.ha	IDs_Cfa.miRNAs	Total.Reads	Significant.rand fold.p.value	Mature.Reads	Loop.Reads	Star.Reads	Mature.Star... Total	Mature.sequence	Star.sequence	Precursor.sequ ence	Precursor.coordinate
	irpin											
chr16_7801	11000	cfa-miR-486	21755	yes	21693	0	62	21755	uccuguaucugagucgcccga	cggggcagcucagucaggau	uccuguaucugagucgccc	chr16:24691913..24691976:-
chr16_7446	10000	cfa-miR-486	21272	yes	21137	0	135	21272	uccuguaucugagucgcccga	cggggcagcucagucaggau	uccuguaucugagucgccc	chr16:24691914..24691978:-
chr20_12000	1900	cfa-miR-191	3731	yes	3727	0	4	3731	caacggaaucaccaagcagc	gucgucgucgucgucguc	caacggaaucaccaagcagc	chr20:40551203..40551266:+
chr9_31496	1500	cfa-miR-144	3080	yes	2741	0	339	3080	uacaguuuagagauaguuacu	ggauuucaucauuuaguuag	ggauuucaucauuuaguu	chr9:43002959..43003016:-
chr22_14054	1200	cfa-miR-20a	2537	yes	2534	0	3	2537	uaaagugucuuuagugcaggua	acugcuaauuagagcacuuaa	uaaagugucuuuagugcag	chr22:42617625..42617683:+
chr6_26598	1000	cfa-miR-93	2029	yes	2002	0	27	2029	caaaugucgucucgucaggua	acugcuaauuagagcacuuaa	caaaugucgucucgucagg	chr6:9445017..9445079:+
chr22_13904	1000	cfa-miR-16	1969	yes	1968	0	1	1969	uagcagcacgcuaaauuagggcu	ccaguuuuuacugucugcaga	uagcagcacgcuaaauuag	chr22:1814294..1814358:+
chr34_21928	990	cfa-miR-16	1947	yes	1937	0	10	1947	uagcagcacgcuaaauuagggcu	accuuuuuuuugucugucuu	uagcagcacgcuaaauuag	chr34:26647895..26647958:+
chr9_30705	780	cfa-miR-423a	1532	yes	641	0	891	1532	agcucggucugagggcccccaguu	uagggggcagagagcagaguu	uagggggcagagagcagaga	chr9:44126095..44126154:-
chr9_31615	730	cfa-miR-126	1430	yes	640	0	790	1430	cauuuuuuuuuugggucgucg	ucgucacgucaguuuuuuuag	cauuuuuuuuuugggucac	chr9:48879029..48879088:-
chr9_30624	690	cfa-miR-21	1356	yes	1344	0	12	1356	uagcuaucacagacugauugac	caacagcagucagugggucuu	uagcuaucacagacugauug	chr9:34306260..34306321:+
chr22_14048	530	cfa-miR-106a	1047	yes	998	0	49	1047	caaaugucuuuacagucagggu	acugcuaauuagagcacuuaa	caaaugucuuuacagucagg	chr22:42617176..42617233:+
chr6_26596	490	cfa-miR-106b	960	yes	944	0	16	960	uaaagugucagacagucagau	ccgacuuuuggguaucuguc	uaaagugucagacagucagg	chr6:9444795..9444853:+
chr20_11968	420	cfa-let-7g	825	yes	824	0	1	825	uaggguaaguuuuuuuagaguu	cuguaacgacacugucgucuc	uaggguaaguuuuuuuagac	chr20:37908499..37908577:+
chr14_5899	400	cfa-miR-29a	785	yes	777	0	8	785	uagcaccacuuuagaaucgggu	acuguuuuuuuuuugggucag	acuguuuuuuuuuuggguc	chr14:6013044..6013103:+
chr22_14050	390	cfa-miR-18a	790	no	789	0	1	790	uaaagugucuuuacugucagau	acugcuaauuagggucuuuag	uaaagugucuuuacugucag	chr22:42617314..42617378:+
chr20_12171	380	cfa-miR-27a	756	yes	712	0	44	756	uucacagugggcuagaauucgg	gggcuuagcucuuuugagac	gggcuuagcucuuuugagag	chr20:49019127..49019187:+
chr1_1383	370	cfa-let-7a	741	yes	736	0	5	741	uaggguaaguuuuuuuaguuuu	cuuuuacuuuuuuuuuuuuu	uaggguaaguuuuuuuaguu	chr1:98627157..98627230:-
chr10_2687	370	cfa-let-7a	739	yes	734	0	5	739	uaggguaaguuuuuuuaguuuu	cuuuuacuuuuuuuuuuuuu	uaggguaaguuuuuuuaguu	chr10:20982640..20982709:-
chr1_1381	300	cfa-let-7f	601	yes	597	0	4	601	uaggguaaguuuuuuuaguuuu	cuuuuacuuuuuuuuuuuuu	uaggguaaguuuuuuuaguu	chr1:98626774..98626852:-

Εικόνα 20 MiRDeep2. Παράδειγμα πίνακα αναφοράς γνώστων miRNA από δείγμα εξωσωμάτων σκύλου

Οι πίνακες αναφοράς όλων των δειγμάτων αντιγράφηκαν σε φύλλα Excel και στη συνέχεια μετατράπηκαν σε csv αρχεία για την περαιτέρω ανάλυση τους μέσω της γλώσσας προγραμματισμού R.

3.2.7 Μεθοδολογία ανάλυσης των δεδομένων από το MiRDeep2 με την χρήση της R

Για την ανάλυση και τη στατιστική επεξεργασία των δεδομένων, χρησιμοποιήθηκε η γλώσσα προγραμματισμού R. Η R είναι ένα ισχυρό εργαλείο που προσφέρει πολλές δυνατότητες για στατιστικές αναλύσεις και απεικόνιση δεδομένων. Διατίθεται πληθώρα βιβλιοθηκών και πακέτων που, διευκολύνουν την εκτέλεση πολύπλοκων αναλύσεων με σαφή και αποτελεσματικό τρόπο. Η ανάλυση πραγματοποιήθηκε μέσω του RStudio, το οποίο παρέχει ένα φιλικό προς τον χρήστη περιβάλλον για την εκτέλεση του κώδικα και τη διαχείριση των δεδομένων.



Ο κώδικας που χρησιμοποιήθηκε, καθώς και ο αναλυτικός τρόπος χρήσης των παραπάνω πακέτων της R για την ταυτοποίηση προφίλ έκφρασης των miRNA στα εξωσώματα του ορού 6 σκύλων - 3 εκ των οποίων ήταν μολυσμένοι με το παράσιτο *L. Infantum* (A6, B1, B8) και οι υπόλοιποι 3 (C16, D16, B6) ήταν υγιείς και χρησιμοποιήθηκαν ως ομάδα ελέγχου-, παρατίθεται παρακάτω.

3.2.7.1 Εγκατάσταση πακέτων και φόρτωση αρχείων στην R

Αρχικά εγκαταστάθηκαν στο RStudio μέσω του CRAN (Πλήρες δίκτυο αρχείων R) (Comprehensive R Archive Network) τα πακέτα “ggplot2”, “reshape2” και “dplyr”

Κώδικας στην R: `install.packages(c("ggplot2", "reshape2", "dplyr"))`

Μέσω του εξειδικευμένου αποθετηριακού χώρου Bioconductor εγκαταστάθηκαν τα πακέτα “apeglm” και “DESeq2” με την χρήση του κώδικα στην R:

```
if (!requireNamespace("BiocManager", quietly = TRUE))
  install.packages("BiocManager")
BiocManager::install(c("apeglm", "DESeq2"))
```

Στη συνέχεια τα πακέτα φορτώθηκαν στο περιβάλλον εργασίας του RStudio με τον κώδικα:

```
library(ggplot2)
library(reshape2)
library(dplyr)
library(DESeq2)
```

```
library(apeg1m)
```

- Το ggplot2 είναι ένα πακέτο της R το οποίο δίνει την δυνατότητα δημιουργίας απλών αλλά και πολύπλοκων διαγραμμάτων απεικόνισης δεδομένων.
- Το dplyr είναι ένα πακέτο της R, που χρησιμοποιείται για την διαχείριση δεδομένων διευκολύνοντας την μετατροπή τους και τον καθαρισμό τους.
- Το reshape2 είναι ένα πακέτο της R, που χρησιμοποιείται για την αναδόμηση και την μετατροπή των δεδομένων.
- Το DESeq2 και το apeg1m είναι πακέτα που χρησιμοποιούνται για την ανάλυση διαφορικής έκφρασης, που θα εξεταστούν αργότερα.

Στη συνέχεια, η γραμμή κώδικα “options(scipen = 999)” χρησιμοποιείται για τον ορισμό των επιλογών για την αποφυγή του επιστημονικού συμβολισμού των αριθμών στα αποτελέσματα.

```
options(scipen = 999)
```

Στη συνέχεια φορτώνονται στο RStudio τα csv αρχεία τα οποία δημιουργήθηκαν μετά την ανάλυση MiRDeep2 (κεφ. 3.2.6) και μετατρέπονται σε πλαίσια δεδομένων (data frames), με τον εξής τρόπο:

```
# File paths for the CSV files
file_path_SAMPLE_X <- "filepath"
# Read the CSV files into data frames
data_SAMPLE_X <- read.csv(file_path_SAMPLE_X)
```

όπου “filepath” είναι η συγκεκριμένη θέση αποθήκευσης του κάθε αρχείου csv στον υπολογιστή.

Στην παρούσα εργασία φορτώθηκαν στο RStudio, ως πλαίσια δεδομένων (dataframes), τα αρχεία csv που περιέχουν τα δεδομένα των miRNA από το MiRDeep2 για κάθε ένα από τα δείγματα (A6, B1, B8, C16, D16, B6) με το τρόπο που φαίνεται παρακάτω:

```
# File paths for the CSV files
file_path_A6 <- "C:/Users/kdoun/Documents/Pasteur
Institute/Bioinformatics/miRDeep2/mirdeep2_A6.csv"
file_path_B1 <- "C:/Users/kdoun/Documents/Pasteur
Institute/Bioinformatics/miRDeep2/mirdeep2_B1.csv"
file_path_B8 <- "C:/Users/kdoun/Documents/Pasteur
Institute/Bioinformatics/miRDeep2/mirdeep2_B8.csv"
file_path_C16 <- "C:/Users/kdoun/Documents/Pasteur
Institute/Bioinformatics/miRDeep2/mirdeep2_C16.csv"
```

```

file_path_D16 <- "C:/Users/kdoun/Documents/Pasteur
Institute/Bioinformatics/miRDeep2/mirdeep2_D16.csv"
file_path_B6 <- "C:/Users/kdoun/Documents/Pasteur
Institute/Bioinformatics/miRDeep2/mirdeep2_B6.csv"
# Read the CSV files into data frames
data_A6 <- read.csv(file_path_A6)
data_B1 <- read.csv(file_path_B1)
data_B8 <- read.csv(file_path_B8)
data_C16 <- read.csv(file_path_C16)
data_D16 <- read.csv(file_path_D16)
data_B6 <- read.csv(file_path_B6)

```

Στη συνέχεια, έγινε δημιουργία στήλης που περιέχει τον αριθμό των ώριμων reads χαρακτηρισμένη ξεχωριστά για κάθε δείγμα δηλαδή τα δεδομένα της στήλης Mature.Reads σε κάθε δείγμα SAMPLE_X, αντιγράφονται σε μια στήλη που για λόγους σαφήνειας και ευκολίας στην περαιτέρω διαχείριση ονομάζεται Mature.Reads_SAMPLE_X :

```

# Characterisation of mature reads per sample
data_SAMPLE_X$Mature.Reads_SAMPLE_X <- (data_SAMPLE_X$Mature.Reads)

```

Στην παρούσα εργασία, για κάθε δείγμα, ο κώδικας που χρησιμοποιήθηκε ήταν ο παρακάτω:

```

# Characterisation of mature reads per sample
data_A6$Mature.Reads_A6 <- (data_A6$Mature.Reads)
data_B1$Mature.Reads_B1 <- (data_B1$Mature.Reads)
data_B8$Mature.Reads_B8 <- (data_B8$Mature.Reads)
data_C16$Mature.Reads_C16 <- (data_C16$Mature.Reads)
data_D16$Mature.Reads_D16 <- (data_D16$Mature.Reads)

```

3.2.7.2 Φιλτράρισμα και αφαίρεση του δείγματος B6

Για την εστίαση σε miRNA υψηλής αξιοπιστίας, επιλέχθηκαν για περαιτέρω ανάλυση μόνο τα miRNAs τα οποία παρουσιάζουν Significant randfold.p.value =yes και MiRDeep2 Score > 4 (βλ. στήλες του πίνακα της Εικόνας 20) [52] [51]. Αυτή η επιλογή διασφαλίζει ότι μόνο τα miRNAs με στατιστικά σημαντικές τιμές και καλύτερη ποιότητα εντάσσονται στην ανάλυση.

```

Κώδικας της R για το φιλτράρισμα:# Filter out miRNAs with scores < 4

```

```
data_SAMPLE_X <- data_SAMPLE_X %>%
  filter(Score.the.log.odds.score.for.hairpin > 4,
         Significant.randfold.p.value == "yes")
```

Στην παρούσα εργασία ο κώδικας της R για το φιλτράρισμα των δειγμάτων, είναι:

```
# Filter out miRNAs with scores < 4
data_A6 <- data_A6 %>%
  filter(Score.the.log.odds.score.for.hairpin > 4,
         Significant.randfold.p.value == "yes")
data_B1 <- data_B1 %>%
  filter(Score.the.log.odds.score.for.hairpin > 4,
         Significant.randfold.p.value == "yes")
data_B8 <- data_B8 %>%
  filter(Score.the.log.odds.score.for.hairpin > 4,
         Significant.randfold.p.value == "yes")
data_C16 <- data_C16 %>%
  filter(Score.the.log.odds.score.for.hairpin > 4,
         Significant.randfold.p.value == "yes")
data_D16 <- data_D16 %>%
  filter(Score.the.log.odds.score.for.hairpin > 4,
         Significant.randfold.p.value == "yes")
data_B6 <- data_B6 %>%
  filter(Score.the.log.odds.score.for.hairpin > 4,
         Significant.randfold.p.value == "yes")
```

Το δείγμα ελέγχου B6 όπως είχε παρατηρηθεί και κατά τον ποιοτικό έλεγχο FastQC έχει ένα πολύ μικρό αριθμό αλληλουχιών με τον αριθμό τους να φθάνει μόλις τις 356, ενώ περαιτέρω ανάλυση του δείγματος μέσω του MiRDeep2 έδειξε μόλις 6 γνωστά miRNA με το κάθε ένα να έχει μόλις ένα read και score < 4. Η αφαίρεση αυτού του δείγματος διασφαλίζει ότι η ανάλυση επικεντρώνεται σε δείγματα με επαρκή δεδομένα και υψηλότερη ποιότητα, εξασφαλίζοντας την εγκυρότητα των ευρημάτων.

Στη συνέχεια, για την διευκόλυνση της διαχείρισης των πινάκων δεδομένων, αφαιρέθηκαν ορισμένες στήλες και διατηρήθηκαν εκείνες υψηλότερου ενδιαφέροντος ("IDs_Cfa.miRNAs", "Mature.Reads_SAMPLE_X", "Score.the.log.odds.score.for.hairpin", "Precursor.coordinate"). Η διαδικασία αυτή πραγματοποιήθηκε με τον παρακάτω κώδικα:

```
# Remove Columns

filtered_data_SAMPLE_X<- data_ SAMPLE_X[,
c("IDs_Cfa.miRNAs", "Mature.Reads_SAMPLE_X", "Score.the.log.odds.score
.for.hairpin", "Precursor.coordinate")]
```

Παρατίθεται παρακάτω ο κώδικας που χρησιμοποιήθηκε στην παρούσα εργασία:

```
# Remove Columns

filtered_data_A6<- data_A6[,
c("IDs_Cfa.miRNAs", "Mature.Reads_A6", "Score.the.log.odds.score.for.h
airpin", "Precursor.coordinate")]

filtered_data_B1<- data_B1[,
c("IDs_Cfa.miRNAs", "Mature.Reads_B1", "Score.the.log.odds.score.for.h
airpin", "Precursor.coordinate")]

filtered_data_B8<- data_B8[,
c("IDs_Cfa.miRNAs", "Mature.Reads_B8", "Score.the.log.odds.score.for.h
airpin", "Precursor.coordinate")]

filtered_data_C16<- data_C16[,
c("IDs_Cfa.miRNAs", "Mature.Reads_C16", "Score.the.log.odds.score.for.
hairpin", "Precursor.coordinate")]

filtered_data_D16<- data_D16[,
c("IDs_Cfa.miRNAs", "Mature.Reads_D16", "Score.the.log.odds.score.for.
hairpin", "Precursor.coordinate")]
```

	IDs_Cfa.miRNAs	Mature.Reads_A6	Score.the.log.odds.score.for.hairpin	Precursor.coordinate
1	cfa-miR-486	21693	11000.0	chr16:24691913..24691976:-
2	cfa-miR-486	21137	10000.0	chr16:24691914..24691978:+
3	cfa-miR-191	3727	1900.0	chr20:40551203..40551266:+
4	cfa-miR-144	2741	1500.0	chr9:43002959..43003016:-
5	cfa-miR-20a	2534	1200.0	chr22:42617625..42617683:+
6	cfa-miR-93	2002	1000.0	chr6:9445017..9445079:+
7	cfa-miR-16	1968	1000.0	chr22:1814294..1814358:+
8	cfa-miR-16	1937	990.0	chr34:26647895..26647958:+
9	cfa-miR-423a	641	780.0	chr9:44126095..44126154:+
10	cfa-miR-126	640	730.0	chr9:48879029..48879088:-
11	cfa-miR-21	1344	690.0	chr9:34306260..34306321:+
12	cfa-miR-106a	998	530.0	chr22:42617176..42617233:+

Εικόνα 21 Μέρος πίνακα με την διατήρηση στηλών υψηλότερου ενδιαφέροντος στο περιβάλλον εργασίας του RStudio

3.2.7.3 Συγχώνευση δεδομένων και εξαγωγή σε αρχείο Excel

Για την δημιουργία ενός πίνακα που θα περιέχει τα ώριμα reads για κάθε miRNA κάθε δείγματος, έπρεπε να πραγματοποιηθεί μια συγχώνευση των διαφορετικών πινάκων που είχαν δημιουργηθεί για κάθε δείγμα (Εικόνα 21).

Τα κριτήρια σύνδεσης αυτών των πινάκων είναι οι στήλες "IDs_Cfa.miRNAs" και "Precursor.coordinate", ώστε η ένωση να πραγματοποιηθεί μόνο μεταξύ των miRNA που έχουν το ίδιο όνομα και τον ίδιο γονιδιακό τόπο του precursor τους. Αυτό είναι σημαντικό, καθώς μερικά miRNA μπορεί να έχουν το ίδιο όνομα αλλά διαφορετικό τόπο προέλευσης του precursor, και συνεπώς θεωρούνται διαφορετικά.

Ο κώδικας που χρησιμοποιείται για την σύζευξη των πινάκων είναι:

```
# Merge datasets
merged_data <- filtered_data_SAMPLE_A %>%
  full_join(filtered_data_SAMPLE_B, by = c("IDs_Cfa.miRNAs",
"Precursor.coordinate")) %>%
...
  full_join(filtered_data_SAMPLE_Z, by = c("IDs_Cfa.miRNAs",
"Precursor.coordinate"))
```

Ο τελεστής Pipe (Pipe operator) "%>%" παίρνει το αποτέλεσμα μιας διαδικασίας και το περνάει σε μια άλλη ως παράμετρο(argument). Αυτό μας επιτρέπει να συνδέσουμε μια ακολουθία βημάτων ανάλυσης.

Για την επεξεργασία των αποτελεσμάτων των δειγμάτων της παρούσας εργασίας χρησιμοποιήθηκε ο παρακάτω κώδικας στην R:

```
# Merge datasets
merged_data <- filtered_data_A6 %>%
  full_join(filtered_data_B1, by = c("IDs_Cfa.miRNAs",
"Precursor.coordinate")) %>%
  full_join(filtered_data_B8, by = c("IDs_Cfa.miRNAs",
"Precursor.coordinate")) %>%
  full_join(filtered_data_C16, by = c("IDs_Cfa.miRNAs",
"Precursor.coordinate")) %>%
  full_join(filtered_data_D16, by = c("IDs_Cfa.miRNAs",
"Precursor.coordinate"))
```

Γίνεται καθαρισμός του αρχείου από μη επιθυμητές στήλες με τον κώδικα:

```
merged_data <- merged_data[,
c("IDs_Cfa.miRNAs", "Mature.Reads_A6", "Mature.Reads_B1", "Mature.Reads_B8", "Mature.Reads_C16", "Mature.Reads_D16", "Precursor.coordinate")]
```

	IDs_Cfa.miRNAs	Mature.Reads_A6	Mature.Reads_B1	Mature.Reads_B8	Mature.Reads_C16	Mature.Reads_D16	Precursor.coordinate
31	cfa-miR-151	185	600	1744	1162	262	chr13:35967806..35967862:-
32	cfa-let-7d	157	379	945	242	74	chr1:98624888..98624964:-
33	cfa-miR-505	151	164	591	217	66	chrX:110850706..110850766:-
34	cfa-miR-22	126	315	329	132	46	chr9:45798565..45798625:-
35	cfa-miR-192	115	NA	NA	NA	26	chr18:52786985..52787047:+
36	cfa-miR-331	106	NA	NA	NA	NA	chr15:35529801..35529858:+
37	cfa-miR-590	68	165	200	131	30	chr6:6103176..6103237:-
38	cfa-miR-7	93	122	638	365	56	chr1:76090714..76090778:+

Εικόνα 22: Μέρος πίνακα συγχωνευμένου αρχείου με τα reads ώριμων miRNA για κάθε δείγμα, στο περιβάλλον εργασίας του RStudio

Για την εξαγωγή του παραπάνω αρχείου στο excel εγκαθίσταται και χρησιμοποιείται το πακέτο “writexl” το οποίο αποκτήθηκε από το CRAN και για την εφαρμογή του ο κώδικας της R είναι:

```
# Extract merged df to excel
install.packages("writexl")
library(writexl)
write_xlsx(merged_data, path = "C:/Users/kdoun/Documents/Pasteur
Institute/Bioinformatics/R_Outputs/merged_final_method.xlsx")
```

Όπου ως “path” ορίζεται η θέση που επιθυμούμε να αποθηκεύσουμε το αρχείο excel που θα δημιουργηθεί.

Η εξαγωγή του αρχείου στο Excel πραγματοποιήθηκε με σκοπό την απομάκρυνση των πολλαπλών εμφανίσεων των miRNA, τα οποία καταγράφηκαν με πολύ μικρή διαφορά στις συντεταγμένες του precursor. Ως μη σημαντική θεωρήθηκε η διαφορά συντεταγμένων μέχρι 10 βάσεις από την αρχή ή το τέλος των αλληλουχιών των precursors για miRNA με το ίδιο όνομα και χρωμόσωμα. Η ένωση πραγματοποιήθηκε χειροκίνητα, λαμβάνοντας υπόψη το MiRDeep2 score κάθε γονιδιακής θέσης και τα reads τοποθετήθηκαν στη θέση με το υψηλότερο score. Η στήλη “precursor coordinate” αναλύεται σε 3 στήλες: “Chromosome” στο οποίο αναφέρεται το χρωμόσωμα στο οποίο χαρτογραφείται το εκάστοτε pre-miRNA, “Start” και “End” όπου αναφέρονται οι συντεταγμένες αρχής και τέλους του.

IDs_Cfa.miRNAs	Mature.Reads_A6	Mature.Reads_B1	Mature.Reads_B8	Mature.Reads_C16	Mature.Reads_D16	Chromosome	Start	End	Δ
cfa-let-7a	736	1398	4704	1388	305	chr1	98627157	98627230	-77644521
cfa-let-7a	734	1403	4709	1381	303	chr10	20982640	20982709	77642255
cfa-let-7d	157	379	945	242	74	chr1	98624888	98624964	7393882
cfa-let-7e				10		chr1	106018781	1.06E+08	-7391994
cfa-let-7f	597	1067	2744	1443		chr1	98626774	98626852	-53034962
cfa-let-7f		1090	2818	1490	248	chrX	45591818	45591890	-7683313
cfa-let-7g	824		3341			chr20	37908499	37908577	1
cfa-let-7g		1105		1917		chr20	37908499	37908578	4708655
cfa-miR-106a	998	1658	6641	2900	544	chr22	42617176	42617233	63203486
cfa-miR-106a			370			chrX	105820661	1.06E+08	-1
cfa-miR-106a					28	chrX	105820659	1.06E+08	-96375865
cfa-miR-106b	944		4046	1541	463	chr6	9444795	9444853	2
cfa-miR-106b		1254				chr6	9444795	9444855	8476072
cfa-miR-107				1114		chr24	17920867	17920927	6872695
cfa-miR-10a	22					chr9	24793558	24793622	0
cfa-miR-10a		114	273		28	chr9	24793560	24793622	-3025205
cfa-miR-1249	10		29			chr10	21768358	21768417	-8025432
cfa-miR-125b		87				chr31	13742923	13742985	-1
cfa-miR-125b			253			chr31	13742923	13742984	-1
cfa-miR-125b				62		chr31	13742923	13742983	35136105
cfa-miR-126	640	2177	5272	3493	774	chr9	48879029	48879088	-11038464
cfa-miR-1271	12			14	4	chr4	37840564	37840624	0

Εικόνα 23. Μέρος πίνακα - παράδειγμα miRNAs με σχεδόν πανομοιότυπες συντεταγμένες τα οποία αφορούν το ίδιο miRNA σε excel spreadsheet (σε κόκκινο πλαίσιο), στα αποτελέσματα.

Μέσω του Excel πέρα από την απομάκρυνση των πολλαπλών εμφανίσεων πανομοιότυπων miRNA, προστέθηκε στη στήλη "IDs_Cfa.miRNAs" το "Precursor.coordinate" ως στοιχείο αναγνώρισης των miRNA που έχουν το ίδιο όνομα αλλά διαφορετικό γονιδιακό τόπο προέλευσης του precursor τους.

IDs_Cfa.miRNAs	Raw_A6	Raw_B1	Raw_B8	Raw_C16	Raw_D16	Chromosome	Start	End
chr1:98627157..98627230_cfa-let-7a	736	1398	4704	1388	305	chr1	98627157	98627230
chr10:20982640..20982709_cfa-let-7a	734	1403	4709	1381	303	chr10	20982640	20982709
chr1:98624888..98624964_cfa-let-7d	157	379	945	242	74	chr1	98624888	98624964
chr1:106018781..106018846_cfa-let-7e	0	0	0	10	0	chr1	106018781	1.06E+08
chr1:98626774..98626852_cfa-let-7f	597	1067	2744	1443	0	chr1	98626774	98626852
chrX:45591818..45591890_cfa-let-7f	0	1090	2818	1490	248	chrX	45591818	45591890
chr20:37908499..37908577_cfa-let-7g	824	1105	3341	1917	0	chr20	37908499	37908577
chr22:42617176..42617233_cfa-miR-106a	998	1658	6641	2900	544	chr22	42617176	42617233
chrX:105820661..105820719_cfa-miR-106a	0	0	370	0	28	chrX	105820661	1.06E+08
chr6:9444795..9444853_cfa-miR-106b	944	1254	4046	1541	463	chr6	9444795	9444853
chr24:17920867..17920927_cfa-miR-107	0	0	0	1114	0	chr24	17920867	17920927
chr9:24793560..24793622_cfa-miR-10a	22	114	273	0	28	chr9	24793560	24793622
chr10:21768358..21768417_cfa-miR-1249	10	0	29	0	0	chr10	21768358	21768417
chr31:13742923..13742984_cfa-miR-125b	0	87	253	62	0	chr31	13742923	13742984
chr9:48879029..48879088_cfa-miR-126	640	2177	5272	3493	774	chr9	48879029	48879088
chr4:37840564..37840624_cfa-miR-1271	12	22	0	14	4	chr4	37840564	37840624

Εικόνα 24. Μέρος πίνακα - παράδειγμα ένωσης των reads του miRNA cfa-let-7g και της προσθήκης του "Precursor.coordinate" στην στήλη "IDs_Cfa.miRNAs" (σε κόκκινο πλαίσιο), στα αποτελέσματα.

Μετά την ολοκλήρωση της επεξεργασίας μέσω του Excel χρησιμοποιήθηκε ο παρακάτω κώδικας για την φόρτωση του αρχείου ξανά στο περιβάλλον εργασίας του RStudio.

```
#After excel editing and filtering
merged_data_1 <- read.csv("file_path")
```

Όπου ως "file_path" χρησιμοποιείται η θέση που βρίσκεται το αρχείο στον υπολογιστή.

Στη συνέχεια οι στήλες "Chromosome", "Start" και "End" αφαιρέθηκαν από τον πίνακα και τα ονόματα των δειγμάτων μετονομάστηκαν για την διευκόλυνση της ανάλυσης:

```
merged_data_1 <- merged_data_1 %>%  
  select(-Start, -Chromosome, -End)
```

Τα δείγματα μολυσμένα με *Leishmania Infantum* μετονομάστηκαν από (A6, B1, B8) σε (L1, L2, L3) αντίστοιχα, ενώ τα δείγματα ελέγχου μετονομάστηκαν από (C16, D16) σε (C1, C2) αντίστοιχα, ως εξής:

```
#Rename merged_data  
colnames(merged_data_1) <-  
c("miRNA", "SAMPLE_A", "SAMPLE_B", ..., "SAMPLE_Z")
```

Για τα δείγματα της παρούσας εργασίας χρησιμοποιήθηκε ο παρακάτω κώδικας:

```
#After excel editing and filtering  
merged_data_1 <- read.csv("C:/Users/kdoun/Documents/Pasteur  
Institute/Bioinformatics/R_Outputs/merged_data2_methodology.csv")  
merged_data_1 <- merged_data_1 %>%  
  select(-Start, -Chromosome, -End)  
#Rename merged_data  
colnames(merged_data_1) <- c("miRNA", "L1", "L2", "L3", "C1", "C2")
```

3.2.7.4 Ανάλυση διαφορικής έκφρασης μέσω της R

Η ανάλυση διαφορικής έκφρασης στοχεύει στην ανίχνευση ποσοτικών αλλαγών στα επίπεδα γονιδιακής έκφρασης μεταξύ πειραματικών ομάδων. Για την πραγματοποίησή της, απαιτείται η χρήση κανονικοποιημένων δεδομένων αριθμού reads και η εφαρμογή κατάλληλων στατιστικών μεθόδων. Στην παρούσα εργασία η ανάλυση διαφορικής έκφρασης χρησιμοποιείται, για να αναδειχτούν οι διαφοροποιήσεις έκφρασης των miRNA σε μολυσμένα δείγματα όταν αυτά συγκρίνονται με δείγματα ελέγχου υγείων σκύλων.

Για την πραγματοποίηση αυτής της ανάλυσης χρησιμοποιήθηκαν τα πακέτα DESeq2 και areglm στο RStudio.

Το DESeq2 αποτελεί ένα ευρέως χρησιμοποιούμενο εργαλείο βιοπληροφορικής για την στατιστική ανάλυση δεδομένων RNA-seq. Τα κύρια χαρακτηριστικά του DESeq2 περιλαμβάνουν:

- Την κανονικοποίηση των reads, χρησιμοποιώντας την διάμεσο (median) των λόγων των παρατηρούμενων μετρήσεων (counts) για τον υπολογισμό των συντελεστών μεγέθους. Ο συντελεστής αυτός θα χρησιμοποιηθεί για την σύγκριση των μετρήσεων διαφορετικών δειγμάτων, με διαφορετικά βάθη αλληλούχησης. Έτσι οι συγκρίσεις αντικατοπτρίζουν βιολογικές διαφορές και όχι τεχνικές διαφορές.
- Την εκτίμηση διασποράς, μέσω της μοντελοποίησης της διασποράς στις μετρήσεις των miRNA μεταξύ των επαναλήψεων, ώστε να ληφθεί υπόψη το γεγονός ότι τα miRNAs με χαμηλό αριθμό μετρήσεων έχουν συχνά υψηλότερη μεταβλητότητα.
- Τον υπολογισμό των αναδιπλούμενων αλλαγών και των τιμών p-values, με τη χρήση στατιστικού μοντέλου σε miRNAs που διαφέρουν σημαντικά μεταξύ της ομάδας των μολυσμένων και της ομάδας ελέγχου.
- Τον έλεγχο ποσοστού ψευδούς ανακάλυψης (False Discovery Rate) (FDR) εφαρμόζοντας τη διόρθωση Benjamini-Hochberg και προσαρμόζοντας τις τιμές p-values “p_{adj}”.

Τα αποτελέσματα του DESeq2 περιέχουν τις εκτιμήσεις λογαριθμικής αναδιπλούμενης αλλαγής (log fold change) (LFC) της έκφρασης των miRNA μεταξύ των ομάδων των μολυσμένων και των υγείων σκύλων, τις τιμές p-values αλλά και τις προσαρμοσμένες τιμές p-values “p_{adj}” [53].

Το `apeglm` χρησιμοποιείται σε συνδυασμό με το DESeq2 για την προσαρμογή των εκτιμήσεων LFC χρησιμοποιώντας μια προσέγγιση συρρίκνωσης (shrinkage) κατά Bayes. Ο σκοπός του είναι η αποφυγή της υπερεκτίμησης των LFC για miRNA με υψηλή μεταβλητότητα ή χαμηλό αριθμό έκφρασης, οδηγώντας σε πιο σταθερές και αξιόπιστες εκτιμήσεις [54].

Το DESeq2 για την δημιουργία μίας βάσης δεδομένων DESeq (DESeqDataSet) απαιτεί την εισαγωγή δύο βασικών πινάκων δεδομένων.

1. **Coldata** το οποίο περιέχει μεταπληροφορίες σχετικά με τα δείγματα στο σύνολο των δεδομένων. Περιγράφει τις πειραματικές συνθήκες ή άλλους σχετικούς παράγοντες για κάθε δείγμα, οι οποίοι χρησιμοποιούνται για το σχεδιασμό του στατιστικού μοντέλου για την ανάλυση της διαφορικής έκφρασης (Εικόνα 25).
2. **Count Data** το οποίο περιέχει τα μη κανονικοποιημένα read counts για τα miRNA κάθε δείγματος (Εικόνα 26).

Ο πίνακας Count Data δημιουργείται από τα ήδη φορτωμένα δεδομένα, αφού αφαιρεθεί η πρώτη στήλη που περιέχει τα περιγραφικά ονόματα των miRNA και οριστούν τα ονόματα αυτά ως λειτουργική ονομασία σειρών, ώστε να μπορέσει να γίνει η περαιτέρω ανάλυση.

Ο πίνακας Coldata δημιουργείται αντιστοιχίζοντας την ιδιότητα κάθε δείγματος ως «Μολυσμένο» ή «Δείγμα Ελέγχου».

Ο κώδικας στην R για την δημιουργία των πινάκων δεδομένων Coldata και Count Data είναι ο εξής:

```
# Create a count matrix (remove the first column which contains
miRNA names)

count_data <- as.matrix(merged_data_1[, 2:6]) # L1, L2, L3, C1, C2
columns

# Set rownames to miRNA names for identification
rownames(count_data) <- merged_data_1$`miRNA`

# Create colData to describe the samples
colData <- data.frame(
  row.names = colnames(count_data),
  condition = c("Infected","Infected","Infected","Control",
"Control"))
```

	condition
L1	Infected
L2	Infected
L3	Infected
C1	Control
C2	Control

Εικόνα 5. Πίνακας Coldata ο οποίος δείχνει την ομαδοποίηση των δειγμάτων (L1, L2, L3) σε μολυσμένα ("Infected") και τα δείγματα (C1, C2) σε δείγματα ελέγχου ("Control").

	L1	L2	L3	C1	C2
chr16:24691913..24691976_cfa-miR-486	21693	20753	71477	7686	3992
chr20:40551203..40551267_cfa-miR-191	3727	3970	18076	5139	1552
chr9:43002959..43003016_cfa-miR-144	2741	1515	5395	4580	335
chr22:42617625..42617684_cfa-miR-20a	2534	3584	13389	5258	0
chr6:9445017..9445079_cfa-miR-93	2002	2541	9905	2991	996
chr22:1814294..1814358_cfa-miR-16	1968	0	19113	0	0
chr34:26647895..26647958_cfa-miR-16	1937	3729	18797	11906	2388
chr9:34306260..34306319_cfa-miR-21	1344	3074	0	1910	667
chr22:42617176..42617233_cfa-miR-106a	998	1658	6641	2900	544
chr6:9444795..9444853_cfa-miR-106b	944	1254	4046	1541	463
chr20:37908499..37908577_cfa-let-7g	824	1105	3341	1917	0
chr14:6013044..6013103_cfa-miR-29a	777	768	2380	659	170
chr1:98627157..98627230_cfa-let-7a	736	1398	4704	1388	305
chr10:20982640..20982709_cfa-let-7a	734	1403	4709	1381	303
chr20:49019126..49019187_cfa-miR-27a	712	1575	3929	2673	406
chr9:44126095..44126154_cfa-miR-423a	641	941	2324	1010	331

Αφού έχουν δημιουργηθεί οι πίνακες Coldata και Count Data (Εικόνα 25, 26), καλούμε το DESeq2 να δημιουργήσει ένα πίνακα δεδομένων (dataset) με βάση τους πίνακες αυτούς, χρησιμοποιώντας την εντολή:

```
# Create DESeq2 dataset

dds <- DESeqDataSetFromMatrix(countData = count_data, colData =
colData, design = ~condition)
```

και να εκτελεστεί η ανάλυση διαφορικής έκφρασης DESeq2 όπως παρακάτω:

```
# Run the DESeq normalisation and analysis

dds <- DESeq(dds)
```

Για να εξαχθούν τα αποτελέσματα της ανάλυσης, δηλαδή οι διαφορικά εκφρασμένες τιμές οι οποίες περιλαμβάνουν στατιστικά όπως Fold Change, p-value, και adjusted p-value, εισάγουμε την εντολή:

```
res <- results(dds)

resultsNames(dds)
```

Στη συνέχεια πραγματοποιούμε συρρίκνωση του Log Fold Change, με τη χρήση του apleglm, η οποία μειώνει το θόρυβο στις εκτιμήσεις και δίνει πιο αξιόπιστα αποτελέσματα για

```
resLFC <- lfcShrink(dds, coef=2, type="apeglm")
```

```
resLFC_df <- as.data.frame(resLFC)
```

Στον πίνακα δεδομένων “resLFC_df” (Εικόνα 27) αναγράφονται τα αποτελέσματα της DESeq2 ανάλυσης, τα οποία περιλαμβάνουν το p-value και το padj για κάθε miRNA καθώς και το “log2FoldChange” δηλαδή τον λογάριθμο (base 2) της διαφοράς έκφρασης μεταξύ των συνθηκών («Infected” vs. “Control”). Θετικές τιμές του “log2FoldChange” υποδεικνύουν υψηλότερη έκφραση, ενώ αρνητικές τιμές υποδεικνύουν χαμηλότερη έκφραση.

	baseMean	log2FoldChange	lfcSE	pvalue	padj
chr1:98627157..98627230_cfa-let-7a	1165.8923353	0.322486526	0.3200415	0.2935405765506238	0.73349348592191
chr10:20982640..20982709_cfa-let-7a	1164.1368941	0.330362564	0.3211876	0.2826160083223133	0.73349348592191
chr1:98624888..98624964_cfa-let-7d	255.3703178	0.465484349	0.3157393	0.1195969567281636	0.54905875588839
chr1:106018781..106018846_cfa-let-7e	1.5974125	-0.207738541	1.0478307	0.2644166240927474	NA
chr1:98626774..98626852_cfa-let-7f	773.2886922	0.106110540	0.9528472	0.7891831507631066	0.98206274531913
chrX:45591818..45591890_cfa-let-7f	745.0494200	-0.085365987	0.9455917	0.7925179162783568	0.98206274531913
chr20:37908499..37908577_cfa-let-7g	964.8213774	0.083732986	0.9508965	0.8342344364793283	0.98206274531913
chr22:42617176..42617233_cfa-miR-106a	1795.7068146	-0.171666015	0.4135083	0.6454365581443965	0.98206274531913
chrX:105820661..105820719_cfa-miR-106a	39.5502820	0.002269755	0.9894402	NA	NA
chr6:9444795..9444853_cfa-miR-106b	1282.7483755	-0.012980447	0.3199049	0.9617729256622243	0.98206274531913
chr24:17920867..17920927_cfa-miR-107	177.9517475	-0.221359901	1.0566154	0.0000000024120176	0.00000006090344
chr9:24793560..24793622_cfa-miR-10a	59.8418239	0.143251401	0.9512045	0.7092054825617565	0.98206274531913
chr10:21768358..21768417_cfa-miR-1249	4.9863841	0.194461702	1.0433698	0.2053510398412346	0.64813921949890
chr31:13742923..13742984_cfa-miR-125b	40.5732058	0.097670829	0.9854888	0.7446603244040262	0.98206274531913
chr9:48879029..48879088_cfa-miR-126	1913.8448410	-0.613680378	0.4707259	0.1422726938655638	0.57478168321688
chr4:37840564..37840624_cfa-miR-1271	12.1188283	0.026150494	0.9264645	0.9543605986623552	0.98206274531913

Εικόνα 27. Μέρος πίνακα “resLFC_df” με τα αποτελέσματα της ανάλυσης διαφορικής έκφρασης DESeq2.

3.2.7.5 Απεικόνιση των αποτελεσμάτων με την χρήση ggplot2

Δημιουργήθηκε αρχικά μια καινούρια στήλη στον πίνακα “resLFC_df” που περιέχει τις ονομασίες των miRNAs, έτσι ώστε να μην περιλαμβάνονται μόνο στις ονομασίες των σειρών.

Για την διαχείριση των μη διαθέσιμων τιμών (NA) στις στήλες “p-value” και “padj” πραγματοποιήθηκε η αντικατάστασή τους με τον αριθμό 1, ώστε να μην δημιουργηθούν προβλήματα στην τελική απεικόνιση. Η επιλογή του αριθμού 1 έγινε ώστε τα αποτελέσματα αυτά να μην εμφανίζονται ως σημαντικά. Οι μη διαθέσιμες τιμές εμφανίζονται στην ανάλυση του DESeq2 όταν ο στατιστικός έλεγχος δεν μπορεί να υπολογίσει το p-value.

Ο κώδικας αντικατάστασης των NA τιμών με 1, παρατίθεται παρακάτω:

```
# Add a miRNA column and handle NA values
resLFC_df$miRNA <- rownames(resLFC_df)
resLFC_df$padj <- ifelse(is.na(resLFC_df$padj), 1, resLFC_df$padj)
```

```
resLFC_df$pvalue <- ifelse(is.na(resLFC_df$pvalue), 1,
resLFC_df$pvalue)
```

“Σημαντικά” δηλαδή διαφορεικά εκφρασμένα, θεωρούνται τα miRNAs τα οποία εμφανίζουν $\text{padj} < 0.05$. Παρόλα αυτά καθώς τα δείγματα προέρχονται από εξωσώματα του ορού των σκύλων, στα οποία ο αριθμός των miRNAs είναι μικρότερος από αυτόν στο σύνολο του κυττάρου, και σε συνδυασμό με τον περιορισμένο αριθμό miRNA που εμφανίζονται ως διαφορετικά εκφρασμένα κατά την ανάλυση, επιλέχθηκε το $p\text{-value} < 0.05$ ως κριτήριο ελέγχου για την ανάλυση [23] [55].

Ο κώδικας στην R που χρησιμοποιήθηκε για να οριστούν τα διαφορετικά εκφρασμένα miRNAs με $p\text{-value} < 0.05$ ως σημαντικά είναι ο παρακάτω:

```
# Add Status column of significance based on p-value
resLFC_df <- resLFC_df %>%
  dplyr::mutate(Status = ifelse(pvalue< 0.05, "Significant", "Not
Significant"))
```

	baseMean	log2FoldChange	lfcSE	pvalue	padj	miRNA	Status
chr1:98627157..98627230_cfa-let-7a	1165.8923353	0.322486526	0.3200415	0.2935405765506238	0.73349348592191	chr1:98627157..98627230_cfa-let-7a	Not Significant
chr10:20982640..20982709_cfa-let-7a	1164.1368941	0.330362564	0.3211876	0.2826160083223133	0.73349348592191	chr10:20982640..20982709_cfa-let-7a	Not Significant
chr1:98624888..98624964_cfa-let-7d	255.3703178	0.465484349	0.3157393	0.1195969567281636	0.54905875588839	chr1:98624888..98624964_cfa-let-7d	Not Significant
chr1:106018781..106018846_cfa-let-7e	1.5974125	-0.207738541	1.0478307	0.2644166240927474	1.0000000000000000	chr1:106018781..106018846_cfa-let-7e	Not Significant
chr1:98626774..98626852_cfa-let-7f	773.2886922	0.106110540	0.9528472	0.7891831507631066	0.98206274531913	chr1:98626774..98626852_cfa-let-7f	Not Significant
chrX:45591818..45591890_cfa-let-7f	745.0494200	-0.085365987	0.9455917	0.7925179162783568	0.98206274531913	chrX:45591818..45591890_cfa-let-7f	Not Significant
chr20:37908499..37908577_cfa-let-7g	964.8213774	0.083732986	0.9508965	0.8342344364793283	0.98206274531913	chr20:37908499..37908577_cfa-let-7g	Not Significant
chr22:42617176..42617233_cfa-miR-106a	1795.7068146	-0.171666015	0.4135083	0.6454365581443965	0.98206274531913	chr22:42617176..42617233_cfa-miR-106a	Not Significant
chrX:105820661..105820719_cfa-miR-106a	39.5502820	0.002269755	0.9894402	1.0000000000000000	1.0000000000000000	chrX:105820661..105820719_cfa-miR-106a	Not Significant
chr6:9444795..9444853_cfa-miR-106b	1282.7483755	-0.012980447	0.3199049	0.9617729256622243	0.98206274531913	chr6:9444795..9444853_cfa-miR-106b	Not Significant
chr24:17920867..17920927_cfa-miR-107	177.9517475	-0.221359901	1.0566154	0.0000000024120176	0.00000006090344	chr24:17920867..17920927_cfa-miR-107	Significant
chr9:24793560..24793622_cfa-miR-10a	59.8418239	0.143251401	0.9512045	0.7092054825617565	0.98206274531913	chr9:24793560..24793622_cfa-miR-10a	Not Significant
chr10:21768358..21768417_cfa-miR-1249	4.9863841	0.194461702	1.0433698	0.2053510398412346	0.64813921949890	chr10:21768358..21768417_cfa-miR-1249	Not Significant
chr31:13742923..13742984_cfa-miR-125b	40.5732058	0.097670829	0.9854888	0.7446603244040262	0.98206274531913	chr31:13742923..13742984_cfa-miR-125b	Not Significant
chr9:48879029..48879088_cfa-miR-126	1913.8448410	-0.613680378	0.4707259	0.1422726938655638	0.57478168321688	chr9:48879029..48879088_cfa-miR-126	Not Significant

Εικόνα 28. Μέρος πίνακα “resLFC_df” έπειτα από την προσθήκη των στηλών “miRNA” και “Status”.

Στη συνέχεια με την χρήση του πακέτου της R ggplot2, δημιουργήθηκε ραβδόγραμμα στο οποίο παρατίθενται οι τιμές των log2 Fold change για κάθε miRNA. Τα miRNA τα οποία θεωρούνται “Σημαντικά” απεικονίζονται με κόκκινη ράβδο ενώ τα “Μη Σημαντικά” με γκρι (Εικόνα 102, κεφ. 4.5)

Ο κώδικας που χρησιμοποιήθηκε για την δημιουργία του ραβδογράμματος είναι ο εξής:

```
p2 <- ggplot(resLFC_df, aes(x = miRNA, y = log2FoldChange, fill =
Status)) +
  geom_bar(stat = "identity", position = "dodge") +
  theme_minimal() +
  labs(title = "Log2 Fold Changes for miRNAs",
```

```

x = "miRNA",
y = "Log2 Fold Change") +
theme(axis.text.x = element_text(angle = 75, hjust = 1)) +
scale_fill_manual(values = c("Significant" = "red", "Not
Significant" = "grey"))
print(p2)

```

Για την δημιουργία ραβδογράμματος, που περιλαμβάνει μόνο τα στατιστικά σημαντικά miRNA με την χρήση του πακέτου ggplot2 δημιουργήθηκε ένα διάνυσμα (Vector) στην R το οποίο περιέχει μόνο τα στατιστικά σημαντικά miRNAs της ανάλυσης.

Κώδικας στην R:

```

# Define the miRNAs you want to plot
miRNAs_of_interest_DEseq2 <- c( "chr14:6012658..6012722_cfa-miR-
29b", "chr13:31440585..31440645_cfa-miR-
30b", "chr24:17920867..17920927_cfa-miR-107",
                                "chr22:1814294..1814358_cfa-miR-16",
"chr12:33982417..33982479_cfa-miR-30c", "chr12_33982417:33982479_cfa-
miR-30c",
                                "chr14:8024123..8024187_cfa-miR-
129", "chrX:39726182..39726247_cfa-miR-
222", "chr6:16001865..16001923_cfa-miR-339",
                                "chr15:35529801..35529858_cfa-miR-
331", "chr7:42324986..42325049_cfa-miR-
92b", "chr1:107675036..107675089_cfa-miR-
150", "chr34:33548044..33548105_cfa-miR-551b",
                                "chr16:24691913..24691976_cfa-miR-
486", "chr7:6416244..6416302_cfa-miR-29c", "chr7:6352404..6352462_cfa-
miR-29c",
                                "chrX:106108237..106108294_cfa-miR-
450a", "chr34:21156856..21156918_cfa-miR-
28", "chr9:33691443..33691505_cfa-miR-
301a", "chrX:57696157..57696208_cfa-miR-374b")

```

Στη συνέχεια, η βάση δεδομένων “resLFC_df” φιλτραρίστηκε με τη χρήση του παραπάνω vector, ώστε να περιέχει μόνο τα στατιστικά σημαντικά miRNAs και με την χρήση του ggplot2 δημιουργήθηκε το ραβδόγραμμα της Εικόνας 103, κεφ 4.5 με τη χρήση του παρακάτω κώδικα στην R:

```

# Filter for miRNAs of interest
filtered_resLFC <- resLFC_df %>%
  filter(miRNA %in% miRNAs_of_interest_DEseq2)
p2 <- ggplot(filtered_resLFC, aes(x = miRNA, y = log2FoldChange,
fill = Status)) +
  geom_bar(stat = "identity", position = "dodge") +
  theme_minimal() +
  labs(title = "Log2 Fold Changes for Specific miRNAs",
    x = "miRNA",
    y = "Log2 Fold Change") +
  theme(axis.text.x = element_text(angle = 75, hjust = 1)) +
  scale_fill_manual(values = c("Significant" = "red", "Not
Significant" = "grey"))
print(p2)

```

4. Αποτελέσματα ανάλυσης ακατέργαστων δεδομένων αλληλούχισης μορίων miRNA από εξωσώματα ορού υγιών και μολυσμένων με λεισμανίαση σκύλων

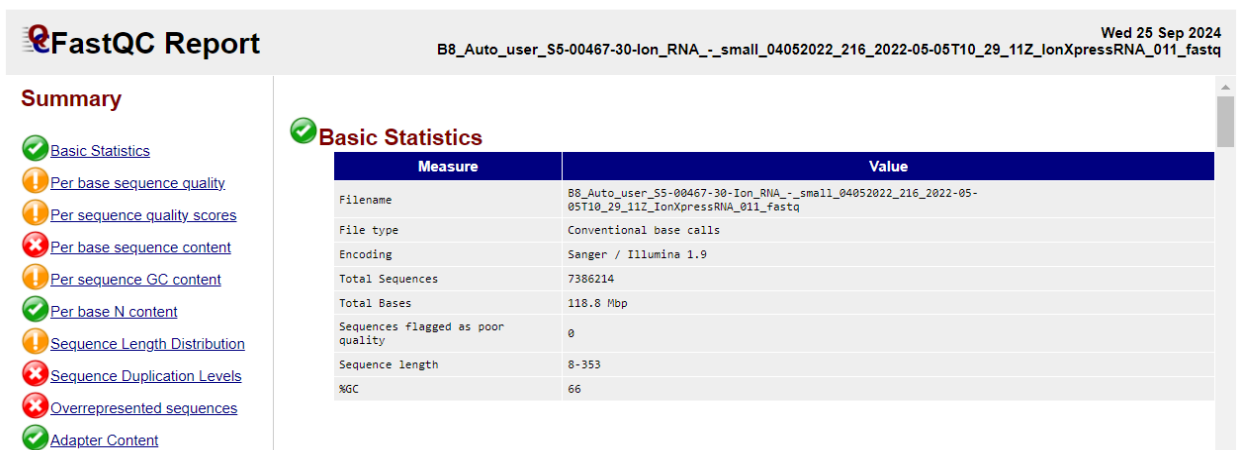
4.1 Δείγματα προς ανάλυση

Πραγματοποιήθηκε αλληλούχιση μικρών RNA (small RNA-seq) για την ταυτοποίηση προφίλ έκφρασης των miRNA στα εξωσώματα του ορού 6 σκύλων, 3 εκ των οποίων ήταν μολυσμένοι με το παράσιτο *L. Infantum* (A6, B1, B8) και οι υπόλοιποι 3 (C16, D16, B6) ήταν υγιείς και χρησιμοποιήθηκαν ως ομάδα ελέγχου.

Η αλληλούχιση πραγματοποιήθηκε στον sequencer Επόμενης Γενιάς Ion GeneStudio S5 και τα δεδομένα που συλλέχτηκαν υπέστησαν βιοπληροφορική επεξεργασία για την ανίχνευση και τον χαρακτηρισμό των miRNAs.

Τα ακατέργαστα δεδομένα μετατράπηκαν σε μορφή FASTQ και ακολούθησε ποιοτικός έλεγχος FASTQC

4.2 Αποτελέσματα ποιοτικού ελέγχου FastQC



Εικόνα 29. Παράδειγμα από περιβάλλον εργασίας FASTQC

Όπως αναφέρθηκε στο κεφάλαιο 3.2.4.2. στην αναφορά του FastQC εμφανίζεται μια περίληψη των ενοτήτων που εκτελέστηκαν και μια γρήγορη αξιολόγηση των αποτελεσμάτων.

Όπου τα αποτελέσματα της ενότητας φαίνονται απολύτως φυσιολογικά, εμφανίζεται ένα πράσινο tik, εάν τα αποτελέσματα είναι ελαφρώς ανώμαλα εμφανίζεται προειδοποίηση με πορτοκαλί θαυμαστικό και εάν τα αποτελέσματα είναι πολύ ασυνήθιστα εμφανίζεται σφάλμα με κόκκινο "X". Τα αποτελέσματα κάθε ενότητας αναλύονται παρακάτω ξεχωριστά.

4.2.1 Βασικές στατιστικές πληροφορίες:

- Δείγμα μολυσμένο με λεισμάνια: A6

Basic Statistics

Measure	Value
Filename	Cutadapt on Adaptors and A6
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	7608989
Total Bases	176.3 Mbp
Sequences flagged as poor quality	0
Sequence length	1-242
%GC	63

Εικόνα 30. Βασικές στατιστικές πληροφορίες: Δείγμα μολυσμένο με λεισμάνια: A6

Οι συνολικές αλληλουχίες που ανιχνεύθηκαν στο μολυσμένο με λεισμάνια δείγμα A6 ήταν 7608989 και το ποσοστό Γουανίνης-Κυτοσίνης (GC%) ήταν 63.

- Δείγμα μολυσμένο με λεισμάνια: B1

Basic Statistics

Measure	Value
Filename	Cutadapt on adaptors and B1_Read 1 Output
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	9091557
Total Bases	180.5 Mbp
Sequences flagged as poor quality	0
Sequence length	1-250
%GC	63

Εικόνα 31. Βασικές στατιστικές πληροφορίες: Δείγμα μολυσμένο με λεισμάνια: B1

Οι συνολικές αλληλουχίες που ανιχνεύθηκαν στο μολυσμένο με λεισμάνια δείγμα B1 ήταν 9091557 και το ποσοστό Γουανίνης-Κυτοσίνης (GC%) ήταν 63.

- Δείγμα μολυσμένο με λεισμάνια: B8

Basic Statistics

Measure	Value
Filename	B8_Auto_user_S5-00467-30-Ion_RNA_-_small_04052022_216_2022-05-05T10_29_11Z_IonXpressRNA_011_fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	7386214
Total Bases	118.8 Mbp
Sequences flagged as poor quality	0
Sequence length	8-353
%GC	66

Εικόνα 32. Βασικές στατιστικές πληροφορίες: Δείγμα μολυσμένο με λεισμάνια: B8

Οι συνολικές αλληλουχίες που ανιχνεύθηκαν στο μολυσμένο με λεισμάνια δείγμα B8 ήταν 7386214 και το ποσοστό Γουανίνης-Κυτοσίνης (GC%) ήταν 66.

- Δείγμα ελέγχου: B6

✓ Basic Statistics

Measure	Value
Filename	B6_Auto_user_S5-00467-30-Ion_RNA_-_small_04052022_216_2022-05-05T10_29_11Z_IonXpressRNA_009_fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	356
Total Bases	6.2 kbp
Sequences flagged as poor quality	0
Sequence length	8-74
%GC	68

Εικόνα 33. Βασικές στατιστικές πληροφορίες: Δείγμα ελέγχου B6

Παρατηρείται ιδιαίτερα χαμηλός αριθμός reads στο δείγμα B6 με μόνο 356 αλληλουχίες και GC% = 68.

- Δείγμα ελέγχου: C16

✓ Basic Statistics

Measure	Value
Filename	Cutadapt on data 83 and data 16042024_ Read 1 Output
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	1634448
Total Bases	45.1 Mbp
Sequences flagged as poor quality	0
Sequence length	1-171
%GC	53

Εικόνα 34. Βασικές στατιστικές πληροφορίες: Δείγμα ελέγχου C16

Οι συνολικές αλληλουχίες που ανιχνεύθηκαν στο μολυσμένο με λεισμάνια δείγμα C16 ήταν 1634468 και το ποσοστό Γουανίνης-Κυτοσίνης (GC%) ήταν 53.

- Δείγμα ελέγχου D16:

✓ Basic Statistics

Measure	Value
Filename	Cutadapt on data 18042024_ Read 1 Output
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	1023952
Total Bases	26.5 Mbp
Sequences flagged as poor quality	0
Sequence length	1-191
%GC	58

Εικόνα 35. Βασικές στατιστικές πληροφορίες: Δείγμα ελέγχου D16

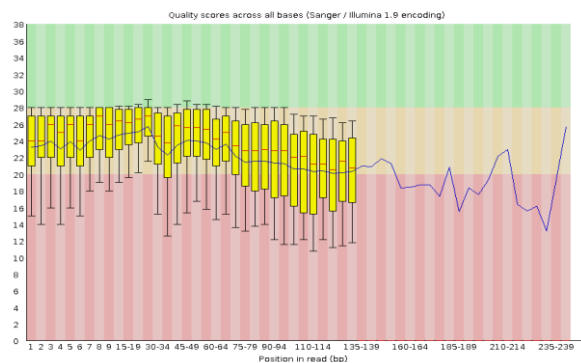
Οι συνολικές αλληλουχίες που ανιχνεύθηκαν στο μολυσμένο με λεισμάνια δείγμα D16 ήταν 1634468 και το ποσοστό Γουανίνης-Κυτοσίνης (GC%) ήταν 53

4.2.2 Διάγραμμα ποιότητας ανά βάση

Απεικονίζεται η ποιότητα ανά βάση (Per base sequence quality), σε ένα διάγραμμα τύπου «boxplot» που δείχνει τη βαθμολογία ποιότητας για κάθε βάση.

A6:

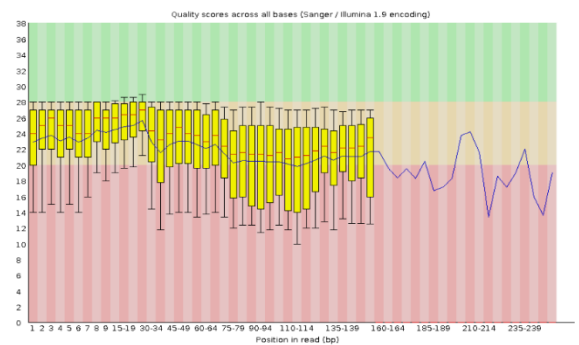
① Per base sequence quality



Εικόνα 36. Ποιότητα ανα βάση δείγματος A6

B1:

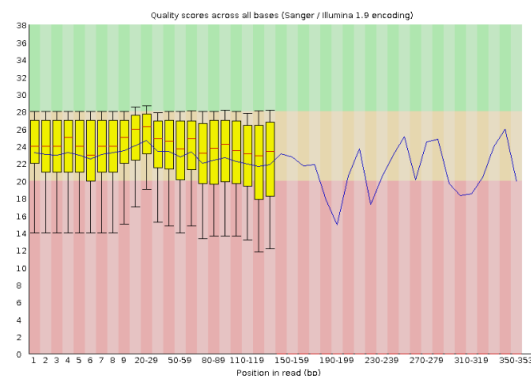
① Per base sequence quality



Εικόνα 37. Ποιότητα ανα βάση δείγματος B1

B8:

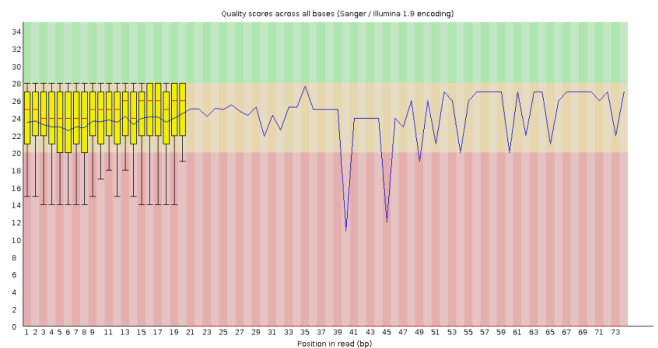
① Per base sequence quality



Εικόνα 38. Ποιότητα ανα βάση δείγματος B8

B6:

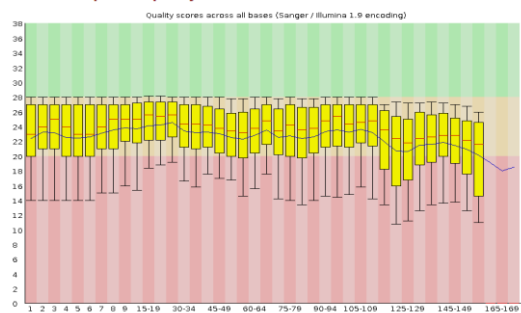
① Per base sequence quality



Εικόνα 39. Ποιότητα ανα βάση δείγματος B6

C16:

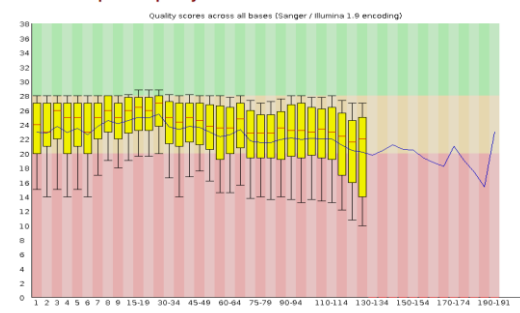
① Per base sequence quality



Εικόνα 40. Ποιότητα ανα βάση δείγματος C16

D16:

① Per base sequence quality



Εικόνα 41. Ποιότητα ανα βάση δείγματος D16

Στα διαγράμματα παρατηρείται πως στις περισσότερες θέσεις των reads ο βαθμός ποιότητας σύμφωνα με το ενδοτεταρτημοριακό εύρος για την κάθε θέση κυμαίνεται από 20 με 28 δηλαδή στην περιοχή της ικανοποιητικής ποιότητας. Ακόμα και οι θέσεις με χαμηλότερους βαθμούς ποιότητας, ο βαθμός είναι >10 και οι περισσότερες διατηρούν μέσο όρο > 20.

Αυτή η ενότητα έχει εμφανίσει ⚠️ προειδοποίηση καθώς η διάμεσος (median) για κάποιες θέσεις είναι χαμηλότερη του 25 σε όλα τα διαγράμματα.

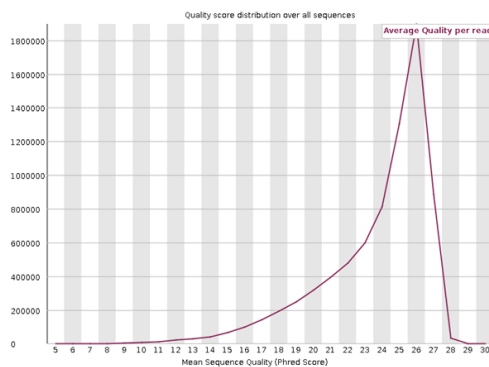
Συμπεραίνουμε ότι οι βαθμοί ποιότητας ανα βάση είναι ικανοποιητικοί.

4.2.3 Βαθμολογία ποιότητας ανά αλληλουχία

Απεικονίζεται η κατανομή της βαθμολογίας ποιότητας σε όλες τις αλληλουχίες.

A6:

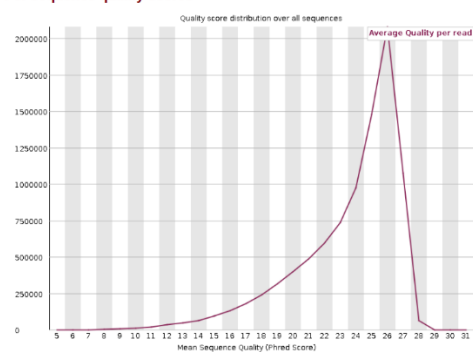
⚠️ Per sequence quality scores



Εικόνα 42. Ποιότητα ανά αλληλουχία δείγματος A6

B1:

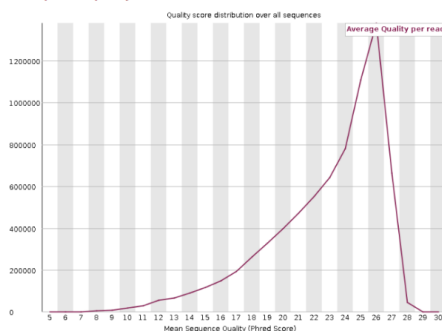
⚠️ Per sequence quality scores



Εικόνα 43. Ποιότητα ανά αλληλουχία δείγματος B1

B8:

⚠️ Per sequence quality scores



Εικόνα 44. Ποιότητα ανά αλληλουχία δείγματος B8

B6:

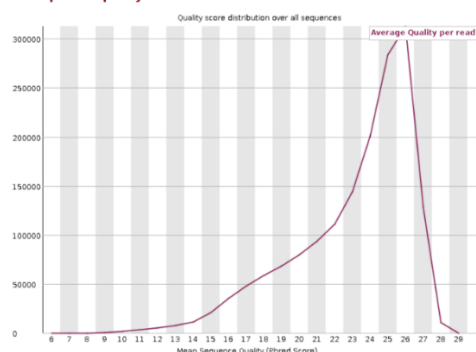
⚠️ Per sequence quality scores



Εικόνα 45. Ποιότητα ανά αλληλουχία δείγματος B6

C16:

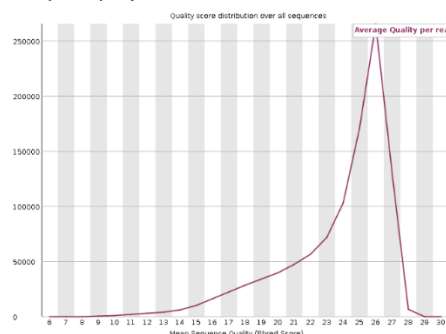
1 Per sequence quality scores



Εικόνα 46. Ποιότητα ανά αλληλουχία δείγματος C16

D16:

1 Per sequence quality scores



Εικόνα 47. Ποιότητα ανά αλληλουχία δείγματος D16

Σε όλα τα διαγράμματα, η πιο συχνά παρατηρούμενη μέση ποιότητα είναι 26 ενώ στο δείγμα ελέγχου B6 παρατηρούνται δύο ακόμα μικρότερες κορυφές στο 18 και στο 22.

Προειδοποίηση: Προειδοποίηση εμφανίζεται εάν η συχνότερα παρατηρούμενη μέση ποιότητα είναι κάτω από 27 το οποίο αντιστοιχεί σε ποσοστό σφάλματος μόλις 0,2%.

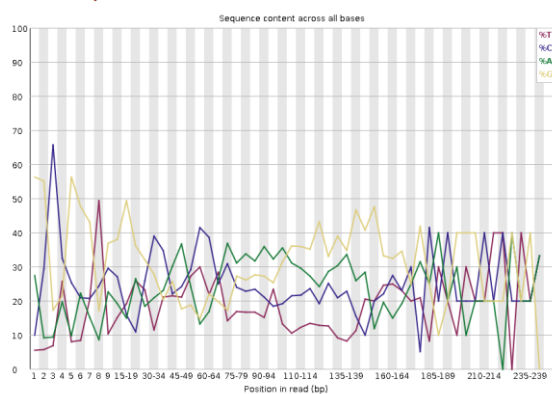
Τα δείγματα είναι ικανοποιητικά να χρησιμοποιηθούν για περαιτέρω ανάλυση.

4.2.4 Περιεχόμενο αλληλουχίας ανά βάση

Τα διαγράμματα δείχνουν τη συχνότητα εμφάνισης των βάσεων (A, T, C, G) σε κάθε θέση.

A6:

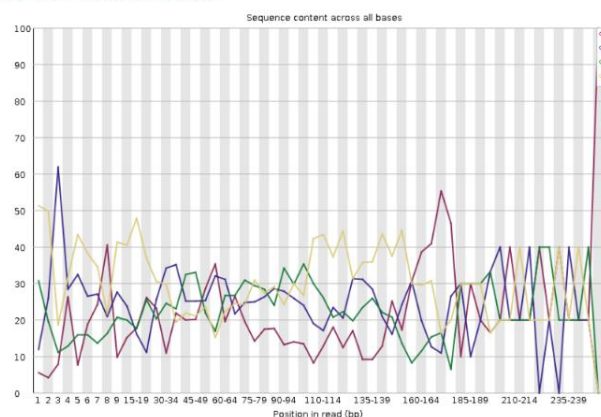
1 Per base sequence content



Εικόνα 48. Περιεχόμενο αλληλουχίας ανα βάση δείγματος A6

B1:

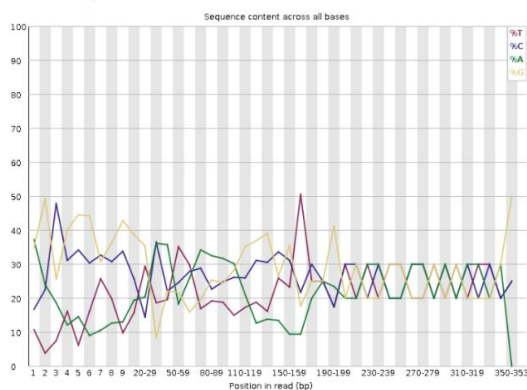
1 Per base sequence content



Εικόνα 49. Περιεχόμενο αλληλουχίας ανα βάση δείγματος B1

B8:

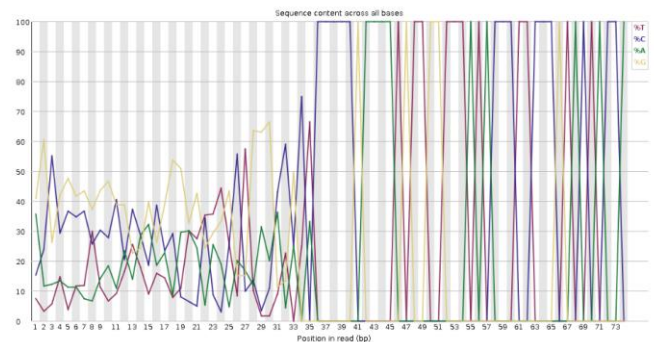
Per base sequence content



Εικόνα 50. Περιεχόμενο αλληλουχίας ανα βάση δείγματος B8

B6:

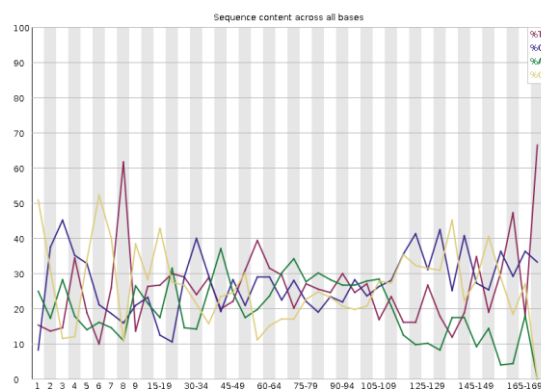
Per base sequence content



Εικόνα 51. Περιεχόμενο αλληλουχίας ανα βάση δείγματος B6

C16

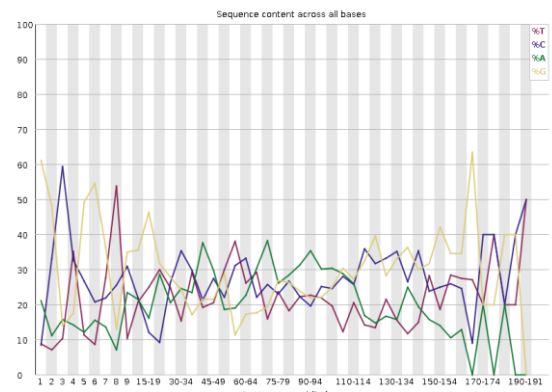
Per base sequence content



Εικόνα 52. Περιεχόμενο αλληλουχίας ανα βάση δείγματος C16

D16

Per base sequence content



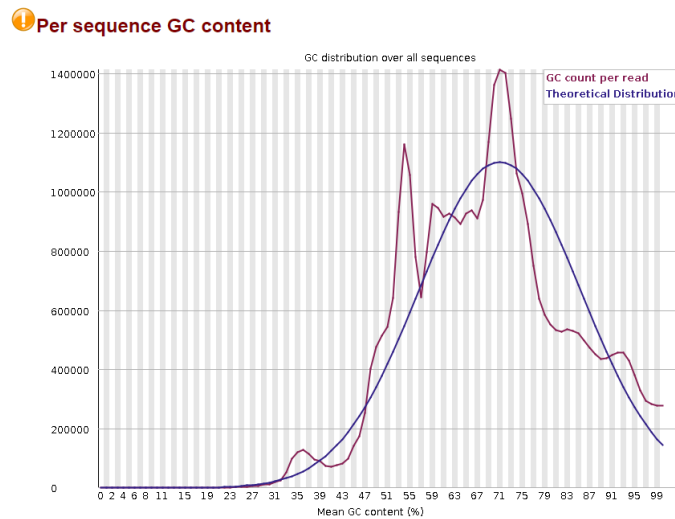
Εικόνα 53. Περιεχόμενο αλληλουχίας ανα βάση δείγματος D16

Αυτό το module αναφέρει αποτυχία εάν η διαφορά μεταξύ A και T ή G και C είναι μεγαλύτερη από 20% σε οποιαδήποτε θέση. Όταν υπάρχουν έντονες κλίσεις που αλλάζουν σε διαφορετικές βάσεις, συνήθως υποδεικνύει μια υπερεκπροσωπούμενη αλληλουχία που μολύνει τη βιβλιοθήκη. Στη παρούσα μελέτη, η αποτυχία της ενότητας στο small RNA-seq οφείλεται στη φυσική προτίμηση των μικρών RNAs σε συγκεκριμένες βάσεις και δεν αποτελεί απαραίτητα ένδειξη προβλήματος ή επιμόλυνσης.

4.2.5 Περιεχόμενο GC ανά βάση

Σε αυτή την ενότητα εμφανίζεται προειδοποίηση εάν το άθροισμα των αποκλίσεων από την κανονική κατανομή αντιπροσωπεύει πάνω από το 15% των reads και αποτυχία εάν το άθροισμα των αποκλίσεων από την κανονική κατανομή αντιπροσωπεύει περισσότερο από το 30% των reads.

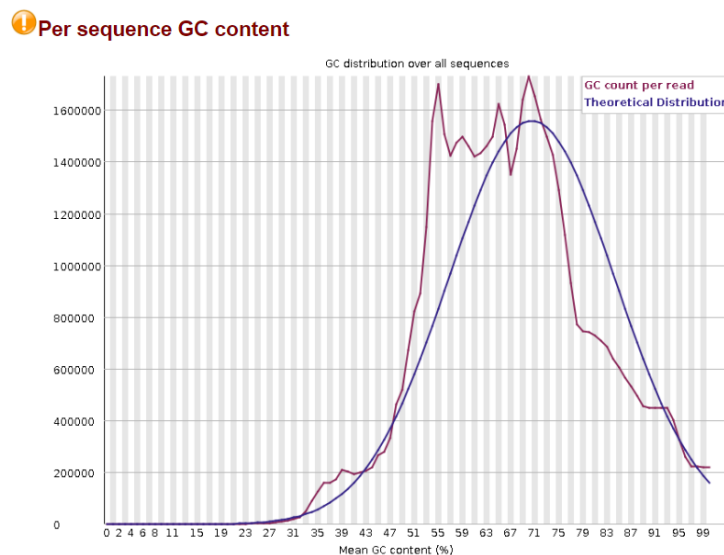
A6:



Εικόνα 54. Περιεχόμενο GC ανά βάση δείγματος A6

Η κατανομή του δείγματος εμφανίζει μία αρκετά υψηλότερη κορυφή στο GC%=71% από την κανονική κατανομή αλλά και στις δύο κορυφές παρατηρούμε πως η υψηλότερη τους κορυφή βρίσκεται στο ίδιο GC content=71%. Επίσης παρατηρούνται δύο ακόμη κορυφές εκτός της κανονικής κατανομής στο GC content ίσες με 54% και στο 59%.

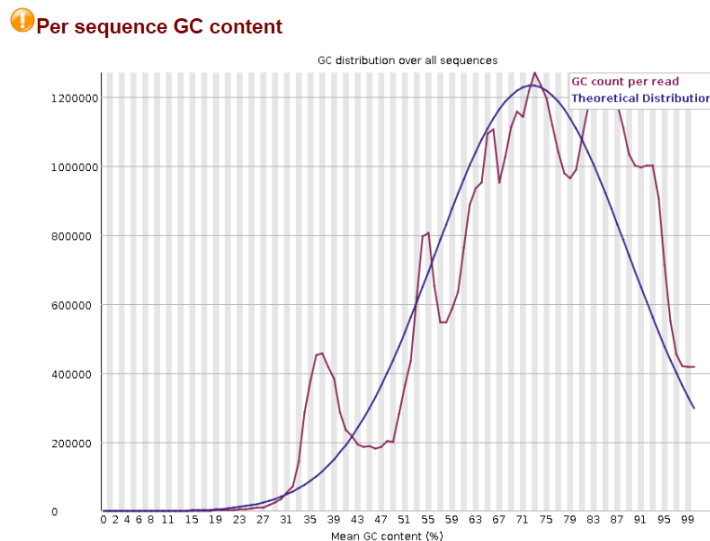
B1:



Εικόνα 55. Περιεχόμενο GC ανά βάση δείγματος B1

Η κατανομή εμφανίζει μια αρκετά υψηλότερη κορυφή στο GC% 70 από την κανονική κατανομή αλλά και στις δύο παρατηρούμε πως η υψηλότερη τους κορυφή βρίσκεται στο ίδιο GC content 70%. Επίσης παρατηρούνται τρεις ακόμα κορυφές εκτός της κανονικής κατανομής στο GC content 55%, στο 59% και στο 65%.

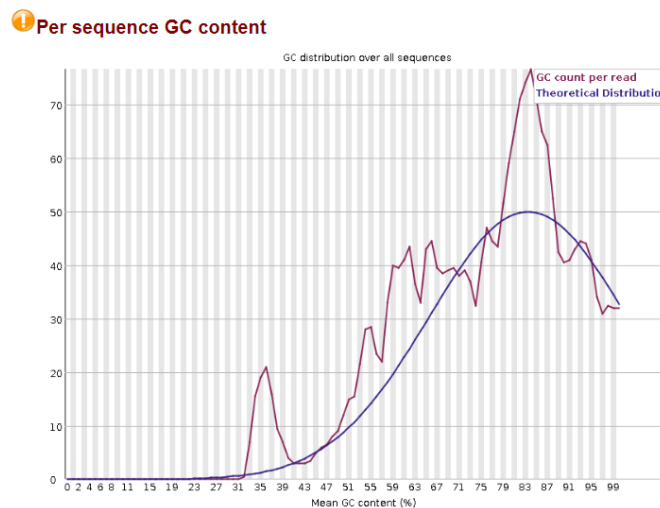
B8:



Εικόνα 56. Περιεχόμενο GC ανά βάση δείγματος B8

Η κατανομή εμφανίζει μία αρκετά υψηλότερη κορυφή στο GC% 73 από την κανονική κατανομή αλλά και στις δύο παρατηρούμε πως η υψηλότερη τους κορυφή βρίσκεται στο ίδιο GC content 73%. Επίσης παρατηρούνται τέσσερις ακόμα κορυφές στο GC content 35%, 55%, 66% και στο 85%.

B6:

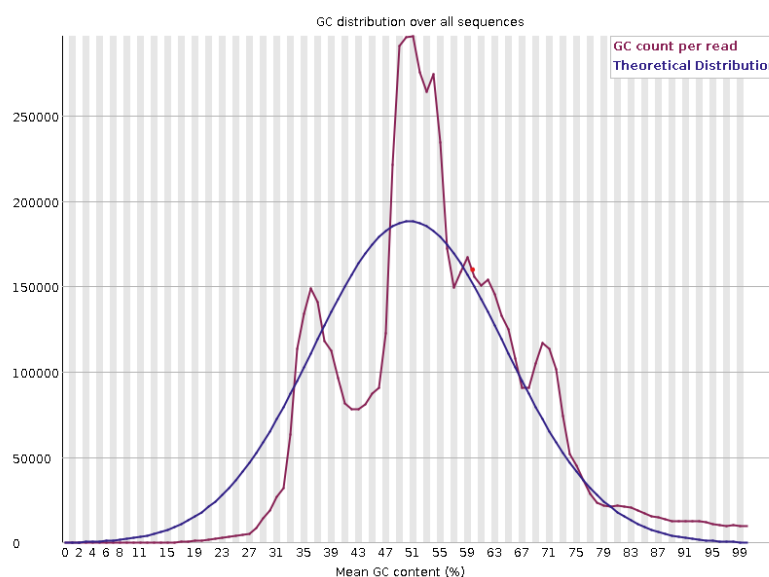


Εικόνα 57. Περιεχόμενο GC ανά βάση δείγματος B6

Αυτό το module υποδεικνύει αποτυχία εάν το άθροισμα των αποκλίσεων από την κανονική κατανομή αντιπροσωπεύει περισσότερο από το 30% των reads. Η κατανομή εμφανίζει ένα αρκετά υψηλότερο peak στο GC% 84 από την κανονική κατανομή αλλά και στις δύο παρατηρείται πως η υψηλότερη τους κορυφή βρίσκεται στο ίδιο GC content 84%. Επίσης παρατηρούνται 6 ακόμα peaks εκτός της κανονικής κατανομής στο GC content 36%, 55%, 62%, 66%, 76% και στο 93%.

C16:

✖ Per sequence GC content

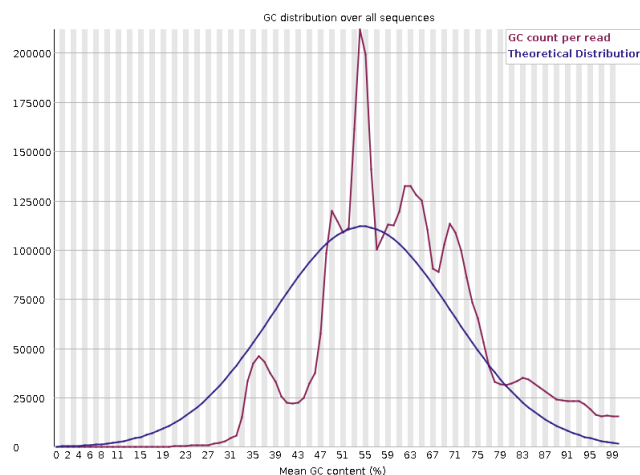


Εικόνα 58. Περιεχόμενο GC ανά βάση δείγματος C16

Η κατανομή εμφανίζει μία αρκετά υψηλότερη κορυφή στο GC% 51 από την κανονική κατανομή αλλά και στις δύο παρατηρείται πως η υψηλότερη τους κορυφή βρίσκεται στο ίδιο GC content 51%. Επίσης παρατηρούνται δύο ακόμα κορυφές εκτός της κανονικής κατανομής στο GC content 36% και στο 72%.

D16:

✖ Per sequence GC content



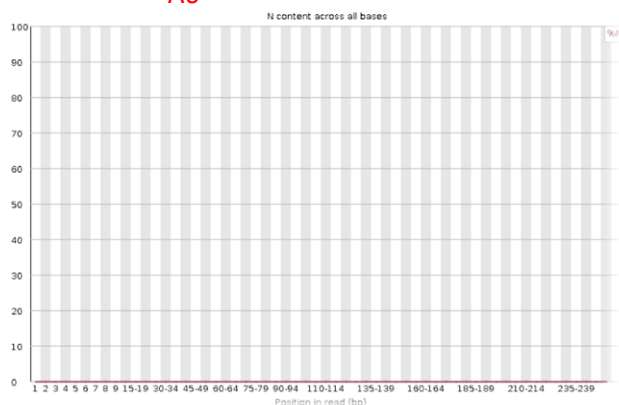
Εικόνα 59. Περιεχόμενο GC ανά βάση δείγματος D16

Η κατανομή εμφανίζει μία αρκετά υψηλότερη κορυφή στο GC% 54 από την κανονική κατανομή αλλά και στις δύο παρατηρούμε πως η υψηλότερη τους κορυφή βρίσκεται στο ίδιο GC content 54%. Επίσης παρατηρούνται τρεις ακόμα κορυφές εκτός της κανονικής κατανομής στο GC content 48%, στο 63% και στο 70%.

4.2.6 Περιεχόμενο N ανά θέση

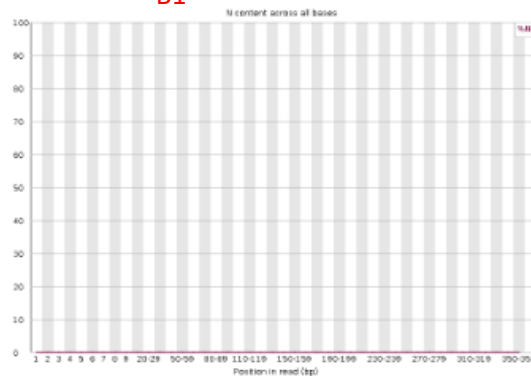
Κατά την αλληλούχιση, όταν δεν μπορεί να αναγνωριστεί με βεβαιότητα μια βάση ως A, T, C ή G, αποδίδεται ένα «N» στη θέση αυτή. Ο έλεγχος καταγράφει τη συχνότητα εμφάνισης της βάσης N (άγνωστη βάση) ανά θέση. Υψηλά επίπεδα περιεκτικότητας σε N μπορεί να υποδηλώνουν προβλήματα με την ποιότητα της αλληλούχισης, όπως κακή ποιότητα ανάγνωσης ή σφάλματα στην επιλογή βάσεων.

✓ Per base N content **A6**



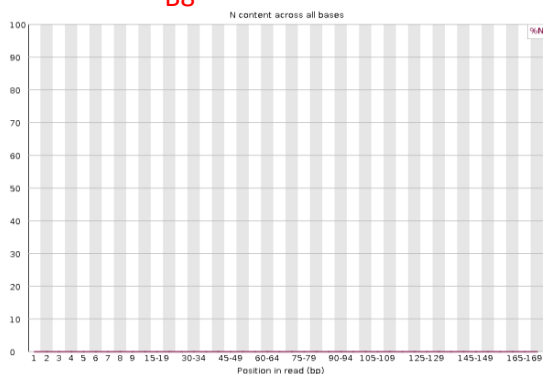
Εικόνα 60. Περιεχόμενο N ανά θέση, δείγμα A6

✓ Per base N content **B1**



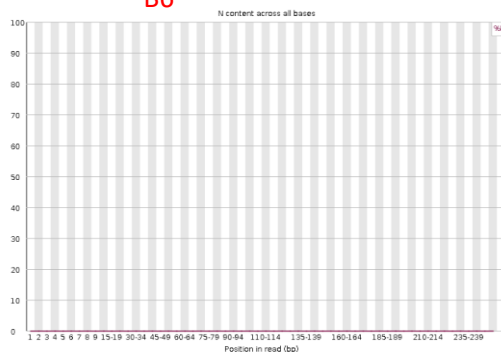
Εικόνα 61. Περιεχόμενο N ανά θέση, δείγμα B1

✓ Per base N content **B8**



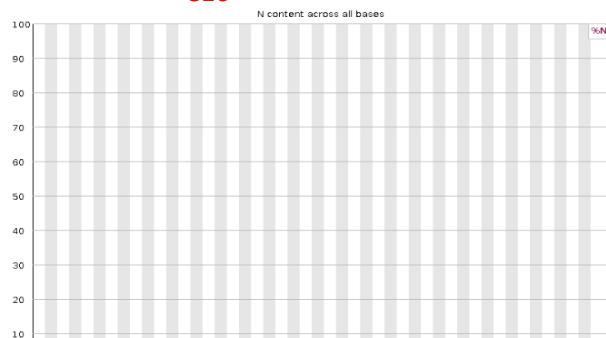
Εικόνα 62. Περιεχόμενο N ανά θέση, δείγμα B8

✓ Per base N content **B6**



Εικόνα 63. Περιεχόμενο N ανά θέση, δείγμα B6

✓ Per base N content **C16**



Εικόνα 64. Περιεχόμενο N ανά θέση, δείγμα C16

✓ Per base N content **D16**



Εικόνα 65. Περιεχόμενο N ανά θέση, δείγμα D16

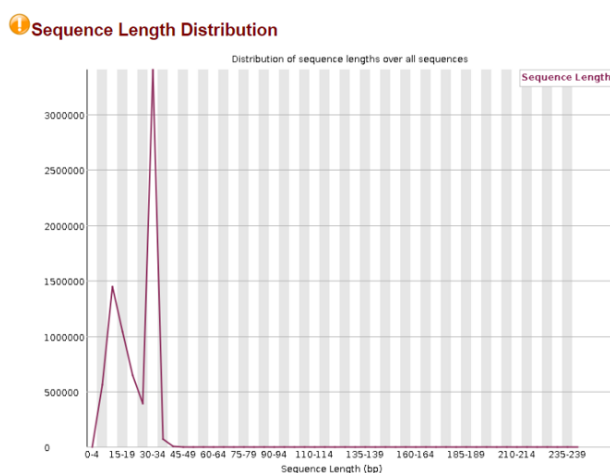
Για όλα τα δείγματα παρατηρείται πως δεν εμφανίζονται N βάσεις κατά την αλληλούχιση και άρα τα αποτελέσματα είναι αξιόπιστα για περαιτέρω ανάλυση

4.2.7 Κατανομή μήκους ακολουθίας

Αυτό το module θα εμφανίσει προειδοποίηση αν όλες οι ακολουθίες δεν έχουν το ίδιο μήκος.

Ορισμένοι αναλυτές αλληλούχισης υψηλής απόδοσης παράγουν θραύσματα αλληλουχίας ομοιόμορφου μήκους, αλλά άλλοι μπορεί να περιέχουν αναγνώσματα με πολύ διαφορετικά μήκη, όπως το Ion GeneStudio S5.

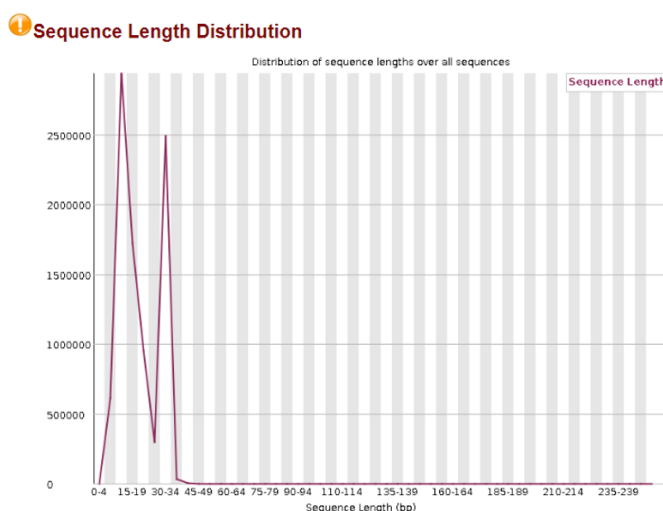
A6:



Εικόνα 67. Κατανομή μήκους ακολουθίας, δείγμα A6

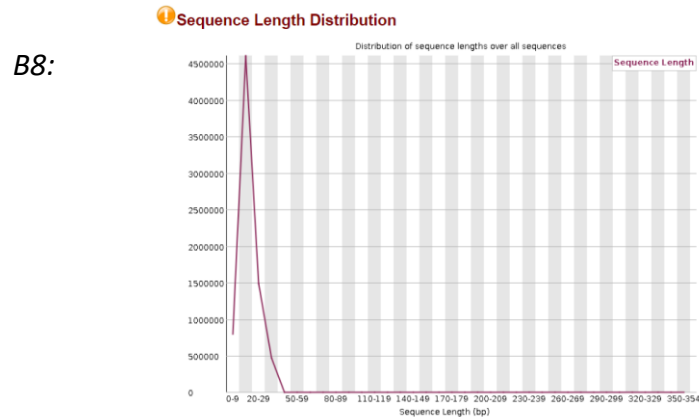
Παρατηρούνται 2 κορυφές στο διάγραμμα του δείγματος A6. Η πρώτη κορυφή βρίσκεται μεταξύ των μηκών 10-15bp και η δεύτερη και μεγαλύτερη μεταξύ των μηκών 30-34 bp.

B1:



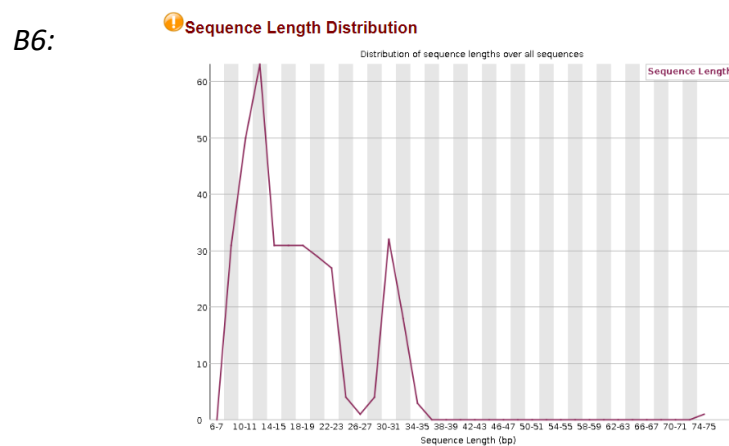
Εικόνα 68. Κατανομή μήκους ακολουθίας, δείγμα B1

Παρατηρούνται 2 peaks, το πρώτο βρίσκεται μεταξύ των μηκών 10-15bp και το δεύτερο και μικρότερο μεταξύ των μηκών 30-34bp.



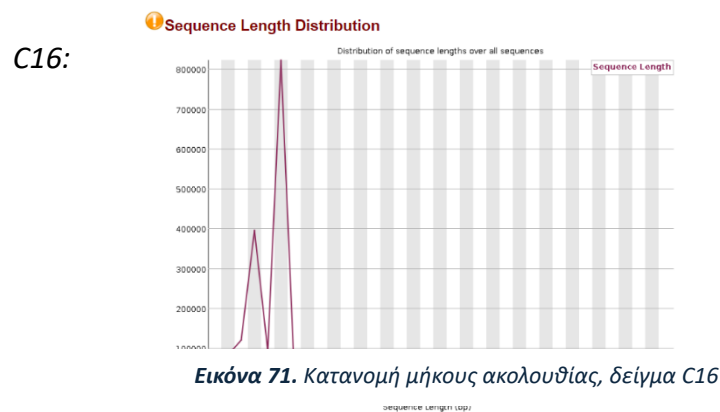
Εικόνα 69. Κατανομή μήκους ακολουθίας, δείγμα B8

Παρατηρείται ένα peak μεταξύ των μηκών 10 και 20 bp.



Εικόνα 70. Κατανομή μήκους ακολουθίας, δείγμα B6

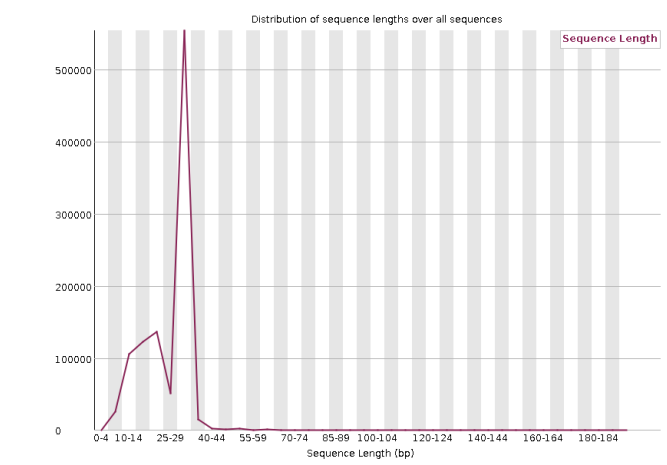
Παρατηρούνται 2 peaks το πρώτο και μεγαλύτερο βρίσκεται μεταξύ των μηκών 11-14bp και το δεύτερο στο μήκος 31 bp.



Εικόνα 71. Κατανομή μήκους ακολουθίας, δείγμα C16

Παρατηρούνται 2 peaks το πρώτο βρίσκεται μεταξύ των μηκών 20-24bp και το δεύτερο και μεγαλύτερο μεταξύ των μηκών 30-34bp.

D16:  **Sequence Length Distribution**



Εικόνα 72. Κατανομή μήκους ακολουθίας, δείγμα D16

Παρατηρούνται 2 peaks το πρώτο βρίσκεται μεταξύ των μηκών 20-24bp και το δεύτερο και μεγαλύτερο μεταξύ των μηκών 30-34bp.

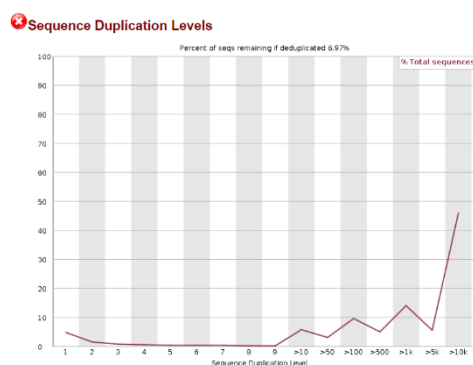
Το μεγάλο εύρος μηκών προκύπτει καθώς, πέρα από τα miRNAs τα οποία έχουν μήκος περίπου 22-nt, η βιβλιοθήκη μπορεί να περιέχει και άλλα μικρά RNAs (π.χ., tRNAs, rRNAs, piRNAs) ή προϊόντα αποικοδόμησης RNA.

4.2.8 Επίπεδο διπλασιασμού αλληλουχίας

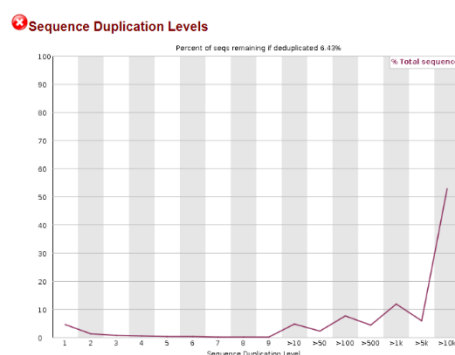
Αυτό το module θα εμφανίσει προειδοποίηση εάν οι μη μοναδικές αλληλουχίες αποτελούν πάνω από το 20% του συνόλου, ενώ θα αναφέρει σφάλμα εάν οι μη μοναδικές ακολουθίες αποτελούν πάνω από το 50% του συνόλου.

A6:

B1:

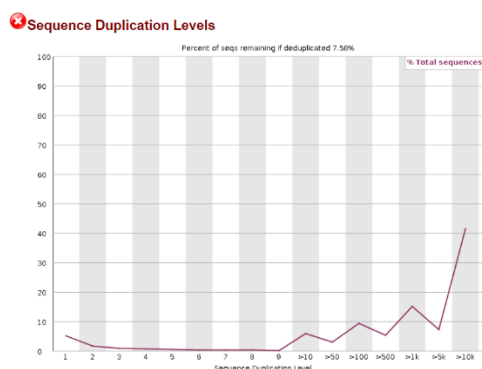


Εικόνα 73. Επίπεδο διπλασιασμού ακολουθίας, δείγμα A6



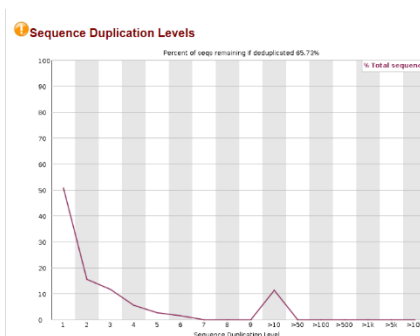
Εικόνα 74. Επίπεδο διπλασιασμού ακολουθίας, δείγμα B1

B8:



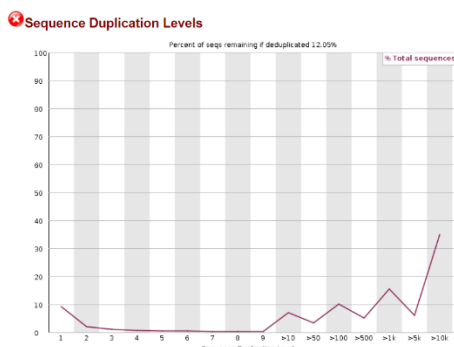
Εικόνα 75. Επίπεδο διπλασιασμού ακολουθίας, δείγμα B8

B6:



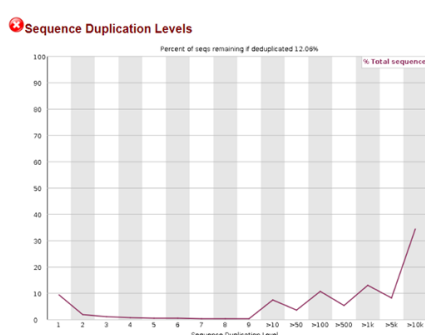
Εικόνα 76. Επίπεδο διπλασιασμού ακολουθίας, δείγμα B6

C16:



Εικόνα 77. Επίπεδο διπλασιασμού ακολουθίας, δείγμα C16

D16:



Εικόνα 78. Επίπεδο διπλασιασμού ακολουθίας, δείγμα D16

Στα δείγματα A6, B1, B8, C16, D16 παρατηρούμε μια αύξηση στην κατηγορία >10 αντιγράφων, η οποία κορυφώνεται στην κατηγορία >10K αντιγράφων. Η αύξηση το οποίο υποδηλώνει ότι έχουμε ένα μεγάλο αριθμό αλληλουχιών με πολύ υψηλά επίπεδα διπλασιασμού. Η αύξηση αυτή δεν είναι ασυνήθιστη για το small RNA-seq, καθώς έχουμε υπερέκφραση συγκεκριμένων miRNAs ή άλλων μικρών μορίων RNA.

Στο δείγμα B6 παρατηρείται μια αύξηση στην κατηγορία των 10 αντιγράφων η οποία στη συνέχεια ελαττώνεται στην κατηγορία των 50. Η κατανομή αυτή είναι αναμενόμενη, δεδομένου ότι το σύνολο των αλληλουχιών στο δείγμα B6, σε σχέση με τα άλλα δείγματα, είναι πολύ μικρό.

4.2.9 Υπερεκπροσωπούμενες Αλληλουχίες

Στους συγκεκριμένους πίνακες αναγράφονται αλληλουχίες που εμφανίζονται πολύ συχνά στα δείγματα, με φθίνουσα σειρά εμφανίσεων.

A6: Overrepresented sequences

Sequence	Count	Percentage	Possible Source
GGCTGGTCCGATGGTAGTGGTTATCAGAAC	367616	4.8313830263127	No Hit
GCTGGTCCGATGGTAGTGGTTATCAGAACT	238797	3.1383538601514607	No Hit
AGCAGAGTGGCGCAGCGGAAGCTGCTGGGC	237803	3.125290363805231	No Hit
GCCCGGATAGCTCAGTGGTAGAGCATCAGAC	235279	3.0921190712721494	No Hit
AGCAGAGTGGCGCAGCGGAAGCTGCTGGGC	123154	1.6185330271866603	No Hit
GCCCGGCTAGCTCAGTGGTAGAGCATGGAC	121512	1.596953287749529	No Hit
GCCCGGCTAGCTCAGTGGTAGAGCATGAGAC	117782	1.5479323205750462	No Hit
CGCGACCTCAGATCAGACG	89946	1.1821018534788261	No Hit
GCATTGGTGGTTAGTGGTAGAATTCTCGCT	84091	1.1051533916003822	No Hit
GCCCGGATAGCTCAGTGGTAGAGCATCAGA	81168	1.0667383012381804	No Hit
AGCAGAGTGGCGCAGCGGAAGCTGCTGGGC	75530	0.9926417294071526	No Hit
AGCAGAGTGGCGCAGCGGAAGCTGCTGGGC	70801	0.9304915541341958	No Hit
AGCAGAGTTGCGCAGCGGAAGCTGCTGGGC	62941	0.8271926796056612	No Hit
CGCGACCTCAGATCAGACGT	59497	0.7819304246595705	No Hit
GGGCCAGT	59058	0.7761609328124932	No Hit
AAACCGTTACCACTACTAGT	51967	0.6829685257791804	No Hit
GGCTGGTCCGAG	49133	0.6457231046069327	No Hit
GTTTCCGTAGTGTAGTGGTTACAGTTCGCC	47583	0.625352461411102	No Hit
GGCTGGTCCGATGGTAGTGGTTATCAGAA	45827	0.6022744940227933	No Hit
AGCAGAGTGGCGCAGCGGAAGCTGCTGGGC	45378	0.5963735786712269	No Hit
GTTTCCGTAGTGTAGTGGTTACAGTTCGCCCT	43500	0.5716922445281495	No Hit
AGCGCCGTTCCGAAGGACGGGCG	41319	0.5430287781990485	No Hit
TTCCCGACGGGAGC	39674	0.5214096117105702	No Hit
AGCAGAGTTGCGCAGCGGAAGCTGCTGGGCC	35098	0.4612702160562987	No Hit

Εικόνα 79. Υπερεκπροσωπούμενες Αλληλουχίες, δείγμα A6

B1: Overrepresented sequences

Sequence	Count	Percentage	Possible Source
CGCGACCTCAGATCAGACG	359625	3.955593084808041	No Hit
GACTCTTAGCGGTG	346132	3.8071806622342024	No Hit
GGCTGGTCCGATGGTAGTGGTTATCAGAAC	293263	3.2256631069903645	No Hit
GACTCTTAGCGGT	174473	1.9190662201499196	No Hit
GGCTGGTCCGATGGTAGTGGTTATCAGAACT	166633	1.8328323740367825	No Hit
AGCAGAGTGGCGCAGCGGAAGCTGCTGGGC	140701	1.547600702497933	No Hit
GCCCGGATAGCTCAGTGGTAGAGCATCAGAC	127796	1.405658189897863	No Hit
AGCAGAGTGGCGCAGCGGAAGCTGCTGGGC	121859	1.3403534730079787	No Hit
GCCCGGATAGCTCAGTGGTAGAGCATCAGA	115008	1.2649978436036864	No Hit
CGCGACCTCAGA	104964	1.1545217172372124	No Hit
AGCGCCGTTCCGAAGGACGGGCG	95591	1.051426064864357	No Hit
GCCGATCGAAAGGAGTCTGGG	93293	1.026149866299029	No Hit
AACGGGTGGG	88056	0.96854697165733	No Hit
GGGCCAGT	76502	0.841462029001193	No Hit
GCCGATCGAAAGGAGTCTGGG	68664	0.7552501733201475	No Hit
GCCCGGCTAGCTCAGTGGTAGAGCATGGAC	63648	0.7000781054324257	No Hit
GCATTGGTGGTTAGTGGTAGAATTCTCGCT	63430	0.6976802763267063	No Hit
CGCGACCTCAG	61989	0.6818304054669682	No Hit
GCGCCGCTGGTGTAGTGGTAGATCATGCAAGA	61128	0.6723608080970345	No Hit
GCCCGGCTAGCTCAGTGGTAGAGCATGAGAC	60968	0.6706002063232953	No Hit
CGCGACCTCAGATCAGACGT	59752	0.657225159562878	No Hit
CGCGACCTCAGAT	59322	0.6524954966459541	No Hit
GACGAGGTGGC	56675	0.6233085716666574	No Hit
GTTTCCGTAGTGTAGTGGTTATCAGGTTCCGC	54430	0.5986873315538802	No Hit
AACGTTAAGAG	52262	0.574841031079715	No Hit
AGCGCCGTTCCGAAGGACGGGC	51303	0.564292782053366	No Hit

Εικόνα 80. Υπερεκπροσωπούμενες Αλληλουχίες, δείγμα A6

B8: Overrepresented sequences

Sequence	Count	Percentage	Possible Source
CGCGACCTCAGATCAGACG	275696	3.732575308540196	No Hit
GGCTGGTCCGATGGTAGTGGTTATCAGAAC	206836	2.8002979604977596	No Hit
AACGGGTGGG	169983	2.301354929602635	No Hit
AAACCGTTACCACTACTAGT	125240	1.6955912731475151	No Hit
AAACCGTTACCACTACTAGT	119368	1.6160918164569833	No Hit
GACTCTTAGCGGTG	97962	1.326281637656315	No Hit
ACGGGTGGG	97526	1.3203787488420997	No Hit
AAACCGTTACCACTACTAGT	91972	1.2451846101399175	No Hit
GGCTGGTCCGATGGTAGTGGTTATCAGAACT	78146	1.0579980487973948	No Hit
AGAAGCCCGGCTC	73004	0.9883818692499297	No Hit
CGCGACCTCAGA	71916	0.9736517246860165	No Hit
GACTCTTAGCGGT	68425	0.9263879979648572	No Hit
CGCGACCTCAGAT	61911	0.8381966728827515	No Hit
AGAAGCCCGGC	52616	0.7123541235063051	No Hit
AGCGCGTTCCGAAGGACGGGC	45200	0.6119508587213964	No Hit
AACGGGTGGGTCCG	43014	0.5823551822354457	No Hit
AGGGGGCGG	41513	0.5620335397810029	No Hit
AGCTCTCGCTGGC	41156	0.5572002110959688	No Hit
AAACCGTTAAGAG	37972	0.5140928762692226	No Hit
GGCGCTGACCCCC	35312	0.47807983900818474	No Hit
GGGGCCCCGGC	32205	0.4360149868389949	No Hit
TCCCGACGGGG	31678	0.4288800730658494	No Hit
TCCCGGGGC	29896	0.4047540458481165	No Hit
CCCGACGGGG	29848	0.4041041865291204	No Hit
GGTGGTGGCTGGGCTGGGCGCGAAGCGGG	28928	0.3916485495816937	No Hit

Εικόνα 81. Υπερεκπροσωπούμενες Αλληλουχίες, δείγμα B8

B6: Overrepresented sequences

Sequence	Count	Percentage	Possible Source
CGCGACCTCAGATCAGACGT	17	4.77528898876404	No Hit
TCGCCCGGAGG	13	3.61685393258427	No Hit
AAACCGTTACCACTACTAGT	11	3.089887640449438	No Hit
ACGGGTGGG	6	1.683393258429662	No Hit
GCTTCCGGAAACCAAG	5	1.4044043820224718	No Hit
CGCTCTCTCTCTCTCC	5	1.4044043820224718	No Hit
AGCAGAGTGGCGCAGCGGAAGCTGCTGGGC	4	1.123595056179776	No Hit
GGCTGGTCCGATGGTAGTGGTTATCAGAAC	4	1.123595056179776	No Hit
GCCCGGCTAGCTCAGTGGTAGAGCATGGAC	4	1.123595056179776	No Hit
GGGGGGGGCGGCCCGCG	4	1.123595056179776	No Hit
AGCGCTCTCGA	4	1.123595056179776	No Hit
GCGCCGCTGGTGTAGTGGTAGATCATGCAAGAT	3	0.8426966292134831	No Hit
GGCTGGGCGCGAAGCGGG	3	0.8426966292134831	No Hit
GGCTGGGCGCGAAGCGGG	3	0.8426966292134831	No Hit
AGCAGAGTGGCGCAGCGGAAGCTGCTGGGC	3	0.8426966292134831	No Hit
AGCAGAGTGGCGCAGCGGAAGCTGCTGGG	3	0.8426966292134831	No Hit
CGCGACCTCAGATCAGACG	3	0.8426966292134831	No Hit
AGCGCGGGGAGC	3	0.8426966292134831	No Hit
AAACCGTTACCACTACTAGT	3	0.8426966292134831	No Hit
GGGGGGGGCGGCCCGCGCT	3	0.8426966292134831	No Hit
AGCAGAGTGGCGCAGCGGAAGCTGCTGGG	3	0.8426966292134831	No Hit
GGGGGGGGCGGCCCGCG	3	0.8426966292134831	No Hit
GACGAGGTGGC	3	0.8426966292134831	No Hit
GGCTGGGCG	3	0.8426966292134831	No Hit
AGGCCGCGGGG	3	0.8426966292134831	No Hit
GCCCGGCTAGCTCAGTGGTAGAGCATGGGA	2	0.5617977528889888	No Hit
GCCCGGCTAGCTCAGTGGTAGAGCATGAGAC	2	0.5617977528889888	No Hit
AGCTCTCGCTGGC	2	0.5617977528889888	No Hit
GCTGGGCGCGAAGCGGG	2	0.5617977528889888	No Hit
ACTCTTAGCGGTGGAT	2	0.5617977528889888	No Hit
AGCGAGTCCGT	2	0.5617977528889888	No Hit
CCCCCGG	1	0.2808988764044944	No Hit
GGGCCGATCGACGCC	1	0.2808988764044944	No Hit
GGGCTGGTCCGATGGTAGTGGTTATCAGAAC	1	0.2808988764044944	No Hit
ATCACCGA	1	0.2808988764044944	ABI Solid3 Adapter B (100% over Bbp)
AGAGGACGGCG	1	0.2808988764044944	No Hit
ACGCGACCTCAGATCAGACG	1	0.2808988764044944	No Hit
AGCGCGCGCGCGCG	1	0.2808988764044944	No Hit
ATCACCGACTGCCCATAGAGAGG	1	0.2808988764044944	ABI Solid3 Adapter B (100% over Bbp)
GGCGGGGGCC	1	0.2808988764044944	No Hit

Εικόνα 82. Υπερεκπροσωπούμενες Αλληλουχίες, δείγμα B6

C16:  Overrepresented sequences

Sequence	Count	Percentage	Possible Source
GCATTGTTGGTTCAGTGGTATGAGATTCTGGCT	97573	3.39573982555306	No Hit
AAACCGTTACCATTAAGTGT	55009	3.3655592530893	No Hit
GGCTGGTCCGATGGTGGTGGTATTCAGAACT	44408	2.71896686585083	No Hit
GGCTGGTCCGATGGTGGTGGTATTCAGAACT	32982	2.8179042232461746	No Hit
AGCAGAGTGGCAGCGAGCGAGCTGCTGGG	32009	1.9587242232461746	No Hit
GGCCGCTGGTGGTGGTGGTATTCAGAACT	30356	1.857240397870773	No Hit
CGGAGCTCAGATCAGAG	29648	1.8139235518835486	No Hit
GTTCCTGATGGTGGTGGTATTCAGCTGGC	24105	1.474793776980889	No Hit
GTTCCTGATGGTGGTGGTATTCAGCTGGC	22278	1.3638213073885326	No Hit
GCATTGTTGGTTCAGTGGTATGAGATTCTGGCT	21751	1.329467965051997	No Hit
AAACCGTTACCATTAAGTGT	20341	1.2449207966668073	No Hit
AAACCGTTACCATTAAGTGT	18333	1.122640364795207	No Hit
GGCCGATGCTCAGTGGTGGTATTCAGAACT	17934	1.0423739673892487	No Hit
GGCCGCTGGTGGTGGTGGTATTCAGAACT	16492	1.00961333033337	No Hit
GTTCCTGATGGTGGTGGTATTCAGCTGGC	16125	0.986595410861536	No Hit
GGCTGGTCCGATGGTGGTGGTATTCAGAACT	15396	0.83382095831635	No Hit
AGCAGAGTGGCAGCGAGCGAGCTGCTGGG	13282	0.812619319517684	No Hit
TCTCTGATGGTGGTGGTGGTATTCAGAACT	12740	0.7794585149438649	No Hit
GGCCGCTGGTGGTGGTGGTATTCAGAACT	12303	0.752723986839921	No Hit
GGCCGCTGGTGGTGGTGGTATTCAGAACT	12033	0.736282838383841	No Hit
GCATTGTTGGTTCAGTGGTATGAGATTCTGGCT	10898	0.666761294886283	No Hit
GCATTGTTGGTTCAGTGGTATGAGATTCTGGCT	10741	0.671577329447941	No Hit
CGGAGCTCAGATCAGAG	10629	0.646951258421236	No Hit
CATTATTAATTTTGGTACG	2851	0.174620605062734	No Hit
AGCAGAGTGGCAGCGAGCGAGCTGCTGGG	2841	0.173838025325295	No Hit
GCATTGTTGGTTCAGTGGTATGAGATTCTGGCT	2741	0.16789862648788476	No Hit
AAACCGTTACCATTAAGTGT	2712	0.1659254886368578	No Hit
AGCAGAGTGGCAGCGAGCGAGCTGCTGGG	2682	0.164080995232827	No Hit
GCATTGTTGGTTCAGTGGTATGAGATTCTGGCT	2685	0.163940995232827	No Hit
GGCCGATGCTCAGTGGTGGTATTCAGAACT	2642	0.1634820897895186	No Hit
TCTCTGATGGTGGTGGTGGTATTCAGAACT	2555	0.15613397689788168	No Hit
ATCAGATTCCAGGAGTTCCAG	2513	0.15375831344382937	No Hit
GGCAGGCTCAGAGAGAACTCTGGTGGG	2512	0.153889314547893	No Hit
GACTCTTAAAGGTT	2477	0.1513477029882417	No Hit
TCTCTGATGGTTCAGTGGTATGAGATTCTGGCT	2434	0.1489109546051613	No Hit
TCTCTGATGGTTCAGTGGTATGAGATTCTGGCT	2416	0.1476526608955047	No Hit
GGTTCATGATGGTGGTGGTATTCAGCTGGC	2330	0.14259482086275966	No Hit
ATCAGAGTGGCAGCGAGCGAGCTGCTGGG	2281	0.139596120472976	ABI Solid3 Adapter B (100% over 23bp)
ATCAGAGTGGCAGCGAGCGAGCTGCTGGG	2184	0.1323978396343764	No Hit

Εικόνα 83. Υπερεκπροσωπούμενες Αλληλουχίες, δείγμα C16

D16:  Overrepresented sequences

Sequence	Count	Percentage	Possible Source
GGCTGGTCCGATGGTGGTGGTATTCAGAACT	90718	8.859387329398533	No Hit
GGCTGGTCCGATGGTGGTGGTATTCAGAACT	29560	2.8887864090564623	No Hit
GCATTGTTGGTTCAGTGGTATGAGATTCTGGCT	25965	2.535703961811685	No Hit
AGCAGAGTGGCAGCGAGCGAGCTGCTGGG	23480	2.293822492714673	No Hit
CGGAGCTCAGATCAGAG	23055	2.251517613694071	No Hit
AAACCGTTACCATTAAGTGT	17136	1.6734767221186746	No Hit
GGCCGATGCTCAGTGGTGGTATTCAGAACT	16029	1.565368719579365	No Hit
GGCCGCTGGTGGTGGTGGTATTCAGAACT	15124	1.4769877419802807	No Hit
GGCCGCTGGTGGTGGTGGTATTCAGAACT	14938	1.4588232536700079	No Hit
GGCTGGTCCGATGGTGGTGGTATTCAGAACT	13278	1.29671087914248	No Hit
GGCCGCTGGTGGTGGTGGTATTCAGAACT	13007	1.2702446151801666	No Hit
AAACCGTTACCATTAAGTGT	12818	1.2517871512613576	No Hit
GCATTGTTGGTTCAGTGGTATGAGATTCTGGCT	12665	1.2368453948139408	No Hit
GGCCGCTGGTGGTGGTGGTATTCAGAACT	11613	1.13418086119205919	No Hit
AGCAGAGTGGCAGCGAGCGAGCTGCTGGG	11353	1.1067173918138706	No Hit
GTTCCTGATGGTGGTGGTGGTATTCAGCTGGC	11106	1.0845957327124655	No Hit
GGCCGATGCTCAGTGGTGGTATTCAGAACT	10767	1.0514894880348757	No Hit
GGCCGCTGGTGGTGGTGGTATTCAGAACT	9607	0.9382055829433503	No Hit
GTTCCTGATGGTGGTGGTGGTATTCAGCTGGC	8259	0.8065618725438878	No Hit
GGCCGCTGGTGGTGGTGGTATTCAGAACT	7528	0.7351734886284522	No Hit
GGCAGGCTCAGAGAGAACTCTGGTGGG	7312	0.71407932362328991	No Hit
GGCTGGTCCGATGGTGGTGGTATTCAGAACT	6677	0.6520668640483762	No Hit

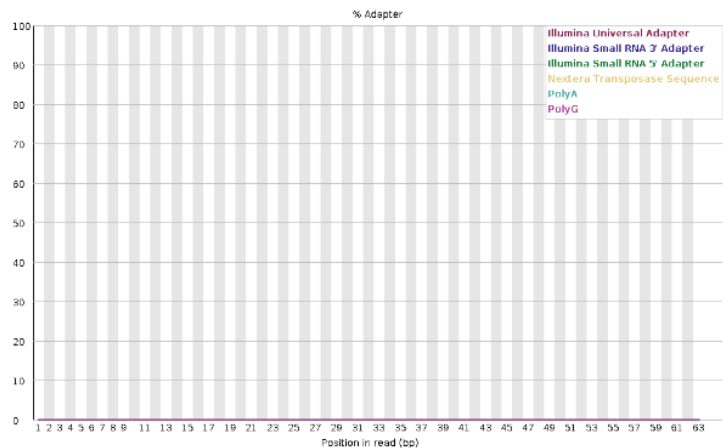
Εικόνα 84. Υπερεκπροσωπούμενες Αλληλουχίες, δείγμα D16

Αυτό το module ενεργοποιείται συχνά όταν χρησιμοποιείται για την ανάλυση βιβλιοθηκών μικρών RNA, όπου οι αλληλουχίες δεν υπόκεινται σε τυχαίο κατακερματισμό και η ίδια αλληλουχία μπορεί να υπάρχει σε σημαντικό ποσοστό της βιβλιοθήκης. Για τις περισσότερες αλληλουχίες το πρόγραμμα δεν βρίσκει αντιστοιχίες (No Hit) εκτός από μία την οποία έχει αναγνωρίσει ως ABI Solid3 Adapter B.

Αυτό το module, θα εμφανίσει σφάλμα εάν διαπιστωθεί ότι κάποια ακολουθία αντιπροσωπεύει περισσότερο από το 1% του συνόλου, όπως συμβαίνει σε όλα τα δείγματα της παρούσας εργασίας. Η υπερεκπροσώπηση κάποιων αλληλουχιών σε δείγματα που περιέχουν miRNAs ή άλλα μικρά μόρια RNA είναι αναμενόμενη, λόγω της υπερέκφρασης ορισμένων εξ αυτών και δεν αποτελεί απαραίτητα ένδειξη επιμόλυνσης των δειγμάτων.

4.2.10 Περιεχόμενο σε Αντάπτορες

A6,B1,B8,B6,C16,D16:  Adapter Content



Εικόνα 85. Περιεχόμενο σε προεπιλεγμένους FastQC αντάπτορες, χαρακτηριστική εικόνα για όλα τα δείγματα

Σε κανένα από τα δείγματα δεν ανιχνεύθηκε κάποιος από τους προεπιλεγμένους adapters που περιλαμβάνονται στη λίστα του προγράμματος FastQC.

4.3 Σύγκριση FastQC πριν και μετά τη χρήση του Cutadapt στα αρχεία fastq

Μετά τη χρήση του Cutadapt, δεν παρατηρήθηκε κάποια μεγάλη αλλαγή μεταξύ των αναφορών FastQC, όπως ήταν αναμενόμενο, καθώς οι περισσότερες ακολουθίες των ανταπτόρων είχαν ήδη αφαιρεθεί από το μηχάνημα Ion GeneStudio S5, με την χρήση του πακέτου λογισμικού του, Torrent Suite.

Στο παράδειγμα του δείγματος B8, μετά τη χρήση του Cutadapt, παρατηρήθηκαν οι παρακάτω διαφορές.

Δείγμα B8 πριν το Cutadapt



Basic Statistics

Measure	Value
Filename	B8_Pre_CA_fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	7386214
Total Bases	118.8 Mbp
Sequences flagged as poor quality	0
Sequence length	8-353
%GC	66

Εικόνα 86. Βασικές στατιστικές πληροφορίες:
Δείγμα ελέγχου B8 πριν το Cutadapt

Δείγμα B8 μετά το Cutadapt



Basic Statistics

Measure	Value
Filename	CA_B8_fastq
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	7385969
Total Bases	118.3 Mbp
Sequences flagged as poor quality	0
Sequence length	1-353
%GC	66

Εικόνα 87. Βασικές στατιστικές πληροφορίες:
Δείγμα ελέγχου B8 μετά το Cutadapt

Οι συνολικές αλληλουχίες του δείγματος B8 πριν τη χρήση του Cutadapt ήταν 7386214 και μετά 7385969.

Παρατηρείται πως μετά το Cutadapt αφαιρέθηκαν συνολικά 245 αλληλουχίες από το συγκεκριμένο δείγμα B8.

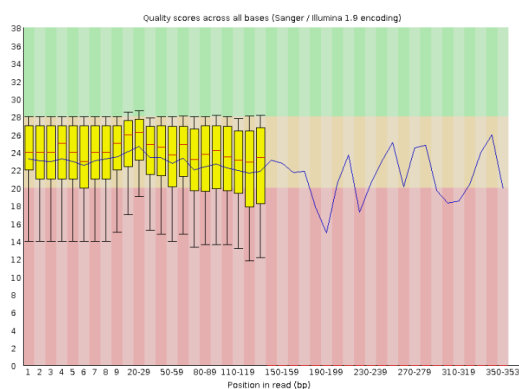
Συνολικά, οι διαφορές που παρατηρήθηκαν στον αριθμό αλληλουχιών σε όλα τα δείγματα της μελέτης, μετά την χρήση του Cutadapt, αναφέρονται στον Πίνακα 1.

Πίνακας 1. Συνολικός αριθμός αλληλουχιών στα δείγματα αυτής της μελέτης πριν και μετά τη χρήση του Cutadapt.

ΔΕΙΓΜΑ	ΠΡΙΝ ΤΟ CUTADAPT	ΜΕΤΑ ΤΟ CUTADAPT	ΑΛΛΗΛΟΥΧΙΕΣ ΠΟΥ ΑΦΑΙΡΕΘΗΚΑΝ ΑΠΟ ΤΟ CUTADAPT
A6	7609317	7608990	327
B1	9092023	9091554	469
B8	7386214	7385969	245
B6	356	356	0
C16	1634468	1634448	20
D16	1023976	1023950	26

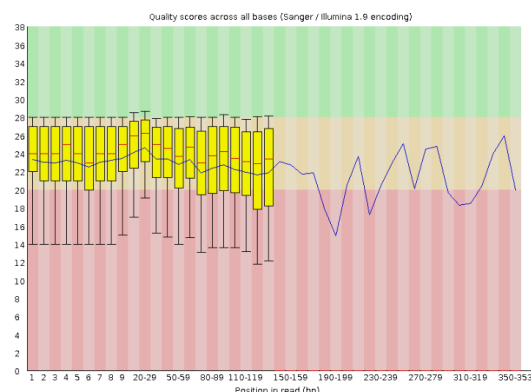
Η εξαιρετικά μικρές διαφορές του συνολικού αριθμού αλληλουχιών πριν και μετά τη χρήση του Cutadapt, της τάξης του 0,00% - 0,005%, όπως αναμενόταν δεν επηρέασαν τις υπόλοιπες παραμέτρους του ποιοτικού ελέγχου FastQC σε κανένα δείγμα (ενδεικτικές Εικόνες 88 – 101)

Per base sequence quality



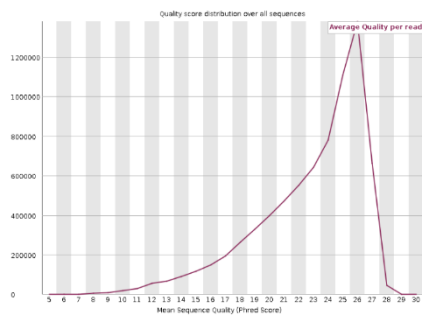
Εικόνα 88. Ποιότητα ανα βάση δείγματος B8 πριν το Cutadapt

Per base sequence quality



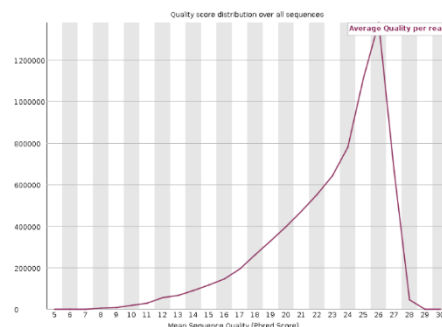
Εικόνα 89. Ποιότητα ανα βάση δείγματος B8 μετά το Cutadapt

Per sequence quality scores



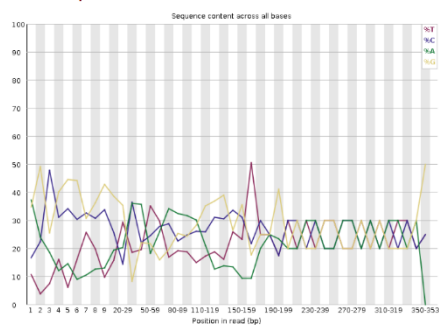
Εικόνα 90. Ποιότητα ανά αλληλουχία, δείγματος B8 πριν το Cutadapt

Per sequence quality scores



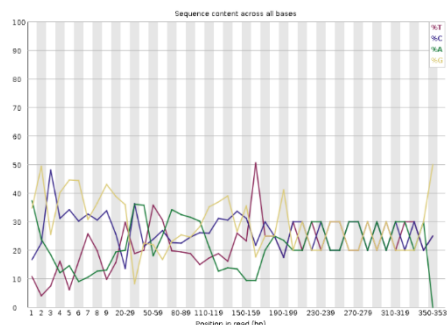
Εικόνα 91. Ποιότητα ανά αλληλουχία, δείγματος B8 μετά το Cutadapt

Per base sequence content



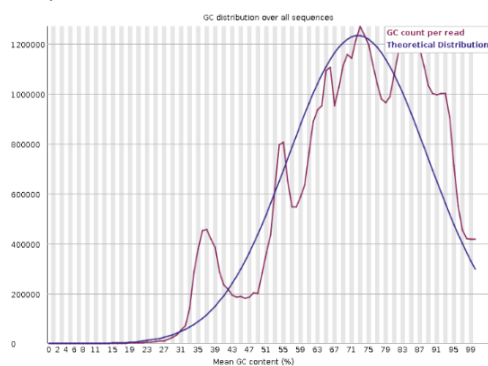
Εικόνα 92. Περιεχόμενο αλληλουχίας ανα βάση δείγματος B8 πριν το Cutadapt

Per base sequence content



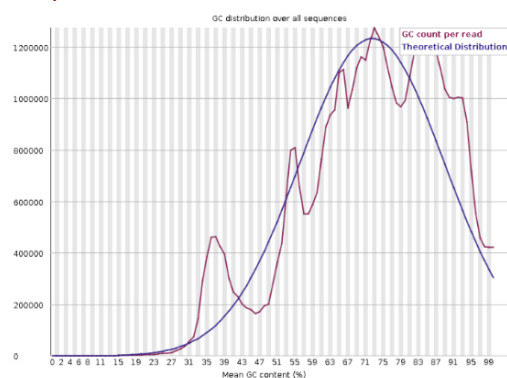
Εικόνα 93. Περιεχόμενο αλληλουχίας ανα βάση δείγματος B8 μετά το Cutadapt

Per sequence GC content



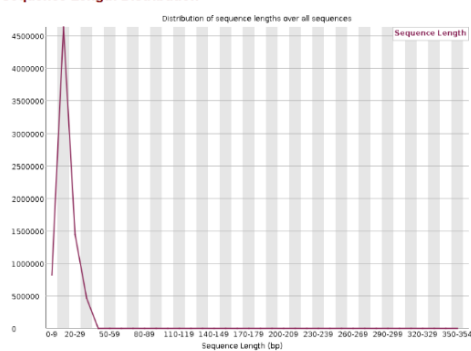
Εικόνα 94. Περιεχόμενο GC ανά βάση δείγματος B8 πριν το Cutadapt

Per sequence GC content



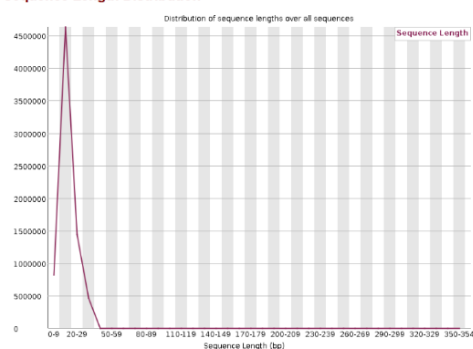
Εικόνα 95. Περιεχόμενο GC ανά βάση δείγματος B8 μετά το Cutadapt

Sequence Length Distribution

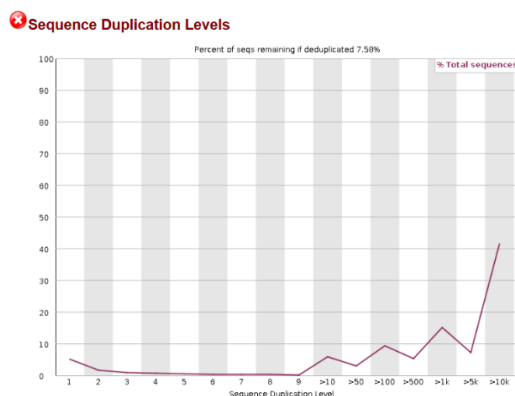


Εικόνα 96. Κατανομή μήκους ακολουθίας δείγματος B8 πριν το Cutadapt

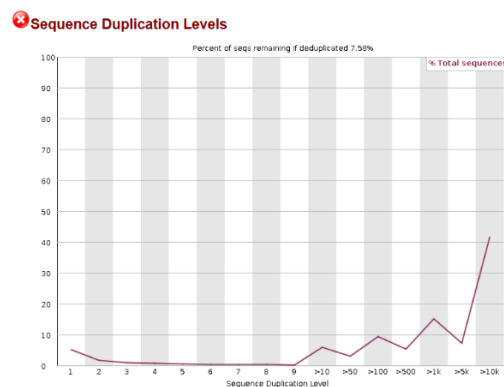
Sequence Length Distribution



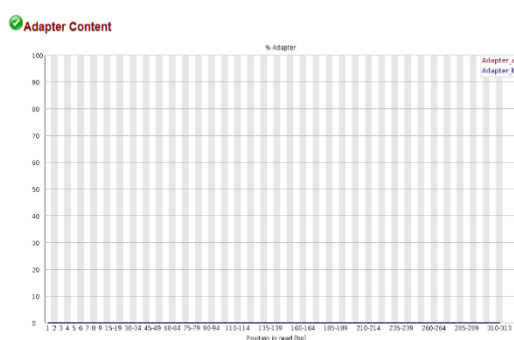
Εικόνα 97. Κατανομή μήκους ακολουθίας δείγματος B8 μετά το Cutadapt



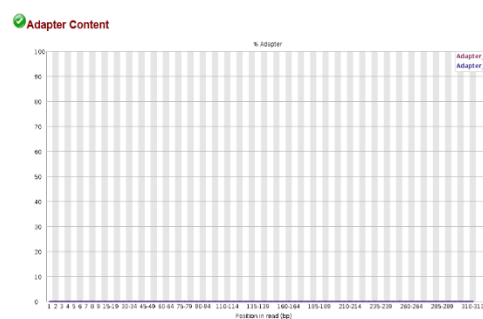
Εικόνα 98. Επίπεδο διπλασιασμού ακολουθίας δείγματος B8 πριν το Cutadapt



Εικόνα 99. Επίπεδο διπλασιασμού ακολουθίας δείγματος B8 μετά το Cutadapt



Εικόνα 100. Περιεχόμενο σε προεπιλεγμένους FastQC αντάπτορες, δείγματος B6 πριν το Cutadapt



Εικόνα 101. Περιεχόμενο σε προεπιλεγμένους FastQC αντάπτορες, δείγματος B6 μετά το Cutadapt

4.4 Αφαίρεση δείγματος B6

Το δείγμα ελέγχου B6 έχει ένα πολύ μικρό αριθμό αλληλουχιών, όπως είχε παρατηρηθεί και κατά τον ποιοτικό έλεγχο FastQC, με τον αριθμό τους να φθάνει μόλις τις 356. (κεφ. 4.2.1)

Περαιτέρω ανάλυση του δείγματος μέσω του MiRDeep2 (κεφ. 3.2.6) έδειξε μόλις 6 γνωστά miRNA με το κάθε ένα να έχει μόλις ένα read και MiRdeep2 score < 4, τη στιγμή που τα υπόλοιπα δείγματα είχαν έως και δεκάδες χιλιάδες reads σε πολλές δεκάδες γνωστά miRNAs το καθένα.

Τα ανωτέρω καθιστούν απαραίτητη την μη συμπερίληψή του στη συνέχεια της ανάλυσης αυτής, καθώς είναι σαφές ότι η ποιότητα του δείγματος δεν είναι αποδεκτή.

Η αφαίρεση αυτού του δείγματος διασφαλίζει ότι η ανάλυση επικεντρώνεται σε δείγματα με επαρκή δεδομένα και υψηλότερη ποιότητα, εξασφαλίζοντας την εγκυρότητα των ευρημάτων.

4.5 Αποτελέσματα ανάλυσης διαφορικής έκφρασης miRNA στα δείγματα

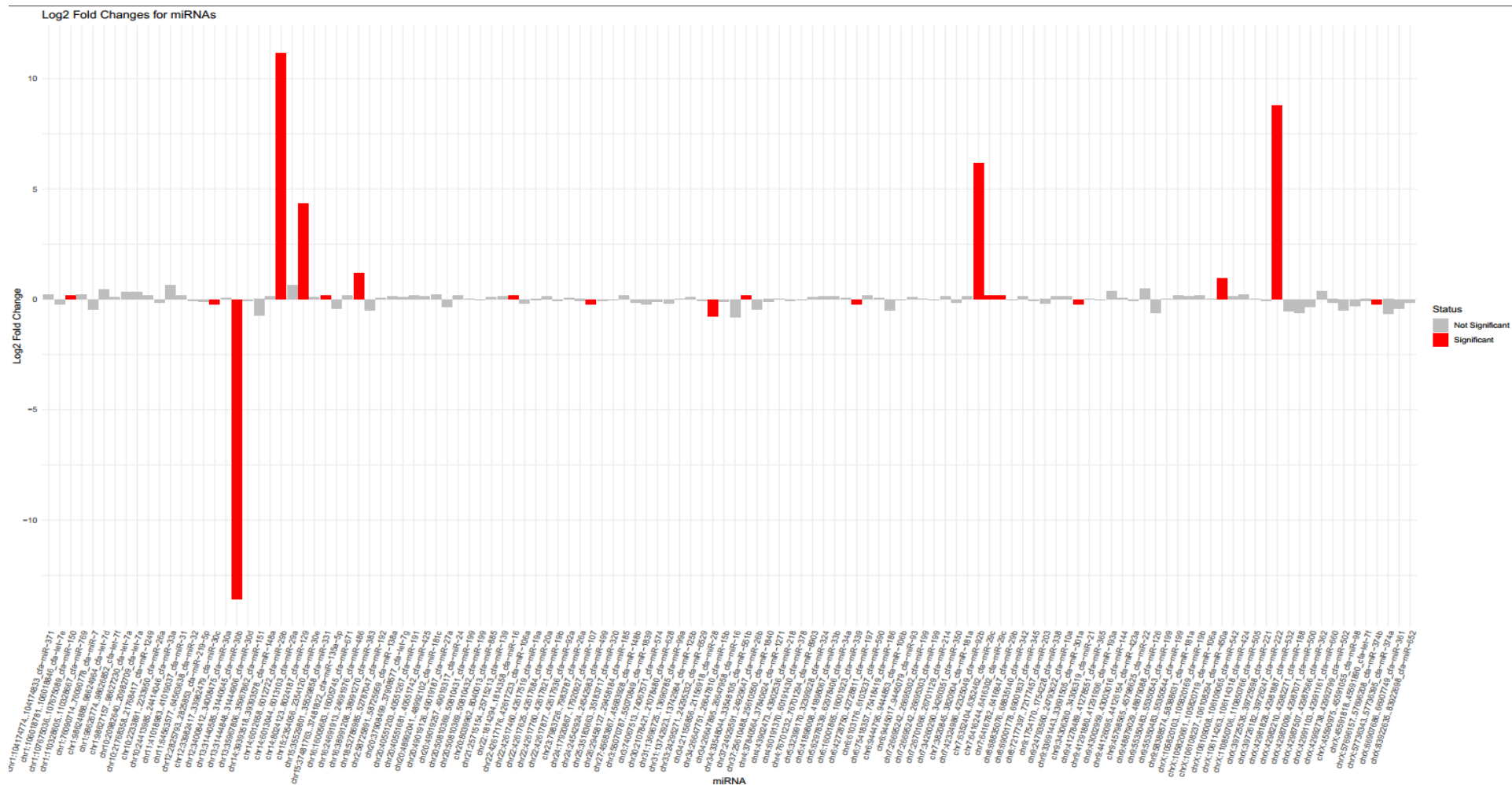
Κατόπιν διαφορικής ανάλυσης σύμφωνα με τη μεθοδολογία, στα δείγματα μολυσμένων με *Leishmania infantum* σκύλων της μελέτης, ταυτοποιήθηκαν τα παρακάτω miRNA, στα οποία παρατηρήθηκαν στατιστικά σημαντικές διαφοροποιήσεις στην έκφραση τους, σε σχέση με την έκφραση των αντίστοιχων miRNA στα δείγματα υγιών σκύλων (Πίνακας 2):

Πίνακας 2. miRNA στα οποία παρατηρήθηκαν στατιστικά σημαντικές διαφοροποιήσεις έκφρασης στα δείγματα μολυσμένων με *Leishmania infantum* σε σχέση με τα αντίστοιχα στα δείγματα υγιών σκύλων, με τις τιμές Log2 Fold Change, p-value και padj.

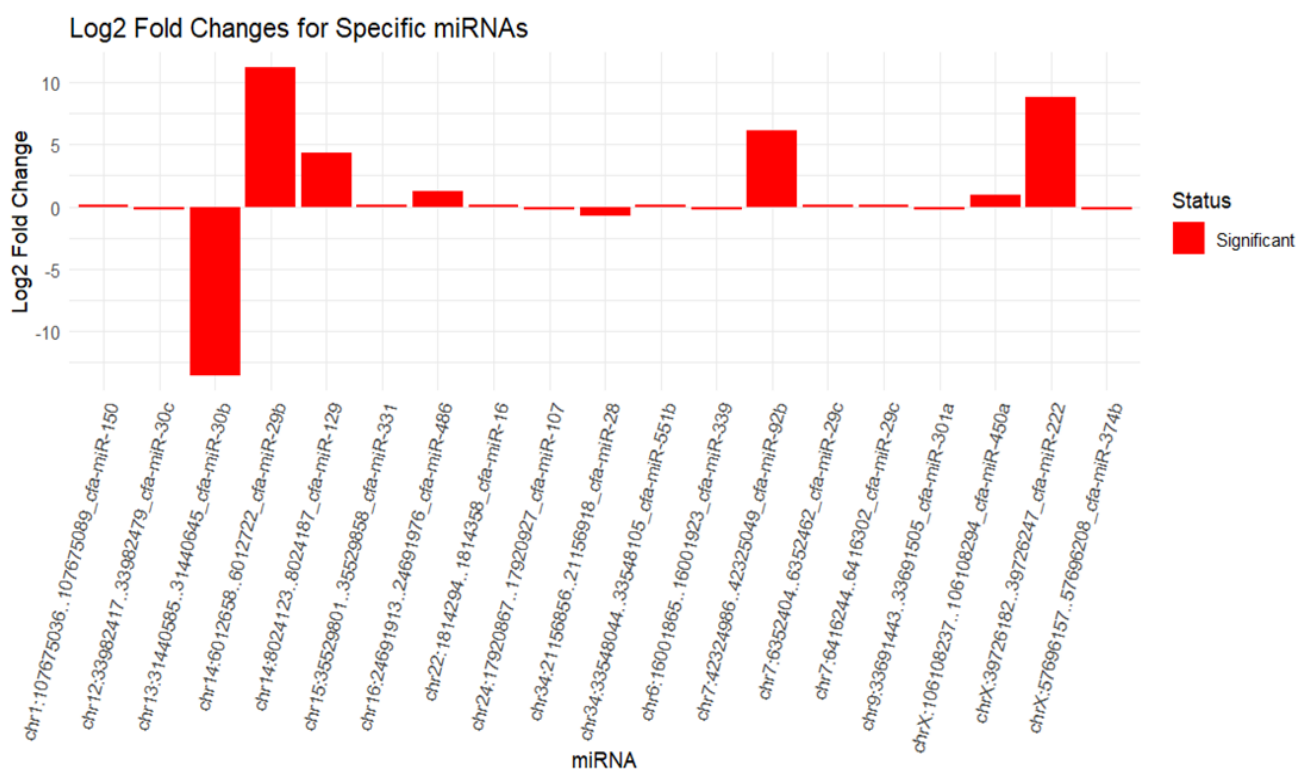
miRNA	log2FoldChange	pvalue	padj	Status	Expression
cfa-miR-107	-0,2213599	0,000000002412	0,000000060903	Significant	Downregulated
cfa-miR-129	4,3301066	0,003213265434	0,032453980888	Significant	Upregulated
cfa-miR-150	0,1650904	0,010663328576	0,089749682180	Significant	Upregulated
cfa-miR-16	0,1688446	0,002363813763	0,026527243344	Significant	Upregulated
cfa-miR-222	8,788353	0,000003201735	0,000046196460	Significant	Upregulated
cfa-miR-28	-0,7722539	0,005605565651	0,051469284615	Significant	Downregulated
cfa-miR-29b	11,1746854	0,000000000137	0,000000013871	Significant	Upregulated
cfa-miR-29c-2	0,17711	0,028075973630	0,166804313921	Significant	Upregulated
cfa-miR-29c-1	0,17711	0,028075973630	0,166804313921	Significant	Upregulated
cfa-miR-301a	-0,2204333	0,044236231724	0,235150494952	Significant	Downregulated
cfa-miR-30b	-13,5672337	0,000000004179	0,000000084425	Significant	Downregulated
cfa-miR-30c	-0,2214598	0,000000000562	0,000000018935	Significant	Downregulated
cfa-miR-331	0,1559962	0,000002593652	0,000043659804	Significant	Upregulated
cfa-miR-339	-0,2206736	0,000000000478	0,000000018935	Significant	Downregulated
cfa-miR-374b	-0,2204333	0,044236231724	0,235150494952	Significant	Downregulated
cfa-miR-450a	0,9428819	0,021969352834	0,155077223073	Significant	Upregulated
cfa-miR-486	1,2013269	0,023031270753	0,155077223073	Significant	Upregulated
cfa-miR-551b	0,1645193	0,012806922045	0,099499932808	Significant	Upregulated
cfa-miR-92b	6,155427	0,000095689568	0,001208080802	Significant	Upregulated

Στην Εικόνα 102 απεικονίζονται όλα τα miRNA τα οποία ταυτοποιήθηκαν στη μελέτη μαζί με τη διαφορική τους έκφραση. Τα στατιστικά σημαντικά miRNA απεικονίζονται με κόκκινο χρώμα.

Το ραβδόγραμμα που απεικονίζει μόνο τα στατιστικά σημαντικά miRNA (Πίνακας 2), δίνεται στην εικόνα 103.



Εικόνα 102. Ραβδόγραμμα Log2 Fold Change miRNAs που ταυτοποιήθηκαν στην ανάλυση (με κόκκινο χρώμα τα στατιστικά σημαντικά)



Εικόνα 103. Ραβδόγραμμα Log2 Fold Change μόνο για τα στατιστικά σημαντικά miRNAs της ανάλυσης

Συγκεντρωτικά τα miRNA που παρουσιάζουν *απορρύθμιση* (ανωρρύθμιση δηλαδή αυξημένη έκφραση ή *κατωρρύθμιση* δηλαδή μειωμένη έκφραση) στους μολυσμένους με λεισμάνια σκύλους, στη μελέτη αυτή, είναι τα παρακάτω:

↑ Αυξημένη έκφραση: cfa-miR-29b, cfa-miR-129, cfa-miR-222, cfa-miR-16, cfa-miR-331, cfa-miR-92b, cfa-miR-29c-1, cfa-miR-29c-2, cfa-miR-150, cfa-miR-486, cfa-miR-551b, cfa-miR-450a

↓ Μειωμένη έκφραση: cfa-miR-30c, cfa-miR-30b, cfa-miR-107, cfa-miR-339, cfa-miR-28, cfa-miR-301a, cfa-miR-374b

5. Συμπεράσματα

Κατά την ανάλυση των miRNA που προέρχονταν από τα εξωσωμάτια υγείων και μολυσμένων με λεϊσμάνια σκύλων, παρατηρήθηκε μία σημαντική διαφορά στον αριθμό των reads μεταξύ των δειγμάτων ελέγχου και των μολυσμένων δειγμάτων. Αυτή η σημαντική διαφορά μπορεί ίσως να εξηγηθεί από τις διαφορές στη συγκέντρωση εξωκυττάρων κυστιδίων (EVs) μεταξύ των δύο ομάδων. Σύμφωνα με μελέτη παρατηρήθηκε σημαντική αύξηση των EVs σε σκύλους μολυσμένους με *Leishmania Infantum* εν συγκρίσει με υγιείς σκύλους [56].

Αυτή η αύξηση πιθανώς αντικατοπτρίζει την ενισχυμένη ανοσολογική δραστηριότητα και μεταγραφική έκφραση που παρατηρείται σε μολυσμένους σκύλους, εξηγώντας των υψηλότερο αριθμό των reads (περίπου 7 εκατομμύρια). Αντίθετα, ο χαμηλός αριθμός των reads που συναντάται στα δείγματα ελέγχου (περίπου 1 εκατομμύριο ή λιγότερο) μπορεί να σχετίζεται με την απουσία του παρασίτου και μειωμένη ανοσολογική απόκριση.

Στην έρευνά μας, μετά την ανάλυση διαφορικής έκφρασης παρατηρήθηκαν 19 miRNA, τα οποία παρουσίασαν σημαντική διαφορά στην έκφραση μεταξύ των μολυσμένων δειγμάτων και των δειγμάτων ελέγχου:

Τα cfa-miR-29b, cfa-miR-129, cfa-miR-222, cfa-miR-16, cfa-miR-331, cfa-miR-92b, cfa-miR-29c-1, cfa-miR-29c-2, cfa-miR-150, cfa-miR-486, cfa-miR-551b, cfa-miR-450a παρουσίασαν αυξημένη έκφραση στα δείγματα των σκύλων μολυσμένα με λεϊσμάνια, ενώ μειωμένη έκφραση παρουσίασαν τα cfa-miR-30c, cfa-miR-30b, cfa-miR-107, cfa-miR-339, cfa-miR-28, cfa-miR-301a, cfa-miR-374b.

Τα cfa-miR-29b, cfa-miR-29c-1 και cfa-miR-29c-2 παρατηρήθηκαν ανωρρυθμισμένα. Ειδικότερα το cfa-miR-29b ήταν ιδιαίτερα αυξημένο στα μολυσμένα δείγματα ($\log_2\text{FoldChange} = 11,1746854$). Όπως περιγράφεται στη μελέτη των R.K.Pandey *et al.*, [57] τα miRNA-29b και miRNA-29c έχουν ως στόχο το γονίδιο *TBX21* (T-box transcription factor 21), το οποίο κωδικοποιεί τον μεταγραφικό παράγοντα T-bet. Ο T-bet είναι βασικός ρυθμιστής της διαφοροποίησης των Th1 κυττάρων και της παραγωγής ιντερφερόνης- γ (IFN- γ), ενός κρίσιμου μορίου για την ανοσολογική απόκριση κατά ενδοκυτταρικών παθογόνων. Η μείωση του μεταγραφικού παράγοντα T-bet, οδηγεί στη μείωση της Th1 απόκρισης και της παραγωγής IFN- γ που θα μπορούσε να διευκολύνει την επιβίωση του παρασίτου.

Η έκφραση του cfa-miR-129 παρατηρήθηκε αυξημένη στην ανάλυση, γεγονός που υποδεικνύει τον πιθανό ρόλο του στην παθογένεση μέσω της ρύθμισης της αυτοφαγίας και άλλων μονοπατιών. Το miR-129 μπορεί να επηρεάζει τη διαδικασία της αυτοφαγίας, η οποία είναι κρίσιμη για την απομάκρυνση των παθογόνων και την αποκατάσταση της κυτταρικής ισορροπίας [58] [59]. Το cfa-miR-222 βρέθηκε ανωρρυθμισμένο στα δείγματα μολυσμένων σκύλων στην ανάλυση. Η ανωρρύθμισή του συγκεκριμένου miRNA έχει επίσης παρατηρηθεί σε ορό σκύλων μολυσμένων με *T. canis* [60] και σε ηπατικά δείγματα ασθενών με *Schistosoma japonicum* [61] γεγονός που πιθανόν να υποδηλώνει έναν κοινό μηχανισμό απόκρισης του οργανισμού σε

παρασιτικές λοιμώξεις. Ο μηχανισμός δράσης του συγκεκριμένου ρυθμιστικού μορίου σε λοίμωξη από *L. Infantum* δεν έχει ακόμη διευκρινιστεί.

Στα δείγματα παρατηρήθηκε επίσης ανωρρύθμιση του cfa-miR-16. Αντίστοιχη ανωρρύθμιση του miR-16 έχει αναφερθεί σε μελέτη εξωσωμάτων από δενδριτικά κύτταρα (DCs) ποντικών που μολύνθηκαν με στελέχη της *Leishmania donovani*, καθώς και σε ανθρώπινα μακροφάγα μολυσμένα με το παράσιτο [62]. Το miR-16 κατέστειλε την έκκριση και την έκφραση προφλεγμονωδών παραγόντων, όπως η ιντερλευκίνη (IL)-6 και ο παράγοντας νέκρωσης όγκων-α (TNF-α), ενώ ενισχύει την έκκριση και την έκφραση της αντιφλεγμονώδους κυτταροκίνης IL-10. [63].

Επιπλέον, αα ακόλουθα miRNAs παρατηρήθηκαν απορρυθμισμένα:

- cfa-miR-331, cfa-miR-486 ανωρρυθμισμένα στα δείγματα μολυσμένων σκύλων και
- cfa-miR-28 κατωρρυθμισμένο στα δείγματα μολυσμένων σκύλων

Τα συγκεκριμένα miRNAs έχουν βρεθεί απορρυθμισμένα σε τουλάχιστον δύο διαφορετικά σημεία μόλυνσης από *Leishmania major* σε ανθρώπινα μακροφάγα. Αυτή τους η απορρύθμιση είναι πιθανό να διαδραματίζει σημαντικό ρόλο στη ρύθμιση των λειτουργιών των μακροφάγων μετά από μόλυνση από το παράσιτο της λεισμάνιας [64].

Το miRNA cfa-miR-92b, στη μελέτη μας βρέθηκε ανωρρυθμισμένο με Log2foldchange 6,155 στα μολυσμένα δείγματα σε σύγκριση με τα υγιή. Η ανωρρύθμιση του miR-92b έχει συνδεθεί με δείγματα εξωσωμάτων από ασθενείς με διάφορους τύπους καρκίνου, όπως το μελάνωμα [65], τον καρκίνο του στομάχου [66], το αρθρικό σάρκωμα [67] και τον μικροκυτταρικό καρκίνο του πνεύμονα [68]. Το MiR-92b-3p έχει επιβεβαιωθεί ότι διαδραματίζει σημαντικό ρόλο στον πολλαπλασιασμό, την εισβολή και τη μετανάστευση του καρκίνου [66]. Ενδιαφέρον παρουσιάζει η πρόσφατη αναφορά της έκφρασης του cfa-miR-92b στον ορό σκύλων με μόλυνση από το παράσιτο *Toxocara canis*. Ωστόσο τα δεδομένα για τον ρόλο του miR-92b σε άλλες παρασιτικές λοιμώξεις, όπως αυτή που προκαλείται από τη *Leishmania infantum*, παραμένουν περιορισμένα [60].

Η έκφραση του cfa-miR-551b παρατηρήθηκε επίσης αυξημένη στα δείγματα μολυσμένων σκύλων. Το miR-551 έχει επίσης παρατηρηθεί αυξημένο σε δενδριτικά κύτταρα μολυσμένα με *L. donovani* όπου η ανωρρύθμιση αυτή μπορεί να παρεμβαίνει στη σηματοδότηση TLR4, η οποία είναι κρίσιμη για την ανοσοαπόκριση στο παράσιτο της λεισμάνιας [69].

Αντιθέτως, το cfa-miR-30c παρατηρήθηκε κατωρρυθμισμένο στα δείγματα μολυσμένων σκύλων. Η μείωση των επιπέδων του miR-30c οδηγεί σε αυξημένη έκφραση του ATG4, καθώς το miR-30c συνήθως καταστέλλει τη δράση του. Η αυξημένη δραστηριότητα του ATG4 ενισχύει τη διαδικασία της αυτοφαγίας, η οποία είναι κρίσιμη για την επιβίωση των παρασίτων *Leishmania* υπό συνθήκες στρες. Αυτό έχει ως αποτέλεσμα τη βελτίωση της επιβίωσης και του πολλαπλασιασμού των

παρασίτων, καθώς η αυτοφαγία προωθεί την ανακύκλωση κυτταρικών συστατικών και μειώνει την απόπτωση (κυτταρικό θάνατο) [70].

Το miRNA miR-30b το οποίο συναντάται απορρυθμισμένο σε λοιμώξεις από λεισμάνια [71], [72], σε αυτή την μελέτη, παρατηρείται να έχει μειωμένη έκφραση στα δείγματα μολυσμένων σκύλων. Το cfa-miR-30b θα μπορούσε ίσως να σχετίζεται με το ρόλο του στη στόχευση της Rab32, η οποία εμπλέκεται στη βιογένεση των λιπιδικών σωμάτων. Αυτά τα λιπιδικά σώματα αξιοποιούνται πιθανώς από το παράσιτο ως πηγή λιπιδίων και ενέργειας για την επιβίωσή του [73] [74].

Το miR-301a ρυθμίζει προς τα κάτω την έκφραση διαφορετικών πρωτεϊνών στο μονοπάτι NF-κΒ στα κύτταρα του ανοσοποιητικού συστήματος, έχοντας ως γονίδιο στόχο το NKRF (NF-κΒ repressing factor) (Κατασταλτικός Παράγοντας NF-κΒ). Επηρεάζει έτσι την παραγωγή φλεγμονωδών μεσολαβητών όπως η ιντερλευκίνη-1β (IL-1β) και ο παράγοντας νέκρωσης όγκου-α (TNF-α).

Σε αυτή την μελέτη, παρατηρήθηκε μείωση της έκφρασης του cfa-miRNA-301a στα δείγματα μολυσμένων σκύλων, το οποίο μπορεί να οδηγήσει στην αύξηση της έκφρασης του γονιδίου NKRF, αναστέλλοντας έτσι την ενεργοποίηση του NF-κΒ και με αποτέλεσμα την καταστολή της έκφρασης των προφλεγμονωδών γονιδίων [75].

Η έκφραση του cfa-miR-374b στα δείγματα μολυσμένων σκύλων βρέθηκε μειωμένη στη παρούσα μελέτη. Η μειωμένη έκφραση του miR-374b έχει παρατηρηθεί ότι μειώνει τα επίπεδα έκφρασης της IFN-γ καθώς και της IL-17A ενώ παράλληλα αυξάνει τα επίπεδα της κυτταροκίνης IL-10, η οποία έχει αντιφλεγμονώδη δράση, στα CD4⁺ T κύτταρα στον σπλήνα ποντικών μολυσμένων με το παράσιτο *Echinococcus granulosus* [76]. Οι συγκεκριμένες κυτταροκίνες διαδραματίζουν σημαντικό ρόλο στην ανοσολογική απόκριση στο παράσιτο της *Leishmania Infantum* [77] [15].

Το cfa-miR-107 παρατηρήθηκε κατωρρυθμισμένο στα δείγματα μολυσμένων σκύλων. Σε έρευνα σε ανθρώπους μολυσμένους με ελονοσία, η χρήση συνθετικών μιμητών του miR-107 έδειξε ότι αναστέλλει τον πολλαπλασιασμό και την ανάπτυξη του παράσιτου στο εσωτερικό των ερυθροκυττάρων [78]. Η παραπάνω παρατήρηση υποδεικνύει ότι η κατωρρύθμισή του miR-107 θα μπορούσε να επηρεάζει μονοπάτια ανοσολογικής σηματοδότησης, χρήζοντας περαιτέρω διερεύνησης.

Συνοψίζοντας, αυτή η μελέτη ενισχύει τη σημασία των miRNAs ως ρυθμιστών στην ανοσολογική απόκριση κατά της *Leishmania*. Η διαφοροποιημένη έκφραση των miRNAs επιβεβαιώνει τον ρόλο τους στις αλληλεπιδράσεις ξενιστή-παρασίτου. Ωστόσο, η έλλειψη παρατήρησης διαφορετικής έκφρασης ορισμένων miRNAs όπως τα cfa-miR-21, cfa-miR-146a και cfa-miR-122, τα οποία έχουν παρατηρηθεί σε αντίστοιχες μελέτες [56] [79], υποδεικνύει τη μεταβλητότητα μεταξύ διαφορετικών πειραματικών πλαισίων. Ο περιορισμένος αριθμός δειγμάτων είναι πιθανόν να έχει περιορίσει την ικανότητα εντοπισμού αυτών των miRNAs σε στατιστικά σημαντικά επίπεδα. Παράμετροι όπως η ατομική διακύμανση έκφρασης των miRNAs, η διαφοροποιημένη έκφραση ανάλογα με την ηλικία, αλλά και ο χρόνος συλλογής δειγμάτων ανάλογα με το στάδιο εξέλιξης της νόσου, αναδεικνύονται ως

σημαντικότεροι λόγω του περιορισμένου αριθμού δειγμάτων. Επιπλέον αναλύσεις μεγαλύτερου αριθμού δειγμάτων ενδείκνυνται, προκειμένου να διασαφηνιστούν πλήρως οι ρόλοι των εξωσωμικών miRNAs και να κατανοηθεί καλύτερα η συμμετοχή τους στη διαδικασία της λοίμωξης από *Leishmania infantum*. Μελλοντική έρευνα θα πρέπει να επικεντρωθεί στην επικύρωση των ρόλων τους ως στόχους θεραπείας και στην αξιολόγηση της διαγνωστικής τους αξίας ως βιοδείκτες.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] K. Ines, A. C. Mark and H. K. James, “New World and Old World Leishmania Infections: A Practical Review,,” pp. 579-593,, 2015.
- [2] M. Y. Manuel , M. S. Joaquina and C. L. Victoriano , “Canine Leishmaniasis: Update on Epidemiology, Diagnosis, Treatment, and Prevention,” *MDPI*, 1 July 2022.
- [3] K. I. Demba, M. S. Ibrahim, B. B. W. André , C. M. Gunter, A. C. Cheick , . T. Bourama , C. B. John και D. Seydou, «Distribution and Diversity of Sand Fly Species (Diptera: Psychodidae, Phlebotominae) in Two Geoclimatic Zones of Chad,» 12 January 2022.
- [4] R. Dr Richard , D. Jean-Claude , L. Hechmi , P. Claude , A. Bruce and B. Simon , “Cutaneous leishmaniasis,” 20 August 2007.
- [5] A. A. Carotenuto, H. R. G. D. and . K. Geelan-Hansen, “Clinical Diagnosis and Management of Mucosal Leishmaniasis in the Context of a Global Pandemic: A Case Report.,” 2022.
- [6] S. D. Scarpini, D. Arianna , T. Camilla, B. Carlotta , . M. Fraia, Z. Daniele , P. Luca , G. Monia , C. Cinzia , . P. Arcangelo and L. Marcello , “Visceral Leishmaniasis: Epidemiology, Diagnosis, and Treatment Regimens in Different Geographical Areas with a Focus on Pediatrics,” 21 September 2022.
- [7] R. R. Ribeiro, M. S. M. Michalick, M. E. da Silva και et al., «Canine Leishmaniasis: An Overview of the Current Status and Strategies for Control,» *BioMed Research International*, p. 12, 2018.
- [8] M. Sarah , F. Katherine , S. Sara , F. H. M. Andrés , N. Sabrina , . R. Poornima και A. S. José, «A Review of Leishmaniasis: Current Knowledge and Future Directions,» *Springer Link*, 17 March 2021.
- [9] S. Jack και . G. Keith, «Shape, form, function and Leishmania pathogenicity: from textbook descriptions to biological understanding,» 2017.
- [10] C. Cinzia , D. T. Filipe, J. N. Matthew και O. Domenico, «The past, present, and future of Leishmania genomics and transcriptomics,» 2015.
- [11] B. Kazemi, «Genomic organization of leishmania species,» *Iranian journal of parasitology*, 2011.
- [12] M. N. Saridomichelakis, «Advances in the pathogenesis of canine leishmaniosis: epidemiologic and diagnostic implications,» 2009.
- [13] O. Martin, J. G. David και . F. Geneviève , «Subversion Mechanisms by Which Leishmania Parasites Can Escape the Host Immune Response: a Signaling Point of View,» *American Society For Microbiology*, τόμ. 18, pp. 293-305, 2005.

- [14] C. L. Barbiéri, «Immunology of canine leishmaniasis,» *Parasite immunology*, p. 329–337, 2006.
- [15] S. Hosein, D. P. Blake και L. Solano-Gallego, «Insights on adaptive and innate immunity in canine leishmaniosis,» *Parasitology*, p. 95–115, 2017.
- [16] B. M. Roatt, J. M. d. O. Cardoso, R. C. F. De Brito, W. Coura-Vital, R. D. de Oliveira Aguiar-Soares και A. B. Reis, «Recent advances and new strategies on leishmaniasis treatment,» 2020.
- [17] P. Swetalina, R. A. Schwartz, A. Patil, M. Goldust και S. Grabbe, «Treatment options for leishmaniasis,» *Clinical and Experimental Dermatology*, τόμ. 47, αρ. 3, p. 516–521, 2022.
- [18] L. Solano-Gallego, . G. Miró, A. Koutinas, L. Cardoso, M. G. Pennisi, L. Ferrer, P. Bourdeau, . G. Oliva, G. Baneth και The LeishVet Group, «LeishVet guidelines for the practical management of canine leishmaniosis,» *Parasites & vectors*, 2011.
- [19] R. Dinc, «Leishmania Vaccines: the Current Situation with Its Promising Aspect for the Future,» *The Korean journal of parasitology*, p. 379–391, 2022.
- [20] E. A. Coelho και M. Christodoulides, «Vaccines for Canine Leishmaniasis,» σε *Vaccines for Neglected Pathogens: Strategies, Achievements and Challenges*, Springer, 2023.
- [21] R. Velez και M. Gállego, «Commercially approved vaccines for canine leishmaniosis: a review of available data on their safety and efficacy,» *Tropical medicine & international health : TM & IH*, p. 540–557, 2020.
- [22] K. Saliminejad, H. R. K. Khorshid, S. F. Fard και S. H. Ghaffari, «An overview of microRNAs: Biology, functions, therapeutics, and analysis methods,» *Cellular Physiology*, τόμ. 234, αρ. 5, pp. 5451-5465, 2019.
- [23] J. O'Brien, H. Hayder, Y. Zayed και C. Peng , «Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation,» *Frontiers in endocrinology*, τόμ. 9, p. 402, 2018.
- [24] M. Rojas-Pirela, . D. Andrade-Alviárez, L. Medina, C. Castillo, A. Liempi, J. Guerrero-Muñoz, Y. Ortega, J. D. Maya, V. Rojas, W. Quiñones, P. A. Michels και U. Kemmerling, «MicroRNAs: master regulators in host-parasitic protist interactions,» *Open biology*, τόμ. 12, αρ. 6, 2022.
- [25] M. Ha και V. N. Kim, «Regulation of microRNA biogenesis,» *Nature reviews. Molecular cell biology*, , p. 509–524, 2014.
- [26] L. A. MacFarlane και P. R. Murphy, «MicroRNA: Biogenesis, Function and Role in Cancer,» *Current genomics*, p. 537–561, 2010.
- [27] M. Fareh, K.-H. Yeom, A. C. Haagsma, S. Chauhan, I. Heo και C. Joo, «TRBP ensures efficient Dicer processing of precursor microRNA in RNA-crowded environments,» *Nature Communications* , 2016.

- [28] J. Krol, I. Loedige και W. Filipowicz, «The widespread regulation of microRNA biogenesis, function and decay,» *Nature Reviews Genetics*, τόμ. 11, pp. 597-610, 2010.
- [29] K. Kim, Y. L. Sik και R. C. W, «Conversion of pre-RISC to holo-RISC by Ago2 during assembly of RNAi complexes,» *RNA*, p. 22–29., 2007.
- [30] A. Vishnoi και S. Rani , «MiRNA Biogenesis and Regulation of Diseases: An Overview,» *Methods in molecular biology*, pp. 1-10, 2017.
- [31] J. E. Braun, E. Huntzinger και E. Izaurralde, «The Role of GW182 Proteins in miRNA-Mediated Gene Silencing,» σε *Ten Years of Progress in GW/P Body Research*, M. J. Fritzler και E. K. L. Chan, Επιμ., New York, NY: Springer, New York,, 2013, pp. 147-163.
- [32] T. Desvignes, P. Batzel, E. Berezikov, . K. Eilbeck, J. T. Eppig, M. S. McAndrews, A. Singer και J. H. Postlethwait, «miRNA Nomenclature: A View Incorporating Genetic Origins, Biosynthetic Pathways, and Sequence Variants,» *Trends in genetics*, p. 613–626, 2015.
- [33] H. Budak, R. Bulut, M. Kantar και B. Alptekin, «MicroRNA nomenclature and the need for a revised naming prescription,» *Briefings in functional genomics*, p. 65–71, 2016.
- [34] H. M. Nail, C.-C. Chiu, C.-H. Leung και et al, «Exosomal miRNA-mediated intercellular communications and immunomodulatory effects in tumor microenvironments,» *Journal of Biomedical Science*, τόμ. 30, αρ. 1, p. 69, 2023.
- [35] S. Rani, A. E. Ryan, M. D. Griffin και T. Ritter, «Mesenchymal Stem Cell-derived Extracellular Vesicles: Toward Cell-free Therapeutic Applications,» *Molecular Therapy*, pp. 812-823, 2015.
- [36] M. Sheta , E. A. Taha , Y. Lu και T. Eguchi, «Extracellular Vesicles: New Classification and Tumor Immunosuppression,» *Biology*, τόμ. 12, p. 1: 110., 2023.
- [37] G. K. Patel, M. A. Khan, . H. Zubair, S. K. Srivastava, . M. Khushman, S. Singh and A. P. Singh, “Comparative analysis of exosome isolation methods using culture supernatant for optimum yield, purity and downstream applications,” *Scientific Reports*, vol. 9, p. 5335, 2019.
- [38] Y. J. Lee, K. J. Shin και Y. C. Chae, «Regulation of cargo selection in exosome biogenesis and its biomedical applications in cancer,» *Experimental & Molecular Medicine* , p. 877–889, 2024.
- [39] J. Gao, . A. Li, J. Hu, . L. Feng, L. Liu και . Z. Shen, «Recent developments in isolating methods for exosomes,» *Frontiers in Bioengineering and Biotechnology*, 2022.
- [40] J. P. Bragato, L. M. Melo, G. L. Venturin, G. T. Rebech, L. E. Garcia, V. M. F. de Lima και V. M. F. de Lima, «Relationship of peripheral blood mononuclear cells miRNA expression and parasitic load in canine visceral leishmaniasis,» *PLOS ONE*, 12 05 2018.
- [41] M. F. Soares, L. M. Melo , J. P. Bragato, A. d. O. Furlan, N. F. Scaramiele, F. L. Lopes και V. M. F. de Lima, «Differential expression of miRNAs in canine peripheral blood

mononuclear cells (PBMC) exposed to *Leishmania infantum* in vitro,» *Research in Veterinary Science*, τόμ. 134, pp. 58-63, January 2021.

- [42] T. Siddika και I. U. Heinemann , «Bringing MicroRNAs to Light: Methods for MicroRNA Quantification and Visualization in Live Cells,» *Frontiers in Bioengineering and Biotechnology*, τόμ. 8, 2021.
- [43] Y. Jiawei, X. Mingcheng, T. Xueke, C. Sheng και Z. Su , «Research advances in the detection of miRNA,» *Journal of Pharmaceutical Analysis*, pp. 217-226, 2019.
- [44] D. Qin, «Next-generation sequencing and its clinical application,» *Cancer biology & medicine*, p. 4–10, February 2019.
- [45] A. K. Gupta και U. G. , «Next generation sequencing and its applications,» σε *Animal Biotechnology (Second Edition)*, 2020, pp. 395-421,.
- [46] L. Yuansheng, S. Xiangzhen, G. Yongshun, L. Yiping, S. Bosheng και Z. Xiangxiang, «Sequence Alignment/Map format: a comprehensive review of approaches and applications,» *Briefings in Bioinformatics*, τόμ. 24, αρ. 5, 2023.
- [47] C. T. Lee και . M. Maragkakis, «SamQL: a structured query language and filtering tool for the SAM/BAM file format,» 2021.
- [48] P. J. Cock, C. J. Fields, N. Goto, M. L. Heuer και P. M. Rice, «The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants,» *Nucleic acids research*,, p. 1767–1771, 2010.
- [49] S. R. Head , K. H. Komori, S. A. LaMere, T. Whisenant, . F. V. Nieuwerburgh, . D. R. Salomon και P. Ordoukhanian, «Library construction for next-generation sequencing: Overviews and challenges,» *BioTechniques*, τόμ. 56(2), pp. 61-4, 66, 68, passim, 2014 .
- [50] M. Martin, «Cutadapt removes adapter sequences from high-throughput sequencing reads,» *EMBnet.journal*, τόμ. 17, pp. 10-12, 2011.
- [51] M. R. Friedländer, . S. D. Mackowiak, N. Li, W. Chen και N. Rajewsky, «miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades,» *Nucleic Acids Research*, τόμ. 40, αρ. 1, pp. 37-52, 1 January 2012.
- [52] L. Unger, V. Jagannathan, A. Pacholewska, T. Leeb και V. Gerber, «Differences in miRNA differential expression in whole blood between horses with sarcoid regression and progression,» *Journal of Veterinary Internal Medicine*, τόμ. 33, pp. 241- 250, 2019.
- [53] M. I. Love, W. Huber και S. Anders, «Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2,» *Genome Biology*, τόμ. 15, p. 550, 2014.
- [54] M. Stephens, «False discovery rates: a new deal,» *Biostatistics*, τόμ. 18, αρ. 2, p. 275–294, 2016.

- [55] W. B. Zheng, L. Cai, Y. Zou, W. W. Gao, Q. Liu και X. Q. Zhu, «Altered miRNA Expression Profiles in the Serum of Beagle Dogs Experimentally Infected with *Toxocara canis*,» *Animals (Basel)*, τόμ. 13, αρ. 2, p. 299, 15 January 2023.
- [56] A. B. da Cruz, F. M. Carneiro, . M. M. Maia και et al., «Dogs with canine visceral leishmaniasis have a boost of extracellular vesicles and miR-21-5p up-expression,» *Parasite Immunology*, τόμ. 45, αρ. 9, 2023.
- [57] R. K. Pandey, S. Sundar και V. K. Prajapati, «Differential Expression of miRNA Regulates T Cell Differentiation and Plasticity During Visceral Leishmaniasis Infection,» *Frontiers in Microbiology*, τόμ. 7, 2016.
- [58] V. Guhe, P. Ingale, A. Tambekar και S. Singh, «Systems biology of autophagy in leishmanial infection and its diverse role in precision medicine,» *Frontiers in Molecular Biosciences*, τόμ. 10, 2023.
- [59] B. Frank, A. Marcu, . A. L. d. O. A. Petersen, H. Weber και et al., «Autophagic digestion of *Leishmania major* by host macrophages is associated with differential expression of BNIP3, CTSE, and the miRNAs miR-101c, miR-129, and miR-210,» *Parasites & Vectors*, 2015.
- [60] . W.-B. Zheng, L. Cai, Y. Zou, W.-W. Gao, Q. Liu και X.-Q. Zhu, «Altered miRNA Expression Profiles in the Serum of Beagle Dogs Experimentally Infected with *Toxocara canis*,» *Animals*, τόμ. 13, 2023.
- [61] S. Paul, L. Ruiz-Manriquez, F. Serrano-Cano και et al., «Human microRNAs in host–parasite interaction: a review,» *3 Biotech*, τόμ. 10, 2020.
- [62] S. Ganguly και et al, «*Leishmania* survives by exporting miR-146a from infected to resident cells to subjugate inflammation.,» *Life science alliance*, τόμ. 5, αρ. 6, 2022.
- [63] L. Yan, M. Liang, X. Hou και et al., «The role of microRNA-16 in the pathogenesis of autoimmune diseases: A comprehensive review,» *Biomedicine & Pharmacotherapy*, τόμ. 112, 2019.
- [64] J. Lemaire, G. Mkannez και et al, «MicroRNA expression profile in human macrophages in response to *Leishmania major* infection.,» *PLoS neglected tropical diseases*, τόμ. 7, αρ. 10, 2013.
- [65] D. Gerloff, S. Kewitz-Hempel, G. Hause και et al., «Comprehensive Analyses of miRNAs Revealed miR-92b-3p, miR-182-5p and miR-183-5p as Potential Novel Biomarkers in Melanoma-Derived Extracellular Vesicles,» *Frontiers in oncology*, τόμ. 12, 2022.
- [66] S. Tang, J. Cheng και et al., «Combination of Four Serum Exosomal MiRNAs as Novel Diagnostic Biomarkers for Early-Stage Gastric Cancer,» *Frontiers in genetics*, τόμ. 11, 2020.
- [67] K. Uotani, T. Fujiwara και et al., «Circulating MicroRNA-92b-3p as a Novel Biomarker for Monitoring of Synovial Sarcoma.,» *Scientific reports*, τόμ. 7, αρ. 1, 2017.

- [68] M. Li, W. Shan και Y. Hua, «Exosomal miR-92b-3p Promotes Chemoresistance of Small Cell Lung Cancer Through the PTEN/AKT Pathway.,» *Frontiers in cell and developmental biology*, τόμ. 9, 2021.
- [69] Y. Wu, S. Leyk, H. Torabi και et al., «Plasmodium falciparum infection reshapes the human microRNA profiles of red blood cells and their extracellular vesicles.,» *iScience*, τόμ. 26, αρ. 7, 2023.
- [70] S. Rashidi, R. Mansouri και et al., «Highlighting the interplay of microRNAs from Leishmania parasites and infected-host cells.,» *Parasitology*, τόμ. 148, αρ. 12, pp. 1434-1446, 2021.
- [71] S. Muxel, M. Laranjeira-Silva, R. Zampieri και et al, «Leishmania (Leishmania) amazonensis induces macrophage miR-294 and miR-721 expression and modulates infection by targeting NOS2 and L-arginine metabolism.,» *Scientific Reports*, τόμ. 7, 2017.
- [72] M. S. Muxel, M. S. Acuña και et al., «Toll-Like Receptor and miRNA-let-7e Expression Alter the Inflammatory Response in Leishmania amazonensis-Infected Macrophages,» *Frontiers in Immunology*, τόμ. 9, 2018.
- [73] Z.-q. Hu, C.-l. Rao και et al., «Rab32 GTPase, as a direct target of miR-30b/c, controls the intracellular survival of Burkholderia pseudomallei by regulating phagosome maturation,» *PLOS Pathogens*, τόμ. 15, αρ. 6, 2019.
- [74] C. Sood, J. Verma, R. Basak, A. Kapoor, S. Gupta και et al, «Leishmania highjack host lipid body for its proliferation in macrophages by overexpressing host Rab18 and TRAPPC9 by downregulating miR-1914-3p expression.,» *PLOS Pathogens*, τόμ. 20, αρ. 2, 2024.
- [75] L. Huang, Y. Liu, L. Wang και et al., «Down-regulation of miR-301a suppresses pro-inflammatory cytokines in Toll-like receptor-triggered macrophages,» *Immunology*, τόμ. 140, αρ. 3, pp. 314-322, 2013.
- [76] D. Li, X. Du, M. Zhu, S. Yang και W. Zhao, «MiR-374b-5p Regulates T Cell Differentiation and Is Associated with rEg.P29 Immunity,» *BioMed research international*, 2020.
- [77] M. S. L. Nascimento, V. Carregaro, D. S. Lima-Júnior και et al., «Interleukin 17A Acts Synergistically With Interferon γ to Promote Protection Against Leishmania infantum Infection,» *The Journal of Infectious Diseases*, τόμ. 211, αρ. 6, p. 1015–1026, 2014.
- [78] S. Tabrez, S. K. Akand και et al., «Leishmania donovani modulates host miRNAs regulating cholesterol biosynthesis for its survival,» *Microbes and Infection*, τόμ. 26, αρ. 8, 2024.
- [79] A. Loria, V. Dattilo, D. Santoro, J. Guccione και et at, «Expression of Serum Exosomal miRNA 122 and Lipoprotein Levels in Dogs Naturally Infected by Leishmania infantum: A Preliminary Study.,» *Animals : an open access journal from MDPI*, τόμ. 10, αρ. 1, 2020.

- [80] L. H. C. Assis, D. Andrade-Silva and et al, "Cell Cycle, Telomeres, and Telomerase in *Leishmania* spp.: What Do We Know So Far?," *Cells*, vol. 10, no. 11, p. 3195, 2021.