



Πτυχιακή εργασία

“Γονιδιωματική μετα-ανάλυση σε ράτσες του είδους *Canis lupus familiaris* σε επίπεδο μιτοχονδριακού DNA”

“Genomic meta-analysis on breeds of *Canis lupus familiaris* species at mitochondrial DNA level”



Εργαστήριο Γενετικής, Συγκριτικής και Εξελικτικής Βιολογίας

Κάτσα Ελένη- Μαρία

Λάρισα, 2016

Μέλη τριμελούς επιτροπής

- **Z. Μαμούρης:** Καθηγητής Γενετικής Ζωικών Πληθυσμών
(Υπεύθυνος καθηγητής)
- **A.Μούτου:** Αναπληρώτρια Καθηγήτρια Βιολογίας Σπονδυλωτών
- **Θ. Σαραφίδου:** Επίκουρος Καθηγήτρια Μοριακής Γενετικής Ζωικών
Οργανισμών

**....Στους γονείς μου, που μου αποδεικνύουν
ότι η προσπάθεια πάντα αξίζει...**

Ευχαριστίες

Η παρούσα πτυχιακή εργασία αποτελεί την κορύφωση των σπουδών μου στο Πανεπιστήμιο Θεσσαλίας, στη Σχολή Επιστημών Υγείας, στο Τμήμα Βιοχημείας και Βιοτεχνολογίας.

Θα ήθελα να εκφράσω τις ευχαριστίες μου στον υπεύθυνο καθηγητή μου κύριο Ζήση Μαμούρη, για την ανάθεση του θέματος.

Ένα μεγάλο ευχαριστώ επίσης οφείλω στα μέλη του εργαστηρίου και περισσότερο στον Γιαννούλη Θεμιστοκλή και στον Τσιπουρλιάνο Ανδρέα για την πολύτιμη βοήθεια τους αλλά και τον χρόνο που διέθεσαν για την διεκπεραίωση αυτής της πτυχιακής εργασίας.

Κλείνοντας, θα ήθελα να ευχαριστήσω την οικογένειά μου και κυρίως τους ανθρώπους στους οποίους οφείλω τα πάντα, τους γονείς μου Νίκο και Αργυρώ. Με στήριζαν και με στηρίζουν με κάθε δυνατό τρόπο σε ό,τι και αν κάνω και είναι αυτοί που με βοηθούν πάντα στο να κάνω τα όνειρά μου πραγματικότητα.

Περιεχόμενα

Περίληψη	6
Abstract	7

Κεφάλαιο 1^ο

Εισαγωγή	8
1.1 Σκύλος (Canis lupus familiaris)	8
1.1.1 Ταξινόμηση	8
1.1.2 Καταγωγή σκύλου	9
1.1.3 Η εξέλιξη των ράτσων	12
1.1.4 Διαίρεση σε ομάδες	13
1.2 Μοριακοί δείκτες	13
1.2.1 Τι είναι Μοριακοί δείκτες;	14
1.2.2 Πώς επιλέγω έναν μοριακό δείκτη;	15
1.2.3 Μιτοχονδριακό DNA (mtDNA)	16
1.2.4 Πυρηνικό DNA	21
1.2.5 Διαφορές μιτοχονδριακών και πυρηνικών δεικτών	23
1.3 Γονιδιωματική	23
1.3.1 Τι είναι γονιδιωματική;	23
1.3.2 Πόσα είδη γονιδιωμάτων υπάρχουν στη γη;	24
1.4 Μετα-ανάλυση	24
1.4.1 Ορισμός Μετα-ανάλυσης	24
1.4.2 Πλεονεκτήματα της Μεθοδολογίας της Μετα-Ανάλυσης	25
1.4.3 Μειονεκτήματα της Μεθοδολογίας της Μετα-Ανάλυσης	25
1.5 Βάσεις δεδομένων	26
1.5.1 Τι είναι οι βάσεις δεδομένων;	26
1.5.2 Κατηγορίες βάσεων δεδομένων	27

1.5.3 Τι είναι η βάση δεδομένων GenBank;.....	29
1.5.3.1 Πρόσβαση στη GenBank	29
1.5.3.2 Χρήση Δεδομένων της GenBank	30
1.5.3.3 Επεξήγηση των σημαντικότερων πεδίων μιας εγγραφής στην GenBank.....	30
1.5.4 Μορφοποιήσεις βάσεων δεδομένων (databaseformats)	32
1.5.5 Εισαγωγή στο Blast.....	35
1.5.5.1 Ο σκοπός του BLAST.....	35
1.5.5.2 Ο αλγόριθμος του BLAST.....	37
 1.6 Σκοπός της παρούσας εργασίας.....	 40

Κεφάλαιο 2^ο

Υλικά και Μέθοδοι	41
--------------------------------	-----------

Κεφάλαιο 3^ο

Αποτελέσματα	47
---------------------------	-----------

Κεφάλαιο 4^ο

Συμπεράσματα – Συζήτηση	67
--------------------------------------	-----------

Κεφάλαιο 5^ο

Βιβλιογραφία.....	71
--------------------------	-----------

Περίληψη

Ο σκύλος (*Canis lupus familiaris*) είναι ένας οργανισμός-μοντέλο που χρησιμοποιείται στην ιατρική έρευνα λόγω της εκτεταμένης γενετικής ποικιλομορφίας και των μορφολογικών διαφορών εντός του είδους, καθώς και λόγω των πρακτικών αναπαραγωγής που οδήγησαν στην ύπαρξη «καθαρών» πληθυσμών σκύλων, που αναφέρονται ως ράτσες. Ο οικόσιτος σκύλος αποτελεί ένα από τα πιο όμορφα γλυπτά γενετικής που διαμορφώθηκε από τη φύση και τον άνθρωπο. Έτσι λοιπόν η μελέτη των φυλών του είδους *Canis lupus familiaris* είναι πολύ σημαντική. Για τη μελέτη της ενδοειδικής ποικιλομορφίας, ο δείκτης επιλογής προέρχεται συνήθως από το μιτοχονδριακό DNA, το οποίο εμφανίζει μητρική κληρονομία και ιδιαίτερα υψηλούς εξελικτικούς ρυθμούς, χαρακτηριστικά τα οποία το καθιστούν χρήσιμο εργαλείο για τη μελέτη της μοριακής εξέλιξης των οργανισμών. Σκοπός της παρούσας μελέτης είναι η ανάπτυξη μοριακών δεικτών για το διαχωρισμό των ρατσών των σκύλων. Για την επίτευξη του, πραγματοποιήθηκε μετά-ανάλυση 381 διαφορετικών μιτοχονδριακών αλληλουχιών DNA σκύλων που ήταν καταχωρημένες σε δημόσιες βάσεις δεδομένων (NCBI) και παρείχαν πληροφορίες για τη ράτσα από την οποία προέρχονται. Στην συνέχεια ακολουθήθηκαν δυο διαφορετικές προσεγγίσεις, με ομαδοποίηση των αλληλουχιών των διαφορετικών ρατσών και χωρίς ομαδοποίηση. Η πρώτη προσέγγιση είναι πιο πληροφοριακή, χωρίς να λαμβάνει υπόψη τον πολυμορφισμό μέσα στις ράτσες ενώ η δεύτερη αντικατοπτρίζει καλύτερα τον πολυμορφισμό σε ατομικό επίπεδο, λαμβάνοντας υπόψη την διαφοροποίηση μέσα σε κάθε ράτσα. Οι δυο διαφορετικές προσεγγίσεις οδήγησαν στην εξεύρεση πολυμορφικών θέσεων. Τα αποτελέσματα της πρώτης προσέγγισης, είναι πληροφοριακά και πρόκειται για πολυμορφισμούς μονού νουκλεοτιδίου (SNPs) οι οποίοι είναι χαρακτηριστικοί για συγκεκριμένες ράτσες, δίνοντας τη δυνατότητα ανάπτυξης μοριακών δεικτών για τον μοριακό προσδιορισμό των ρατσών αυτών. Μελλοντικός στόχος της μελέτης είναι ο σχεδιασμός εκκινητών που να στοχεύουν τις συγκεκριμένες θέσεις και η εφαρμογή τους σε δείγματα σκύλων με γνωστή ράτσα, για να διαπιστωθεί η πιστότητα των αποτελεσμάτων μας *in vitro*.

Abstract

Dog (*Canis lupus familiaris*) is a valuable model organism used in medical research because of the extensive genetic diversity and morphological differences within species and due to the long story of selective breeding that led to the existence of pure dog populations, referred as breeds. The domesticated dog is one of the most beautiful genetic sculptures formed by nature and man. Thus, the study of *Canis lupus familiaris* species is very important. To study the intra-species variation, the selected molecular markers are usually derived from the mitochondrial DNA (mtDNA), which is maternally inherited and it has particularly high evolutionary rates. The characteristics listed above, make mtDNA a useful tool for studying the molecular evolution of organisms. The purpose of this study is to develop molecular markers in order to separate the breeds of dogs. To achieve this, we performed a meta-analysis of 381 different dog's mtDNAs sequences that were registered in public databases (NCBI) and provided information on the breed of each origin. After that, two different approaches were followed, one by grouping sequences of different breeds and one without regrouping. The first approach is more informative, without taking into account the polymorphism within breeds and the second one reflects better the polymorphism in each individual, taking into account the differentiation within each breed. The two different approaches led to finding polymorphic sites. The results of the first approach are informative and produce a list of single nucleotide polymorphisms (SNPs) which are characteristic for specific breeds, allowing the development of molecular markers for the molecular identification of these breeds. Future goal of this study is the design of primers that target the specific locations and the application in dogs with known breeds, to determine the conformity of our results in vitro.

Κεφάλαιο 1^ο

Εισαγωγή

1.1 Σκύλος (*Canis lupus familiaris*)



Εικόνα 1. Huskysiberian

Ο σκύλος (αρχ. ελλ. κύων) είναι το ανώτερο θηλαστικό ζώο του γένους *Canis familiaris* (Κύων ο λύκος ο οικείος) (Wang, 2008) και αποτελεί ένα από πιο κοινά κατοικίδια ζώα. Πιο συγκεκριμένα είναι ένα εξημερωμένο κυνίδιο (*canid*) που έχει αναπαραχθεί επιλεκτικά για χιλιετίες για να εμφανίζει σήμερα διάφορες συμπεριφορές, αισθητήριες ικανότητες, και φυσικά χαρακτηριστικά (Dewey, 2002). Το γεγονός ότι είναι το παλαιότερο εξημερωμένο ζώο και η επιτυχημένη πορεία του σε τόσους πολλούς τομείς οφείλεται στην ικανότητα του να προσαρμόζεται στο διαρκώς εξελισσόμενο ανθρώπινο περιβάλλον. Είναι αξιοσημείωτο ότι τα σκυλιά είχαν θεμελιώδη σημασία και κατάφεραν να διαδραματίσουν πρωταγωνιστικό ρόλο στην εξέλιξη του ανθρώπου όσο κανένα άλλο ζώο που πέρασε από τη Γη.

1.1.1 Ταξινόμηση

Ο σκύλος έχει χαρακτηριστεί ως *Canis lupus familiaris* με την έννοια του βιολογικού είδους και *Canis lupus* με την έννοια του εξελικτικού είδους. (Wang, 2008)
Το 1758, ο Linnaeus δημοσίευσε στην *Systema Naturae*, μια κατηγοριοποίηση των ειδών στην οποία περιλαμβάνονται τα είδη *Canis*. Η *Canis* είναι μια λατινική λέξη που σημαίνει σκύλος (*Online Etymology Dictionary*) και στη λίστα περιλαμβάνει τα σαρκοφάγα που μοιάζουν με σκύλους: τον κατοικίδιο σκύλο, λύκους, αλεπούδες και τσακάλια. Ο σκύλος έχει χαρακτηριστεί ως *Canis familiaris*, (Linnaeus, 1758) που σημαίνει "σκύλος-οικογένεια" ή το οικογενειακό σκυλί (*Oxford Dictionaries*). Στην επόμενη σελίδα κατέγραψε τον λύκο ως *Canislupus*, που σημαίνει «σκύλος-λύκος». Το 1978, μια αναθεώρηση με στόχο τη μείωση του αριθμού των αναγνωρισμένων ειδών *Canis* πρότεινε ότι ως «*Canis Dingo*» θα χαρακτηρίζεται ο άγριος σκύλος κατοικίδιο. Ο χαρακτηρισμός ως *Canis familiaris* χρησιμοποιείται για σκύλους κατοικίδια, αν και ταξινομικά θα πρέπει πιθανώς να είναι συνώνυμη με την «*Canis lupus*» (Van Gelder, 1978).

Το 1982, η πρώτη έκδοση των ειδών των θηλαστικών του κόσμου ταξινομήσε το *Canis familiaris* κάτω από το *Canis lupus* με το σχόλιο: " Πιθανός πρόγονος του και παρόμοιο με τον σκύλο κατοικίδιο, *familiaris* (Honaki ,1982). Αποφεύγει την ταξινόμηση του λύκου ως οικογενειακό σκυλί. Ο σκύλος είναι καταχωρημένος μεταξύ των άλλων λατινικά επωνομασμένων υποειδών του *Canis lupus familiaris* ως *Canis lupus* (Smithsonian-Animal Species of the World database. " *Canis lupus*")

Πίνακας 1. Συστηματική ταξινόμηση του σκύλου

Συστηματική ταξινόμηση σκύλου	
Επικράτεια:	Ευκαρυωτικά
Βασίλειο:	Ζώα (Animalia)
Υποβασίλειο:	Αμφίπλευρα (Bilateria)
Κατώτερο βασίλειο:	Δευτεροστόμια (Deuterostomia)
Συνομοταξία, Φύλο:	Χορδωτά (Chordata)
Υποφύλο:	Σπονδυλωτά (vertebrata)
Κατώτερο φύλο:	Γναθοστόματα (Gnathostomata)
Υπερκλάση:	Τετράποδο (tetrapoda)
Ομοταξία, Τάξη:	Θηλαστικά (Mammalia) Linnaeus, 1758
Υπο-Τάξη:	Theria
Κατώτερη τάξη:	Eutheria
Τάξη:	Σαρκοφάγα (Carnivora)
Υποτάξη:	Κυνοειδή (Caniformia)
Οικογένεια:	Κυνίδες (Canidae)
Γένος:	<i>Canis</i> (Κύων), Linnaeus, 1758
Είδος:	<i>Canis lupus</i> , Linnaeus, 1758
Υποείδος:	<i>Canis lupus familiaris</i> , Linnaeus, 1758

1.1.2 Καταγωγή σκύλου

Δεν υπάρχει αμφιβολία για τη ζωολογική του κατάταξη όπως είδαμε και παραπάνω. Όμως, η καταγωγή του κατοικίδιου σκύλου (*Canis lupus familiaris* ή *Canis familiaris*) δεν είναι σαφής. Γενικά, είναι αποδεκτό ότι μαζί με το λύκο, την αλεπού και το τσακάλι, ανήκει στη σύγχρονη οικογένεια των κυνιδίων (canidae).

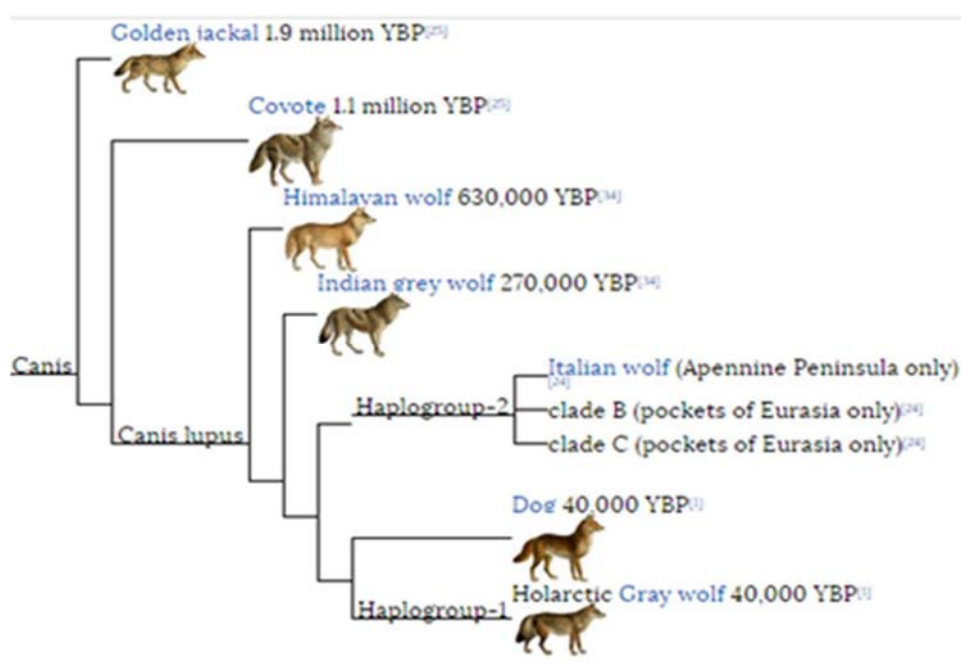
Παλαιότερα, υπήρχαν 3 θεωρίες για την προέλευση του σκύλου:

- ότι κατάγεται από το λύκο
- ότι κατάγεται από το τσακάλι
- ότι κατάγεται από ένα είδος άγριου σκυλιού που τώρα έχει εκλείψει και από το οποίο δεν έχουν βρεθεί ακόμα λείψανα.

Ολόκληρο το γονιδίωμα δείχνει ότι ο σκύλος, ο λύκος και η εξαφάνιση του λύκου *Taymīr* αποκλίνει περίπου την ίδια περίοδο, 27,000-40,000 χρόνια πριν (Skoglund, 2015). Τα τέλη της δεκαετίας του 1990, οι επιστήμονες ανακάλυψαν ότι όλα τα σκυλιά (*canis lupus familiaris*) έχουν γενετική ομοιότητα με τους λύκους (*canis lupus*). Τα σύγχρονα σκυλιά είναι πιο στενά συνδεδεμένα με τα αρχαία

απολιθώματα λύκων και συγκεκριμένα με αυτά των σύγχρονων γκριζων λύκων που έχουν βρεθεί στην Ευρώπη (Freedman,2014).

Ο καλύτερος φίλος του ανθρώπου ήταν αρχικά μάλλον γκριζος λύκος που ενδεχομένως να ήρθε σε επαφή με τους πρώτους ανθρώπους περίπου 33.000 χρόνια πριν, κάπου στη νοτιοανατολική Ασία (Radford,2015). Οι πρώτες μελέτες που διεξήχθησαν το 1993, σε μιτοχονδριακό DNA των αρτίγονων κυνιδίων έδειξαν ότι υπάρχει πολύ στενή σχέση μεταξύ των γκριζων λύκων και των σκύλων με μόλις 0,2% απόκλιση (Wayne, 1993). Αντίθετα, η απόκλιση μεταξύ του γκριζου λύκου και του κογιότ ήταν 4%, περίπου. Άμεσα, ο κατοικίδιος σκύλος, από ξεχωριστό είδος (*Canis familiaris*), αναταξινομήθηκε ως υποείδος του γκριζου λύκου το επονομαζόμενο πλέον ως *Canis lupus familiaris* (Radford, 2015).



Εικόνα 2. Φυλογενετικό δέντρο στο οποίο διαφαιίνεται ότι οι σκύλοι και οι σωζόμενοι λύκοι αποτελούν δύο αδελφούς κλάδους. Φαίνεται ότι ο πληθυσμός ή οι πληθυσμοί των λύκων (indiangreywolf, Italianwolf κ.α) από τους οποίους προήλθε ο σκύλος (dog) πιθανώς να έχουν εξαφανιστεί. (Zhenxin,2016)

Έρευνες, οι οποίες βασίστηκαν στην μελέτη του Πολυμορφισμού Μοναδικού-Νουκλεοτιδίου (Single-Nucleotid Polymorphism, SNP) εντός των γονιδιωμάτων γκριζων λύκων και σκύλων από διαφορετικές περιοχές έδειξαν ότι, οι σκύλοι δεν κατάγονται από τους γκριζους λύκους, αλλά μοιράζονται έναν πρόσφατο κοινό πρόγονο (**Εικόνα 2**). Ο κοινός πρόγονος και των δύο ταξινομικών μονάδων θεωρείται ότι υπήρξε ένας ευμεγέθης «λύκος» που ζούσε στην Ευρώπη, 9.000-34.000 χρόνια, πριν (Viegas,2014). Υπάρχει σημαντικός αριθμός διαγνωστικών χαρακτηριστικών μεταξύ γκριζου λύκου και κατοικίδιου σκύλου.

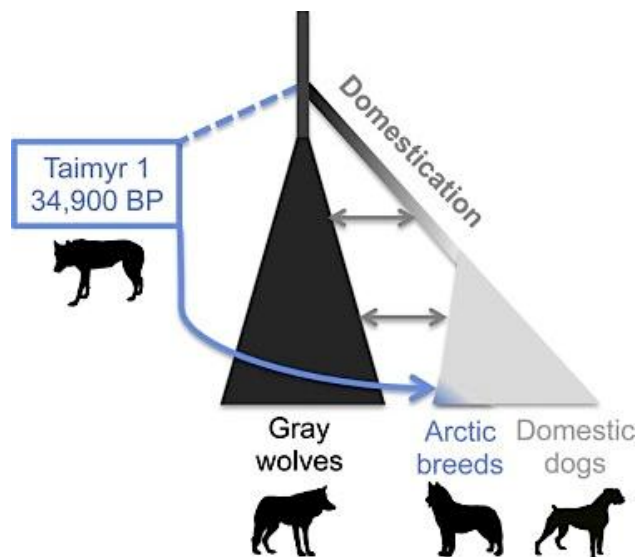


Εικόνα 3. Γκρίζος λύκος



Εικόνα 4. Οστό από την κάτω γνάθο του λύκου Ταιμυγ: πρόγονος σκύλου (*Love Dalén*)

Η ανάλυση του γονιδιώματος από ένα αρχαίο οστό, του οποίου η ηλικία εκτιμήθηκε στα 35.000 έτη αποκαλύπτει ότι ο λύκος Ταιμυγ αντιπροσωπεύει τον πιο πρόσφατο κοινό πρόγονο των σύγχρονων λύκων και σκύλων. Μάλιστα, τα στοιχεία του DNA δείχνουν επίσης ότι τα σύγχρονα σκυλιά Χάσκυ Σιβηρίας και τα σκυλιά ελκήθρων της Γροιλανδίας μοιράζονται έναν ασυνήθιστα μεγάλο αριθμό γονιδίων με τον αρχαίο λύκο του Ταιμυγ. Αυτό υποδεικνύει ότι τα σκυλιά μπορεί να εξημερώθηκαν πολύ νωρίτερα από ότι γενικά είναι αποδεκτό.



Εικόνα 5. Παρουσιάζεται ένα σχέδιο της σειράς του γονιδιώματος ενός λύκου 35.000 ετών από την χερσόνησο Ταιμυγ στη βόρεια Σιβηρία που χρονολογείται στο 34.900 BP. Θεωρούμε ότι αυτό το άτομο ανήκε σε έναν πληθυσμό που απέκλιε από τον κοινό πρόγονο των λύκων και των σκύλων πολύ κοντά στο χρόνο εμφάνισης της γενεαλογίας του κατοικίδιου σκύλου. (Skoglund, 2015)

The origins of man's best friend



Εικόνα 6. Η

Ασιατική προέλευση των σύγχρονων σκύλων.

1.1.3 Η εξέλιξη των ράτσων

Περισσότερο από κάθε άλλο θηλαστικό, ο σκύλος έχει διακλαδωθεί σε μια πλειονότητα από ράτσες. Έχει εκτιμηθεί ότι υπάρχουν περισσότερες από 300 σε όλο τον κόσμο. Πριν αρχίσει η υποδιαίρεση σε ράτσες, η εξέλιξη που έλαβε χώρα στα διαφορετικά μέρη του κόσμου, δημιούργησε τύπους σκύλων πιο κατάλληλους στο εκάστοτε περιβάλλον και συνθήκες. Πέρα από αυτούς τους τύπους σκυλιών που προήλθαν μέσω της εξέλιξης, μερικές ράτσες ήρθαν στην επιφάνεια σαν αποτέλεσμα μεταλλαγών, που ο άνθρωπος εκμεταλλεύτηκε με επιλεκτική αναπαραγωγή.

Διασταυρώνοντας τις «βασικές ράτσες» μεταξύ τους, νέες ράτσες δημιουργήθηκαν κάνοντας μια μεγάλη ποικιλία σκύλων διαφόρων τύπων και μεγεθών. Με την επιλεκτική αναπαραγωγή επιτεύχθηκε το καλύτερο δυνατό αποτέλεσμα σε σχέση με το μέγεθος και την εμφάνιση του σκύλου αλλά και την ικανότητα του στο κυνήγι, τη βοσκή, τη φύλαξη και την προστασία των ανθρώπων. Ενώ όλα τα σκυλιά είναι γενετικά πολύ παρόμοια, (Savolainen, 2002) η φυσική επιλογή και η επιλεκτική αναπαραγωγή έχουν ενισχύσει ορισμένα χαρακτηριστικά σε ορισμένους πληθυσμούς σκύλων, δίνοντας αφορμή την δημιουργία νέων φυλών σκύλων. Οι διάφορες κατηγορίες σκύλων χωρίζονται με βάση τη λειτουργία, τη γενετική, ή χαρακτηριστικά. Οι ράτσες (breeds) λοιπόν, είναι ομάδες ζώων που κατέχουν μια σειρά από κληρονομήσιμα χαρακτηριστικά που τους διακρίνει από τα άλλα ζώα του ίδιου είδους. Τα σύγχρονα είδη σκυλιών είναι μη-επιστημονικές ταξινομήσεις σκύλων που τηρούνται από τα σύγχρονα κυνοτροφεία.

Καθαρόαιμα σκυλιά από την ίδια φυλή, είναι γενετικά διακριτά από καθαρόαιμα σκυλιά άλλων φυλών.

Αναλύσεις μικροδορυφόρων 85 φυλών σκύλων έδειξαν ότι τα σκυλιά χωρίζονται σε τέσσερις κύριους τύπους (Parker, 2004):

1. Τα " σκυλιά παλιού κόσμου » (π.χ., Malamute και shar pei)
2. Τα τύπου- «μαντρόσκυλα» (π.χ. english Mastiff)
3. Τα σκυλιά αγέλης (π.χ., Border Collie)
4. «Όλα τα άλλα» ("σύγχρονα" ή "κυνηγετικού" -τύπου)

Ενώ πολλές φυλές όπως το Siberianhusky ,το chow- chow, shar-peí, akita, shiba-inu και Pekinese είναι πραγματικά αρχαίες, οι περισσότερες από 400 φυλές σκύλων που υπάρχουν στον κόσμο είναι στην συντριπτική τους πλειοψηφία δημιούργημα μας των τελευταίων 200 ετών.

1.1.4 Διαίρεση σε ομάδες

Οι προσπάθειες που έγιναν σε διάφορα μέρη του κόσμου για τη δημιουργία μιας συστηματικής ταξινόμησης των σκυλιών, δεν είναι ούτε λογικές ούτε συνεπείς. Μερικές φορές η ταξινόμηση βασίζεται σε σωματικές ομοιότητες ή κοινή καταγωγή. Μερικές ομάδες αποτελούνται από ράτσες που χρησιμοποιούνται για παρόμοιους σκοπούς.

Το 1965 η Διεθνής Κυνολογική Ομοσπονδία (FCI) καθόρισε ένα σύστημα, που περιλάμβανε δέκα ομάδες και το οποίο χρησιμοποιείται τώρα στις πιο πολλές ευρωπαϊκές χώρες. Αυτές οι ομάδες είναι:

- Ράτσες ποιμενικές
- Ράτσες φυλάκων/αστυνομικών σκυλιών/εργατικών σκυλιών
- Τεριέ
- Ντάτσχουντ
- Κυνηγετικές ράτσες (μεγάλων θηραμάτων)
- Κυνηγετικές ράτσες (μικρών θηραμάτων)
- Κυνηγόσκυλα φέρμας (εκτός από βρετανικές ράτσες)
- Κυνηγόσκυλα φέρμας (βρετανικές ράτσες)
- Άλλα βρετανικά κυνηγόσκυλα
- Σκυλιά-παιχνίδια (μίνι)

1.2 Μοριακοί δείκτες

Για την πιστοποίηση των παραπάνω αποτελεσμάτων είναι αναγκαία η ύπαρξη μοριακών δεικτών. Οι μοριακοί δείκτες θεωρούνται ως ένα αξιόπιστο εργαλείο για την ανίχνευση πολυμορφισμών της αλληλουχίας του DNA και την εξακρίβωση των γενετικών διαφορών οι οποίες θα μπορούσαν να οδηγήσουν στη διάκριση και κατοχύρωση των εμπορικών ποικιλιών σε σχέση με την περιγραφή των μορφολογικών χαρακτηριστικών.

1.2.1 Τι είναι Μοριακοί δείκτες;

Ως μοριακός δείκτης χαρακτηρίζεται ένας πολυμορφικός τύπος που χρησιμοποιείται για την παρατήρηση της γενετικής ποικιλότητας αλλά ο ίδιος δεν αποτελεί αντικείμενο πρωταρχικού ενδιαφέροντος. (Barton, 2007) Ουσιαστικά οι μοριακοί δείκτες, είναι μικρές περιοχές του γονιδιώματος που χρησιμοποιούνται ως δείκτες γενετικής ποικιλομορφίας.

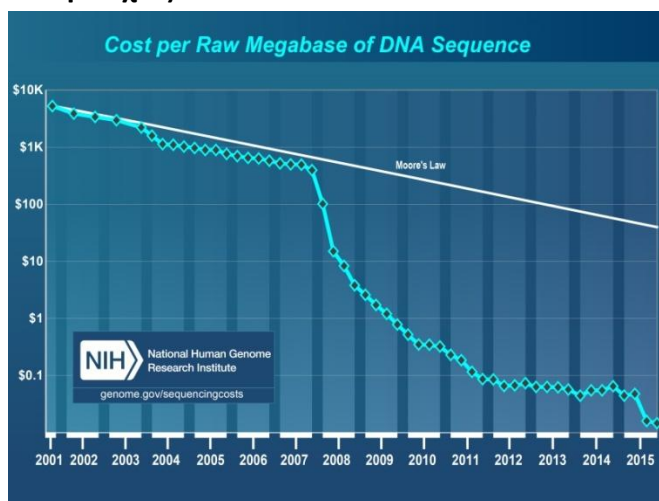
Οι μοριακοί δείκτες χρησιμοποιούνται ευρέως στην εξελικτική βιολογία και στις φυλογενετικές μελέτες ενώ οι μοριακές μέθοδοι ανάλυσης βελτιώνονται με ταχείς ρυθμούς. Είναι χρήσιμοι μόνο όταν είναι πολυμορφικοί στους πληθυσμούς.

Χαρακτηριστικά μοριακών δεικτών που τους καθιστούν σημαντικούς

- Ενυπάρχουν στα άτομα (δεν μπορούν να χαθούν)
- Είναι κληρονομήσιμοι (ταυτοποίηση απογόνων)
- Δεν καταστρέφεται το δείγμα (δεν απαιτείται θανάτωση του ζώου)

Υπάρχει μία πληθώρα μοριακών δεικτών που μελετούν τον πολυμορφισμό είτε σε επίπεδο πρωτεϊνών είτε DNA.

- **Αλλοένζυμα**
- **Πολυμορφισμοί μεγέθους τμημάτων περιορισμού (RFLP)** (Botstein, 1980)
- **Τυχαία ενισχυμένο πολυμορφικό DNA (RAPD)** (Williams., 1990)
- **Αλληλουχίες DNA**



Εικόνα 7. Οι αλληλουχίες DNA χρησιμοποιούνται ευρύτατα στις μέρες καθώς το κόστος αλληλούχησης του DNA έχει μειωθεί σημαντικά (National Human Genome Research Institute).

- **Αλληλουχίες μιτοχονδριακού (mt) DNA**
- **Αλληλουχίες χλωροπλαστικού (cp) DNA**
- **Αλληλουχίες πυρηνικού DNA**
- **Μικροδορυφορικό DNA** (Queller *et al.*, 1993).

Οι μικροδορυφοροείναι τύποι όπου μικρές αλληλουχίες DNA επαναλαμβάνονται στη σειρά η μια αμέσως μετά την άλλη.

1.2.2 Πώς επιλέγω έναν μοριακό δείκτη;

Η χρήση των μοριακών δεικτών είναι ευρέως διαδεδομένη μεταξύ των εξελικτικών βιολόγων και οι μέθοδοι ανάλυσης των γενετικών δεδομένων βελτιώνονται ταχύτατα, ωστόσο υπάρχει έλλειψη ενός οργανωμένου πλαισίου εργασίας με το οποίο να μπορούν να δουλέψουν οι επιστήμονες. Στοιχεία της μοριακής εξέλιξης συνοψίζονται για να εξηγήσουν την προέλευση της μεταβλητότητας στο επίπεδο του DNA, τα όριά της, καθώς και τις σχέσεις που συνδέουν τη γενετική διαφοροποίηση με τις βιολογικές παραμέτρους των υπό μελέτη οργανισμών (Chenuil, 2006).

Οι μοριακοί δείκτες καθορίζονται από δύο συνιστώσες:

- ✓ Την περιοχή του DNA που πρόκειται να «καλύψουν».
- ✓ Την τεχνική που χρησιμοποιείται για την εύρεση της διαφοροποίησης.

Τα κριτήρια για την επιλογή του κατάλληλου μοριακού δείκτη χωρίζονται σε τρεις κατηγορίες:

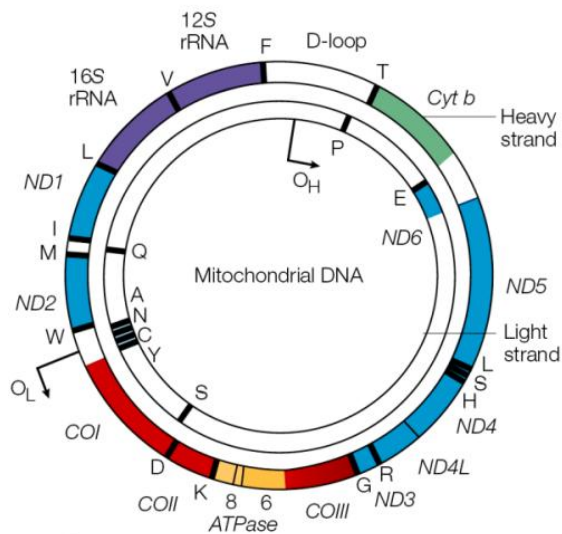
1. Το επίπεδο της διαφοροποίησης.
2. Τη φύση της πληροφορίας, η οποία πρέπει να προσδιορίζεται σύμφωνα με το επιστημονικό ερώτημα.
3. Κάποια πρακτικά κριτήρια, τα οποία έχουν να κάνουν κυρίως με τον εξοπλισμό του εκάστοτε εργαστηρίου και την εμπειρία του ερευνητή.

Σύμφωνα με τον Avise και τους συνεργάτες του (1987), οι ιδανικές και επιθυμητές ιδιότητες ενός μοριακού συστήματος κατάλληλου για φυλογενετική ανάλυση συνοψίζονται στα ακόλουθα:

- α) να αποτελεί διακριτό χαρακτήρα, ευρέως κατανεμημένο ώστε να μπορούν να γίνουν συγκρίσεις ομολογίας σε ένα μεγάλο φάσμα οργανισμών,
- β) να μπορεί να απομονωθεί και να εξεταστεί εύκολα,
- γ) να έχει απλή γενετική δομή χωρίς πολύπλοκα στοιχεία, όπως επαναλαμβανόμενο DNA, μεταθετά στοιχεία, ψευδογονίδια και ιντρόνια,
- δ) να έχει άμεσο τρόπο γενετικής μετάδοσης, χωρίς ανασυνδυασμό ή άλλες γενετικές αναδιατάξεις,
- ε) να παρέχει μια σειρά καταστάσεων ποιοτικού χαρακτήρα, των οποίων οι φυλογενετικές σχέσεις να μπορούν να συναχθούν με βάση λογικά φειδωλά κριτήρια
- στ) να εξελίσσεται με ραγδαίο ρυθμό ώστε να εμφανίζονται νέες καταστάσεις χαρακτήρων ακόμα και κατά τη διάρκεια ζωής ενός είδους, προκειμένου να χρησιμοποιηθεί σε μελέτες μικροεξέλιξης.

Αν και κανένα μοριακό σύστημα δεν είναι τέλειο για φυλογενετικές αναλύσεις, το mtDNA των ανώτερων ζώων πληροί σε μεγάλο βαθμό όλα τα προαναφερόμενα κριτήρια.

1.2.3 Μιτοχονδριακό DNA (mtDNA)



Εικόνα 8. Μιτοχονδριακό DNA (mtDNA)

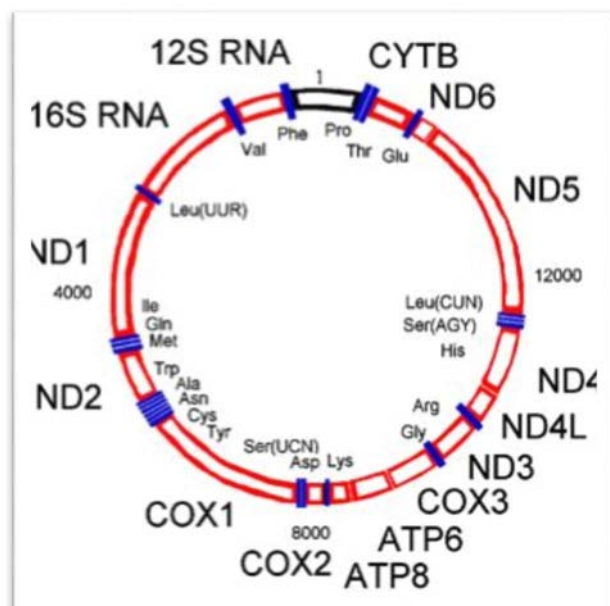
Ένας μοριακό δείκτης ο οποίος χρησιμοποιείται ευρέως στις φυλογενετικές μελέτες, είναι το μιτοχονδριακό DNA (mtDNA). Τα περισσότερα μιτοχονδριακά γονιδιώματα είναι κυκλικά, δίκλινα, υπερελικωμένα μόρια DNA ενώ γραμμικά μόρια εντοπίζονται μόνο σε ορισμένα πρωτόζωα και μύκητες. Τα μιτοχονδριακά γονιδιώματα διαφορετικών ειδών παρουσιάζουν πολλές ομοιότητες ως προς τον αριθμό και τις λειτουργίες των γονιδίων τους.

Το μιτοχονδριακό DNA των ζώων είναι ένα μικρό, εξωχρωμοσωματικό γονιδίωμα, που έχει συνήθως μέγεθος της τάξης των 16-17 kb (στους περισσότερους πολυκύτταρους οργανισμούς) και περιλαμβάνει 37 γονίδια από τα οποία 2 κωδικοποιούν ριβοσωμικά RNAs (rRNAs), 13 κωδικοποιούν πρωτεΐνες και 22 κωδικοποιούν μεταφορικά RNAs (tRNAs) (Εικόνα 8). Επιπλέον, υπάρχει μια μεγάλη μη κωδική περιοχή (D-loop), η οποία είναι γνωστό ότι περιέχει στοιχεία ελέγχου της αντιγραφής και της μεταγραφής (Krzyszewski et al., 2006). Αυτή η περιοχή είναι ταχέως εξελισσόμενη και χρησιμοποιείται περισσότερο για αναλύσεις σε ενδοειδικό επίπεδο (Avise, 1994). Δεν είναι ξεκάθαρο κατά πόσο αυτές οι περιοχές ελέγχου είναι ομόλογες μεταξύ απόμακρων ζωικών ειδών ή, εναλλακτικά, έχουν προκύψει ανεξάρτητα, από διαφορετικές μη κωδικές αλληλουχίες σε ξεχωριστές εξελικτικές γενεαλογίες, δεδομένου ότι δεν παρουσιάζουν νουκλεοτιδική ομολογία εκτός από τις περιπτώσεις που πρόκειται για στενά συγγενικά είδη (Boore, 1999).

Τα προϊόντα αυτών των γονιδίων, μαζί με μόρια RNA και πρωτεΐνες που εισάγονται από το κυτταρόπλασμα, παρέχουν στο μιτοχόνδριο το δικό του σύστημα αντιγραφής και μεταγραφής του DNA και μετάφρασης των πρωτεϊνών.

Η οργάνωση των μιτοχονδριακών γονιδίων φαίνεται να είναι अपαράλλακτη για μεγάλες χρονικές περιόδους της εξελικτικής ιστορίας και είναι πολύ χρήσιμα ως δείκτες λόγω του συμπαγούς πακεταρίσματος των γονιδίων με λίγα μόνο μη μεταφραζόμενα νουκλεοτίδια εντός των γονιδίων (Cantatore & Saccine, 1987). Το

mtDNA έχει το πλεονέκτημα να υπάρχει σε όλο το ζωικό και φυτικό βασίλειο και να έχει εξαιρετικά παρόμοιο γονιδιακό περιεχόμενο (Wilson et al., 1985).



Εικόνα 9.

Η δομή του μιτοχονδριακού DNA, όπου φαίνεται αναλυτικά η διάταξη των γονιδίων. Οι κωδικές περιοχές των γονιδίων φαίνονται με **κόκκινο** χρώμα (ετικέτες εξωτερικά από τον κύκλο). Τα tRNAs, τα γονίδια που κωδικοποιούν αμινοξέα διαφαίνονται με τρία γράμματα για τα αμινοξέα και με **μπλέ** χρώμα (ετικέτες εσωτερικά του κύκλου).

Με βάση την αλληλουχία αναφοράς U96639.2, που ανήκει στο θυληκό boxer, το οποίο χρησιμοποιήθηκε για το πρώτο wholegenomesequencing του είδους, τα γονίδια του μιτοχονδριακού DNA χαρτογραφούνται στις θέσεις που εμφανίζονται στον **πίνακα 2**.

Πίνακας 2. Τα ονόματα και οι περιοχές των γονιδίων του mtDNA.

Όνομα γονιδίου	Περιοχή
12S ribosomal RNA	70..1023
16S ribosomal RNA	1091..2670
NADH dehydrogenase subunit 1 (ND1)	2747..3703
NADH dehydrogenase subunit 2 (ND2)	3914..4957
cytochrome c oxidase subunit I (COX1)	5349..6893
cytochrome c oxidase subunit II (COX2)	7034..7717
ATP synthase F0 subunit 8 (ATP8)	7803..8006
ATP synthase F0 subunit 6 (ATP6)	7964..8644
cytochrome c oxidase III (COX3)	8644..9427
NADH dehydrogenase subunit 3 (ND3)	9496..9842
NADH dehydrogenase subunit 4L (ND4L)	9911..10207
NADH dehydrogenase subunit 4 (ND4)	10201..11578
NADH dehydrogenase subunit 5 (ND5)	11778..13598
NADH dehydrogenase subunit 6 (ND6)	13582..14109
cytochrome b (CYTB)	14183..15322

Πίνακας 3 Γενετικός κώδικας του μιτοχονδρίου

Αμινοξύ	Τριπλέτα κωδικονίων	Αμινοξύ	Τριπλέτα κωδικονίων
Alanine(Ala)	GC*	Leucine(Leu)	TTA TTG CT*
Arginine(Arg)	AGA AGG CG*	Lysine(Lys)	AAA AAG
Asparagine(Asn)	AAC AAT	Methionine(Met)	ATA ATG ATT
Aspartic Acid(Asp)	ATC	Phenylalanine(Phe)	TTC TTT
Cysteine(Cys)	TGC TGT	Valine(Val)	CC*
Glutamic acid(Glu)	GAA GAG	Serine(Ser)	AGC AGT TC*
Glutamine(Gln)	CAA CAG	Threonine(Thr)	AC*
Glycine(Gly)	GG*	Tryptophan(Trp)	GT*
Histidine(His)	CAC CAT	Tyrosine(Tyr)	TAC TAT
Isoleucine(Ile)	GAC GAT	Proline(Pro)	TGG

Σημείωση (1): Ο αστερίσκος (*) υποδηλώνει τις βάσεις A, C, G, ή T.

Σημείωση (2): Για λόγους απλότητας, η βάση «Τ» χρησιμοποιείται εδώ σαν ένα νουκλεοτίδιο του mtDNA. Ωστόσο χρησιμοποιείται μια ουρακίλη (U) στο RNA που μεταγράφεται. Επίσης, τα μιτοχόνδρια χρησιμοποιούν έναν ελαφρώς διαφορετικό γενετικό κώδικα από το πυρηνικό DNA- ATA και ATT κωδικοποιούν για την μεθειονίνη αντί της κανονικής ισολευκίνης. Τα κωδικόνια λήξης μερικές φορές συμπληρώνονται μόνο στην μετα-μεταγραφική επεξεργασία, και έτσι ορισμένα mtDNA γονίδια δεν θα έχουν κανονικό κωδικόνιο λήξης στο τέλος του γονιδίου.

Ουσιαστικά για να εξασφαλίσουμε ότι έχουμε την σωστή πρωτεΐνη θα πρέπει να τροποποιηθούν όλα τα 5'-CUA-3' κωδικόνια στο μιτοχονδριακό γονίδιο σε 5'-CAN-3' (όπου N= A,T,C,G) πριν από τη θέση του γονιδίου σε ένα χρωμόσωμα. Αυτό εξασφαλίζει ότι θα υπάρχει Thr σε όλες τις θέσεις που απαιτείται στην πρωτεΐνη.

Η **έναρξη** και η **λήξη** κωδικής περιοχής συνήθως έχει τριπλέτα κωδικονίων:

Έναρξη: ATA ATG ATT

Λήξη: TAA TAG TG

Πίνακας 4 Οι θέσεις των tRNAs στο mtDNA (Accession number: U96639.2)

Αμινοξύ	Θέση	Αμινοξύ	Θέση
Phenylalanine	1..69	Serine(UCN)	6891..6961
Valine	1024..1090	Aspartic acid	6966..7033
Leucine(UUR)	2671..2744	Lysine	7735..7801
Isoleucine	3703..3771	Glycine	9428..9495
Glutamine	3768..3842	Arginine	9842..9910
Methionine	3844..3913	Histidine	11577..11647
Tryptophan	4956..5023	Serine(AGY)	11648..11707
Alanine	5037..5105	Leucine(CUN)	11708..11777
Asparagine	5107..5178	Glutamic acid	14108..14178
Cysteine	5212..5279	Threonine	15323..15392
Tyrosine	5280..5347	Proline	15392..15457

Σημείωση: Οι δύο περιοχές που κωδικοποιούν για λευκίνη και σερίνη αντιστοιχούν έναντι διαφορετικών αντικωδικονίων.

Παρόλο που μικρές διαφορές στο **μέγεθος** του μιτοχονδριακού γονιδιώματος (σε επίπεδο μερικών ζευγών βάσεων) μπορεί να παρατηρούνται τόσο μεταξύ ατόμων διαφορετικών ειδών όσο και μεταξύ ατόμων του ίδιου είδους αλλά διαφορετικών πληθυσμών (Crews, 1979, Upholt & Dawid, 1977, Greenberg, 1982, Aquardo & Greenberg, 1983) το μέγεθος του μιτοχονδριακού γονιδιώματος είναι σχετικά σταθερό σε επίπεδο ειδών και σε μερικές περιπτώσεις παραμένει σταθερό και μεταξύ ανώτερων ταξινομικών βαθμίδων.

Η παρουσία διαφορετικών μεγεθών mtDNA υποδηλώνει σαφέστατα την ταυτόχρονη ύπαρξη μηχανισμών οι οποίοι είναι υπεύθυνοι για τη δημιουργία ενθέσεων ή ελλείψεων στο γονιδίωμα. Μικρές διαφορές στο μέγεθος του γονιδιώματος του mtDNA εμφανίζονται ως αποτέλεσμα ενθέσεων ή ελλείψεων μερικών ζευγών βάσεων (Crews et al, 1979, Brown et al, 1982, Aquardo & Greenberg 1983, Greenberg et al. 1982). Παρόλο που αυτές οι μεταβολές λαμβάνουν χώρα σε όλες τις περιοχές του γονιδιώματος, η συχνότητα εμφάνισής τους ποικίλλει μεταξύ των διαφόρων περιοχών με τη μικρότερη συχνότητα στις περιοχές εκείνες στις οποίες κωδικοποιούνται πρωτεΐνες.

Είναι σημαντικό ότι τα μιτοχονδριακά γονιδιώματα των ζώων **εξελισσονται** ταχύτατα. Ο γρήγορος εξελικτικός ρυθμός αποδίδεται σε μεταλλάξεις και άρα στον υψηλότερο ρυθμό εισαγωγής σφαλμάτων, οι οποίες συσσωρεύονται λόγω της χαμηλότερης αξιοπιστίας του μηχανισμού αντιγραφής και οδηγούν στην αδυναμία επιδιόρθωσης των λαθών κατά την αντιγραφή (Wilson et al. 1985). Σε αντίθεση με τις πολυμεράσες DNA του πυρήνα, οι πολυμεράσες DNA του μιτοχονδρίου δεν έχουν ικανότητα επιδιορθωτικού ελέγχου.

Το mtDNA εμφανίζεται σε μεγάλο αριθμό αντιγράφων στα κύτταρα, είναι πλήρως χαρακτηρισμένο, ενώ ο **μητρικός** (Kondo et al., 1990; Gyllestein et al., 1991), κυρίως, τρόπος κληρονομικότητας (Zhao X. et al. , 2004) δεν αφήνει περιθώρια για **ανασυνδυασμό** (Clayton, 1982, 1992; Hayashi et al., 1985), με ελάχιστες εξαιρέσεις (Rokas et al. 2003). Ο μητρικός τρόπος κληρονομής των μιτοχονδρίων οδηγεί στο συμπέρασμα ότι οι απόγονοι της ίδιας μητρικής γενιάς έχουν το ίδιο μιτοχονδριακό DNA και μπορούν να διαχωριστούν μέσω αυτού. Υπολογίζεται ότι για τα μιτοχονδριακά γονίδια στα ζώα και τα φυτά, ο πατρικός γενότυπος ανιχνεύεται όχι συχνότερα του 1-10% (Birky et al., 1983). Επειδή όλο το μιτοχονδριακό DNA κληρονομείται ως μία μονάδα, ή απλότυπος, η σύγκριση του μιτοχονδριακού DNA από διαφορετικά άτομα μπορεί να παρουσιαστεί με τη μορφή φυλογενετικού δένδρου.

Οι μελέτες στο επίπεδο του mtDNA δείχνουν ότι οι πληθυσμοί ενός είδους παρουσιάζουν μια γεωγραφικά εντοπισμένη γενεαλογική δομή (Bermingham & Avise, 1986; Avise, 1989). Το πρότυπο της φυλογεωγραφικής δομής του mtDNA εξαρτάται σε μεγάλο βαθμό από την ικανότητα διασποράς του οργανισμού, καθώς και από το βαθμό κατάτμησης του περιβάλλοντος στο οποίο αυτός κατοικεί. Συνεπώς, οργανισμοί με μικρή ικανότητα **μετανάστευσης** ή οργανισμοί που ζουν σε απομονωμένα περιβάλλοντα εμφανίζουν μεγάλη γεωγραφική-γενετική διαφοροποίηση ενώ αντίθετα οργανισμοί με μεγαλύτερη ικανότητα διασποράς ή οργανισμοί που καταλαμβάνουν χώρους με εκτεταμένη γεωγραφική συνέχεια, παρουσιάζουν μικρότερη γενετική διαφοροποίηση.

Περιορισμοί στη **γονιδιακή ροή**, που πιθανώς να σχετίζονται με μικρό δραστικό μέγεθος του πληθυσμού των θηλυκών Nf(e) αλλά και υποδιαιρέσεων των πληθυσμών εξαιτίας ιστορικών δημογραφικών παραγόντων οδηγούν στον γεωγραφικό εντοπισμό και στα φυλογενετικά χάσματα των mtDNA απλοτύπων που συναντάμε ανάμεσα στους πληθυσμούς ενός είδους (Slatkin, 1987; Avise, 1994).

Είναι αξιοσημείωτο ότι η μελέτη του mtDNA παρέχει πολύτιμες πληροφορίες για την γονιδιωματική εξέλιξη. Επιπλέον, η σύγκριση της οργάνωσης των μιτοχονδριακών γονιδίων είναι δυνατόν να οδηγήσει στην εξαγωγή αρχαίων **εξελικτικών σχέσεων**, καθώς η διάταξη των γονιδίων σε απόμακρες ταξινομικά ομάδες είναι μοναδική και έτσι είναι σχεδόν αδύνατο να συμβεί αυτό το γεγονός με τον ίδιο τρόπο σε ανεξάρτητες εξελικτικά γενεαλογίες (Boore, 1999).

Για παράδειγμα, τα δύο γονίδια που κωδικοποιούν για τα rRNAs (12S και 16S) είναι πιο συντηρημένα, οπότε μπορούν να χρησιμοποιηθούν για τον προσδιορισμό των φυλογενετικών σχέσεων σε υψηλά (φύλο, υποφύλο) και μεσαία (οικογένεια, γένος) ταξινομικά επίπεδα (Weisburg *et al.*, 1991, Ballard *et al.*, 1992). Εκτός όμως, από τις πληροφορίες που μπορεί να δώσει η αλληλούχηση των γονιδίων του μιτοχονδριακού DNA, η διάταξη των μιτοχονδριακών γονιδίων είναι δυνατό να παρέχει πληροφορίες για τη «βαθιά» φυλογένεση (Boore *et al.*, 1995). Αν εξαιρέσει κανείς την εκτεταμένη γονιδιακή μετατόπιση σε ορισμένα τάξα (Machida *et al.* 2002), η ταυτοποίηση συγκεκριμένης γονιδιακής διάταξης του μιτοχονδριακού γονιδιώματος οδήγησε στην ισχυρή στήριξη των μέχρι τώρα αβέβαιων ομαδοποιήσεων και την εύρεση νέων φυλογενετικών σχέσεων ανάμεσα σε διαφορετικά φύλα (Podsiadlowski *et al.*, 2008). Τέλος, για τη μελέτη των φυλογενετικών σχέσεων σε κατώτερα ταξινομικά επίπεδα, έχουν χρησιμοποιηθεί γονίδια όπως αυτά που κωδικοποιούν το κυτόχρωμα *b*, την υπομονάδα I της κυτοχρωμικής οξειδάσης, την περιοχή του βρόχου εκτόπισης και την υπομονάδα I της NADH αφυδρογονάσης.

Μερικοί ακόμη λόγοι είναι το μικρό του μέγεθος, ο εύκολος χαρακτηρισμός του (εύρεση γονιδίων και σειράς αυτών, αλληλούχηση) και η απουσία των πολύπλοκων χαρακτήρων που υπάρχουν στο mtDNA όπως τα ιντρόνια, οι επαναλαμβανόμενες αλληλουχίες, οι διαγονιδιακές αλληλουχίες, τα ψευδογονίδια και τα μεταθετά

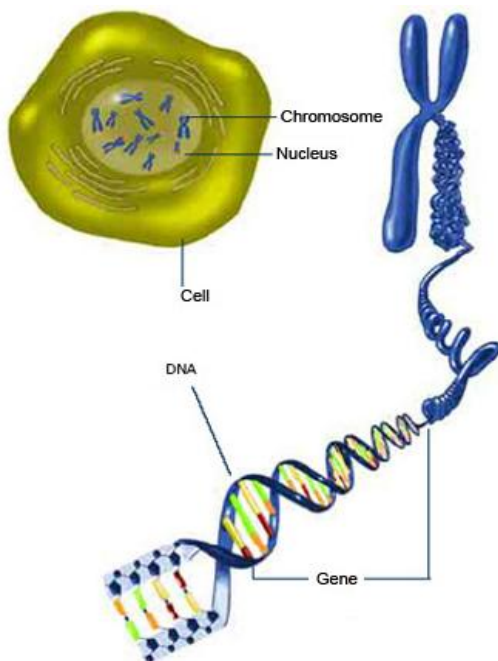
στοιχεία.

Η αξία του μιτοχονδριακού DNA σε μελέτες μοριακής οικολογίας και φυλογεωγραφίας είναι ανεκτίμητη. Ωστόσο, τα τελευταία χρόνια, υπάρχει η αντίληψη ότι η αναφορά σε ένα μόνο μόριο δεν είναι πολλές φορές ικανοποιητική για την απάντηση σε διάφορα ερωτήματα. Υποστηρίζεται συνεπώς από πολλούς ερευνητές ότι είναι καιρός να επανεκτιμηθεί η αξιοπιστία του mtDNA στη φυλογεωγραφία.

Όλα τα παραπάνω έκαναν το mtDNA αναπόσπαστο κομμάτι της συστηματικής και της γενετικής πληθυσμών (Avisé, 1986, Awadalla et al., 1999, Simon et al., 2006). Όμως, παρόλα τα πλεονεκτήματα που παρουσιάζει του ως πηγή μοριακών δεικτών για φυλογενετικές μελέτες, το γεγονός ότι κληρονομείται ως απλότυπος, έχει ως συνέπεια να μην είναι δυνατόν να θεωρηθούν τα γονίδια που περιλαμβάνει ως ανεξάρτητες πηγές φυλογενετικής πληροφορίας.

Έτσι λοιπόν, εναλλακτικά από τη χρήση του mtDNA θα μπορούσαν να χρησιμοποιηθούν καιάλλοι γενετικοί μάρτυρες, όπως γονίδια από το πυρηνικό γονιδίωμα. Παρόλο που υφίστανται κάποιοι περιορισμοί, όπως η μητρική κληρονομία, επειδή το μιτοχονδριακό DNA κληρονομείται ανεξάρτητα από το πυρηνικό, η χρήση δεικτών και από τα δύο γονιδιώματα μπορεί να προσφέρει βαθύτερη κατανόηση των γενετικών σχέσεων μεταξύ των ειδών ενός συμπλέγματος (Muraji and Nakahara, 2001).

1.2.4 ΠυρηνικόDNA



Το πυρηνικό γονιδίωμα περιλαμβάνει περιοχές που κωδικοποιούν για πρωτεΐνες, RNA, ενώ φέρει και μη κωδικές περιοχές και προσφέρει μια πληθώρα ανεξάρτητων δεικτών, οι οποίοι εξελίσσονται με διαφορετικούς ρυθμούς. Ωστόσο, η χρήση των πυρηνικών γονιδίων στις φυλογενετικές αναλύσεις είναι πιο δύσκολη σε σχέση με τα μιτοχονδριακά. Οι μη κωδικές περιοχές, καθώς και οι βρόγχοι των rRNA γονιδίων, εξελίσσονται γρηγορότερα, οπότε μπορούν πιθανώς να χρησιμοποιηθούν για τη μελέτη των σχέσεων μεταξύ συγγενών ειδών (Dolman & Phillips, 2004, Willows– Munro et al., 2005). Οι περιοχές αυτές, όμως, έχουν

την τάση να ποικίλλουν σε μήκος δυσκολεύοντας πολύ τη στοίχιση των αλληλουχιών, ειδικά σε υψηλότερα ταξινομικά επίπεδα (Sequeira et al., 2006).

Οι πυρηνικοί γενετικοί τύποι που κωδικοποιούν πρωτεΐνες συνήθως δεν εμφανίζουν πολυμορφισμό μεγέθους ή, εάν εμφανίσουν, αυτός αντιστοιχεί σε αριθμό

πολλαπλάσιο των τριών νουκλεοτιδίων, με αποτέλεσμα η ομοπαράθεση των αλληλουχιών να μην παρουσιάζει δυσκολίες.

Τα πλεονεκτήματα αυτά καθιστούν τους πυρηνικούς γενετικούς τόπους πολύ καλή εναλλακτική λύση για τις φυλογενετικές αναλύσεις, ειδικά σε ανώτερα επίπεδα φυλογένεσης.

Ωστόσο, οι συγκεκριμένοι γενετικοί δείκτες είναι δυνατόν να δημιουργήσουν δυσκολίες κατά την επεξεργασία των αποτελεσμάτων, λόγω των ιντρονίων που περιέχουν (Downie & Gullan, 2004) και για αυτό το λόγο, σε ορισμένες φυλογενετικές μελέτες χρησιμοποιούνται μεμονωμένες περιοχές εξονίων. Αυτοί οι δείκτες αξιοποιούνται συνήθως σε φυλογενετικές μελέτες υψηλότερων ταξινομικών επιπέδων. Από την άλλη πλευρά, οι αλληλουχίες και η τοπολογία των ιντρονίων είναι δυνατόν να παρέχουν χρήσιμες πληροφορίες για την εξελικτική ιστορία και τη φυλογένεση μιας ομάδας ειδών. Έτσι, οι αλληλουχίες των ιντρονίων, οι οποίες εξελίσσονται ταχέως, χρησιμοποιούνται συχνά για φυλογενετικές μελέτες που αφορούν στενά σχετιζόμενα είδη ενώ η εξέλιξη της γονιδιακής οργάνωσης, μέσω απώλειας ή προσθήκης ιντρονίων, γεγονός που φαίνεται να συμβαίνει με πολύ αργούς ρυθμούς σε ορισμένες γενεαλογίες, χρησιμοποιείται συχνά για τη διερεύνηση βαθύτερων φυλογενετικών σχέσεων για τις οποίες από μόνη της η αλληλουχία, λόγω μεγάλων αποκλίσεων, δεν παρέχει αξιόπιστα συμπεράσματα (Irimia&Roy, 2008).

Τέλος, στις φυλογενετικές αναλύσεις χρησιμοποιούνται επίσης οι γενετικοί τόποι που κωδικοποιούν τη σύνθεση του rRNA. Οι γενετικοί αυτοί τόποι βρίσκονται συγκεντρωμένοι σε γειτονικές περιοχές του γονιδιώματος και οργανώνονται σε μια ενιαία λειτουργική μονάδα. Κάθε τέτοια μονάδα περιλαμβάνει γονίδια τα οποία κωδικοποιούν τη μεγάλη (28S) και τη μικρή (18S) υπομονάδα του ριβοσώματος, καθώς και το 5,8S rRNA, τα οποία χωρίζονται από μη κωδικές περιοχές (ETS, ITS1, ITS2) .

Άλλοι πυρηνικοί γενετικοί τόποι που μπορούν να χρησιμοποιηθούν για μελέτες φυλογένεσης είναι οι μικροδορυφορικές αλληλουχίες. Πρόκειται για επαναλαμβανόμενες αλληλουχίες ενός έως πέντε νουκλεοτιδίων που υπάρχουν σε σημαντικό ποσοστό στο γονιδίωμα των ευκαρυωτικών οργανισμών. Εμφανίζουν υψηλά επίπεδα πολυμορφισμού, λόγω του υψηλού ρυθμού μεταλλάξεων. Η αξιοποίησή τους σε μελέτες πληθυσμιακής γενετικής συνέβαλαν στον καθορισμό της γενετικής δομής πολλών ειδών, συγκρίνοντας τις εξελικτικές σχέσεις και μελετώντας την πρόσφατη ιστορία σε επίπεδο πληθυσμών (Bubuguzan *et al.*, 2007).

Το κύριο πλεονέκτημα που καθιστά το πυρηνικό ριβοσωμικό DNA κατάλληλο φυλογενετικό δείκτη, είναι το μεγάλο μέγεθος της αλληλουχίας που συνεπάγεται και μεγάλη ποσότητα φυλογενετικής πληροφορίας. Η επιλογή του κατάλληλου γενετικού δείκτη στην εκάστοτε φυλογενετική μελέτη εξαρτάται από τον αντίστοιχο εξελικτικό ρυθμό ενώ, κατά περίπτωση, οι εναλλακτικοί γενετικοί δείκτες μπορούν να χρησιμοποιηθούν και συνδυαστικά (Olsen & Woese, 1993).

1.2.5 Διαφορές μιτοχονδριακών και πυρηνικών δεικτών

Τα γονιδιώματα των οργανιδίων αντιγράφονται και κληρονομούνται με διαφορετικό τρόπο απ' ό τι τα γονίδια του πυρήνα και ως εκ τούτου η δυναμική των νουκλεοτιδικών αντικαταστάσεων είναι πολύ διαφορετική. Συχνά σε αυτό διαδραματίζουν σημαντικό ρόλο οι υψηλότερες συγκεντρώσεις μεταλλαξιγόνων, όπως οι ελεύθερες ρίζες οξυγόνου, οι οποίες προκύπτουν από μεταβολικές διεργασίες που διεξάγονται στα μιτοχόνδρια.

Ο μέσος ρυθμός συνώνυμων αντικαταστάσεων στα μιτοχονδριακά γονίδια των θηλαστικών είναι περίπου $5,7 \times 10^{-8}$ ανά θέση ανά έτος, είναι δηλαδή περίπου δεκαπλάσιος από το μέσο ρυθμό συνώνυμων αντικαταστάσεων των γονιδίων του πυρήνα. Ο ρυθμός μη συνώνυμων αντικαταστάσεων ποικίλλει σημαντικά μεταξύ των γονιδίων του μιτοχονδρίου αλλά σε κάθε περίπτωση είναι σημαντικά υψηλότερος από το μέσο ρυθμό συνώνυμων αντικαταστάσεων που παρατηρείται στα πυρηνικά γονίδια.

Επιπλέον, είναι πιθανόν, η πίεση της φυσικής επιλογής, που φυσιολογικά εξαλείφει πολλές μεταλλαγές στα γονίδια του πυρήνα, να είναι λιγότερο έντονη στα μιτοχόνδρια, επειδή τα περισσότερα κύτταρα περιέχουν πολλές δεκάδες μιτοχόνδρια καθένα από τα οποία περιέχει μέχρι και δώδεκα αντίγραφα του μιτοχονδριακού γονιδιώματος.

Σε πολλές περιπτώσεις, το περιεχόμενο του mtDNA σε γουανίνες/κυτοσίνες (GC) διαφέρει σημαντικά από το πυρηνικό DNA. Το mtDNA δεν συνδέεται με ιστόνες ή παρόμοιες πρωτεΐνες (Russell, 2009)

Το mtDNA προτιμάται λόγω ευκολίας απομόνωσης σε σχέση με το πυρηνικό DNA. Η ευκολία αυτή έγκειται στην ασυνήθιστα μικρή πυκνότητά του σε κλίση CsCl, στο μεγάλο αριθμό αντιγράφων του και την παρουσία του σ' ένα οργανίδιο διαφορετικό του πυρήνα.

Το μικρό λοιπόν δραστικό μέγεθος του mtDNA σε σχέση με τα πυρηνικά γονίδια και ο υψηλός ρυθμός νουκλεοτιδικών υποκαταστάσεων, καθιστούν το mtDNA πολύτιμο εργαλείο στη μελέτη φυλογενετικών σχέσεων σε διάφορα ταξινομικά επίπεδα, ανάλογα πάντα με το υπό μελέτη γονίδιο, ή συνδυασμό γονιδίων.

1.3 Γονιδιωματική

1.3.1 Τι είναι γονιδιωματική:

Η **Γονιδιωματική** είναι μια υποκατηγορία της γενετικής που μελετά το γονιδίωμα των οργανισμών. Ουσιαστικά είναι ο τομέας της γενετικής που εφαρμόζει την τεχνολογία του ανασυνδυασμένου DNA, μεθόδους αλληλούχισης DNA (DNA sequencing methods) και βιοπληροφορικής (bioinformatics) για την αλληλούχιση, την οργάνωση και ανάλυση της δομής και της λειτουργίας γονιδιωμάτων. (Pevsner et al., 2009).

Η Υπηρεσία Προστασίας Περιβάλλοντος Ηνωμένων Πολιτειών έχει έναν κάπως ευρύτερο **ορισμό**:

- Ο όρος "**γονιδιωματική**" περιλαμβάνει ένα ευρύτερο πεδίο της επιστημονικής έρευνας συσχετιζόμενων τεχνολογιών από όταν η γονιδιωματική αρχικά θεωρήθηκε.
- Ένα **γονιδίωμα** είναι το σύνολο όλων των γονιδίων ενός μεμονωμένου οργανισμού.
- Έτσι, η **γονιδιωματική** είναι η μελέτη όλων των γονιδίων ενός κυττάρου, ή ιστού, στα επίπεδα του DNA (γονότυπος), mRNA (μεταγραφικό), ή της πρωτεΐνης (πρωτέωμα)".
- **Γονιδιωματική** λοιπόν πιο συγκεκριμένα είναι η καταγραφή, χαρτογράφηση και μελέτη του συνόλου των γονιδίων ενός οργανισμού.

1.3.2 Πόσα είδη γονιδιωμάτων υπάρχουν στη γη:

- ο Προκαρυωτικά γονιδιώματα
 - ο Ευκαρυωτικά γονιδιώματα
1. Γονιδιώματα του πυρήνα
 2. Γονιδιώματα του **μιτοχονδρίου**
 3. Γονιδιώματα του χλωροπλάστη

1.4 Μετα-ανάλυση

Το μεγάλο βήμα της τεχνολογίας των υπολογιστών είναι η διασύνδεσή τους μέσω Διαδικτύου. Μέσα σε λίγα δευτερόλεπτα μπορούν να μεταδοθούν με μεγάλη ταχύτητα πληροφορίες απ' όλο τον κόσμο στον υπολογιστή μας. Η εξέλιξη αυτή ανέδειξε ακόμα περισσότερο την αξία της πληροφορίας στις σύγχρονες κοινωνίες, ώστε να μιλάμε πια για Κοινωνία της Πληροφορίας.

Αν λοιπόν μέχρι σήμερα το ζητούμενο ήταν η πρόσβαση στην πληροφορία, στην σημερινή εποχή το ζητούμενο είναι η βέλτιστη διαχείριση και δυναμική επεξεργασία του τεράστιου όγκου ψηφιοποιημένης πληροφορίας που τείνει να κατακλύσει τον πολίτη των δικτύων (netizen). Σε αυτό το σημείο είναι απαραίτητο να αναφερθούμε στην έννοια της μετα-ανάλυσης.

1.4.1 Ορισμός Μετα-ανάλυσης

Μετα-ανάλυση (meta-analysis) είναι μια αντικειμενική και ποσοτική μεθοδολογία που χρησιμοποιείται για την σύνθεση (π.χ. συνδυασμό και σύνοψη) ερευνητικών μελετών που έχουν γίνει στο παρελθόν για κάποιο συγκεκριμένο θέμα, ώστε να οδηγήσουν σε ένα συνολικό συμπέρασμα. (Panaretos, J. και Xekalaki, E., 2000)

1.4.2 Πλεονεκτήματα της Μεθοδολογίας της Μετα-Ανάλυσης

Τα κυριότερα πλεονεκτήματα της μεθοδολογίας της μετα-ανάλυσης είναι τα εξής:

1. Παρέχει μια αντικειμενική σύνθεση πολλών σχετικών μελετών ή πειραμάτων για ένα συγκεκριμένο πρόβλημα προκειμένου να οδηγηθεί η έρευνα σε ένα μοναδικό συνολικό συμπέρασμα.

2. Επιτρέπει μια αξιολόγηση της μεροληπτικότητας των υπαρχουσών μελετών με τον υπολογισμό του αριθμού των μελετών που οδήγησαν στο συμπέρασμα ύπαρξης μη σημαντικής επίδρασης (που για τον λόγο αυτό, είναι ενδεχόμενο να μην έχουν πάρει ευρεία δημοσιότητα) και οι οποίες είχαν πραγματοποιηθεί πριν ο ερευνητής καταλήξει στο γενικό συμπέρασμα που θα ανατρέπει την προηγούμενη πεποίθηση διαπιστώνοντας ότι υπάρχει κάποια σημαντική επίδραση.

3. Επιτρέπει μια αντικειμενική ερμηνεία της αξίας των μεθόδων που είχαν χρησιμοποιηθεί σε περιορισμένες μελέτες δίνοντας την ευκαιρία στον ερευνητή να δώσει σταθμικά βάρη στις στατιστικές συναρτήσεις ελέγχου των μελετών αυτών πριν συγκεντρώσει και συνθέσει τα αποτελέσματα.

4. Δίνει στον ερευνητή την δυνατότητα να δώσει διαφορετική βαρύτητα σε κάθε μια από τις προηγούμενες μελέτες με βάση το μέγεθος του δείγματος που κάθε μια από αυτές είχε χρησιμοποιήσει.

5. Παρέχει διάφορες δυνατότητες συνδυασμού των ελέγχων έτσι ώστε να αξιοποιηθούν τα αποτελέσματα των προηγούμενων ερευνών που χρησιμοποίησαν μόνο την στατιστική συνάρτηση t (π.χ. τύπος του Winner (Winner's formula)) ή εκείνες για τις οποίες μόνο οι p -τιμές είχαν αναφερθεί ανεξάρτητα από την στατιστική συνάρτηση που είχε χρησιμοποιηθεί [π.χ. τύπος του Fisher (Fisher's formula)].

6. Τέλος, επιτρέπει στον ερευνητή να μετρήσει το μέγεθος της επίδρασης, δηλαδή την ποσότητα (σε τυποποιημένες μονάδες) κατά την οποία η πραγματική εναλλακτική υπόθεση (δηλαδή το αποτέλεσμα της επίδρασης) διαφέρει από την ελεγχθείσα μηδενική υπόθεση, έτσι ώστε να καθορίσει την πρακτική σημασία μιας διαπίστωσης που είναι στατιστικά σημαντική. (Panaretos, J. και Xekalaki, E. 2000)

1.4.3 Μειονεκτήματα της Μεθοδολογίας της Μετα-Ανάλυσης

Οι επικριτές της μεθοδολογίας της μετα-ανάλυσης προβάλλουν τις εξής αντιρρήσεις για την μεθοδολογία αυτή:

1. Η μεθοδολογία της μετα-ανάλυσης δεν είναι, κατ' αυτούς, τίποτα περισσότερο παρά μια "ποικιλία", ένας συνδυασμός των αποτελεσμάτων μελετών που έχουν γίνει κάτω από διαφορετικές συνθήκες, ενδεχομένως έχουν χρησιμοποιήσει διαφορετικές μεθόδους, διαφορετικούς λειτουργικούς ορισμούς, διαφορετικές κλίμακες μέτρησης και διαφορετικά είδη αντικειμένων, οπότε δεν είναι και δυνατόν να εξαχθούν λογικά συμπεράσματα από τον συνδυασμό τέτοιων μελετών.

2. Η σωστή εφαρμογή της επιστήμης απαιτεί την δυνατότητα επανάληψης ενός πειράματος. Η μεθοδολογία της μετα-ανάλυσης μπορεί να αιτιολογηθεί θεωρητικά και, όπως έδειξε το παράδειγμα που προηγήθηκε, να οδηγήσει σε μια κατάλληλη στάθμιση, όταν οι συγκεκριμένες μελέτες οι οποίες χρησιμοποιούνται και έχουν γίνει στο παρελθόν έχουν χρησιμοποιήσει την ίδια μέθοδο, περιέχουν τους ίδιους λειτουργικούς ορισμούς και κλίμακες μετρήσεων και αναφέρονται στα ίδια είδη αντικειμένων. Παρ' όλα αυτά, στην πράξη, αυτό συμβαίνει σπάνια.

3. Η μετα-ανάλυση είναι πολύ συχνά μεροληπτική προς την κατεύθυνση ενός γενικού συμπεράσματος στατιστικής σημαντικότητας, επειδή δεν έχει την δυνατότητα να πάρει υπόψη της τα αποτελέσματα μελετών που έχουν οδηγήσει σε μη σημαντικά στατιστικά συμπεράσματα διότι δεν έχουν δημοσιευθεί.

4. Η μετα-ανάλυση χρησιμοποιεί συχνά πολλαπλά συμπεράσματα από την ίδια μελέτη στην σύνθεση των αποτελεσμάτων και επομένως καθιστά αναξιόπιστα τα συνολικά συμπεράσματα, δεδομένου ότι παραβιάζει την υπόθεση της ανεξαρτησίας των παρατηρήσεων.

5. Η μετα-ανάλυση συχνά οδηγεί σε μία υπεραπλούστευση κάποιων ερευνητικών προβλημάτων δεδομένου ότι επικεντρώνει την προσοχή της σε μια συνολική μέση επίδραση αντί να μελετήσει τους ενδιάμεσους παράγοντες.

6. Στην πιο ανεπτυγμένη της μορφή, η μεθοδολογία της μετα-ανάλυσης δεν εξασφαλίζει την αντικειμενικότητα που οι υποστηρικτές της μεθόδου υποστηρίζουν ότι έχει δεδομένου ότι εξαρτάται σε μεγάλο βαθμό από την προσωπική κρίση του ερευνητή όσο αφορά την επιλογή των ερευνών που έχουν δημοσιευθεί προηγουμένως στο συγκεκριμένο θέμα και στον καθορισμό της σημασίας των ερευνών αυτών (δηλαδή στον καθορισμό των βαρών) πριν γίνει η σύνθεση των αποτελεσμάτων αυτών, όπως επίσης και από την ικανότητα του ερευνητή να εντοπίσει μη δημοσιευθείσες μελέτες. (Panaretos, J. και Xekalaki, E. 2000)

1.5 Βάσεις δεδομένων

1.5.1 Τι είναι οι βάσεις δεδομένων;

Οι βάσεις δεδομένων λειτουργούν ως χώρος αποθήκευσης και άντλησης πληροφορίας, ενώ έχουν και τη δυνατότητα αναζητήσεων και ανταλλαγής δεδομένων με άλλες βάσεις. Ο αριθμός των διαθέσιμων αλληλουχιών αυξάνει ταχύτατα.

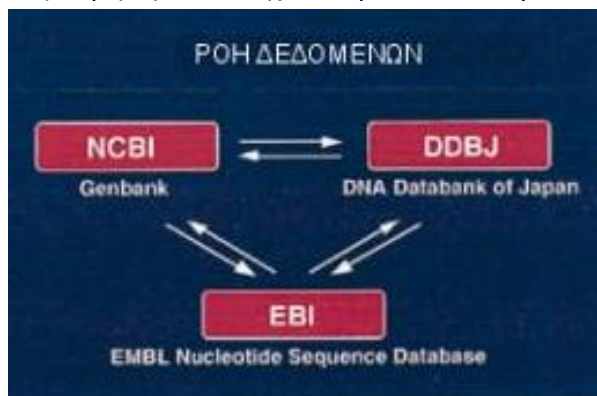
Οι πρωτογενείς βάσεις δεδομένων περιλαμβάνουν κυρίως πειραματικά αποτελέσματα (με κάποια ερμηνεία), αλλά δεν υφίστανται επιμέλεια και αναθεώρηση. Επιμέλεια και αναθεώρηση πραγματοποιείται στις δευτερογενείς βάσεις δεδομένων.

1.5.2 Κατηγορίες βάσεων δεδομένων

A. Βάσεις δεδομένων νουκλεοτιδικών ακολουθιών

Βάσεις νουκλεοτιδικών ακολουθιών ελεύθερα διαθέσιμες, οι οποίες συνεργάζονται μεταξύ τους ανταλλάσσοντας εγγραφές και δημιουργώντας κοινούς κανόνες για την ταξινόμηση και το σχολιασμό των δεδομένων.

Οι κυριότερες πηγές δεδομένων για νουκλεοτιδικές ακολουθίες είναι οι Βάσεις Δεδομένων που συμμετέχουν στο International Nucleotide Sequence Database Collaboration (DDBJ, EMBL και GenBank). Αυτές οι τρεις, ανταλλάσσουν δεδομένα κάθε 24ηπου κατατίθενται ανεξάρτητα σε κάθε βάση δεδομένων έχοντας θέσει παράλληλα και κοινούς κανόνες για την ταξινόμηση και το σχολιασμό των δεδομένων.



Εικόνα 10. Η ροή της πληροφορίας ανάμεσα στις βάσεις δεδομένων.

Αυτή η μεταφορά διευκολύνεται καθώς χρησιμοποιούν κοινές μορφοποιήσεις δεδομένων για τα είδη πληροφορίας που περιέχουν. Οι εγγραφές νουκλεοτιδικών ακολουθιών των DDBJ/EMBL/GenBank είναι η συχνότερη πηγή πληροφοριών αλληλούχησης και βιολογικών δεδομένων, από την οποία παράγονται οι εγγραφές άλλων βάσεων δεδομένων.

Πισυγκεκριμένα



DNA Data Bank of Japan (DDBJ)
(<http://www.ddbj.nig.ac.jp/>)



Submit to GenBank

GenBank

(<http://www.ncbi.nlm.nih.gov/genbank/>)

στο National Center for Biotechnology

Information (NCBI). Βρίσκεται υπό την αιγίδα του Εθνικού Ινστιτούτου Υγείας των Η.Π.Α. Η κύρια πηγή πληροφορίας προέρχεται από απευθείας υποβολές δεδομένων ερευνητών. Σε τακτά χρονικά διαστήματα τα δεδομένα επανεξετάζονται.



EMBL_Bank

(<http://www.ebi.ac.uk/embl/index.html>)

στο European Bioinformatics Institute (EBI).

Είναι η μεγαλύτερη βάση νουκλεοτιδικών ακολουθιών στην Ευρώπη και βρίσκεται υπό την αιγίδα του Ευρωπαϊκού Εργαστηρίου Μοριακής Βιολογίας (EMBL).

Β. Εξειδικευμένη βάση δεδομένων που συνδυάζει τα δεδομένα γονιδιωματικών ακολουθιών και το σχολιασμό τους με άλλα στοιχεία για τα συγκεκριμένα είδη.



Ensembl

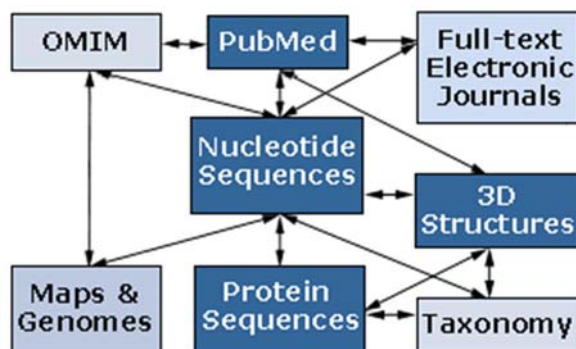
(<http://www.ensembl.org/index.html>)

Το Ensembl είναι ένα πρόγραμμα συνεργασίας ανάμεσα στο EMBL – EBI και το Wellcome Trust Sanger Institute (WTSI). Αποτελεί ένα λογισμικό που παράγει και διατηρεί “automatic annotation” σε επιλεγμένα ευκαρυωτικά γονιδιώματα. Παράγει βάσεις δεδομένων του γονιδιώματος σπονδυλωτών και άλλων ευκαρυωτικών ειδών. Όλες οι πληροφορίες του είναι διαθέσιμες δωρεάν στο δίκτυο.

Δ. Ολοκληρωμένα συστήματα ανάκτησης πληροφοριών από βάσεις δεδομένων.

- Το SRS είναι ένα ισχυρό, εύχρηστο σύστημα διαχείρισης δεδομένων το οποίο διατίθεται από την εταιρία LION Bioscience. Το SRS μέσω ενός φιλικού προς το χρήστη γραφικού περιβάλλοντος δίνει την δυνατότητα αναζήτησης και ανάκτησης δεδομένων από περισσότερες από 400 βάσεις δεδομένων οι οποίες μπορεί να είναι αποθηκευμένες στον ίδιο κεντρικό υπολογιστή.

- Το Entrez αποτελεί ένα σύστημα διαχείρισης για την αναζήτηση και ανάκτηση πληροφοριών ανάλογο του SRS όλων των βάσεων δεδομένων που περιέχονται στο NCBI (National Center for Biotechnology Information) των ΗΠΑ. Το Entrez δίνει τη δυνατότητα αναζήτησης σε βάσεις δεδομένων νουκλεοτιδικών και πρωτεϊνικών ακολουθιών, δομές βιομορίων, γονιδιωμάτων και στη βάση για την αναζήτηση βιβλιογραφίας MEDLINE μέσω του ίδιου γραφικού περιβάλλοντος επιτρέποντας και πιο πολύπλοκες αναζητήσεις ανάμεσα στα στοιχεία τους.



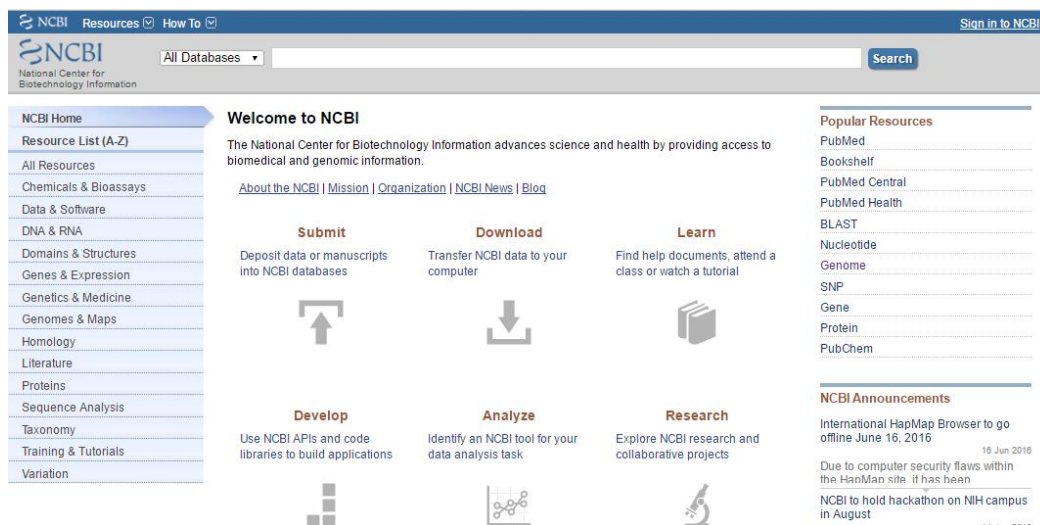
Εικόνα 11. Σχεδιάγραμμα των αλληλεπιδράσεων μεταξύ των βάσεων που διαχειρίζεται το σύστημα Entrez.

1.5.3 Τι είναι η βάση δεδομένων GenBank;



Η GenBank[®] είναι η βάση δεδομένων γενετικής αλληλουχίας του Εθνικού Ινστιτούτου Υγείας (NIH: National Institutes of Health). Είναι μια σχολιασμένη συλλογή όλων των διαθέσιμων στο κοινό DNA αλληλουχιών (Nucleic Acids Research, 2013 Jan.). Η

GenBank αποτελεί μέρος της Διεθνούς Συνεργασίας των Βάσεων Δεδομένων νουκλεοτιδικών ακολουθιών, που περιλαμβάνει την DNA DataBank της Ιαπωνίας (DDBJ), το Ευρωπαϊκό Εργαστήριο Μοριακής Βιολογίας (EMBL), και την GenBank στο NCBI.



Εικόνα 12. Αρχική σελίδα NCBI (<http://www.ncbi.nlm.nih.gov/>)

1.5.3.1 Πρόσβαση στη GenBank

Υπάρχουν διάφοροι τρόποι για να αναζητήσουμε και να ανακτήσουμε δεδομένα από την GenBank.

- Αναζήτηση στην GenBank για ακολουθίες αναγνώρισης και υποσημειώσεις με προεπιλογή nucleotide (νουκλεοτίδιο), η οποία χωρίζεται σε τρία τμήματα: Core Nucleotide (η κύρια συλλογή), dbEST (Εκφρασμένες ετικέτες αλληλουχίας, Nature Genetics), και dbGSS (Genome Survey Sequences Database).
- Αναζήτηση και align (ευθυγράμμιση) των ακολουθιών της GenBank σε μια ακολουθία επερώτησης χρησιμοποιώντας το BLAST (Basic Local Alignment Search Tool). Το BLAST αναζητά Core Nucleotide, dbEST, και dbGSS ανεξάρτητα.
- Αναζήτηση, σύνδεση, και download ακολουθιών κάποιου προγράμματος χρησιμοποιώντας NCBI e-utilities.
- Τα formatASN.1 και Flatfile είναι διαθέσιμα στο ανώνυμο serverFTP του NCBI είναι :

<ftp://ftp.ncbi.nlm.nih.gov/ncbi-asn1> και <ftp://ftp.ncbi.nlm.nih.gov/genbank>

1.5.3.2 Χρήση Δεδομένων της GenBank

Η βάση δεδομένων GenBank έχει σχεδιαστεί για να παρέχει και να ενθαρρύνει την πρόσβαση στο εσωτερικό της επιστημονικής κοινότητας για τις πιο ενημερωμένες και περιεκτικές πληροφορίες αλληλουχιών DNA. Ως εκ τούτου, η NCBI δεν θέτει κανέναν περιορισμό όσον αφορά τη χρήση ή τη διανομή των δεδομένων της GenBank.

1.5.3.3 Επεξήγηση των σημαντικότερων πεδίων μιας εγγραφής στην GENBANK

Canis familiaris mitochondrion, complete genome

GenBank: U96639.2

[FASTA](#) [Graphics](#)

Go to:

```
LOCUS          CFU96639          16727 bp    DNA     circular MAM 11-APR-2000
DEFINITION     Canis familiaris mitochondrion, complete genome.
ACCESSION     U96639
VERSION       U96639.2  GI:7534303
KEYWORDS      .
SOURCE        mitochondrion Canis lupus familiaris (dog)
ORGANISM      Canis lupus familiaris
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Laurasiatheria; Carnivora; Caniformia; Canidae;
               Canis.
REFERENCE     1 (bases 1 to 16727)
AUTHORS      Kim,K.S., Lee,S.E., Jeong,H.W. and Ha,J.H.
TITLE        The complete nucleotide sequence of the domestic dog (Canis
               familiaris) mitochondrial genome
JOURNAL      Mol. Phylogenet. Evol. 10 (2), 210-220 (1998)
PUBMED      9878232
REFERENCE     2 (bases 1 to 16727)
AUTHORS      Kim,K.S., Lee,S.E., Jeong,H.W., Jeong,S.Y., Sohn,H.S. and Ha,J.H.
TITLE        Direct Submission
JOURNAL      Submitted (07-APR-1997) Genetic Engineering, Animal Genetics, 1370
               Sankyuk-dong, Pukgu, Taegu 702-701, Korea
REFERENCE     3 (bases 1 to 16727)
AUTHORS      Kim,K.S., Lee,S.E., Jeong,H.W., Jeong,S.Y., Sohn,H.S. and Ha,J.H.
TITLE        Direct Submission
JOURNAL      Submitted (11-APR-2000) Genetic Engineering, Animal Genetics, 1370
               Sankyuk-dong, Pukgu, Taegu 702-701, Korea
REMARK       Sequence update by submitter
COMMENT      On Apr 11, 2000 this sequence version replaced gi:4154170.
FEATURES     Location/Qualifiers
               source
                 1..16727
                 /organism="Canis lupus familiaris"
                 /organelle="mitochondrion"
                 /mol_type="genomic DNA"
                 /sub_species="familiaris"
                 /db_xref="taxon:9615"
```

Εικόνα 13. Αποτελέσματα μετά από αναζήτηση για το ολικό μιτοχονδριακό γονιδίωμα του *Canis familiaris* στην βάση δεδομένων NCBI Genome

- ✓ LOCUS: Περιέχει ένα μικρό όνομα για τον χαρακτηρισμό της εγγραφής.
- ✓ DEFINITION: Μια λεπτομερής περιγραφή της ακολουθίας.
- ✓ ACCESSION: Κωδικός που αποκτά μια νεοεισερχόμενη εγγραφή χαρακτηριστικός για την GENBANK. Ο κωδικός παραμένει σταθερός
- ✓ VERSION: Ειδικός κωδικός που απαρτίζεται από το πρωταρχικό Accession Number, ακολουθεί το σύμβολο της τελείας και στη συνέχεια ένας αριθμός που δηλώνει την έκδοση της παρούσας εγγραφής.

- ✓ KEYWORDS: Χαρακτηριστικές λέξεις-κλειδιά που σχετίζονται με την νουκλεοτιδική αλληλουχία και τις ιδιότητες των προϊόντων της.
- ✓ SOURCE: Βιολογική πηγή της ακολουθίας όπου αναφέρεται ο οργανισμός από τον οποίο έχει απομονωθεί με τα ιδιαίτερα χαρακτηριστικά του (πιθανές μεταλλάξεις, πλασμίδια κ.α.).
- ✓ ORGANISM: Οργανισμός απ' όπου προήλθε η ακολουθία. Ακολουθείται η διώνυμη ονομασία κατά Λινναίο. Επίσης παρατίθεται και η συστηματική ταξινόμηση του οργανισμού.
- ✓ REFERENCE: Περιέχει τον αριθμό της αναφοράς καθώς και το μήκος της ακολουθίας που έχει προσδιοριστεί στην παρούσα εργασία.
- ✓ AUTHORS: Αναφέρονται οι συμμετέχοντες στην διεξαγωγή της παρούσας εργασίας.
- ✓ TITLE: Τίτλος της δημοσιευμένης εργασίας.
- ✓ JOURNAL: Περιέχει λεπτομερειακά στοιχεία για την αναζήτηση της αναφοράς όπως είναι ο τίτλος του περιοδικού που εκδόθηκε, τεύχος, ημερομηνία έκδοσης και σελίδες που καταλαμβάνει στο συγκεκριμένο τεύχος.
- ✓ MEDLINE: Κωδικός για την βιβλιογραφική αναφορά στην βάση δεδομένων MEDLINE.
- ✓ COMMENT: Περιέχει κάποιες γενικές παρατηρήσεις, ή αναφορές και σε άλλες βάσεις.
- ✓ FEATURES: Πίνακας που περιέχει πληροφορίες σχετικά με τα προϊόντα της ακολουθίας όπως πολυπεπτιδικές αλυσίδες (από μετάφραση) και RNA (από μεταγραφή) και στοιχεία από πειραματικά δεδομένα που καταδεικνύουν τη βιολογική της σημασία.
- ✓ BASE COUNT: Αριθμητική ανάλυση της ακολουθίας στα επιμέρους συστατικά της. Περιέχει το σύνολο καταλοίπων Αδενίνης, Γουανίνης, Κυτοσίνης, Θυμίνης.
- ✓ ORIGIN: Θέση της πρώτης βάσης της κατατεθειμένης ακολουθίας σε σχέση με το γονιδίωμα από το οποίο έχει απομονωθεί.

Ακριβώς από κάτω παρατίθεται η ακολουθία της παρούσας εγγραφής.

Η αναπαράσταση της ακολουθίας είναι της μορφής:

ORIGIN

```
1 gttaatgtag cttattaat aaagcaaggc actgaaaatg ccaagatgag tcgcacgact
61 ccataaacat aaaggtttgg tcctagcctt cctattagtt ttagtagac ttacacatgc
.....(συνεχίζεται)
```



```

tRNA      1..69
           /product="tRNA-Phe"
rRNA      70..1023
           /product="s-rRNA"
           /note="12S ribosomal RNA"
tRNA      1024..1090
           /product="tRNA-Val"
rRNA      1091..2670
           /product="l-rRNA"
           /note="16S ribosomal RNA"
STS       2341..2534
           /standard_name="RH136367"
           /db_xref="UniSTS:210357"
tRNA      2671..2744
           /product="tRNA-Leu"
           /codon_recognized="UUR"
gene      2747..3702
           /gene="ND1"
           /db_xref="GeneID:804476"
CDS       2747..3702
           /gene="ND1"
           /note="TAA stop codon is completed by the addition of 3' A
           residues to the mRNA"
           /codon_start=1
           /transl_except=(pos:3701..3702,aa:TERM)
           /transl_table=2
           /product="NADH dehydrogenase subunit 1"
           /protein_id="NP_008471.1"
           /db_xref="GI:5835653"
           /db_xref="GeneID:804476"
           /translation="MFFINIISLIIPILLAVAF LTLVERKVLGYMQLRKGPNIVGPGY
           LLQPIADAVKLF TKEPLRPLTSSMSMFI LAPILALS LALTMWIPLMPYPLINMNLGV
           LFM LAMSSLAVYSILWSGWASNSKYALIGALRAVAQTISYEVT LAIILL SVLLMNGSF
           TLSTLIITQEHWLIFPAWPLAMWFI STLAETNRAPFDL TEGESELVSGFNVEYAAG
           PFALFFLAEYANIIMMNI LTTILFFGAFHNPFMP ELYSINFMTKLLLTICFLWIRAS
           YPRFRYDQLMHL LKKNFLPLTLALCMW HVALPIITASIPPQT"
tRNA      3703..3771
           /product="tRNA-Ile"
tRNA      complement(3768..3842)
           /product="tRNA-Gln"
tRNA      3844..3913
           /product="tRNA-Met"

```

Εικόνα 14. Διακρίνουμε τα αρχικά από τα 37 γονίδια του mtDNA τα οποία κωδικοποιούν ριβοσωμικά RNAs (rRNAs), πρωτεΐνες και μεταφορικά RNAs (tRNAs). Επίσης διαφαίνεται η θέση τους στο μιτοχονδριακό DNA καθώς και ποιο γονίδιο κωδικοποιούν. Όσον αφορά τις CDS (κωδικές περιοχές) μπορούμε να αντλήσουμε περισσότερες πληροφορίες όπως την θέση έναρξης του κωδικονίου, την μετάφραση της κ.α.

1.5.4 Μορφοποιήσεις βάσεων δεδομένων (databaseformats)

Οι βασικές πληροφορίες είναι κατατεθειμένες στις DDBJ/EMBL/GenBank με μορφή επιπέδων αρχείων (flatfiles). Η ύπαρξη αντιστοιχίας μεταξύ των διάφορων μορφοποιήσεων ενός "επιπέδου" (flat) αρχείου διευκολύνει την ανταλλαγή δεδομένων μεταξύ των αντίστοιχων βάσεων δεδομένων: στις περισσότερες περιπτώσεις, τα πεδία των διάφορων μορφών ενός αρχείου έχουν μεταξύ τους αντιστοιχία ένα προς ένα. Ο βαθμός επιτυχίας μιας ορισμένης μορφοποίησης εξαρτάται από την χρησιμότητά της σε μια ποικιλία περιεχομένων, καθώς και οι δυνατότητές της να συμπεριλαμβάνει αποτελεσματικά τις μορφές βιολογικής

πληροφορίας που πρέπει να αρχειοθετηθούν και να διατεθούν στην επιστημονική κοινότητα.

Στην απλούστερη μορφή της, μια εγγραφή ακολουθίας μπορεί να αναπαρασταθεί ως μια σειρά νουκλεοτιδίων με μια βασική ετικέτα (tag) ή αναγνωριστικό (identifier). Η πιο δημοφιλής από αυτές τις μορφές είναι η FASTA, η οποία προσφέρει έναν εύκολο τρόπο χειρισμού των πρωτογενών δεδομένων, τόσο από ανθρώπους όσο και από υπολογιστές.

1.5.4.1 Ο αλγόριθμος FASTA

Ο αλγόριθμος FASTA δημιουργήθηκε από τους Lipman & Pearson το 1985 και βασίζεται στην κεντρική ιδέα της αναζήτησης μικρών λέξεων (words ή k-tuples) που εμφανίζονται και στις δυο ακολουθίες. Στην περίπτωση πρωτεϊνικών ακολουθιών το μήκος των λέξεων είναι 1-2 residues ενώ για ακολουθίες DNA το μήκος μιας λέξης μπορεί να φθάνει τις 6 βάσεις. Ο αλγόριθμος χρησιμοποιεί ευρετικές μεθόδους για να δημιουργήσει περιοχές που περιέχουν κοινές λέξεις. Η στοίχιση που προκύπτει περιλαμβάνει διαφορές ανάμεσα σε κοινές λέξεις.

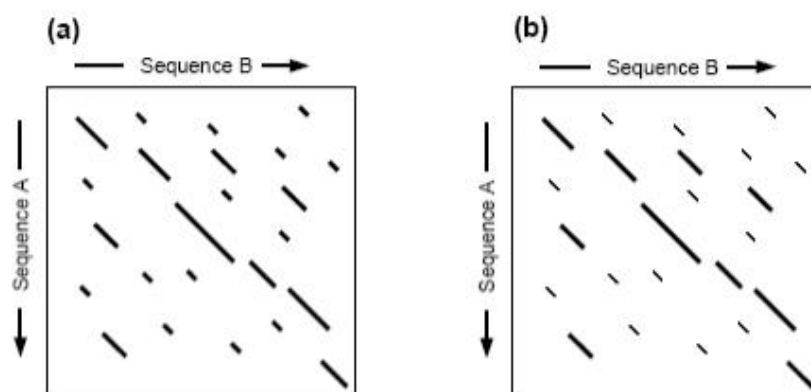
Canis familiaris mitochondrion, complete genome

GenBank: U96639.2

[GenBank](#) [Graphics](#)

```
>gi|7534303|gb|U96639.2|CFU96639 Canis familiaris mitochondrion, complete genome  
GTTAATGTAGCTTAATTAATAAAGCAAGGCACTGAAAATGCCAAGATGAGTCGCACGACTCCATAAACAT
```

Εικόνα 15. Η αρχή του αλγορίθμου fasta. Ο χαρακτήρας ">" υποδεικνύει την αρχή μιας νέας εγγραφής ακολουθίας. Η γραμμή που ακολουθεί το ">" ονομάζεται γραμμή ορισμού. Ένα αναγνωριστικό (συνήθως ο αριθμός εισαγωγής, αριθμός δημοσίευσης) που ακολουθεί το ">" ακολουθείται από την ακολουθία DNA, είτε σε κεφαλαία είτε σε πεζά γράμματα (συνήθως δίνονται 60 χαρακτήρες ανά γραμμή). (Baxevanis,2016)

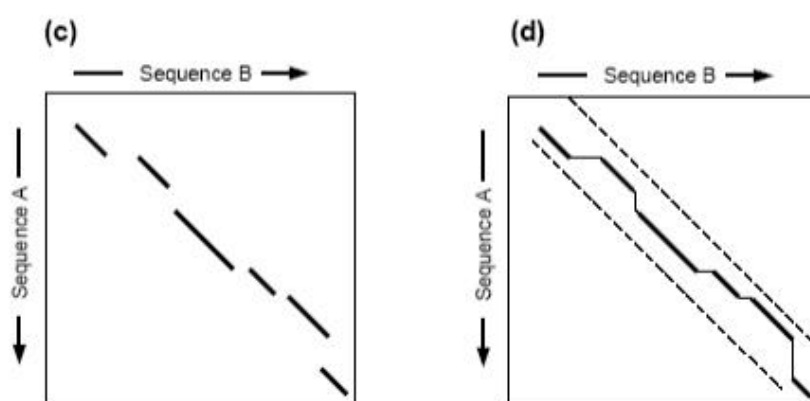


Εικόνα 16 Βήμα 1^ο και 2^ο αλγορίθμου Fasta

Βήμα 1^ο: Σχηματίζονται δύο λίστες λέξεων μήκους k (k-mers), οι οποίες προκύπτουν αντίστοιχα από την προς αναζήτηση (query) και από την πρότυπη ακολουθία

ξεκινώντας από την αρχή και μετατοπίζοντας κάθε φορά το πλαίσιο ανάγνωσης κατά 1 χαρακτήρα.

Βήμα 2^ο: Γίνεται ακριβής αναζήτηση για τις παραπάνω λέξεις στην πρότυπη ακολουθία. Η έννοια "ακριβής" χρησιμοποιείται σχετικά ευρέως, αφού πρακτικά μας ενδιαφέρουν ακολουθίες που εμφανίζουν ένα υψηλό ποσοστό ομοιότητας, το οποίο με την σειρά του υπολογίζεται βάσει πινάκων αντιστοίχισης. Οι αντιστοιχίσεις απεικονίζονται με τη μορφή διαγωνίων (**Εικόνα 17**). Μόνο οι καλύτερες 10 αντιστοιχίσεις επιλέγονται για περαιτέρω εξέταση.

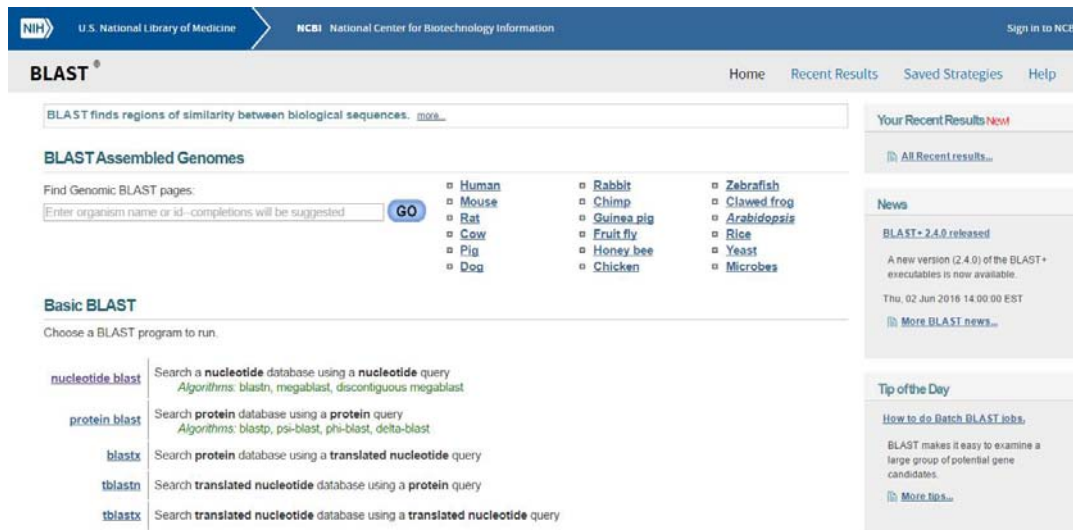


Εικόνα 17 Βήμα 3^ο και 4^ο αλγορίθμου Fasta

Βήμα 3^ο: Ελέγχεται αν κάποιες από τα τμήματα της query ακολουθίας που ευθυγραμμίστηκαν με την πρότυπη μπορούν να συνενωθούν για να σχηματίσουν μεγαλύτερες ευθυγραμμίσεις, έστω προσεγγιστικές. Σε αυτό το στάδιο επιτρέπεται και η εισαγωγή κενών, ωστόσο κάθε τέτοια κίνηση αντιστοιχεί σε αρνητικό score. Τελικά υπολογίζονται τα scores των νέων (μεγαλύτερων σε μήκος) ευθυγραμμίσεων και οι ευθυγραμμίσεις με τα υψηλότερα scores αποθηκεύονται για τη συνέχεια.

Βήμα 4^ο: Χρησιμοποιείται ο Smith - Waterman (δυναμικός προγραμματισμός) για να υπολογιστούν οι βέλτιστες τιμές ομοιότητας (scores) για κάθε ευθυγράμμιση. Οι ευθυγραμμίσεις αξιολογούνται με κριτήριο την αναμενόμενη κατανομή τυχαίων ομοιοτήτων, εξάγεται δηλαδή συμπέρασμα ως προς το κατά πόσο η ομοιότητα βασίζεται σε πραγματική σχέση μεταξύ των ακολουθιών ή έχει προέλθει τυχαία.

1.5.5 Εισαγωγή στο Blast



Εικόνα 18. Αρχική σελίδα Blast

Στην Επιστήμη της Γενετικής, έχουν δημιουργηθεί διάφορα εργαλεία για την σύγκριση γενετικών αλληλουχιών. Το πιο γνωστό μεταξύ αυτών είναι το BLAST (Basic Local Alignment Search Tool). Αυτό το λογισμικό πρόγραμμα βρίσκει περιοχές ομοιότητας μεταξύ των γενετικών αλληλουχιών. Συγκρίνει νουκλεοτιδικές ή πρωτεϊνικές αλληλουχίες με αλληλουχίες που βρίσκονται ήδη σε μία βάση δεδομένων, και υπολογίζει τη στατιστική σημαντικότητα των αποτελεσμάτων. Το πρόγραμμα BLAST μπορεί να χρησιμοποιηθεί για την εκτίμηση της συγγένειας και της εξελικτικής σχέσης μεταξύ διαφορετικών αλληλουχιών. Το πρόγραμμα χρησιμοποιεί μια βάση δεδομένων η οποία περιέχει γενετικές ακολουθίες και συγκρίνει βάση προς βάση μια ακολουθία νουκλεοτιδίων που θα ορίσει ο χρήστης, με μία άλλη που προέρχεται από τη βάση δεδομένων (alignment).

Το Blast είναι μια μηχανή έρευνας που αναπτύχθηκε και συντηρήθηκε από το NCBI (National Centre for Biotechnology Information).

Το σύνολο προγραμμάτων BLAST έχει σχεδιαστεί για να βρίσκει τοπικές ευθυγραμμίσεις ανάμεσα σε ακολουθίες, χωρίς να θυσιάζει την ταχύτητα τέτοιων ερευνών. Το BLAST χρησιμοποιεί έναν αλγόριθμο που αναζητά τοπικές ευθυγραμμίσεις αντίθετες των παγκοσμίων και επομένως είναι ικανό να αναζητήσει σχέσεις ανάμεσα σε ακολουθίες που μοιράζονται μόνο απομονωμένες περιοχές ομοιότητας. Η πρώτη έκδοση του BLAST, εκδόθηκε το 1990 και επέτρεπε την εκτέλεση μόνο έρευνες χωρίς κενά. Η δεύτερη έκδοση του BLAST, εκδόθηκε το 1997 και επέτρεπε έρευνες με κενά.

1.5.5.1 Ο σκοπός του BLAST

Γιατί το BLAST είναι τόσο χρήσιμο για τους βιολόγους? Ο στόχος της έρευνας βάσεων δεδομένων γενικά και του BLAST συγκεκριμένα είναι να αποκαλυφθεί η ύπαρξη ομοιότητας ανάμεσα στην ακολουθία, που ο χρήστης θέλει να βρει

περισσότερα στοιχεία γι' αυτήν (η ακολουθία λέγεται ακολουθία ερώτησης) και άλλες ακολουθίες που είναι αποθηκευμένες σε μια βιολογική βάση δεδομένων (στοχευμένες ακολουθίες). Αυτό είναι συνήθως το πρώτο βήμα που ένας ερευνητής κάνει για να καθορίσει την βιολογική σημασία μιας άγνωστης ακολουθίας. Δεδομένου ότι το μέγεθος των βάσεων δεδομένων των βιολογικών ακολουθιών συντηρείται από το NCBI, η έρευνα βάσεων δεδομένων συνήθως αποκαλύπτει ακολουθίες που έχουν κάποιο βαθμό ομοιότητας με την ακολουθία ερώτησης. Οι ακολουθίες από την βάση δεδομένων που έχουν ομοιότητες με την ακολουθία που ο χρήστης εισάγει αναφέρονται κοινώς σαν hits. Όταν βρεθούν τέτοια hits, οι ερευνητές/χρήστες μπορούν να βγάλουν συμπεράσματα για την θεωρούμενη μοριακή χρήση της ακολουθίας ερώτησης. Ένας βασικός κανόνας για της διεξαγωγή συμπερασμάτων είναι ότι δύο ακολουθίες που μοιράζονται περισσότερο από 50% ομοιότητα είναι συνήθως παρόμοιες σε δομή και λειτουργία. Κάτω από τέτοιες συνθήκες, τα βασικά χαρακτηριστικά δυο ακολουθιών μπορούν εύκολα να ευθυγραμμιστούν και να αναγνωριστούν. Εάν υπάρχει μόνο 25% ομοιότητα, μπορεί να υπάρχει κάποια δομική ομολογία, αν και σε τέτοιες καταστάσεις, η αντιστοιχία ανάμεσα σε δυο πρωτεΐνες δεν είναι εύκολα εμφανής. Επίσης είναι γενικά αποδεκτό ότι ακολουθίες που είναι σημαντικές για τη λειτουργία (και επομένως για την επιβίωση ενός οργανισμού ή ενός είδους) είναι γενικά προστατευμένες.

Στην πραγματικότητα το BLAST είναι ένα σύνολο προγραμμάτων - η συγκεκριμένη επιλογή προγράμματος ή προγραμμάτων εξαρτάται από την ακολουθία εισαγωγής και στον τύπο της βάσης δεδομένων που ερευνάται. Το πιο γνωστά ερευνητικά προγράμματα που χρησιμοποιούνται και οι εφαρμογές τους περιγράφονται στον **πίνακα 5**.

Πίνακας 5. Προγράμματα BLAST

Πρόγραμμα	Σύγκριση	Εφαρμογή
BLASTN	DNA κατά DNA. Συγκρίνει μια ακολουθία αναζήτησης νουκλεοτιδίου με μια βάση δεδομένων ακολουθίας νουκλεοτιδίου.	Βρίσκει ακολουθίες DNA που ταιριάζουν στην αναζήτηση.
BLASTP	Πρωτεΐνη κατά πρωτεΐνης. Συγκρίνει μια ακολουθία αναζήτησης αμινοξυ με μια βάση δεδομένων ακολουθίας πρωτεΐνης.	Βρίσκει ταυτόσημες πρωτεΐνες
BLASTX	DNA κατά πρωτεΐνης. Συγκρίνει μια ακολουθία αναζήτησης νουκλεοτιδίου μεταφρασμένη σε όλα τα πλαίσια ανάγνωσης με μια βάση δεδομένων ακολουθίας πρωτεΐνης.	Βρίσκει για ποια πρωτεΐνη κωδικοποιείται η ακολουθία αναζήτησης.
TBLASTN	Πρωτεΐνη κατά DNA. Συγκρίνει μια ακολουθία αναζήτησης πρωτεΐνης με μια βάση δεδομένων ακολουθίας νουκλεοτιδίου, δυναμικά μεταφρασμένη σε όλα τα πλαίσια ανάγνωσης.	Βρίσκει γονίδια σε άγνωστες ακολουθίες DNA.

TBLASTX	DNA κατά DNA. Συγκρίνει τις μεταφράσεις έξι πλαισίων μιας ακολουθίας αναζήτησης νουκλεοτιδίου με τις μεταφράσεις έξι πλαισίων μιας βάσης δεδομένων ακολουθίας νουκλεοτιδίου.	Ανακαλύπτει τη δομή των γονιδίων. (Βρίσκει το βαθμό ομολογίας ανάμεσα στην κωδικοποιημένη περιοχή της ακολουθίας αναζήτησης και στα γνωστά γονίδια στη βάση δεδομένων.)
---------	--	---

Συνοψίζοντας, οι διαθέσιμες επιλογές BLAST είναι:

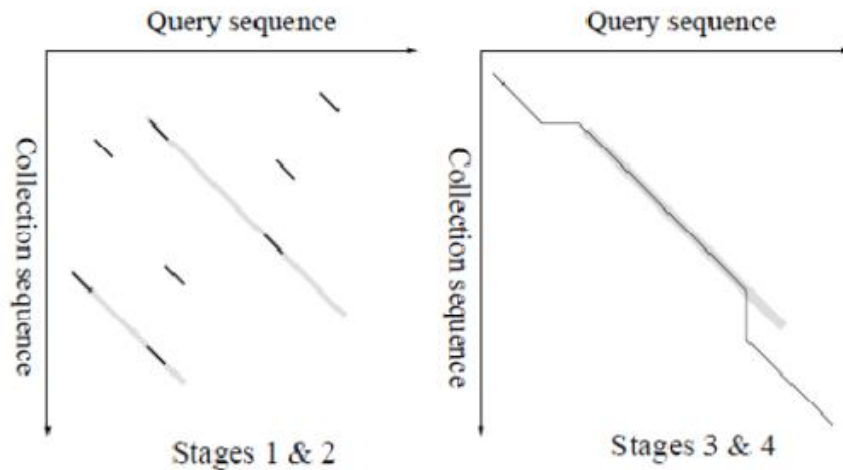
1. Για ακολουθίες nucleotide: BLASTN, BLASTX, TBLASTX
2. Για ακολουθίες amino acid: BLASTP, TBLASTN

Στην απλούστερη περίπτωση, χρειαζόμαστε τις ακόλουθες πληροφορίες για να εκτελέσουμε μια BLAST έρευνα χρησιμοποιώντας την υπηρεσία NCBI:

- 1) Μια ακολουθία εισαγωγής δεδομένων (μπορεί να είναι είτε nucleotide είτε amino acid).
- 2) Μια βάση δεδομένων για έρευνα (μπορεί να είναι είτε nucleotide είτε protein).
- 3) Ένα πρόγραμμα έρευνας βάσης δεδομένων (Μπορεί να είτε ένα οποιοδήποτε από τις 5 BLAST επιλογές). Επιπλέον παράμετροι όπως το δίκτυο κυκλωμάτων και οι e-values πρέπει επίσης να τεθούν. Η εισαγωγή BLAST αποτελείται από μια επικεφαλίδα που δίνει πληροφορίες σε συγκεκριμένες παραμέτρους BLAST, το ζητούμενο ID για την έρευνα, την έκταση της ακολουθίας και τη βάση δεδομένων που χρησιμοποιείται.

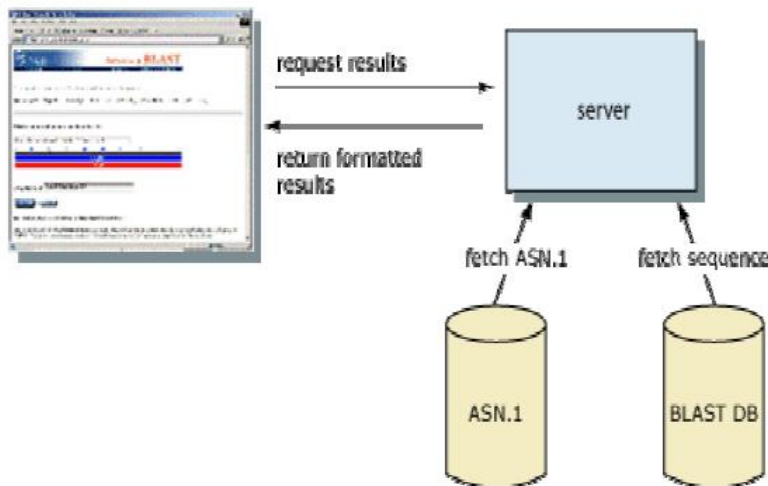
1.5.5.2 Ο αλγόριθμος του προγράμματος BLAST

Ο αλγόριθμος του προγράμματος BLAST είναι ένα λογισμικό πρόγραμμα το οποίο είναι ικανό να λειτουργεί με μερικές συντομεύσεις, για να πραγματοποιεί την αναζήτηση γρηγορότερα. Ο αλγόριθμος εκτελεί «τοπικές» συγκρίσεις ακολουθιών. Αν το BLAST ξεκινούσε προσπαθώντας να ευθυγραμμίσει δύο νουκλεοτιδικές αλληλουχίες σε ολόκληρο το μήκος τους (καθολική ευθυγράμμιση), πολύ λιγότερες ομοιότητες θα είχαν εντοπιστεί μεταξύ των συγκρινόμενων ακολουθιών. Όταν μια εξέταση - έλεγχος υποβάλλεται μέσω μιας ιστοσελίδας του προγράμματος BLAST, η νουκλεοτιδική ακολουθία μαζί με οποιαδήποτε άλλη πληροφορία (όπως η βάση δεδομένων μέσα στην οποία θα γίνει η αναζήτηση, το μέγεθος της ακολουθίας, η αναμενόμενη τιμή) δίνονται στον αλγόριθμο του δικτυακού υπολογιστή. Η αρχή λειτουργίας του BLAST βασίζεται στη δημιουργία ενός πίνακα ελέγχου όλων των «λέξεων» (μικρές ακολουθίες) καθώς και των «γειτονικών λέξεων» (παρόμοιες ακολουθίες στην αλληλουχία αναζήτησης).



Εικόνα 19. Η φιλοσοφία του αλγόριθμου blast.

Στη συνέχεια η νουκλεοτιδική βάση δεδομένων σαρώνεται, έτσι ώστε να μπορέσουν να βρεθούν όλα αυτά τα σημαντικά σημεία (λέξεις, γειτονικές λέξεις). Αφού ο αλγόριθμος έχει ελέγξει όλες τις πιθανές «λέξεις» από την εξεταζόμενη αλληλουχία και τις προεκτείνει μέγιστα, συγκεντρώνει τις καλύτερες ευθυγραμμίσεις νουκλεοτιδικών ακολουθιών (alignments) για κάθε ζεύγος ακολουθιών που συγκρίνει, και στη συνέχεια γράφει αυτή την πληροφορία σε μια «Seq Align» δομή δεδομένων. Η δομή δεδομένων «Seq Align» δεν περιέχει πληροφορία για την ακολουθία, αλλά αναφέρεται στις ακολουθίες οι οποίες βρίσκονται στη βάση δεδομένων αναζήτησης του προγράμματος BLAST (Εικόνα 20)



Εικόνα 20. Η διάρθρωση των αποτελεσμάτων του προγράμματος BLAST.

Τη στιγμή που το πρόγραμμα βρει μια παρόμοια ακολουθία με αυτή του χρήστη κατά την αναζήτηση μέσα στη βάση δεδομένων, είναι καλό ο ερευνητής να έχει μια ιδέα εάν το αποτέλεσμα της ευθυγράμμισης (alignment) αντικατοπτρίζει μια πιθανή βιολογική σχέση, ή εάν η ομοιότητα που βρέθηκε οφείλεται στη τύχη. Το πρόγραμμα BLAST χρησιμοποιεί τη στατιστική θεωρία για να παράγει ένα

αποτέλεσμα (bit score) και την αναμενόμενη τιμή (Expectation value), για κάθε ζεύγος ακολουθιών που συγκρίνεται. Η τιμή E είναι ο αριθμός των διαφορετικών συγκρίσεων ακολουθιών (alignments), για τις οποίες η ομοιότητα οφείλεται στη τύχη. Συνεπώς, όσο χαμηλότερη είναι η τιμή E, τόσο πιο σημαντικό είναι το αποτέλεσμα από την εξελικτική και βιολογική σκοπιά.

Sequences producing significant alignments:

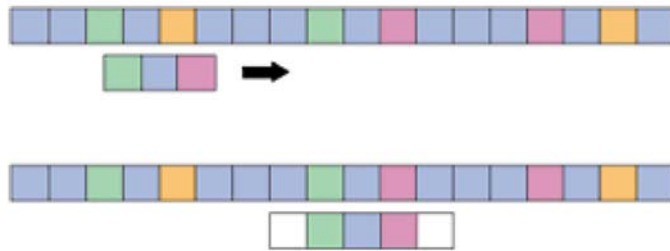
Select: All None Selected: 0

Alignments Download GenBank Graphics Distance tree of results

Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/> Canis familiaris mitochondrion complete genome	30166	30166	100%	0.0	100%	U96639.2
<input type="checkbox"/> Canis lupus familiaris isolate 439H mitochondrion complete genome	30151	30151	100%	0.0	99%	KU290670.1
<input type="checkbox"/> Canis lupus familiaris isolate 422H mitochondrion complete genome	30148	30148	100%	0.0	99%	KU290649.1
<input type="checkbox"/> Canis lupus familiaris isolate 386H mitochondrion complete genome	30148	30148	100%	0.0	99%	KU290591.1
<input type="checkbox"/> Canis lupus familiaris isolate 338H mitochondrion complete genome	30142	30142	100%	0.0	99%	KU290542.1
<input type="checkbox"/> Canis lupus familiaris isolate 364H mitochondrion complete genome	30139	30139	100%	0.0	99%	KU290567.1
<input type="checkbox"/> Canis lupus familiaris isolate 270H mitochondrion complete genome	30133	30133	100%	0.0	99%	KU290496.1

Εικόνα 21. Η τιμή E στη σελίδα των αποτελεσμάτων του προγράμματος BLAST.

Η **τιμή E** (Evalue) λοιπόν είναι μία ένδειξη στατιστικής σημαντικότητας των αποτελεσμάτων ευθυγράμμισης των ακολουθιών (alignments), και αντικατοπτρίζει ως ένα βαθμό το μέγεθος της βάσης δεδομένων αναζήτησης καθώς και το σύστημα αποτελεσμάτων (scoring system) που έχει χρησιμοποιηθεί. Όσο πιο χαμηλή είναι η τιμή E, τόσο πιο σημαντική είναι η ομοιότητα μεταξύ των ακολουθιών. Ένα 12 αποτέλεσμα ευθυγράμμισης (sequence alignment) που έχει μια τιμή E 0.05, δείχνει ότι αυτή η ομοιότητα μεταξύ των ακολουθιών έχει μία πιθανότητα 5% να οφείλεται στην τύχη. Εάν και ένας στατιστικολόγος μπορεί να θεωρήσει αυτό το αποτέλεσμα σημαντικό, εντούτοις μπορεί να μην αντικατοπτρίζει ένα σημαντικό αποτέλεσμα από βιολογική, εξελικτική και φυλογενετική σκοπιά. Το αποτέλεσμα (bit score) δίνει μία ένδειξη για το πόσο καλή είναι η σύγκριση - ευθυγράμμιση των ακολουθιών κατά ζεύγη (alignment). Όσο πιο υψηλή είναι η τιμή, τόσο καλύτερο είναι το αποτέλεσμα της ευθυγράμμισης (Εικόνα 22). Σε γενικές γραμμές, αυτή η τιμή υπολογίζεται από έναν τύπο που λαμβάνει υπ' όψιν την ευθυγράμμιση παρόμοιων ή ταυτόσημων νουκλεοτιδικών κομματιών (μεταξύ συγκρινόμενων ακολουθιών), καθώς επίσης και τα κενά που εισάγονται για να ευθυγραμμιστούν οι ακολουθίες. Ένα στοιχείο κλειδί σε αυτό τον υπολογισμό είναι η «μήτρα αντικατάστασης» (substitution matrix), η οποία αναθέτει μία τιμή για την ευθυγράμμιση κάθε πιθανού ζεύγους συγκρινόμενων νουκλεοτιδικών ακολουθιών.



Εικόνα 22. Πάνω: το πρόγραμμα προσπαθεί να στοιχίσει τρεις μονάδες της υπό μελέτη αλληλουχίας με αλληλουχίες από τη βάση δεδομένων.

Κάτω: Εφόσον έχει γίνει η πρώτη στοίχιση, συμπληρώνεται η υπό μελέτη αλληλουχία για να ελεγχθεί αν η στοίχιση και των υπόλοιπων μονάδων είναι ακριβής.

Έτσι λοιπόν, όσον αφορά τις διάφορες βάσεις δεδομένων (databases) τα συγκεκριμένα βήματα που ακολουθούνται είναι:

- 1) Ανάκτηση δεδομένων από βάσεις δεδομένων
- 2) Αξιολόγηση των δεδομένων (datacleaning)
- 3) Οργάνωση/Ομαδοποίηση δεδομένων
- 4) Εξαγωγή συμπερασμάτων

1.6 Σκοπός της παρούσας εργασίας

Σκοπός της παρούσας μελέτης είναι η αξιολόγηση της γενετικής ποικιλότητας εντός-φυλής, μέσω ανάπτυξης μοριακών δεικτών σε ένα μεγάλο σύνολο φυλών σκύλων. Πιο συγκεκριμένα, μέσω της χρήσης του μιτοχονδριακού DNA, που δρα ως μοριακός δείκτης, πραγματοποιείται σύγκριση όσον αφορά τις θέσεις στα διαφορετικά γονίδια του. Συγκεκριμένα αναζητούνται SNPs (single nucleotide polymorphisms) με απώτερο στόχο την εύρεση μοναδικών βάσεων που διαφοροποιούν τις διάφορες ράτσες μεταξύ τους.

Κεφάλαιο 2^ο

Υλικά και Μέθοδοι

Στην παρούσα εργασία, αρχικά αναζητήθηκε η αλληλουχία του μιτοχονδριακού DNA του σκύλου στη διαδικτυακή βάση δεδομένων NCBI(<http://www.ncbi.nlm.nih.gov/genome/>), χρησιμοποιώντας τις λέξεις κλειδιά : "*Canislupusfamiliaris* mitochondrion, completegenome". Έτσι λοιπόν επιλέχθηκε το πρώτο αποτέλεσμα με AccessionnumberU96639.2 που αντιστοιχεί στο μιτοχονδριακό DNA ενός θηλυκού μπόξερ, το οποίο ήταν και το πρώτο του είδους *Canislupusfamiliaris* του οποίου έγινε αλληλούχηση ολόκληρου του γονιδιώματος.



Εικόνα 23.Boxer

Canis familiaris mitochondrion, complete genome

GenBank: U96639.2

[FASTA](#) [Graphics](#)

Go to:

LOCUS	CFU96639	16727 bp	DNA	circular	MAM 11-APR-2000
DEFINITION	Canis familiaris mitochondrion, complete genome.				
ACCESSION	U96639				
VERSION	U96639.2 GI:7534303				
KEYWORDS	.				
SOURCE	mitochondrion Canis lupus familiaris (dog)				
ORGANISM	Canis lupus familiaris Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Carnivora; Caniformia; Canidae; Canis.				
REFERENCE	1 (bases 1 to 16727)				
AUTHORS	Kim,K.S., Lee,S.E., Jeong,H.W. and Ha,J.H.				
TITLE	The complete nucleotide sequence of the domestic dog (<i>Canis familiaris</i>) mitochondrial genome				
JOURNAL	Mol. Phylogenet. Evol. 10 (2), 210-220 (1998)				
PUBMED	9878232				
REFERENCE	2 (bases 1 to 16727)				
AUTHORS	Kim,K.S., Lee,S.E., Jeong,H.W., Jeong,S.Y., Sohn,H.S. and Ha,J.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (07-APR-1997) Genetic Engineering, Animal Genetics, 1370 Sankyuk-dong, Pukgu, Taegu 702-701, Korea				
REFERENCE	3 (bases 1 to 16727)				
AUTHORS	Kim,K.S., Lee,S.E., Jeong,H.W., Jeong,S.Y., Sohn,H.S. and Ha,J.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (11-APR-2000) Genetic Engineering, Animal Genetics, 1370 Sankyuk-dong, Pukgu, Taegu 702-701, Korea				
REMARK	Sequence update by submitter				
COMMENT	On Apr 11, 2000 this sequence version replaced gi:4154170 .				
FEATURES	Location/Qualifiers				
source	1..16727 /organism="Canis lupus familiaris" /organelle="mitochondrion" /mol_type="genomic DNA" /sub_species="familiaris" /db_xref="taxon:9615"				

Εικόνα 24. Αποτελέσματα από αναζήτηση του "*Canislupusfamiliaris*mitochondrion, completegenome" στην Βάση δεδομένων NCBINucleotide.

Το μιτοχονδριακό του DNA όπως βλέπουμε και στην **εικόνα 25** είναι κυκλικό και αποτελείται από 16,727 bp.

```

/sub_species="familiaris"
/db_xref="taxon:9615"
tRNA
1..69
/product="tRNA-Phe"
rRNA
70..1023
/product="12S ribosomal RNA"
tRNA
1024..1090
/product="tRNA-Val"
rRNA
1091..2670
/product="16S ribosomal RNA"
tRNA
2671..2744
/product="tRNA-Leu"
/note="codons recognized: UUR"
gene
2747..3703
/gene="ND1"
CDS
2747..3703
/gene="ND1"
/codon_start=1
/transl_table=2
/product="NADH dehydrogenase subunit 1"
/protein_id="AAD04763.1"
/db_xref="GI:4154171"
/translation="MFFINIISLIIPILLAVAFLLVERKVLGYMQLRKGNIVGPGY
LLQPIADAVKLFTEPLRPLTSSMSMFIAPILALSLALTMWIPLMPYPLINMNLGV
LFMLAMSSLAVYSILWSGWASNSKYALIGALRAVAQTISYEVTLAIIILSVLLMNGSF
TLSTLIITQEHMWLIFPAWPLAMMWFISTLAETNRAPFDLTEGESELVSGFNVEYAAG
PFALFFLAEYANIIMINIILTTILFFGAFHNPFMPPELYSINFMTKTLTLLTICFLWIRAS
YPRFRYDQLMHLLWKNFLPLTLALCMHWVALPIITASIPPQT"
tRNA
3703..3771
/product="tRNA-Ile"
tRNA
complement(3768..3842)
/product="tRNA-Gln"
tRNA
3844..3913
/product="tRNA-Met"
gene
3914..4957
/gene="ND2"

```

Εικόνα 25. Διακρίνουμε τα αρχικά από τα 37 γονίδια του mtDNA του *canislupusfamiliaris*.

Στην συνέχεια, χρησιμοποιώντας την επιλογή FASTA ανακτήθηκε η αλληλουχία στη μορφή fasta, η οποία είναι και η μορφή η οποία αναγνωρίζεται από τα περισσότερα προγράμματα Βιοπληροφορικής που χρησιμοποιούνται στη συνέχεια.

Στη συνέχεια, πραγματοποιήθηκε BLASTN χρησιμοποιώντας ως αλληλουχία επερώτησης (query) την παραπάνω (θηλυκό boxer-U96639.2), χρησιμοποιώντας τα ακόλουθα κριτήρια του αλγορίθμου :

The screenshot shows the BLASTN web interface. The 'Enter Query Sequence' section contains a text area with a DNA sequence: TACACTATTATTAGTCCCGCCAAACCCAAAACAGGACTAAGTGCATACAATACTCACAAGCTTTATTAAATATATACAAATGATTTGCTACTCTAGTTAACTTAACACACAGTCTTACACGCATTTGGTCTCGTAGTCTATCTATAGATAGCATTCCCTTTTTCCTCTCATATTACTATGTATTTATTTATACGCACACACAAATTCAGTATAA. Below this is a 'Job Title' field with the text 'boxer'. The 'Choose Search Set' section has 'Database' set to 'Nucleotide collection (nr/nt)' and 'Organism' set to 'Canis lupus familiaris (taxid:9615)'. Under 'Program Selection', 'Highly similar sequences (megablast)' is selected.

Εικόνα 26. Αρχική σελίδα του BLASTN

Πιο συγκεκριμένα λοιπόν (εικόνα 26), σαν οργανισμός (Organism) επιλέχθηκε ο *Canislupusfamiliaris* (taxid: 9615).

Όσον αφορά το "programselection" και συγκεκριμένα το "optimizefor", επιλέχθηκε ο highlysimilarsequences (megablast) καθώς είναι καλύτερη για ορισμένες εργασίες, όπως η ταυτοποίηση του ερωτήματος εισόδου και η αναζήτηση με μεγάλη γενωμική ερώτηση.


The screenshot shows the 'Algorithm parameters' section of the BLASTN interface. The 'General Parameters' section includes: 'Max target sequences' (1000), 'Short queries' (Automatically adjust parameters for short input sequences checked), 'Expect threshold' (1e-10), 'Word size' (28), and 'Max matches in a query range' (0). The 'Scoring Parameters' section includes: 'Match/Mismatch Scores' (1,-2) and 'Gap Costs' (Linear). The 'Filters and Masking' section includes: 'Filter' (Low complexity regions checked, Species-specific repeats for: Homo sapiens (Human)), and 'Mask' (Mask for lookup table only checked, Mask lower case letters unchecked). At the bottom, there is a 'BLAST' button and a checkbox for 'Show results in a new window'.

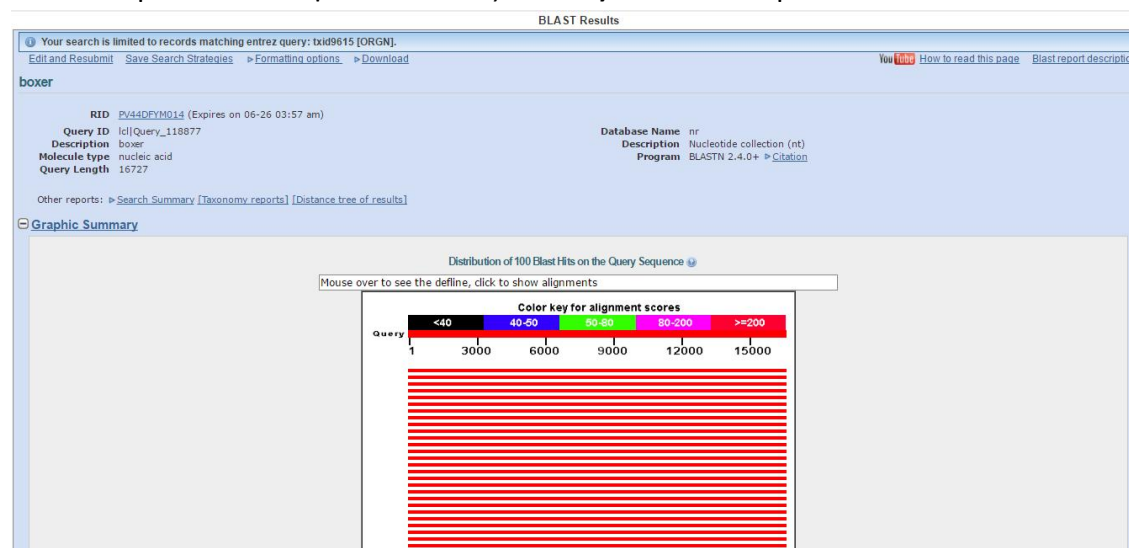
Εικόνα 27. BLASTN με επιλεγμένο το algorithmparameters

Στην συνέχεια επιλέχθηκαν περισσότερες παράμετροι μέσω του "Link" "Algorithm parameters" όπως General parameters, scoring parameters, filters και Masking.

Όσον αφορά τις "General parameters", επιλέξαμε το 1000 στο "Max target sequence" ενώ στο "expect threshold" γράψαμε $1e-10$ (10^{-10}) στην περιοχή αναζήτησης. Τα "scoring parameters" είχαν τις προεπιλεγμένες επιλογές (εικόνα 27). Πιο συγκεκριμένα, όσον αφορά το "Match/Mismatch Scores" επιλέχθηκε το 1,-2 ενώ όσον αφορά το "Gap Costs" επιλέχθηκε το linear.

ενώ στο "filter" έπρεπε να ήταν επιλεγμένο το lowcomplexityregions.

Στην συνέχεια για να γίνει η σύγκριση πατήθηκε το κουμπί BLAST.  Και τα αποτελέσματα του διαφαινονται στις εικόνες 28 και 29 παρακάτω.



Εικόνα 28. Τα αποτελέσματα είναι της παραπάνω μορφής.

Description	Max score	Total score	Query cover	E value	Ident	Accession
<input checked="" type="checkbox"/> Canis familiaris mitochondrion, complete genome	30890	30890	100%	0.0	100%	U96639.2
<input checked="" type="checkbox"/> Canis lupus familiaris isolate 439H mitochondrion, complete genome	30873	30873	100%	0.0	99%	KU290670.1
<input checked="" type="checkbox"/> Canis lupus familiaris isolate 422H mitochondrion, complete genome	30867	30867	100%	0.0	99%	KU290649.1
<input checked="" type="checkbox"/> Canis lupus familiaris isolate 386H mitochondrion, complete genome	30867	30867	100%	0.0	99%	KU290591.1
<input checked="" type="checkbox"/> Canis lupus familiaris isolate 338H mitochondrion, complete genome	30862	30862	100%	0.0	99%	KU290542.1
<input checked="" type="checkbox"/> Canis lupus familiaris isolate 364H mitochondrion, complete genome	30856	30856	100%	0.0	99%	KU290567.1
<input checked="" type="checkbox"/> Canis lupus familiaris isolate 270H mitochondrion, complete genome	30851	30851	100%	0.0	99%	KU290496.1
<input checked="" type="checkbox"/> Canis lupus familiaris isolate 858H mitochondrion, complete genome	30845	30845	100%	0.0	99%	KU291049.1
<input checked="" type="checkbox"/> Canis lupus familiaris isolate 857H mitochondrion, complete genome	30845	30845	100%	0.0	99%	KU291047.1
<input checked="" type="checkbox"/> Canis lupus familiaris isolate 795H mitochondrion, complete genome	30845	30845	100%	0.0	99%	KU290965.1
<input checked="" type="checkbox"/> Canis lupus familiaris isolate 602H mitochondrion, complete genome	30845	30845	100%	0.0	99%	KU290830.1
<input checked="" type="checkbox"/> Canis lupus familiaris isolate 134H mitochondrion, complete genome	30845	30845	100%	0.0	99%	KU290429.1
<input checked="" type="checkbox"/> Canis lupus familiaris isolate 120H1 mitochondrion, complete genome	30845	30845	100%	0.0	99%	KU290417.1

Εικόνα 29. Διαφορετικές καταχωρήσεις που βρέθηκαν στις οποίες διαφάνεται το E value, το querycover και το ident. με βάση τα οποία θα απορριφθούν ή όχι οι καταχωρήσεις.

Ακολούθησε "selectall" αλλά στην πράξη δεν επιλέξαμε όλες διαφορετικές καταχωρήσεις απείχαν πολύ από το 100% του query cover, το 100% του ident και το 0.0 του E value.

Όσον αφορά τα παραπάνω:

Query cover: Τι ποσοστό της ακολουθίας επερώτησης (query) έχει ευθυγραμμιστεί, είναι το ποσοστό της αλληλουχίας επερώτησης (query) που αλληλεπικαλύπτεται με την αλληλουχία subject). Άρα απαιτείται υψηλό ποσοστό (περίπου 100%).

E value: Αριθμός των ακολουθιών (hits) που περιμένει κανείς να δει καθαρά από τύχη με το ίδιο σκορ ή υψηλότερο. Για χαμηλές τιμές E έχουμε μεγαλύτερη ομοιότητα μεταξύ των ακολουθιών. Εάν έχουμε υψηλές τιμές E τότε έχουμε τυχαία ευθυγράμμιση των ακολουθιών. Για $evalue=0$ τυπικά δεν υπάρχει (σχεδόν) καμία πιθανότητα η ευθυγράμμιση να προκύψει καθαρά από τύχη. Το $E < 0.05$ είναι περισσότερο αποδεκτό.

Maximum Identity: Είναι το ποσοστό ομοιότητα μεταξύ της αλληλουχίας επερώτησης (query) και αλληλουχίας subject σε όλο το μήκος της περιοχής αλληλεπικάλυψης σε μια στοίχιση νουκλεοτιδίου-προς-νουκλεοτίδιο. Άρα θέλουμε υψηλά ποσοστά (σχεδόν 100%).

Έτσι λοιπόν δεν επιλέχθηκαν και οι 1000 ακολουθίες που ανακτήθηκαν από τις databases μέσω του BLAST αλλά αυτές που τηρούν τα παραπάνω κριτήρια.

Στη συνέχεια, ανακτήθηκαν οι αλληλουχίες του μιτοχονδριακού DNA (mtDNA) των διάφορων ρατσών σε μορφή fasta στο σημειωματάριο μέσω επιλογή Download και συγκεκριμένα του FASTA (complete sequence). Έτσι λοιπόν με αυτόν τον τρόπο είχαμε όλες τις ολικές αλληλουχίες μιτοχονδριακού DNA των ρατσών που ήταν καταχωρημένες και ήταν υψηλά παρόμοιες με την αλληλουχία επερώτησης του θηλυκού Boxer.

Στην συνέχεια αποθηκεύτηκε το κάθε γονίδιο, από τα 37 που περιέχει το mtDNA, σε διαφορετικό σημειωματάριο. Ακολούθησε ένα δεύτερο BLAST που αυτή την φορά είχε σαν ακολουθία επερώτησης (query sequence) το κάθε γονίδιο που αποθηκεύσαμε παραπάνω και σαν υποκειμενική ακολουθία στην βάση δεδομένων (subject sequence) το συνολικό αρχείο με τις ολικές αλληλουχίες του μιτοχονδριακού DNA των ρατσών που ανακτήθηκε μέσω του 1^{ου} BLAST. Τα διάφορα κριτήρια ήταν ίδια με του 1^{ου} BLAST με μια μικρή διαφορά όσον αφορά το "programselection" και συγκεκριμένα το "optimize for". Στο 2^ο BLAST επιλεγόταν σε αυτό το κομμάτι είτε το "somewhat similar sequences (blastn)" αν η αλληλουχία επερώτησης ήταν το tRNA λόγω του μικρού του μεγέθους είτε το "highly similar sequences (megablast)" σε όλες τις άλλες περιπτώσεις (rRNA, gene).

Έτσι με τον ίδιο τρόπο κάναμε download τις αλληλουχίες του μιτοχονδριακού DNA σε διαφορετικό αρχείο με βάση το γονίδιο επερώτησης.

Οι διαφορετικές αλληλοθυχίες που βρέθηκαν ήταν 550 και από αυτές μόνο οι 381 είχαν καταχωρημένη την ράτσα τους. Όπως είπαμε και παραπάνω, στην παρούσα μελέτη χρησιμοποιήθηκαν μόνο σκυλιά με γνωστές ράτσες.

Για την οπτικοποίηση των αποτελεσμάτων χρησιμοποιήσαμε το πρόγραμμα Bioedit. Ακολούθως, χρησιμοποιήσαμε το πρόγραμμα BlastClust (<https://toolkit.tuebingen.mpg.de/blastclust>) που έχει την ιδιότητα να ομαδοποιεί

τις όμοιες αλληλουχίες μεταξύ των διαφορετικών γονιδίων και έγινε αντιγραφή των αποτελεσμάτων σε διαφορετικά φύλλα του excel.

Στην συνέχεια καταγράφηκαν οι θέσεις που εμφάνιζαν ποικιλομορφία νουκλεοτιδικών βάσεων καθώς και οι διάφορες βάσεις που αντιστοιχούν σε κάθε μια από αυτές τις θέσεις. Για παράδειγμα έστω ότι για το γονίδιο COX1 όλα τα σκυλιά είχαν στις θέσεις 1,2,3 μόνο A, μόνο C και μόνο G αντιστοίχα κάτι που μας οδηγεί σε έλειψη πολυμορφισμού και άρα μη επιλογή αυτών των θέσεων. Στην συνέχεια όμως στην θέση 4 κάποια σκυλιά περιείχαν A και άλλα T μας οδήγησε να την επιλέξουμε σαν πολυμορφική θέση. Έτσι λοιπόν καταγράφηκαν σε excel οι διάφορες θέσεις και έγινε ένα προφίλ πολυμορφικών θέσεων.

Στην συνέχεια έγιναν 2 μορφές ανάλυσης:

Σε πρώτη φάση, αναζητήθηκαν όλα τα μοναδικά SNPs στις διαφορετικές θέσεις των γονιδίων και καταγράφηκαν σε διαφορετικό excel τα SNPs για κάθε γονίδιο. Έτσι συγκρίνοντας τις βάσεις κάθε θέσης μεταξύ τους, σε κάθε γονίδιο, βρέθηκαν τα άτομα που είχαν βάσεις που ήταν μοναδικές στην εκάστοτε θέση (SNPs) και έτσι διαφοροποιήθηκαν από τα υπόλοιπα. Αυτή η μελέτη περιελάμβανε όλα τα άτομα και όλες τις ράτσες. Όμως καταγράφηκαν μόνο τα άτομα που είχαν μοναδική πολυμορφική βάση σε μια θέση αλλά δεν ήταν μέλη μιας ευρύτερης ράτσας. Αν περιλαμβάνουμε και αυτά τα άτομα, δεν θα μιλούσαμε για πολυμορφισμό μεταξύ ρατσών καθώς κάποια άλλα άτομα της ίδιας ράτσας θα είχαν διαφορετική βάση στην ίδια θέση.

Σε δεύτερη φάση, έγινε μελέτη σε ράτσες που είχαν παραπάνω από 1 άτομα.

Έτσι λοιπόν, καταγράφηκαν σε excel τα άτομα της κάθε ράτσας (πχ bordercollie 1-5) καθώς και το προφίλ μόνο των διαφορετικών βάσεων στις συγκεκριμένες πολυμορφικές θέσεις που πραγματοποιήθηκε ακριβώς πριν από τις 2 μορφές ανάλυσης. Στο τέλος λοιπόν καταγράφηκαν οι θέσεις που ήταν όμοιες για τα άτομα της ίδιας ράτσας ενώ τοποθετήθηκε "N" στα κελιά που για κάθε άτομο της ράτσας δεν είχαν την ίδια βάση στην ίδια θέση. Έτσι λοιπόν στο τέλος μέσω συγκρίσεων πάρθηκε ένα προφίλ βάσεων που ήταν κοινό για όλα τα άτομα της κάθε ράτσας. Για παράδειγμα, η ράτσα bordercollie (η οποία περιλαμβάνει 5 άτομα) όσον αφορά το γονίδιο ATP8 είχε προφίλ : N AAA N A που αντιστοιχεί στις θέσεις 15, 35, 112, 114, 121, 146 (δηλαδή όλα τα σκυλιά είχαν στις θέσεις 15 και 121 παραπάνω από μια επιλογές νουκλεοτιδίων ενώ είχαν όλα A στις θέσεις 35, 112, 114 και 146). Επίσης να αναφερθεί ότι τα σκυλιά του παραδείγματος είχαν ίδια όλα τα υπόλοιπα νουκλεοτίδια για τις υπόλοιπες θέσεις στο γονίδιο ATP8. Το ίδιο έγινε για όλες τις υπόλοιπες ράτσες για όλα τα γονίδια.

Στο τέλος κατασκευάστηκαν πίνακες (**πίνακας 8 και πίνακας 9**) που περιείχαν τα αποτελέσματα των δύο παραπάνω διαφορετικών αναλύσεων.

Κεφάλαιο 3^ο

Αποτελέσματα

Στο κεφάλαιο αυτό καταγράφονται τα αποτελέσματα σε μορφή πινάκων. Στον πρώτο πίνακα (**Πίνακας 6**) καταγράφονται οι διάφορες ράτσες σκύλων που χρησιμοποιήθηκαν στην παρούσα έρευνα. Αυτές όπως αναφέρθηκε και παραπάνω ανακτήθηκαν με την βοήθεια του προγράμματος BLAST. Για το λόγω αυτό χρησιμοποιήθηκε ως αλληλουχία επερώτησης (query) η αλληλουχία του σκύλου που για τον οποίο πραγματοποιήθηκε το πρώτο whole genome sequencing του είδους, ενός θηλυκού boxer (U96639.2). Στον πίνακα αυτό λοιπόν, διακρίνουν τις ράτσες καθώς και τον αριθμό των διαφόρων ατόμων της κάθε ράτσας. Στην συνέχεια με βέλη διακρίνουμε εικόνες με κάποιες ράτσες σκύλων (κυρίως αυτές με τα περισσότερα άτομα) μέσω των οποίων διαφαίνεται η τεράστια ποικιλομορφία μεταξύ των διαφόρων ειδών. Τα σκυλιά εμφανίζουν μια τεράστια ποικιλία διαφορετικών εξωτερικών χαρακτηριστικών (μεγεθών, χρωμάτων, τριχωμάτων) παρόλο που ανήκουν όλα στο ίδιο είδος (*Canis lupus familiaris*).

Ακολουθεί ο **πίνακας 7** στον οποίο καταγράφονται οι διαφορετικές ράτσες που χρησιμοποιήθηκαν στην παρούσα μελέτη και δίπλα τους παρατίθενται τα accession numbers των ατόμων αυτών που ανακτήθηκαν από την βάση δεδομένων GenBank.

Η ποικιλομορφία των ατόμων των διαφορετικών ρατσών που όμως ανήκουν στο ίδιο είδος διαφαίνεται στην αλληλουχία των βάσεων του DNA τους και συγκεκριμένα οι γενετικοί δείκτες (μιτοχονδριακό DNA, mtDNA) είναι αυτοί που αντανακλούν τις διαφορές σε επίπεδο DNA.

Τέλος, στους **πίνακες 8,9** παρατηρούμε τα αποτελέσματα των δυο διαφορετικών αναλυσεων που εξηγήθηκαν στο προηγούμενο κεφάλαιο. Στην ουσία, αναπαρίστανται οι διάφοροι SNPs πολυμορφισμοί μεταξύ των ρατσών οι οποίοι προσδίδουν μοναδικότητα σε ένα γονίδιο, για μια θέση, για μια συγκεκριμένη ράτσα.



πίνακα 6. Καταγράφονται όλες οι ράτσες που βρέθηκαν μέσω της μηχανή αναζήτησης BLAST (Basic Local Alignment Search Tool) και ανήκουν στο είδος *Canis lupus familiaris* καθώς και ο αριθμός των ατόμων της κάθε φυλής που χρησιμοποιήθηκε στην παρούσα εργασία.

Διαφορετικές ράτσες	Αριθμόςδειγμάτων
Afghan	2
Airedale terrier	2
Akita	2
Alaskan Malamute	1
American cocker spaniel	2
American Eskimo Dog	1
American Staffordshire Terrier	3
Australian shepherd	3
australian terrier	1
Bali	1
Basenji	4
Basset hound	6
Beagle	2
Bearded Collie	1
Bedlington Terrier	1
Belgian Shepherd Malinois	5
Belgian Shepherd Tervueren	1
Belgian Tervuren	1
Belgium Malinois	1
Bernese Mountain Dog	3
Bichon frise	3
black russian terrier	2
Bloodhound	1
blue heeler	1
bolognese	1
Border Collie	6
Borzoi	1
Boston Terrier	1
Bouvier	2
Bouvier Des Ardennes	1
Bouvier Des Flandres	1
boxer	3
Braque	1
brittany spaniel	2



Brussels Griffon	1	
Bull Terrier	2	→
Bullmastiff	1	
cairn terrier	1	
cardigan corgi	1	
Carin Terrier	1	
Catahoula Leopard Dog	1	
cavalier king charles spaniel	4	→
Chihuahua	9	→
chinese crested	1	
Chinese indigenous dog	1	
Chinese Shar Pei	1	
Chinook	1	→
Chow chow	2	
coated retriever	1	
cockapoo	2	→
cocker spaniel	8	→
Collie	1	
Collie, smooth	1	
corgi	1	
Coton de Tulear	1	
dachshund	4	→
dachshund miniature	3	
dalmatian	2	→
Dingo	21	
doberman	2	→
Doberman Pinscher	1	
Dogue de Bordeaux	1	
English Cocker Spaniel	1	
english mastiff	1	
English Setter	2	→
English Shepard	2	
English Springer Spaniel	3	→
epagneul breton	1	
Epagneul Picard/Picardy Spaniel	1	
Fila Brasileiro	1	
Finnish Hound	1	
Finnish Lapphund	1	→
Flatcoated retriever	1	
fox terrier	3	
French Bull Dog	4	→
german pointer	1	
German Shephard	6	→
German Shorthaired Pointer	1	
golden retriever	10	→

gordon setter	1	
great dane	4	→
great pyrenees	2	
Greyhound (galgo)	4	→
Hamilton-stövare	1	
Harrier	1	
havanese	2	→
Husky siberian	4	→
Iceland Dog	1	
irish setter	4	→
irish soft coated wheaten terrier	1	
Irish Wolfhound	1	
Italian Greyhound	1	
jack russel terrier	6	→
Jagdterrier	2	
Jämthund (Swedish elkhound)	7	→
Japanese Chin	1	
Japanese Spitz	1	
Jindo	1	
Kalimantan	1	
Kangal	1	
keeshond	2	→
Kelpie	1	
kerry blue terrier	1	
Kisyu	1	
Kyi Leo	1	
labradoodle	1	
Labrador Retriever	7	→
Laika, East/ West Siberian	3	→
leonberger	2	→
Lhasa Apso	1	
Maltese	2	→
Manchester Terrier	1	
Maremma Sheepdog	1	
Mastiff	1	
Miniature Pinscher	1	
Miniature Schnauzer	1	
Mountain Cur	1	
neapolitan mastiff	3	→
new foundland	2	→
New Guinea Singing Dog	1	
Norfolk Terrier	1	
norwegian elkhound	2	→
Nova Scotia Duck Trolling Retriever	1	
Offspring Shar-pei/ Gunagdong indigenous dog	1	



Old English Sheepdog	3	→	
Old English Shepherd	1		
Owczarek Podkalanski	1		
Papillion	1		
Pekingese	2	→	
pincher	1		
pitbull terrier	1		
Plott Hound	1		
pomeranian	1	↗	
Poodle	4		
Portuguese Water Dog	1		
pug	1	↗	
Pungsan	4	→	
Pyrenean Mastiff	2	→	
Pyrenean Mountain dog	2	→	
Rat Terrier	1		
Rottweiler	5	→	
saint bernard	4	→	
Saluki	3	↘	
sapsaree	2	↘	
schipperke	1	↘	
Scottish Deerhound	1		
Scottish Terrier	1		
Shar pei	2		
shepherd	3		
Shetland sheepdog	5		
Shiba Inu	1		
shih tzu	3	↘	
Shikoku	1		
Silky Terrier	1		
Simao chinese	1		
Sloughi	2	→	
Small Munsterlander	1		
spanish cocker spaniel	1		
Spinone Italiano	1	↗	
spitz	2		
standard schnauzer	3	→	
Tazi	1		
Irish terrier/(soft coated wheaten)	2	↘	
tibetan mastiff	5	→	
tibetan spaniel	1		
tibetan terrier	1		
toy poodle	1		
Treeing Walker Coonhound	1		
Viszla	2	→	

Wairau Bar dog	14	
walker hound	1	
Weimaraner	2	
welsh springer springer	1	
Welsh Terrier	1	
West highlandterrier	4	
Whippet	1	
White Shepherd	1	
yorkie chihuahua	1	
Yorkshire terrier	4	

Πίνακας 7. Στον παρακάτω πίνακα καταγράφονται τα διάφορα accessionnumbers των αλληλουχιών που χρησιμοποιήθηκαν στην παρούσα μελέτη τα οποία έχουν αναλυτικές πληροφορίες στο NCBI.

κωδικοί	Breed
gi 189493898 gb EU789758.1 gi 333945480 gb JF342859.1	Afghan
gi 50301959 gb AY656748.1 gi 333945005 gb JF342825.1	Airedale terrier
gi 166065515 gb EU408245.1 gi 333945187 gb JF342838.1	Akita
gi 333944823 gb JF342812.1	Alaskan Malamute
gi 166065529 gb EU408246.1 gi 647736901 gb KJ637109.1	American cocker spaniel
gi 333944809 gb JF342811.1	American Eskimo Dog
gi 647736117 gb KJ637053.1 gi 647736943 gb KJ637112.1 gi 647736579 gb KJ637086.1	American Staffordshire Terrier
gi 166065571 gb EU408249.1 gi 166065557 gb EU408248.1 gi 333946022 gb JF342898.1	Australian shepherd
gi 166065543 gb EU408247.1	Australian terrier
gi 393007562 gb JX088690.1	Bali
gi 50301805 gb AY656737.1 gi 189493912 gb EU789759.1 gi 333944879 gb JF342816.1 gi 558478234 gb KF661036.1	Basenji
gi 166065669 gb EU408256.1 gi 166065655 gb EU408255.1	basset hound

gi 166065641 gb EU408254.1 gi 333945620 gb JF342869.1 gi 647736985 gb KJ637115.1 gi 647736971 gb KJ637114.1	
gi 647737069 gb KJ637121.1 gi 51872055 gb AY729880.1	Beagle
gi 333945941 gb JF342892.1	Bearded Collie
gi 333945899 gb JF342889.1	Bedlington Terrier
gi 647736789 gb KJ637101.1 gi 647736677 gb KJ637093.1 gi 647735809 gb KJ637031.1 gi 647737251 gb KJ637134.1 gi 647737265 gb KJ637135.1	Belgian Shepherd Malinois
gi 333944767 gb JF342808.1	Belgian Tervuren
gi 583829614 gb KF926378.1	Belgium Malinois
gi 647736103 gb KJ637052.1 gi 333945746 gb JF342878.1	Bernese Mountain Dog
gi 166065585 gb EU408250.1 gi 333944795 gb JF342810.1 gi 333945550 gb JF342864.1	Bichon frise
gi 91702451 gb DQ480493.1 gi 723429741 gb KM061561.1	black russian terrier
gi 333945159 gb JF342836.1	Bloodhound
gi 166065599 gb EU408251.1	blue heeler
gi 166065613 gb EU408252.1	bolognese
gi 647736159 gb KJ637056.1 gi 189492848 gb EU789683.1 gi 333946036 gb JF342899.1 gi 333944837 gb JF342813.1 gi 625293726 gb KJ472767.1	Border collie
gi 333945578 gb JF342866.1	Borzoi
gi 333944907 gb JF342818.1	Boston Terrier
gi 189493996 gb EU789765.1 gi 189493926 gb EU789760.1	Bouvier
gi 333945061 gb JF342829.1	Bouvier Des Ardennes
gi 333945201 gb JF342839.1	Bouvier Des Flandres
gi 647736271 gb KJ637064.1 gi 166065627 gb EU408253.1 gi 723429517 gb KM061545.1	boxer
gi 189494066 gb EU789770.1	Braque
gi 166065683 gb EU408257.1 gi 333946078 gb JF342902.1	brittany spaniel
gi 333945718 gb JF342876.1	Brussels Griffon

gi 647736565 gb KJ637085.1 gi 333946064 gb JF342901.1	Bull Terrier
gi 333945103 gb JF342832.1	Bullmastiff
gi 166065781 gb EU408264.1	cairn terrier
gi 166065725 gb EU408260.1	cardigan corgi
gi 333946134 gb JF342906.1	Carin Terrier
gi 333944865 gb JF342815.1	Catahoula Leopard Dog
gi 647735949 gb KJ637041.1 gi 647735963 gb KJ637042.1 gi 166065767 gb EU408263.1	cavalier king charles spaniel
gi 647736327 gb KJ637068.1 gi 647736285 gb KJ637065.1 gi 647736019 gb KJ637046.1 gi 166065739 gb EU408261.1 gi 647737139 gb KJ637126.1 gi 647736859 gb KJ637106.1 gi 166065753 gb EU408262.1	Chihuahua
gi 50302043 gb AY656754.1	chinese crested
gi 577029811 gb KF907309.1	Chinese indigenous dog
gi 333945676 gb JF342873.1	Chinese Shar Pei
gi 333945326 gb JF342848.1	Chinook
gi 647736803 gb KJ637102.1 gi 333946092 gb JF342903.1	Chow chow
gi 91702409 gb DQ480490.1	coated retriever
gi 166065711 gb EU408259.1 gi 166065697 gb EU408258.1	cockapoo
gi 647735879 gb KJ637036.1 gi 189494206 gb EU789780.1 gi 166065809 gb EU408266.1 gi 91702479 gb DQ480495.1 gi 166065837 gb EU408268.1 gi 647737279 gb KJ637136.1 gi 166065823 gb EU408267.1	cocker spaniel
gi 189492414 gb EU789652.1	Collie
gi 333945145 gb JF342835.1	Collie, smooth
gi 166065795 gb EU408265.1	corgi
gi 333945980 gb JF342895.1	Coton de Tulear
gi 166065893 gb EU408272.1 gi 166065865 gb EU408270.1 gi 116119247 gb DQ983934.1 gi 723429699 gb KM061558.1	dachshund

gi 166066089 gb EU408286.1 gi 189494164 gb EU789777.1 gi 91702521 gb DQ480498.1	dachshund miniature
gi 647737405 gb KJ637145.1 gi 333945760 gb JF342879.1	dalmatian
gi 393007478 gb JX088684.1 gi 393007520 gb JX088687.1 gi 393007506 gb JX088686.1 gi 558478248 gb KF661037.1 gi 393007492 gb JX088685.1 gi 393007464 gb JX088683.1 gi 393007450 gb JX088682.1 gi 393007436 gb JX088681.1 gi 393007422 gb JX088680.1 gi 393007408 gb JX088679.1 gi 393007394 gb JX088678.1 gi 393007352 gb JX088675.1 gi 393007380 gb JX088677.1 gi 393007366 gb JX088676.1 gi 393007325 gb JX088673.1 gi 393007311 gb JX088672.1 gi 393007604 gb JX088693.1 gi 393007297 gb JX088671.1 gi 393007590 gb JX088692.1 gi 393007576 gb JX088691.1 gi 393007534 gb JX088688.1	Dingo
gi 647735865 gb KJ637035.1 gi 647737391 gb KJ637144.1	doberman
gi 166065851 gb EU408269.1 gi 333945173 gb JF342837.1 gi 723429825 gb KM061567.1 gi 723429811 gb KM061566.1 gi 723429447 gb KM061540.1	Doberman pinscher
gi 333944949 gb JF342821.1	Dogue de Bordeaux (French Mastiff)
gi 333945871 gb JF342887.1	English Cocker Spaniel
gi 166065921 gb EU408274.1	english mastiff
gi 333944753 gb JF342807.1 gi 333945592 gb JF342867.1	English Setter
gi 333945606 gb JF342868.1 gi 166065907 gb EU408273.1	English Shepard
gi 333945298 gb JF342846.1 gi 50301903 gb AY656744.1 gi 50301917 gb AY656745.1	English Springer Spaniel
gi 647736341 gb KJ637069.1	epagneul breton

gi 333945354 gb JF342850.1	Epagneul Picard (Picardy Spaniel)
gi 333944921 gb JF342819.1	Fila Brasileiro
gi 333945396 gb JF342853.1	Finnish Hound
gi 189492470 gb EU789656.1	Finnish Lapphund
gi 189494276 gb EU789785.1	Flatcoated retriever
gi 647736145 gb KJ637055.1 gi 647736089 gb KJ637051.1 gi 647736831 gb KJ637104.1	fox terrier
gi 166065935 gb EU408275.1 gi 647735921 gb KJ637039.1 gi 333946120 gb JF342905.1 gi 647735907 gb KJ637038.1	French Bull Dog
gi 189493282 gb EU789714.1 gi 189492456 gb EU789655.1	Galgo Español
gi 647736313 gb KJ637067.1 gi 577029797 gb KF907307.1 gi 189494290 gb EU789786.1 gi 91702395 gb DQ480489.1 gi 647735823 gb KJ637032.1 gi 723429713 gb KM061559.1 gi 723429363 gb KM061534.1	German pointer
gi 333945466 gb JF342858.1	German Shorthaired Pointer
gi 647737041 gb KJ637119.1 gi 723429405 gb KM061537.1 gi 225543916 gb FJ817362.1 gi 225543902 gb FJ817361.1 gi 225543888 gb FJ817360.1 gi 225543874 gb FJ817359.1 gi 225543860 gb FJ817358.1 gi 225543944 gb FJ817364.1 gi 225543930 gb FJ817363.1 gi 189493478 gb EU789728.1	golden retriever
gi 50302001 gb AY656751.1	gordon setter
gi 166065949 gb EU408276.1 gi 333945634 gb JF342870.1 gi 723429461 gb KM061541.1 gi 723429419 gb KM061538.1	great dane
gi 166065977 gb EU408278.1 gi 333945270 gb JF342844.1	great pyrenees
gi 647736439 gb KJ637076.1 gi 333945994 gb JF342896.1	greyhound
gi 189493170 gb EU789706.1	Hamilton-stövare
gi 333945508 gb JF342861.1	Harrier

gi 166065991 gb EU408279.1 gi 333944851 gb JF342814.1	havanese
gi 189494150 gb EU789776.1 gi 294774361 dbj AB499817.1 gi 189492498 gb EU789658.1	Husky siberian
gi 333945690 gb JF342874.1	Iceland Dog
gi 50302029 gb AY656753.1 gi 50301861 gb AY656741.1 gi 189494262 gb EU789784.1 gi 91702423 gb DQ480491.1	irish setter
gi 91702493 gb DQ480496.1	irish soft coated wheaten terrier
gi 333945536 gb JF342863.1	Irish Wolfhound
gi 333945312 gb JF342847.1 gi 224813811 gb EU408280.2	Italian Greyhound
gi 647736761 gb KJ637099.1 gi 647737223 gb KJ637132.1 gi 647737083 gb KJ637122.1 gi 647737307 gb KJ637138.1	jack russel terrier
gi 166066019 gb EU408281.1 gi 50301819 gb AY656738.1	jack russell
gi 723430063 gb KM061584.1 gi 723429377 gb KM061535.1	Jagdterrier
gi 189494108 gb EU789773.1 gi 91702577 gb DQ480502.1 gi 91702437 gb DQ480492.1 gi 189494248 gb EU789783.1 gi 189494122 gb EU789774.1 gi 189493982 gb EU789764.1	Jämthund (Swedish elkhound)
gi 333945019 gb JF342826.1	Japanese Chin
gi 189493856 gb EU789755.1	Japanese Spitz
gi 189493954 gb EU789762.1	Jindo
gi 393007548 gb JX088689.1	Kalimantan
gi 189492442 gb EU789654.1	Kangal
gi 166066033 gb EU408282.1 gi 647735781 gb KJ637029.1	keeshond
gi 333945452 gb JF342857.1	Kelpie
gi 50301847 gb AY656740.1	kerry blue terrier
gi 294774347 dbj AB499816.1	Kisyu
gi 333945047 gb JF342828.1	Kyi Leo
gi 166066075 gb EU408285.1	labradoodle
gi 647736705 gb KJ637095.1 gi 647736425 gb KJ637075.1 gi 647736957 gb KJ637113.1	Labrador Retriever

gi 647737013 gb KJ637117.1	Labrador Retriever -golden retriever
gi 189493674 gb EU789742.1 gi 189494038 gb EU789768.1 gi 189493254 gb EU789712.1	Laika, East/ West Siberian
gi 50301987 gb AY656750.1 gi 333945648 gb JF342871.1	leonberger 1 leonberger 2
gi 333945382 gb JF342852.1	Lhasa Apso
gi 723429349 gb KM061533.1 gi 647736033 gb KJ637047.1	Maltese
gi 333945117 gb JF342833.1	Manchester Terrier
gi 333945830 gb JF342884.1	Maremma Sheepdog
gi 333945788 gb JF342881.1	Mastiff
gi 723429321 gb KM061531.1	Miniature Pinscher
gi 723429503 gb KM061544.1	Miniature Schnauzer
gi 333945494 gb JF342860.1	Mountain Cur
gi 166066145 gb EU408290.1 gi 166066131 gb EU408289.1 gi 333944963 gb JF342822.1	neapolitan mastiff
gi 166066103 gb EU408287.1 gi 333945340 gb JF342849.1	new foundland
gi 393007339 gb JX088674.1	New Guinea Singing Dog
gi 333945704 gb JF342875.1	Norfolk Terrier
gi 166066117 gb EU408288.1 gi 333944991 gb JF342824.1	norwegian elkhound
gi 333945774 gb JF342880.1	Nova Scotia Duck Trolling Retriever
gi 50301875 gb AY656742.1 gi 333946106 gb JF342904.1	Old English Sheepdog
gi 333944781 gb JF342809.1	Old English Shepherd
gi 333945913 gb JF342890.1	Owczarek Podkalanski
gi 333946008 gb JF342897.1	Papillion
gi 333944935 gb JF342820.1 gi 723429755 gb KM061562.1	Pekingese
gi 647736355 gb KJ637070.1	Pincher – Chihuahua
gi 166066187 gb EU408293.1	pitbull terrier
gi 333944893 gb JF342817.1	Plott Hound
gi 166066159 gb EU408291.1	pomeranian
gi 189493338 gb EU789718.1 gi 166066173 gb EU408292.1 gi 50301833 gb AY656739.1 gi 91702465 gb DQ480494.1	Poodle
gi 333945284 gb JF342845.1	Portuguese Water Dog

gi 333945033 gb JF342827.1 gi 647735893 gb KJ637037.1 gi 166066201 gb EU408294.1 gi 3172423 gb AF064568.1	pug
gi 189493940 gb EU789761.1 gi 189492554 gb EU789662.1 gi 189494094 gb EU789772.1 gi 189493212 gb EU789709.1	Pungsan
gi 333944977 gb JF342823.1	Pyrenean Mastiff
gi 647735977 gb KJ637043.1 gi 647737153 gb KJ637127.1	Pyrenean Mountain dog
gi 647736509 gb KJ637081.1	Rat Terrier
gi 647736201 gb KJ637059.1 gi 647736187 gb KJ637058.1 gi 166066215 gb EU408295.1 gi 333945802 gb JF342882.1 gi 723429307 gb KM061530.1	Rottweiler
gi 50301973 gb AY656749.1 gi 647736663 gb KJ637092.1 gi 50301889 gb AY656743.1 gi 333945075 gb JF342830.1	saint bernard
gi 189493044 gb EU789697.1 gi 189492792 gb EU789679.1 gi 333945242 gb JF342842.1	Saluki
gi 50302057 gb AY656755.1 gi 7534303 gb CU96639.2	sapsaree 1 sapsaree 2
gi 166066243 gb EU408297.1	schipperke
gi 723429769 gb KM061563.1	Schnauzer
gi 333945844 gb JF342885.1	Scottish Deerhound
gi 333945256 gb JF342843.1	Scottish Terrier
gi 723429475 gb KM061542.1 gi 723429489 gb KM061543.1	Shar pei
gi 3172422 gb AF064567.1	shepherd
gi 189494136 gb EU789775.1 gi 333945662 gb JF342872.1 gi 91702549 gb DQ480500.1 gi 647736593 gb KJ637087.1 gi 647737027 gb KJ637118.1	Shetland sheepdog
gi 333945368 gb JF342851.1	Shiba Inu
gi 647735991 gb KJ637044.1 gi 647737055 gb KJ637120.1 gi 647736929 gb KJ637111.1	shih tzu
gi 333945089 gb JF342831.1	Shikoku

gi 91702535 gb DQ480499.1	siberian Husky
gi 333945410 gb JF342854.1	Silky Terrier
gi 656985138 gb KJ522809.1	Simao chinese
gi 189493268 gb EU789713.1 gi 189492890 gb EU789686.1	Sloughi
gi 333945816 gb JF342883.1	Small Munsterlander
gi 647737335 gb KJ637140.1	spanish cocker spaniel
gi 333945215 gb JF342840.1	Spinone Italiano
gi 647736495 gb KJ637080.1 gi 647736047 gb KJ637048.1	spitz
gi 50301931 gb AY656746.1 gi 50302015 gb AY656752.1	standard schnauzer
gi 91702563 gb DQ480501.1	swedish elkhound
gi 189493884 gb EU789757.1	Tazi
gi 189494234 gb EU789782.1	terrier Black Russian
gi 189494192 gb EU789779.1	terrier Irish (soft coated wheaten)
gi 189493002 gb EU789694.1 gi 166066285 gb EU408300.1 gi 294679605 gb HM048871.1 gi 333945522 gb JF342862.1 gi 189094573 gb EU740414.1	tibetan mastiff
gi 166066299 gb EU408301.1	tibetan spaniel
gi 647737111 gb KJ637124.1	tibetan terrier
gi 166066313 gb EU408302.1	toy poodle
gi 333945424 gb JF342855.1	Treeing Walker Coonhound
gi 166066355 gb EU408305.1 gi 333946050 gb JF342900.1	Viszla
gi 943496475 gb KT591870.1 gi 918056648 gb KT168370.1 gi 918056634 gb KT168369.1 gi 918056816 gb KT168382.1 gi 918056788 gb KT168380.1 gi 918056774 gb KT168379.1 gi 918056760 gb KT168378.1 gi 918056746 gb KT168377.1 gi 918056732 gb KT168376.1 gi 918056718 gb KT168375.1 gi 918056704 gb KT168374.1 gi 918056690 gb KT168373.1 gi 918056676 gb KT168372.1 gi 918056662 gb KT168371.1	Wairau Bar dog
gi 166066383 gb EU408307.1	walker hound

gi 647735851 gb KJ637034.1 gi 333945131 gb JF342834.1	Weimaraner
gi 50301945 gb AY656747.1	welsh springer springer
gi 333945438 gb JF342856.1	Welsh Terrier
gi 91702507 gb DQ480497.1 gi 333945885 gb JF342888.1 gi 189494178 gb EU789778.1 gi 647737195 gb KJ637130.1	west highland terrier
gi 333945927 gb JF342891.1	Whippet
gi 647737377 gb KJ637143.1	White Shepherd
gi 166066397 gb EU408308.1	yorkie chihuahua
gi 647736817 gb KJ637103.1 gi 647736215 gb KJ637060.1 gi 647736551 gb KJ637084.1 gi 723429391 gb KM061536.1	yorkshire terrier

Πίνακας 8. Στον παρακάτω πίνακα φαίνονται οι βάσεις που είναι μοναδικές για μια συγκεκριμένη θέση (SNPs) σε ένα συγκεκριμένο γονίδιο για μια συγκεκριμένη ράτσα ως προς τις υπόλοιπες που συμπεριλήφθηκαν στην μελέτη.

Σκυλιά (breeds)	γονίδιο	Φυσιολογική βάση/θέση/ SNP
Harrier	12S ribosomal RNA	G 692 A
Owczarek Podkalanski	12S ribosomal RNA	A 857 G
Bullmastiff	12S ribosomal RNA	T 29 G
Wairau Bar dog 1-14	12S ribosomal RNA	G 764 A
American escimo dog	16S ribosomal RNA	C 495 T C 773 T
Braque	16S ribosomal RNA	G 44 A
Brittany Spaniel 1-2	16S ribosomal RNA	C 100 T
Chinese indigenous dog	16S ribosomal RNA	T 499 C
Kangal	16S ribosomal RNA	C 1108 T C 1121 T
Kisyu	16S ribosomal RNA	G 463 A T 688 C
Finnish Hound	ND1 gene	T 93 A
Bali Dog	ND1 gene	A 435 G
Japanese chin	ND1 gene	G 235 C A 236 T T 237 A N 240 A C 282 A

Kangal	ND1 gene	T 372 C
Kisyu	ND1 gene	T 463 C
Manchester terrier	ND1 gene	T 240 A
New guinea singing dog	ND1 gene	T 394 G
Tazi	ND1 gene	G 612 A
Corgi	tRNA-Ile	T 42 C
Bolognese	tRNA-Glu	C 37 T
American escimo dog	ND2 gene	A 846 G C 1036 T
Bali Dog	ND2 gene	A 722 C C 833 T
Kangal	ND2 gene	A 42 G T 879 C
Kerry Blue Terrier	ND2 gene	G 796 A
Kisyu	ND2 gene	T 841 C
Simao chinese	ND2 gene	T 228 C
Manchester Terrier	ND2 gene	C 486 T
Irish terrier/(soft coated wheaten)	tRNA trp	A 50 G
Silky Terrier (australian terrier)	tRNA ala	A 6 G
Simao chinese	tRNA-Asn	A 55 G C 56 T
Borzoï	Cox1 gene	G 576 A
Boston terrier	Cox1 gene	G 1260 A
Brittany Spaniel 1-2	Cox1 gene	C 693 T
Norwegian elkhound / jamthund	Cox1 gene	A 1533 G C 348 T
Rat terrier	Cox1 gene	C 40 G
Simao chinese	Cox1 gene	T 68 G
Shikoku	Cox1 gene	C 564 A
Spinone italiano	Cox1 gene	A 21 G
Tazi	Cox1 gene	G 18 A T 1181 C
Wairau Bar dog	Cox1 gene	T 201 C
Shikoku	tRNA-Asp	C 58 T
Norfolk Terrier	tRNA-Asp	A 4 G
Chinese crested	Cox2 gene	A 161 G
Fila Brasileiro	Cox2 gene	C 560 T
Kalimantan Dog	Cox2 gene	A 181 G
Manchester Terrier	Cox2 gene	A 577 T T 578 A
New Guinea Singing Dog	Cox2 gene	C 37 G G 260 T A 340 G
Offspring Shar-pei/ Gunagdong indigenous dog	Cox2 gene	T 303 C
Blue heeler	ATP8 gene	G 35 A
Treeing Walker Coonhound	ATP8 gene	T 146 A
American escimo dog	ATP6 gene	G 456 A
Jindo	ATP6 gene	C 402 T
Bali Dog	ATP6 gene	T 381 C

Chinese shar pei	ATP6 gene	T 186 C
Borzoi	ATP6 gene	G 383 T A 395 G
Bouvier Des Ardennes	ATP6 gene	T 26 A,G
Australian (silky terrier)	ATP6 gene	G 340 A
Manchester terrier	ATP6 gene	G 669 A
Bali dog	Cox3 gene	T 159 C
Bouvier des flandres	Cox3 gene	G 220 A
Kisyu	Cox3 gene	G 355 A C 723 T
Saint bernard	Cox3 gene	G 686 C
Chinese indigenous dog	Cox3 gene	C 732 T
Manchester Terrier	tRNA-gly	G 54 A
Borzoi	ND3 gene	G 93 A
Treeing Walker Coonhound	ND4L gene	G 49 A
Bearded Collie	ND4L gene	G 33 A
Borzoi	ND4L gene	C 18 T
Bali (kalimantan)	ND4 gene	C 957 T
Bedlington	ND4 gene	T 780 C
Bloodhound	ND4 gene	C 270 T
Brittany spaniel 1-2	ND4 gene	T 630 C
Chinese indigenous fog	ND4 gene	G 699 T
Dogue de Bordeaux/French mastiff	ND4 gene	T 1348 A
Flatcoated retriever	ND4 gene	T 921 A
Kangal	ND4 gene	T 1089 C G 1341 A
Munsterlander small	ND4 gene	C 498 T
Owczarek podkalanski	ND4 gene	T 705 C
Pointer german	ND4 gene	C 753 T
Pomeranian	ND4 gene	G 51 A
Carin terrier	tRNA-his	G 47 A
Dogue de Bordeaux/French Mastiff	tRNA-his	C 17 A
Kalimantan Dog	tRNA-leu	C 19 T G 21 A
American escimo dog	ND5 gene	T 611 C A 843 G T 1269 C
Bali dog	ND5 gene	T 1584 C
Belgian Tervuren	ND5 gene	T 681 C
Brittany spaniel 1-2	ND5 gene	C 1737 T
Chinese indigenous dog	ND5 gene	G 1729 A
Fila brasileiro	ND5 gene	C 910 T
Kangal	ND5 gene	A 474 G G 1493 T
Kisyu	ND5 gene	A 1195 G
Kyi leo	ND5 gene	A 177 G
New guinea singing dog	ND5 gene	G 1014 A T 1042 C
Offspring shar pei/gunagdog indigenous	ND5 gene	C 1554 T
Pointer	ND5 gene	T 1015 C

Toy poodle	ND5 gene	C 780 T
Welsh Springer Spaniel	ND5 gene	C 403 A
American escimo dog	ND6 gene	T 405 C
Belgian Tervuren	ND6 gene	G 21 C
Kisyu	ND6 gene	C 361 T
Walker hound	ND6 gene	C 58 T
American escimo dog	CYTB gene	G 1102 A
Basset Hound	CYTB gene	T 618 C
Chow chow	CYTB gene	A 93 G
Corgi	CYTB gene	G 195 A
Finnish Hound	CYTB gene	T 702 C
Kalimantan	CYTB gene	C 684 T
Kangal	CYTB gene	C 1112 T
Treeing walker coonhound	CYTB gene	T 124 A
Wairau bar dog	CYTB gene	T 990 C
White shepherd	CYTB gene	C 834 T

Η πρώτη μετάλλαξη αναπαρίσταται ως G 692 **A** όπου το G συμβολίζει τη βάση που υπάρχει στην θέση 692 στις υπόλοιπες ράτσες ενώ στη Harrier στην συγκεκριμένη θέση υπάρχει η βάση **A** στο γονίδιο που κωδικοποιεί για το 12S ribosomal RNA. Αντίστοιχα ισχύει και για τις υπόλοιπες ράτσες.

Πίνακας 9. Στον παρακάτω πίνακα φαίνονται οι διάφορες ράτσες που εμπεριέχουν πάνω από 2 άτομα καθώς και οι διαφορετικές μοναδικές βάσεις που εμφανίζουν (SNPs) για μια συγκεκριμένη θέση και ένα συγκεκριμένο γονίδιο ως προς τις υπόλοιπες ράτσες .

Σκυλιά (breeds)	γονίδιο	Φυσιολογική βάση /θέση/ SNP
Black russian terrier 1,2	ND1	T 748 C
Jamthund 1-6 + elkhound (norwegian+swedish)	ND1	A 882 G
Leonberger 1-2	ND2	C 222 T
Chow chow 1-2 Leonberger 1-2	ND2	A 256 G
Black Russian Terrier 1-2	ND2	C 321 T
Black Russian Terrier 1,2	ND2	A 590 G
Galgo espanol 1-2	ND2	T 624 C
Chow chow 1-2 Leonberger 1-2	ND2	T 659 C
Elkhound (jamthund)	COX1	C 171 T
Black Russian Terrier 1-2	COX1	G 276 A
Chow chow 1-2	COX1	C 507 T

Leonberger 1-2		
Chow chow 1-2 Elkhound (Jamthund)	COX1	C 589 T
Brittany Spaniel 1-2	COX1	T 693 C
Elkhound (jamthund)	COX1	C 705 T
Elkhound (jamthund)	COX1	G 744 A
chow chow 1-2	COX1	T 984 C
Chow chow 1-2 Leonberger 1-2	COX1	G 1170 A
Rottweiler 1-5	COX1	T 1281 C
Black Russian Terrier 1-2	COX1	T 1363 A
Black Russian Terrier 1-2	COX1	G 1392 A
Norwegian elkhound (Jamthund)	COX1	G 1419 A
Afghan 1 +hound Norwegian elkhound (Jamthund)	COX1	G 1512 A
Black Russian Terrier 1-2	COX1	C 1515 T
Elkhound (Jamthund)	COX1	G 1533 A
Shepard 1-2	COX2	C 153 A
Bernese Mountain Dog 1	COX2	A 350 G
Leonberger Chow-chow	ATP8	T 121 C
Viszla 1-2	ATP6	C 368 T
Galgo espanol	COX3	C 27 T
Sapsaree 1-2	COX3	A 164 G
Black Russian Terrier 1-2	COX3	C 579 T
Black Russian Terrier 1-2	ND3	C 213 T
Jamthund 1-6 (Sweden)	ND3	A 340 G
Chow chow 1-2 Leonberger 1-2	ND3	G 343 A
Pointer german+shorthaired Rottweiler 1-5	ND4L	C 255 T
Leonberger 1-2	ND4	C 146 T
Chow chow 1-2 Leonberger 1-2	ND4	A 413 G
Brittany Spaniel 1 -2	ND4	C 630 T
Chow chow 1-2 Leonberger 1-2	ND4	A 972 G
Black Russian Terrier 1-2	ND5	C 182 T
Chow chow 1-2 Leonberger 1-2	ND5	G 286 A
Black Russian Terrier 1-2	ND5	C 345 T
Bull terrier 1-2	ND5	C 423 T
Black Russian Terrier 1-2	ND5	T 495 C
Black Russian Terrier 1-2	ND5	A 553 G
Black Russian Terrier 1-2	ND5	T 569 A
Chow chow 1-2 Leonberger 1-2	ND5	T 624 C

Black Russian Terrier 1-2	ND5	T 859 C
Sapsaree 1-2	ND5	A 1522 T
Brittany Spaniel 1-2	ND5	T 1737 C
Black Russian Terrier 1-2	ND6	T 210 C
Chow chow 1-2	CYTB	G 93 A
Saluki 1-3	CYTB	C 207 T
Black Russian Terrier 1-2 jamthund 1-6	CYTB	T 465 C
Black Russian Terrier 1-2	CYTB	T 1003 C

Οι θέσεις που ενώ εμφανίζουν μοναδικές βάσεις σε όλα τα άτομα μιας συγκεκριμένης ράτσας αλλά εμφανίζουν την συγκεκριμένη βάση και σε κάποια άλλη ράτσα χωρίς να είναι χαρακτηριστική της ράτσας αυτής για την κάθε θέση συμβολίζονται με N.

Έτσι λοιπόν, στον πίνακα αυτό συμπεριλήφθηκαν με συγκεκριμένα χρώματα.

Πιο συγκεκριμένα λοιπόν:

- ✓ Με **κόκκινο** χρώμα παρουσιάζονται τα σκυλιά που έχουν 0N δηλαδή αυτά που έχουν υψηλή διακριτική ικανότητα καθώς το SNP εμφανίζεται σε μια μόνο ράτσα και είναι χαρακτηριστικό για αυτή για μια συγκεκριμένη θέση
- ✓ Με **Μαύρο** χρώμα παρουσιάζονται τα σκυλιά που έχουν 1-6N και άρα έχουν μέτρια διακριτική ικανότητα
- ✓ Με **Πράσινο** χρώμα παρουσιάζονται τα σκυλιά που έχουν 6-12N και άρα έχουν χαμηλή διακριτική ικανότητα απλά συμπεριλήφθηκαν και αυτά στα αποτελέσματα.

4^ο Κεφάλαιο

Συμπεράσματα –Συζήτηση

Όπως γνωρίζουμε, μέχρι σήμερα το ζητούμενο ήταν η πρόσβαση στην πληροφορία, όμως στην σημερινή εποχή το ζητούμενο είναι η βέλτιστη διαχείριση και δυναμική επεξεργασία του τεράστιου όγκου ψηφιοποιημένης πληροφορίας που τείνει να κατακλύσει τον πολίτη των δικτύων (netizen). Έτσι λοιπόν λόγω των πολυάριθμων εύκολα προσβάσιμων πηγών και συγκεκριμένα των βάσεων δεδομένων υπάρχουν μεγάλα ποσοστά κατατεθειμένης πληροφορίας στο διαδίκτυο. Αυτές οι πληροφορίες στην παρούσα φάση είναι οι αλληλουχίες του μιτοχονδριακού DNA του *Canis lupus familiaris*, που ενώ είναι κατατεθειμένες δεν έχει πραγματοποιηθεί σύγκρισή τους σε μεγάλο βαθμό.

Μέσω της επιλεκτικής αναπαραγωγής, ο σκύλος έχει εξελιχθεί σε εκατοντάδες ποικίλες ράτσες, και δείχνει περισσότερη μορφολογική μεταβολή από οποιοδήποτε άλλο θηλαστικό γη (SpradyTC, 2008). Αυτό μπορούμε να το διακρίνουμε και μέσω των εικόνων του **πίνακα 6**. Τα σκυλιά εμφανίζουν μια τεράστια ποικιλία διαφορετικών εξωτερικών χαρακτηριστικών (μεγεθών, χρωμάτων, τριχωμάτων) παρόλο που ανήκουν όλα στο ίδιο είδος (*Canis lupus familiaris*).

Έτσι λοιπόν όπως είναι αποδεκτό, αυτή η διαφορά διαφαίνεται στην αλληλουχία των βάσεων του DNA τους. Συγκεκριμένα, οι γενετικοί δείκτες είναι αυτοί που αντανακλούν διαφορές σε επίπεδο DNA και συγκεκριμένα το μιτοχονδριακό DNA (mtDNA).

Γι' αυτό το λόγο, η παρούσα ανάλυση πραγματοποιήθηκε μέσω μετα-ανάλυσης των διαθέσιμων δεδομένων, και συγκεκριμένα των διαφόρων mtDNA σκύλων, που ήταν καταχωρημένα στην βάση δεδομένων NCBI και είχαν γνωστή ράτσα. Η μελέτη αποσκοπούσε στην εύρεση SNPs (singlenucleotidepolymorphisms) τα οποία να είναι μοναδικά μεταξύ των διάφορων ρατσών για τις διάφορες θέσεις στα διάφορα γονίδια.

Αυτό πραγματοποιήθηκε μέσω δυο διαφορετικών αναλυσεων. Τα αποτελέσματα τους (**πίνακες 8,9**) είναι διάφοροι SNPs πολυμορφισμοί μεταξύ των ρατσών οι οποίοι προσδίδουν μοναδικότητα σε ένα γονίδιο, για μια θέση, για μια συγκεκριμένη ράτσα. Στην πρώτη συμπεριλήφθηκαν όλα τα άτομα των διαφόρων ρατσών και αναζητήθηκαν τα SNPs που ήταν μοναδικά για μια συγκεκριμένη ράτσα. Έτσι λοιπόν, βρέθηκαν 135 SNPs, που καταγράφονται στον **πίνακα 8**, και με αυτόν τον τρόπο μπόρεσαν να διαχωρίσουν κάποιες ράτσες. Στην δεύτερη μελέτη συμπεριλήφθηκαν μόνο οι ράτσες που περιείχαν παραπάνω από 1 άτομα. Όπως αναφέρθηκε και παραπάνω, έγινε ομαδοποίηση των ατόμων που ανήκουν στην ίδια ράτσα και φτιάχτηκε ένα πρότυπο βάσεων το οποίο είναι χαρακτηριστικό για όλα τα άτομα της ίδιας ράτσας. Σε αυτή την μελέτη βρέθηκαν 55 SNPs, που καταγράφονται στον **πίνακα 9**, τα όποια είχαν διαφορετική αξιοπιστία. Πιο συγκεκριμένα με

κόκκινο διακρίνουμε τα SNPs με 100% αξιοπιστία, με **μαύρο** την μέτρια αξιοπιστία ενώ με **πράσινο** αυτά με χαμηλή αξιοπιστία. Σε αυτές με 100% αξιοπιστία (**κόκκινο**) εμφανιζόταν μια διαφορετική βάση (SNP) σε όλα τα άτομα μιας συγκεκριμένης ράτσας που ήταν διαφορετική σε σχέση με όλες τις υπόλοιπες ράτσες σε μια συγκεκριμένη θέση και ένα συγκεκριμένο γονίδιο. Έτσι λοιπόν, λόγω αυτής της μοναδικότητας, κατάφεραν να διακριθούν κάποιες ράτσες οι οποίες καταγράφονται παρακάτω.

Μέσω της **πρώτης** μελέτης κατάφεραν να διακριθούν με 100% ακρίβεια αλλά σε μικρό αριθμό ατόμων από κάθε ράτσα οι:

American escimo dog, Australian (silky terrier), Bali (kalimantan), Basset Hound, Bearded Collie, Bedlington, Belgian Tervuren, Bloodhound, Blue heeler, Bolognese, Borzoi, Boston terrier, Bouvier Des Ardennes, Braque, Brittany Spaniel 1-2, Bullmastiff, Carin terrier, Chinese crested, Chinese indigenous dog, Chinese shar pei , Chow chow, Corgi, Dogue de Bordeaux/French mastiff, Fila Brasileiro, Finnish Hound, Flatcoated retriever, Harrier, Irish terrier/(soft coated wheaten), Japanese chin, Jindo, Kalimantan, Kangal, Kerry Blue Terrier, Kisyu, Kyi leo, Manchester terrier, Munsterlander small, New guinea singing dog, New Guinea Singing Dog , Norfolk Terrier, Norwegian elkhound / jamthund, Offspring shar pei/gunagdog indigenous, Owczarek Podkalanski, Pointer german, Pomeranian, Rat terrier, Saint Bernard, Shikoku, Silky Terrier (australian terrier), Simao chinese, Spinone italiano , Tazi, Toy poodle, Treeing walker coonhound, Wairau Bar dog, Walker hound, Welsh, Springer Spaniel, White shepherd.

Μέσω της **δεύτερης** μελέτης κατάφεραν να διακριθούν οι παρακάτω ράτσες με μεγαλύτερη ή μικρότερη ακρίβεια όσον αφορά συγκεκριμένες θέσεις σε συγκεκριμένα γονίδια αλλά με μεγαλύτερο αριθμό ατόμων από κάθε ράτσα:

Afghan, Bernese Mountain Dog , Black Russian Terrier, Brittany Spaniel, Bull terrier, Chow chow, Elkhound (jamthund), Galgo espagnol, jamthund, Leonberger, Norwegian elkhound (Jamthund), german pointer, Rottweiler, Saluki, Sapsaree , Shepard, Vizsla.

Πολλές ράτσες από αυτές είναι αρκετά εμπορικές και η εύρεση δεικτών που τις διαχωρίζουν από άλλες είναι εξαιρετικά σημαντικό.

Είναι σημαντικό να αναφερθεί ότι στις παραπάνω δυο μελέτες συμπεριλήφθηκαν όλα τα γονίδια εκτός από την D-100p ή περιοχή ελέγχου. Η D-100p εμφανίζεται στην κύρια μη-κωδικοποιητική περιοχή του μιτοχονδριακού DNA. Ορισμένες βάσεις εντός της περιοχής D-100p είναι συντηρημένες, όμως το μεγαλύτερο μέρος της είναι εξαιρετικά μεταβλητό. Επίσης για την σύγκριση των διαφόρων μιτοχονδριακών γονιδίων πρέπει να υπάρχει σταθερό μήκος κάτι που δεν ισχύει στην περίπτωση της D-100p και αυτό μαζί με τα παραπάνω αποτελεί ένα μεγάλο μειονέκτημα για την ανάλυσή του στην παρούσα μελέτη.

Εύρεση διαφορετικών SNPs πραγματοποιήθηκε και από τονVerscheureκαι τους συνεργάτες του (2014) όπου μελέτησαν έναν συγκεκριμένο πληθυσμό σκύλων από το Βέλγιο και συγκεκριμένα 119 διαφορετικές μιτοχονδριακές αλληλουχίες. Αυτοί κατέληξαν σε 33 SNPs που χωρίζουν με επιτυχία τους 6 πιο κοινούς απλοτύπων στην περιοχή ελέγχου σε 32 ομάδες των μιτοχονδριακών DNA αλληλουχιών.

Μελέτη του μιτοχονδριακού DNAπραγματοποιήθηκε από την Kristen και τους συνεργάτες της (2009). Στην παραπάνω μελέτη βρέθηκαν 79 σκυλιά και 356 SNPs στο μιτοχονδριακό DNA εκτός από την d-100p. Στις περισσότερες μελέτες μετά την εύρεση των SNPs κατασκευάστηκαν φυλογενετικά δέντρα για την επιβεβαίωση των αποτελεσμάτων όπως στην παραπάνω μελέτη.

Ο Jun-Hui Zhao και ο Wei Liu (2015) πραγματοποίησαν ενίσχυση και αλληλούχηση του mtDNA του Simao Chinese indigenous dog αλλά σε πειραματικό επίπεδο όχι in silico όπως στην συγκεκριμένη μελέτη.

Οι Matthew και οι συνεργάτες του (2014), μελέτησαν τη γενετική του *Canis lupusfamiliaris*, λόγω της στενής του σχέσης με το canislupus. Μελέτησαν πληθυσμούς λύκων στη Νοτιοανατολική Αλάσκα και βρήκαν κάποια SNPs τα οποία χρησιμοποιήθηκαν για την ποσοτικοποίηση της γενετικής διαφοροποίησης των λύκων αυτών και των λύκων, κογιότ και σκύλων από άλλες περιοχές της Βόρεια Αμερικής. Βρέθηκε ότι υπάρχει διαφοροποίηση των συχνοτήτων SNP μεταξύ των ειδών (λύκοι, κογιότ, και σκυλιά) και μεταξύ των γεωγραφικών περιοχών. Στην παρούσα μελέτη, τα αποτελέσματα ενισχύθηκαν και μέσω της κατασκευής φυλογενετικών δέντρων για το πυρηνικό και mtDNA.

Μια άλλη μελέτη, αυτή τη φορά για την σύγκριση λύκων με σκύλους πραγματοποιήθηκε από τους η Sindić M.και τους συνεργάτες της (2011).Ο στόχος αυτής της εργασίας ήταν να παρέχει ένα σύνολο δεδομένων όσον αφορά το μιτοχονδριακό DNA της περιοχής ελέγχου του σκύλου από την Κροατία και του λύκου για την εύρεση εάν αυτές οι ακολουθίες μπορούν να χρησιμοποιηθούν ως σημείο αναφοράς για την αναγνώριση των ειδών.

Πληθυσμιακή μελέτη, για την αναζήτηση της γεωγραφικής προέλευση του σκύλου , πραγματοποιήθηκε από τον Jun-Feng Pang και του συνεργάτες του (2009). Οι παραπάνω,βέβεια δεν αναζητούσαν συγγένεια μεταξύ ρατσών όπως στην δική μα μελέτη. Γεωγραφική προέλευση μέσω ανάλυσης μιτοχονδρικού DNA πραγματοποίησε και ο Thalmann με τους συνεργάτες του (2013).

Οι περισσότερες αναλύσεις, όπως παρατηρούμε και από τι παραπάνω μελέτες, πραγματοποιήθηκαν σε **πληθυσμούς** σκύλων. Στην παρούσα εργασία όμως, οι φυλές εκπροσωπήθηκαν από λίγα άτομα, αυτά που ήταν καταχωρημένα στο διαδίκτυο. Για να επιτευχθεί μια ολοκληρωμένη εικόνα της συσχέτισης των mtDNA των διάφορων φυλών- ρατσών και για την εξαγωγή περισσότερο αξιόπιστων αποτελεσμάτων θα ήταν επιθυμητό οι μελέτες να ήταν μεγαλύτερης κλίμακαςκαι να περιείχαν περισσότερα άτομα για κάθε ράτσα.Επιπρόσθετα, θα ήταν αποτελεσματικότερη η εφαρμογή σε πραγματικό πληθυσμό καθώς ο αριθμός των

διαφορετικών σκυλιών θα ήταν μεγάλος και θα οδηγούσε στην μεγαλύτερη ακρίβεια των αποτελεσμάτων μας.

Όπως αναφέρθηκε και παραπάνω, οι **πυρηνικοί γενετικοί τόποι** που κωδικοποιούν πρωτεΐνες συνήθως δεν εμφανίζουν πολυμορφισμό μεγέθους ή, εάν εμφανίσουν, αυτός αντιστοιχεί σε αριθμό πολλαπλάσιο των τριών νουκλεοτιδίων, με αποτέλεσμα η ομοπαράθεση των αλληλουχιών να μην παρουσιάζει δυσκολίες. Το κύριο πλεονέκτημα που καθιστά το πυρηνικό ριβοσωμικό DNA κατάλληλο φυλογενετικό δείκτη, είναι το μεγάλο μέγεθος της αλληλουχίας που υποδηλώνει το μεγάλο ποσό της φυλογενετικής πληροφορίας που φέρει. Κατά συνέπεια λοιπόν, μια εναλλακτική λύση θα αποτελούσε η μετά-ανάλυση του πυρηνικού DNA σε δείγματα με γνωστές ράτσες.

Μελλοντικός στόχος της μελέτης είναι ο σχεδιασμός **εκκινήτων** που στοχεύουν συγκεκριμένες θέσεις και η εφαρμογή τους σε δείγματα σκύλων με γνωστή ράτσα, για να διαπιστωθεί η πιστότητα των αποτελεσμάτων μας *in vitro*. Με αυτόν τον τρόπο θα μπορούσαν οι εκτροφείς, εφόσον βρεθούν και καταχωρηθούν περισσότερα SNPs χαρακτηριστικά για κάθε ράτσα, να βρίσκουν ποιες ράτσες είναι "καθαρόαιμες" και αυτό θα απόφερε ένα μεγάλο κέρδος σε αυτούς.

Μέσω της εύρεση καταγραφής διαφόρων SNPs που διαχωρίζουν τις ράτσες μεταξύ τους καθώς και των διαφορετικών εκκινήτων σε κάθε περίπτωση, θα ήταν εφικτό, μελλοντικά, με την ύπαρξη μόνο μιας τρίχας ενός σκύλου να βρούμε και να χαρακτηρίσουμε τη ράτσα του. Έτσι το δείγμα θα απαιτούσαν σαν μελέτη θα ήταν **ελάχιστο** και κάτι που θα αντικαθιστούσε την λήψη αίματος που ταλαιπωρεί το ζώο. Στην ουσία, η παραπάνω διαδικασία αποτελεί μια μη επεμβατική μέθοδο η οποία είναι εύκολη, δεν θέτει σε κίνδυνο την ζωή του ζώου και εμφανίζει μεγάλη ακρίβεια αποτελεσμάτων.

Συμπερασματικά λοιπόν, από την παρούσα μελέτη συμπεραίνουμε ότι πρέπει να γίνουν πληθυσμιακές μελέτες σε μεγαλύτερη έκταση έτσι ώστε να βρεθούν περισσότερα και πιο αξιόπιστα SNPs που θα μπορούσαν να χρησιμοποιηθούν στο μέλλον προς όφελος του ανθρώπου

Κεφάλαιο 5°

Βιβλιογραφία

- ❖ Akaike H. (1974) A new look at the statistical model identification, IEEE TRANSACTIONS ON AUTOMATIC CONTROL, 19:716–723.
- ❖ Avise J.C. & D. Ellis (1986) Mitochondrial DNA and the evolutionary genetics of higher animals. *Philosophical Transactions of the Royal Society London B* 312:325-342.
- ❖ Avise J.C., Arnold J & Ball RM (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Reviews Inc, Ecology System* 18, 489–522
- ❖ Avise J C. (1989) A role for molecular genetics in the recognition and conservation of endangered species. *Trends in Ecology and Evolution*, 4, 279–281.
- ❖ Avise J.C. (1994) *Molecular Markers, Natural History, and Evolution*. Chapman & Hall, New York
- ❖ Aquadro, C.F., & Greenberg B. D. (1983) Human mitochondrial DNA variation and evolution: analysis of nucleotide sequences from seven individuals. *Genetics*, 103, 287–312.
- ❖ Awadalla P., Eyre-Walker A.& Smith JM. (1999) Linkage Disequilibrium and Recombination in Hominid Mitochondrial DNA. *Science*. 286 (5449): 2524–2525.
- ❖ Bairoch A.& Apweiler R. (2000) The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res.* 28: 45-48
- ❖ Ballard J.W., Olsen J. G., Faith P. D., Odgers A.W., Rowell M.D., Atkinson W.P.(1992). Evidence from 12S ribosomal RNA sequences that onychophorans are modified arthropods. *Science* 258: 1345-1348.
- ❖ Barton Derek E.G. Briggs, Jonathan A. Eisen, David B. Goldstein & Nipam H. Patel (2007), *Evolution*, Cold spring Harbor Laboratory Press, New York, p. G-6 Glossary
- ❖ Baxevanis D. A. (2016) *Bioinformatics: A Practical Guide for the analysis of genes and proteins*, 3rd edition, scientific publications PARISIANOU, 2016 Chapter 12 and 14
- ❖ Benson D.A. et al. (2007). GenBank. *Nucleic Acids Res.* 35: D21-25.
- ❖ Benson D.A., Karsch-Mizrachi I., Lipman D.J., Ostell J., Wheeler D.L. (2007) GenBank. *Nucleic Acids Res.* 35: D21-25.
- ❖ Benson D.A., Karsch-Mizrachi, Clark K., Lipman D.J., Ostell J., Sayers E.W. (2012) GenBank, Oxford University Press, p.48-53
- ❖ Benson D.A., Karsch-Mizrachi, I., Lipman D.J., Ostell, J., Ouellette B.F.F., Rapp B.A. & Wheeler D.L. (2000) GenBank. *Nucleic Acids Res.*, 28, 15–18.

- ❖ Birky C. W., Jr., Maruyama T., & Fuerst P.(1983) An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts and some results. *Genetics* 103:5 13-527.
- ❖ Boore J. L., T. M. Collins, D. Stanton, L. L. Daehler, W. M. Brown, 1995 Deducing the pattern of arthropod phylogeny from mitochondrial DNA rearrangements *Nature* 376:163-165
- ❖ Boore, L.J. (1999) Animal Mitochondrial Genomes. *Nucleic Acids Research* 27(8): 1767-1780.
- ❖ Bosa H. D. & David P.(2004). Using models of nucleotide evolution to build, *Developmental and Comparative Immunology* 29, 211–227
- ❖ Botstein, D., White, K.L., Skolnick, M. & Davis, R .W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32: 3 14-33
- ❖ Buburuzan L. Gorgan L, Bara I. (2007). Types of DNA used in speciation and phylogeny studies. *TOM VIII*.
- ❖ Burnham P. Kenneth & Anderson R. David (2002). Model Selection and Multimodel, *SOCIOLOGICAL METHODS & RESEARCH*, Vol. 33, 261-304
- ❖ Burns K.J. (1997). Molecular systematics of tanagers (Thraupinae): evolution and biogeography of a diverse radiation of neotropical birds. *Mol. Phylogenet. Evol.* 8: 334–348.
- ❖ Calisher, CH (2007). "Taxonomy: what's in a name? Doesn't a rose by any other name smell as sweet?" *Croatian Medical Journal* 48 (2): 268–270
- ❖ Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* 17, 540-552.
- ❖ Chenuil A. (2006). Choosing the right molecular genetic markers for studying biodiversity: from molecular evolution to practical aspects. *Genetica* 127: 101-120.
- ❖ Clayton D.A. (1982) Replication of animal mitochondrial DNA. *Cell* 28: 693-705.
- ❖ Clayton D.A. (1991) Replication and transcription of vertebrate mitochondrial DNA. *Annu. Rev. Cell Biol.* 7: 453-478.
- ❖ Crews S., Ojala D., Posakony J., Nishiguchi J., Attardi G.(1979) Nucleotide sequence of a region of human mitochondrial DNA containing the precisely identified origin of replication. *Nature* 277: 192–198
- ❖ Cronin A. M., Cánovas A., Bannasch L.D., Oberbauer M.A. , Medrano F. J.(2014) Single Nucleotide Polymorphism (SNP) Variation of Wolves (*Canis lupus*) in Southeast Alaska and Comparison with Wolves, Dogs, and Coyotes in North America, *The American Genetic Association*, p. 26-36.
- ❖ Davis P.H.& V.H.Heywood. (1973) *Principle of Angiosperm Taxonomy*. Robert, E. Krieger Publishing company , Huntington, New York.p.17

- ❖ Dewey T. & Bhagat. S. (2002) *Canis lupus familiaris*, Animal Diversity Web. Retrieved 6 January 2009.
- ❖ Dixon M.T. & Hillis D.M. (1993). Ribosomal RNA Secondary Structure: Compensatory Mutations and Implications for Phylogenetic Analysis, The University of Chicago, *Molecular Biology Evolution* 10(1), p:256-267,
- ❖ Dolman G. & Phillips B. (2004). Single copy nuclear DNA markers characterized for comparative phylogeography in Australian wet tropics rainforest skinks. *Molecular Ecology Notes* 4: 185–187.
- ❖ Downie, D.A. & Gullan, P.J. (2004) Phylogenetic analysis of mealybugs (Hemiptera: Coccoidea: Pseudococcidae) based on DNA sequences from three nuclear genes, and a review of the higher classification. *Syst. Entomol.* 29: 238–259.
- ❖ Felsenstein J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution* 17: 368–376.
- ❖ Freedman A.H., Gronau I., Schweizer R.M., Vecchyo D., Han E., Silva P.M., Galaverni M., Fan Z., Marx P., Lorente-Galdos B. (2014) Genome sequencing highlights the dynamic early history of dogs
- ❖ Garrigan, D. & Hammer, M.F. (2006) Reconstructing human origins in the genomic era. *Nat. Rev. Genet.* 7 (9): 669–80.
- ❖ Guliaeva N.A., Kuznetsova E.A.& Gaziev Al. (2006). Proteins associated with mitochondrial DNA protect it against X-rays and hydrogen peroxide. *Biophysics* 51(4): 620–623.
- ❖ Greenberg B.D., Newbold J.E.& Sugino A.(1983) Intraspecific nucleotide sequence variability surrounding the origin of replication in human mitochondrial DNA, Published by Elsevier B.V, *Gene* (1983) 21:33-49.
- ❖ Gyllensten, U., Wharton, D., Josefsson, A. & Wilson, A. C. (1991) Paternal inheritance of mitochondrial DNA in mice. *Nature* 352:255-257.
- ❖ Hall & E. Raymond (1981) *The Mammals of North America*, vols. I and II, 2nd ed. John Wiley and Sons, Inc, New York, New York, USA, p.1181
- ❖ Hao Bl., Qi J& Wang B(2004). Prokaryote phylogeny without sequence alignment:a K-string composition approach.*Journal of Molecular Evolution* , Volume 58, Issue 1, pp 1-11
- ❖ Hasegaw M., Kishino H. & Yano T. (1985). Dating of human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, 22(2):160-74.
- ❖ Hayashi J.I., Tagashira Y., & Yoshida M. C (1985) Absence of extensive recombination between inter- and intraspecies mitochondrial DNA in mammalian cells. *Exp. Cell Res.* 160:387-395.
- ❖ Heise P. J., Maxson L. R., Dowling H. G., & Hedges S. B. (1995). Higher level snake phylogeny inferred from mitochondrial-DNA sequences of 12S ribosomal-RNA and 16S ribosomal-RNA genes. *Mol. Biol. Evol.* 12: 259–265.

- ❖ Honaki J. (1982). Mammal species of the world: A taxonomic and geographic reference (First Edition). Allen Press and the Association of Systematics Collections. p. 245
- ❖ Huelsenbeck P. & Crandall, K.A (1997). Phylogeny estimation and hypothesis testing using maximum likelihood. *Annu. Rev. Ecol. Syst.* 28: 437–466.
- ❖ Irimia M. & Roy S.W. (2008) Spliceosomal introns as tools for genomic and evolutionary analysis. *Nucleic Acids Res.* 36 (5): 1703-12.
- ❖ Jermin LS, Jayaswal V, Ababneh F, Robinson J. (2008). Phylogenetic Model Evaluation. *Bioinformatics* 452: 16. *Journal of Molecular Evolution* 16: 111–120.
- ❖ Jukes T.H. & Cantor C.R. (1969). Evolution of Protein Molecules, pp. 21–132.
- ❖ Jun-Hui Zhao & Wei Liu (2015) The complete mitochondrial genome of the Simao Chinese indigenous dog, 14;27(1):545-6
- ❖ Jun-Feng Pang, Cornelya Kluetsch, Xiao-Ju Zou, Ai-bing Zhang, Li-Yang Luo, Helen Angleby, Arman Ardalan, Camilla Ekström, Anna Skölleremo, Joakim Lundeborg, Shuichi Matsumura, Thomas Leitner, Ya-Ping Zhang & Peter Savolainen (2009) mtDNA Data Indicate a Single Origin for Dogs South of Yangtze River, Less Than 16,300 Years Ago, from Numerous Wolves, *Molecular Biology Evolution*, 26 (12): 2849-2864.
- ❖ Kimura M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences *Journal of molecular evolution*, 16(2):111-20.
- ❖ Kondo, R., Satta, Y., Matsuura, E.T., Ishiwa, H., Takahata, N. & Chigusa, S.I. (1990). Incomplete maternal transmission of mitochondrial DNA in *Drosophila*. *Genetics* 126:657-663.
- ❖ Webb K. M. & Allard M. W. (2009), Mitochondrial Genome DNA Analysis of the Domestic Dog: Identifying Informative SNPs Outside of the Control Region*. *Journal of Forensic Sciences*, 54: 275–288.
- ❖ Krzywinski J., Grushko O.G., Besansky N.J.(2006) Analysis of the complete mitochondrial DNA from *Anopheles funestus*: an improved dipteran mitochondrial genome annotation and a temporal dimension of mosquito evolution. *Molecular phylogenetics and evolution*, 39(2): 417-23
- ❖ Machida RJ, Miya MU, Nishida M & Nishida S (2002). Complete mitochondrial DNA sequence of *Tigriopus japonicus* (crustacea: copepoda). *Marine Biotechnology* 4:406–417.
- ❖ Maggiris Vaggelis (1997) Carl Linnaeus (1707-1778): The organizer of the Scientific Nomenclature, *Periscope of Science*, Volume 210, , pp. 10-11
- ❖ Malmström H., Storå J., Dalén L., Holmlund G., Götherström A. (2005) Extensive human DNA contamination in extracts from ancient dog bones and teeth. *Molecular Biology and Evolution* 22: 2040-2047.

- ❖ Muraji M. & Nakahara S. 2001. Phylogenetic relationships among fruit flies, *Bactrocera* (Diptera, Tephritidae), based on the mitochondrial rDNA sequences. *Insect Mol Biol.* 10 (6): 549-59
- ❖ Ronald M. Nowak (1991) , *Walker's Mammals of the World*, vol. 1, 5th ed.,The Johns Hopkins University Press, Baltimore and London, p.1629
- ❖ Ogden T.H.& Rosenberg M.S.(2006). Multiple sequence alignment accuracy and phylogenetic inference, *systematic biology*,55(2):314-28.
- ❖ Olsen G.J. & Woese C.R. (1993) *Ribosomal RNA: a key to phylogeny*. FASEB, Baltimore, Maryland, USA p.xviii + 1207, Smithsonian Institution Press , Johns Hopkins University Press *Journal* 7: 113-123. P.214
- ❖ Page Roderic D. M & Holmes Edward C. (1998) *Molecular Evolution : A Phylogenetic Approach*, Evolutionary genetics
- ❖ Panaretos J. & Xekalaki E. (2000) *Introduction to Statistical Thinking*, vol. II (Estimation and Testing of Hypothesis), Athens.
- ❖ Parker H.G., Kim L.V., Sutter N.B., Carlson S., Lorentzen T.D., Malek T.B., Johnson G.S., DeFrance H.B., Ostrander E.A. & Kruglyak L. (2004). Genetic structure of the purebred domestic dog. *Science*, 304(5674):1160-4.
- ❖ Pevsner J. (2009). *Bioinformatics and functional genomics*, 2nd publication, John Wiley and Sons(951) p. 211-227.
- ❖ Podsiadlowski L., Braband A.& Mayer G. (2008). The Complete Mitochondrial Genome of the Onychophoran *Epiperipatus biolleyi* reveals a Unique Transfer RNA Set and Provides Further Support for the Ecdysozoa Hypothesis. *Molecular Biology Evolution* 25(1):42–51.
- ❖ Pontus Skoglund, Erik Ersmark, Eleftheria Palkopoulou, Love Dalen (2015) *Ancient Wolf Genome Reveals an Early Divergence of Domestic Dog Ancestors and Admixture into High-Latitude Breeds*, Elsevier Ltd. Published by Elsevier Inc., *Current Biology* 25, 1515–1519
- ❖ Queller D.C., Strassmann J.E., Hughes C.R. (1993). Microsatellites and Kinship. *Trends in Ecology and Evolution*, 8: 285–288 rearrangements. *Nature* 376:163–165.
- ❖ King C. R., Mulligan P., & Stansfield W. (2006). *A dictionary of genetics*, Oxford University Press p. 336
- ❖ Rokas A., Ladoukakis E. & Zouros E. (2003). Animal mitochondrial DNA recombination revisited. *TRENDS in Ecology and Evolution* 18(8): 411-417.
- ❖ Rowan Hooper, (2005) *Boxer is first dog to have full genome revealed* DAILY NEWS
- ❖ Russell P. J. (2009) *i-Genetics: a mendelian approach*, academic publishers, p.234
- ❖ Saitou N. & Nei M. (1987). The Neighbor-Joining method - a new method for reconstructing phylogenetic trees,*Molecular Biology and Evolution* 4:406-425.

- ❖ Savolainen P., Zhang Y.P., Luo J., Lundeberg J. & Leitner T. (2002) Genetic evidence for an East Asian origin of domestic dogs, *Science* 298 (5598): 1610-3.
- ❖ Sequeira F., Ferrand N. & Harris J. (2006). Assessing the phylogenetic signal of the nuclear beta-fibrinogen intron 7 in salamandrids (Amphibia: Salamandridae). *Amphibia-Reptilia* 27: 409–418.
- ❖ SIMON C., T. R. BUCKLEY, F. FRATI, J. B. STEWART & A. T. BECKENBACH (2006) Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. *Annu. Rev. Ecol. Evol. Syst.* 37: 545–579.
- ❖ Sindicic, M., Gomercic, T., Galov, A., Polanc, P., Huber, D., & Slavica, A. (2012). Repetitive sequences in Eurasian lynx (*Lynx lynx* L.) mitochondrial DNA control region. *Mitochondrial DNA* 23(3): 201-207.
- ❖ Skoglund P., Ersmark E. & Palkopoulou E., Dalén L. (2015) Ancient wolf genome reveals an early divergence of domestic dog ancestors and admixture into high-latitude breeds. *Current Biology*, 25: 1515–1519
- ❖ Slatkin M. (1987) Gene flow and the geographic structure of natural populations. *Science*, 236, 787–792.
- ❖ Spady T.C. & Ostrander E.A. (2008) Canine Behavioral Genetics: Pointing Out the Phenotypes and Herding up the Genes, *American Journal of Human Genetics* 82 (1): 10–8.
- ❖ Talavera G. & Castresana J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology* 56, 564-577.
- ❖ Tamura K. (1992). Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C content biases. *Molecular Biology and Evolution* 9: 678–687.
- ❖ Tamura K., Dudley J., Nei M. & Kumar S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0, *Molecular Biology Evolution* (8):1596-9
- ❖ Tamura K. & Nei M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution* 10: 512–526.
- ❖ Tavaré S. (1986). Some probabilistic and statistical problems in the analysis of DNA sequences. In *American Mathematical Society: Lectures on Mathematics in the Life Sciences*, Vol. 17 (1986), pp. 57-86
- ❖ Thalmann O., Shapiro B., Cui P., Schuenemann V.J., Sawyer S.K., Greenfield D.L., Germonpré M.B., Sablin M.V., López-Giráldez F., Domingo-Roura X., Napierala H., Uerpmann H.P., Loponte D.M., Acosta A.A., Giemsch L., Schmitz R.W., Worthington B., Buikstra J.E., Druzhkova A., Graphodatsky A.S., Ovodov N.D., Wahlberg N., Freedman A.H., Schweizer R.M., Koepfli K.P., Leonard J.A.,

- Meyer M., Krause J., Pääbo S., Green R.E. & Wayne R.K. (2013) Complete Mitochondrial Genomes of Ancient Canids Suggest a European Origin of Domestic Dogs, *Science*, 342(6160):871-4.
- ❖ Tim Radford (2015) Dog DNA study reveals the incredible journey of man's best friend, the guardian *Transactions on Automatic Control* 19(6): 716–723.
 - ❖ Upholt W.B. & Dawid I.B. (1977) Mapping of mitochondrial DNA of individual sheep and goats: rapid evolution in the D loop region, *cell*, (3):571-83
 - ❖ Van Gelder & Richard G.(1978) A Review of Canid Classification, AMERICAN MUSEUM OF NATURAL HISTORY CENTRAL PARK WEST AT 79TH STREET, NEW YORK, N.Y. 10024, pp. 1-10
 - ❖ Viegas J. (2014), Dogs Not as Close Kin to Wolves as Thought, *Discovery News*
 - ❖ Verscheure S., Bäckeljau T. & Desmyter S. (2014) Dog mitochondrial genome sequencing to enhance dog mtDNA discrimination power in forensic casework, *forensic science of international genetics*, 12:60-8
 - ❖ Xiaoming Wang, Richard H. Tedford, Mauricio Antón (2008) *Dogs: Their Fossil Relatives and Evolutionary History*, New York : Columbia University Press,
 - ❖ Wayne, R.K. (1993.) Molecular evolution of the dog family. *Theoretical & Applied Genetics*, *Trends in genetics*, 218-24.
 - ❖ Weisburg W.G., Barns S.M., Pelletier D.A., Lane DJ. (1991). 16S ribosomal DNA amplification for phylogenetic study, *Journal of Bacteriology*, 173(2): 697-703.
 - ❖ Williams JG., Kubelik AR., Livak KJ. , Rafalski JA. & Tingey SV. (1990) DNA polymorphisms amplified by arbitrary primers are useful genetic markers. *Nucleic Acids Research*, 18: 6531-6535..
 - ❖ Willows-Munro S., Robinson T.J., Matthee C.A. (2005). Utility of nuclear DNA intron markers at lower taxonomic levels: phylogenetic resolution among nine *Tragelaphus* spp. *Molecular Phylogenetics and Evolution*, 35: 624–636.
 - ❖ Wilson, Don E., & DeeAnn M. Reeder, eds., (1993) , *Mammal Species of the World: A Taxonomic and Geographic Reference*, 2nd ed., 3rd printing
 - ❖ Wilson, Don E., & F. Russell Cole (2000) *Common Names of Mammals of the World*, xiv + 204 , Smithsonian Institution Press, Washington, DC, USA
 - ❖ Wilson, Don E., & DeeAnn M. Reeder (2005) , *Mammal Species of the World: A Taxonomic and Geographic Reference*, 3rd ed., vols. 1 & 2
 - ❖ Wu CH, Huang H, Arminski L, Castro-Alvear J, Chen Y, Hu ZZ, Ledley RS, Lewis KC, Mewes HW, Orcutt BC, Suzek BE, Tsugita A, Vinayaka CR, Yeh LS, Zhang J, Barker WC (2002) The Protein Information Resource: an integrated public resource of functional annotation of proteins, *Nucleic Acids Research*, 30(1): 35–37.
 - ❖ YANAN WANG (2015) The 33,000-year-old origin story of man's best friend, *The Washington Post*

- ❖ Yang Z. (1996). The among-site rate variation and its impact on phylogenetic analyses. *Trends Ecol. Evol.* 11: 367-372.
- ❖ Zhao X., Li N., Guo W., Hu X., Liu Z., Gong G., Wang A., Feng J., Wu C. (2004). Further evidence for paternal inheritance of mitochondrial DNA in the sheep (*Ovis aries*). *Heredity* 93: 399–403.
- ❖ Fan Z., Silva P., Gronau I., Wang S., Armero AS., Schweizer RM., Ramirez O., Pollinger J., Galaverni M., Ortega Del-Vecchio D., Du L., Zhang W., Zhang Z., Xing J., Vilà C., Marques-Bonet T., Godinho R., Yue B. & Wayne RK. (2016) Worldwide patterns of genomic variation and admixture in gray wolves, *Genome research*, p. 163-73