

ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΥΓΕΙΑΣ
ΤΜΗΜΑ ΒΙΟΧΗΜΕΙΑΣ ΚΑΙ ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ



ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**«ΑΝΑΛΥΣΗ ΓΕΝΕΤΙΚΗΣ ΔΟΜΗΣ ΕΚΤΡΕΦΟΜΕΝΩΝ
ΠΛΗΘΥΣΜΩΝ ΨΑΡΙΩΝ: ΕΦΑΡΜΟΓΕΣ ΣΕ ΔΙΑΧΕΙΡΙΣΗ ΚΑΙ
ΑΠΟΔΟΣΗ ΠΑΤΡΟΤΗΤΑΣ»**



ΣΤΑΜΟΥΛΑΚΗ ΜΑΡΙΑΝΘΗ

Λάρισα 2018

Ευχαριστίες

Η παρούσα διπλωματική ερευνητική εργασία αποτελεί αντικείμενο μελέτης εκτρεφόμενων πληθυσμών ψαριών για την ανάλυση του τρόπου διαχείρισης ιχθυοκαλλιεργειών. Η συγκεκριμένη εργασία πραγματοποιήθηκε στο εργαστήριο Συγκριτικής και Εξελικτικής Βιολογίας του τμήματος Βιοχημείας και Βιοτεχνολογίας του Πανεπιστημίου Θεσσαλίας.

Για την βοήθειά τους στην πραγματοποίηση της μελέτης θα ήθελα να ευχαριστήσω την Αναπληρώτρια Καθηγήτρια και επιβλέπουσα της διπλωματικής εργασίας Μούτου Αικατερίνη που με βοήθησε στην καλύτερη οργάνωση και παρουσίαση των αποτελεσμάτων της ερευνητικής μελέτης. Επίσης ευχαριστώ τον Γιαννούλη Θέμη, υποψήφιο διδάκτορα του τμήματος Βιοχημείας και Βιοτεχνολογίας για την πολύτιμη βοήθειά του σε όλη τη διαδικασία πραγματοποίησης της μελέτης αλλά και στην επίλυση κάθε απορίας.

Περίληψη

Οι μικροδορυφόροι αποτελούν σημαντικό εργαλείο στην ανάλυση της γενετικής πληθυσμών με ιδιαίτερη ανάπτυξη τα τελευταία χρόνια. Χρησιμοποιούνται ευρέως σε αναλύσεις λόγω της εύκολης αξιοποίησής τους και του εύκολου τρόπου εφαρμογής τους. Στη συγκεκριμένη διπλωματική εργασία χρησιμοποιούνται τέτοια δεδομένα από μικροδορυφόρους γενετικών τόπων για τη γενετική ανάλυση εκτρεφόμενων πληθυσμών λαβρακιού και τσιπούρας. Σε συνδυασμό με προγράμματα βιοπληροφορικής προσφάτως εξελιγμένα, τα δεδομένα αυτά αναλύθηκαν και εξήχθησαν αποτελέσματα που αξιοποιήθηκαν στην εκτίμηση και απόδοση πατρότητας, την ταυτοποίηση γενετικής δομής και την ανακατασκευή γενεαλογίας των υπό μελέτη πληθυσμών.

Οι αναλύσεις αυτές έδωσαν συγκεκριμένα στοιχεία ώστε να είναι εφικτή η σύγκριση μεταξύ των πληθυσμών καθώς και η εκτίμηση των αρχικών πληθυσμών από τους οποίους προήλθαν. Αυτά τα συμπεράσματα μπορούν να αξιοποιηθούν στην περεταίρω μελέτη για τον τρόπο διαχείρισης των εκτρεφόμενων πληθυσμών και των ιχθυοκαλλιεργειών γενικότερα με απώτερο σκοπό τη βελτίωση των γενετικών τους χαρακτηριστικών.

Abstract

Microsatellites are an invaluable tool in the study of population genetics. They are increasingly utilised due to their technical simplicity and ease of analysis. This dissertation utilises data from microsatellites of gene loci for the genetic analysis of cultivated populations of european bass (*Dicentrarchus Labrax*) and gilt-head (sea) bream (*Sparus aurata*). The data were analysed with last generation bioinformatic platforms. The data were used to determine paternity, the identification of genetic structure and the pedigree reconstruction of the studied populations.

The analysed data allowed the comparison between populations and the determination of the initial populations they originated from. The conclusions drawn from this analysis can be utilised to further study the management strategies for the farmed populations studied, in particular, and fisheries in general with the ultimate goal of improving the genetic characteristics of the farmed populations.

Περιεχόμενα

I.	ΕΙΣΑΓΩΓΗ.....	9
1	Εκτρεφόμενα είδη πληθυσμών.....	9
1.1	Λαβράκι (<i>Dicentrarchus Labrax</i>).....	10
1.1.1	Ιστορική και γεωγραφική προέλευση του είδους.....	10
1.1.2	Βιολογία του είδους.....	10
1.1.3	Σύστημα καλλιέργειας.....	11
1.1.4	Αναπαραγωγή σε ιχθυοκαλλιέργεια.....	12
1.1.5	Επεξεργασία φωτοπεριόδου της ιχθυοκαλλιέργειας.....	12
1.1.6	Ορμονική θεραπεία και επεξεργασία.....	12
1.1.7	Εφαρμογές ιχθυοκαλλιέργειας λαβρακιού στο εμπόριο.....	12
1.2	Τσιπούρα (<i>Sparus aurata</i>).....	13
1.2.1	Προέλευση του είδους - ιστορική και γεωγραφική προέλευση.....	13
1.2.2	Βιολογία της τσιπούρας.....	13
1.2.3	Συστήματα εκτροφής τσιπούρας σε ιχθυοκαλλιέργειες.....	14
1.2.4	Εφαρμογές ιχθυοκαλλιέργειας της τσιπούρας στο εμπόριο.....	15
1.3	Ιχθυοκαλλιέργεια στην Ελλάδα – λαβράκι και τσιπούρα.....	16
1.4	Ανάλυση γενετικής δομής εκτρεφόμενων πληθυσμών με μικροδορυφόρους.....	16
2	Δεδομένα που χρησιμοποιήθηκαν.....	17
2.1	Πλεονεκτήματα των μικροδορυφόρων.....	18
2.2	Μειονεκτήματα των μικροδορυφόρων.....	19
2.3	Μηδενικά αλληλόμορφα σε μικροδορυφορικές αναλύσεις.....	19
2.4	Δοκιμές για χρήση βέλτιστων μικροδορυφόρων.....	20
3	Βιοπληροφορική ανάλυση δεδομένων.....	20
3.1	Συγγενικές σχέσεις μελών ενός πληθυσμού.....	20
3.2	Εκτίμηση πατρότητας.....	21
3.3	Ανάλυση ετεροζυγωτίας και στατιστικών στοιχείων.....	22
3.4	Ταυτοποίηση γενετικής δομής και ανακατασκευή γενεαλογίας πληθυσμών.....	22
4	Προγράμματα Βιοπληροφορικής στην γενετική ανάλυση πατρότητας.....	23
4.1	Colony.....	23
4.2	Kinalyzer.....	24
4.3	Pedigree.....	25
4.4	Άλλα προγράμματα βιοπληροφορικής για συγγενικές σχέσεις μελών πληθυσμών.....	25
4.5	Microsatellite Analyzer (MSA).....	25
4.6	Άλλα προγράμματα στατιστικής ανάλυσης δεδομένων πληθυσμών.....	26

4.7	Structure	27
II.	ΥΛΙΚΑ – ΜΕΘΟΔΟΙ.....	29
1.	Δεδομένα που χρησιμοποιήθηκαν	29
2	Συγγενικές σχέσεις μελών ενός πληθυσμού – Εκτίμηση πατρότητας.....	29
2.1	Colony.....	30
2.1.1	Παράμετροι που εισάγονται.....	30
2.1.2	Αποτελέσματα με τη χρήση του Colony	31
2.2	Kinalyzer	32
2.2.1	Η μέθοδος που χρησιμοποιείται	32
2.2.2	Παράμετροι που απαιτούνται	32
2.2.3	Αποτελέσματα που εξάγονται	33
2.3	Pedigree	33
2.3.1	Μέθοδος που χρησιμοποιεί.....	33
2.3.2	Αποτελέσματα ανάλυσης.....	33
3	Ανάλυση ετεροζυγωτίας και στατιστικών στοιχείων	34
3.1	Microsatellite Analyzer (MSA).....	34
3.1.1	Ανάλυση δεδομένων με το MSA	34
3.1.2	Εισαγωγή παραμέτρων	34
3.1.3	Εξαγωγή δεδομένων.....	34
3.1.4	Έλεγχος σφαλμάτων.....	35
4	Ταυτοποίηση γενετικής δομής και ανακατασκευή γενεαλογίας πληθυσμών	35
4.1	Structure software.....	35
4.1.1	Η μέθοδος που χρησιμοποιεί.....	35
4.1.2	Παράμετροι που είναι απαραίτητο να οριστούν.....	36
4.1.3	Αρχείο δεδομένων (data file – input file)	36
4.1.4	Επιλογή μοντέλου γενεαλογίας.....	37
4.1.5	Μοντέλα για την συχνότητα αλληλομόρφων.....	37
4.1.6	Επισκόπηση του Structure software.....	38
III.	ΑΠΟΤΕΛΕΣΜΑΤΑ	39
1	Συγγενικές σχέσεις μελών ενός πληθυσμού	39
1.1	Ανάλυση συγγένειες με τα προγράμματα Colony, Pedigree και Kinalyzer	39
1.2	Σύγκριση των παραπάνω προγραμμάτων	39
2	Γενετικές αποστάσεις μεταξύ γονιδιακών δεξαμενών – ανάλυση ετεροζυγωτίας.....	41
3	Ταυτοποίηση γενετικής δομής – αρχικοί πληθυσμοί γονιδιακών δεξαμενών.....	45
4	Ανακατασκευή γενεαλογίας – γενετικές αποστάσεις αρχικών πληθυσμών.....	48
IV.	ΣΥΜΠΕΡΑΣΜΑΤΑ – ΣΥΖΗΤΗΣΗ.....	51

1	Σύγκριση των προγραμμάτων Colony, Pedigree και Kinalyzer.....	51
2	Δείκτης πολυμορφίας πληθυσμών.....	51
	Βιβλιογραφία.....	52

I. ΕΙΣΑΓΩΓΗ

Δεδομένα από ιχθυοκαλλιέργειες στην Ελλάδα αναλύθηκαν με στόχο τη βελτίωση του τρόπου διαχείρισής τους και την καλύτερη αξιοποίησή τους στο εμπόριο για τη βελτίωση των διατροφικών συνθηκών του Μεσογειακών πληθυσμών. Η κατανάλωση ψαριού από ιχθυοκαλλιέργεια αποτελεί ποιοτικό χαρακτηριστικό της μεσογειακής διατροφής διότι τα εκτρεφόμενα ψάρια, λόγω της ευαισθησίας του είδους, τρέφονται με ιχθυέλαια και φυτικές πρώτες ύλες. Κρίνεται σημαντικός ο ορθότερος τρόπος διαχείρισης των ιχθυοκαλλιεργειών για τη βελτίωση αυτών των ποιοτικών χαρακτηριστικών των ψαριών και κατ' επέκταση της βελτίωση της διατροφής.

Ένας ακόμα σημαντικός λόγος που κρίνεται αναγκαία η βελτίωση της διαχείρισης των ελληνικών ιχθυοκαλλιεργειών, είναι και η πρώτη θέση που κατέχει το ψάρι στα εξαγωγικά προϊόντα της χώρας. Η άνοδος της παραγωγής τους και η βελτίωση των ποιοτικών τους χαρακτηριστικών το καθιστά πιο προσιπό στην εγχώρια αγορά και το διεθνές εμπόριο.

1 Εκτρεφόμενα είδη πληθυσμών

Ο κλάδος της ιχθυοκαλλιέργειας είναι ο πιο δυναμικά αναπτυσσόμενος στην Ελληνική οικονομία την τελευταία δεκαετία. Η ιχθυοκαλλιέργεια ξεκίνησε στην Ελλάδα την δεκαετία του '80, αλλά γνώρισε εξαιρετική άνοδο την δεκαετία του '90. Στην ανάπτυξη αυτή συνέβαλλαν οι κλιματολογικές και γεωμορφολογικές συνθήκες, η μείωση των αποθεμάτων σε ψάρια (κυρίως λόγω νομοθετικών και περιβαλλοντικών περιορισμών), η ανάπτυξη της τεχνογνωσίας εκτροφής ψαριών, καθώς και οι επιχορηγήσεις από την πολιτεία προς αυτή την κατεύθυνση. Επίσης, η χαμηλή τιμή διάθεσης των ψαριών, η τάση των καταναλωτών για πιο υγιεινή διατροφή και η προσπάθεια πιστοποίησης της ποιότητας των ψαριών, η οποία έχει ξεκινήσει, δημιουργούν θετικές προοπτικές για τη ζήτηση των ψαριών ιχθυοκαλλιέργειας. Αξίζει να αναφερθεί πως η τσιπούρα μαζί με το λαβράκι από πλευράς διατροφικής αξίας ανήκουν στα πιο πολύτιμα ψάρια της Μεσογείου, καθώς είναι πλούσια στα λιπαρά οξέα ω-3. Είναι από τα κύρια ψάρια για τις ιχθυοκαλλιέργειες και το πιο εκτροφεύσιμο είδος της Μεσογείου.

Πηγή: <http://www.gaiapedia.gr/gaiapedia/index.php/>

Σαν γεννήτορες στους Ιχθυογεννητικούς σταθμούς (ιχθυοκαλλιέργειες) επιλέγονται ψάρια που έχουν δείξει καλή ανάπτυξη στους ιχθυοκλωβούς και γενικά ανθεκτικότητα σε ασθένειες και καλά οργανοληπτικά χαρακτηριστικά. Τα ψάρια αυτά διατηρούνται σε ελεγχόμενες συνθήκες σε μικρές ομάδες μεικτών πληθυσμών και διατρέφονται πολύ προσεκτικά ενώ γίνεται προσεκτικός έλεγχος της σεξουαλικής αναλογίας. Για την προετοιμασία και επίτευξη της ωορρηξίας και της γονιμοποίησης μέσα στο νερό οι γεννήτορες προετοιμάζονται ειδικά σε δεξαμενές όπου μπορούμε να ελέγξουμε τη θερμοκρασία του θαλασσινού νερού και την φωτοπερίοδο ώστε να μιμηθούμε τις περιβαλλοντικές συνθήκες στις οποίες αυτά τα είδη αναπαράγονται στη φύση. Στη φύση η τσιπούρα διαφέρει από

το λαβράκι στο ότι είναι πρωτο-ανδρικό ερμαφρόδιτο είδος που σημαίνει ότι όλες οι τσιπούρες γεννιούνται και παραμένουν αρσενικά άτομα για τα πρώτα 2-3 χρόνια ενώ γίνονται όλα θηλυκά μετά από αυτό το στάδιο. Η τσιπούρα φτάνει στη φύση τα 50-60 εκατοστά ενώ το λαβράκι το ένα μέτρο. Στην παρούσα εργασία αναλύονται αυτά τα δύο είδη ιχθύων με σκοπό την ανάλυση της γενετικής δομής των πληθυσμών τους και τη διαχείριση τις καλλιέργειας.

Πηγή: <http://www.gaiapedia.gr/gaiapedia/index.php/>

1.1 Λαβράκι (Dicentrarchus Labrax)

1.1.1 Ιστορική και γεωγραφική προέλευση του είδους

Το ένα είδος ψαριού που αξιοποιείται στην παρούσα εργασία είναι το λαβράκι *Dicentrarchus Labrax* το οποίο ιστορικά προέρχεται από καλλιέργεια σε παράκτιες λιμνοθάλασσες και παλιρροιακές δεξαμενές πριν από την σηματοδότηση της έναρξης για ανάπτυξη της μαζικής παραγωγής νεαρών ζώων που ξεκίνησε στα τέλη της δεκαετίας του 1960. Η ιχθυοκαλλιέργεια αρχικά συσχετίστηκε με την παραγωγή αλατιού σε παράκτιες πεδιάδες εξάτμισης και έλη. Το αλάτι συλλέχθηκε κατά τη διάρκεια της υψηλής περιόδου εξάτμισης του καλοκαιριού και του φθινοπώρου και τα ψάρια καλλιεργήθηκαν κατά τη διάρκεια του χειμώνα και της άνοιξης. Η ενίσχυση αυτής της ιχθυοκαλλιέργειας προήλθε από την παγίδευση πληθυσμών ψαριών που ζούσαν σε αυτές τις εκβολές.

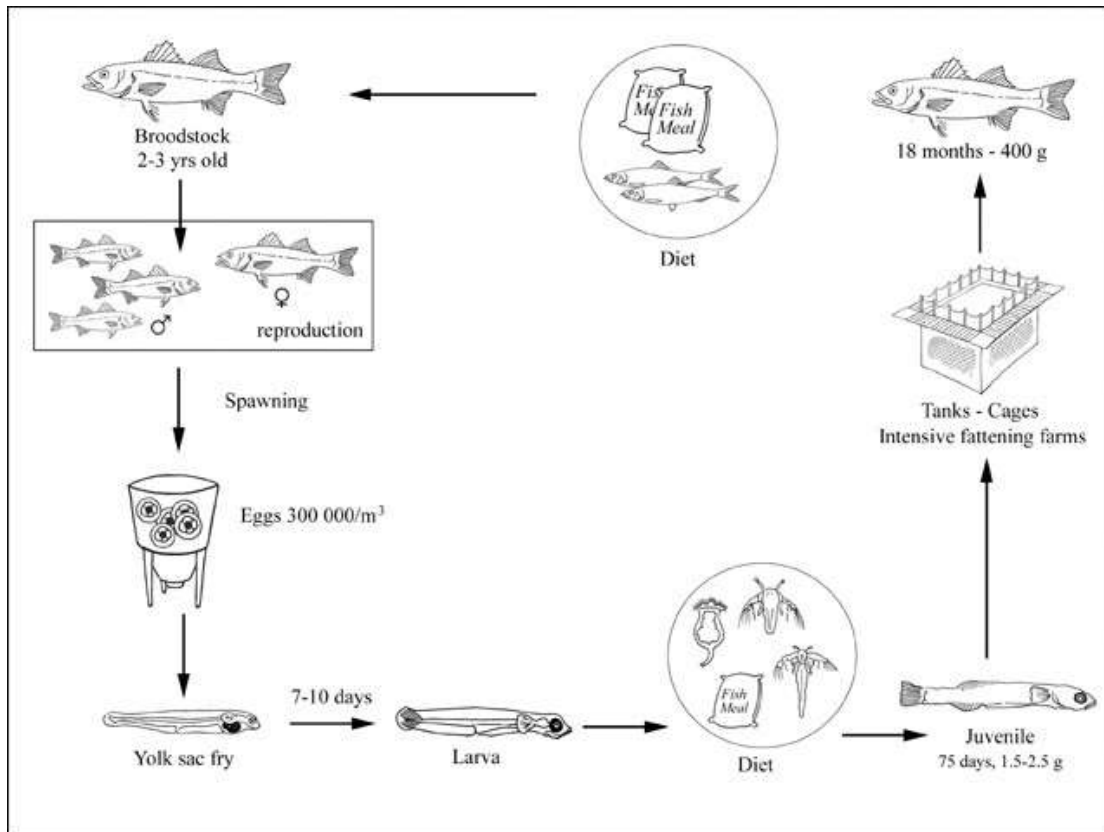
Στα τέλη της δεκαετίας του '60, η Γαλλία και η Ιταλία ανταγωνίστηκαν να αναπτύξουν αξιόπιστες τεχνικές μαζικής παραγωγής για το νεαρό λαβράκι και, μέχρι τα τέλη της δεκαετίας του 1970, οι τεχνικές αυτές αναπτύχθηκαν αρκετά καλά στις περισσότερες μεσογειακές χώρες για να δώσουν εκατοντάδες χιλιάδες προνύμφες ψαριών του συγκεκριμένου είδους. Το ευρωπαϊκό λαβράκι ήταν το πρώτο θαλάσσιο μη σχετικό με σολομό είδος που καλλιεργήθηκε εμπορικά στην Ευρώπη και σήμερα είναι το σημαντικότερο εμπορικό ψάρι που καλλιεργείται ευρέως στις μεσογειακές χώρες. Η Ελλάδα, η Τουρκία, η Ιταλία, η Ισπανία, η Κροατία και η Αίγυπτος αποτελούν χώρες με τη μεγαλύτερη παραγωγή σε λαβράκι.

Πηγή: http://www.fao.org/fishery/culturedspecies/Dicentrarchus_labrax/en#tcN80050

1.1.2 Βιολογία του είδους

Το λαβράκι της Ευρώπης ζει σε περιβάλλοντα με συνθήκες ευρέος φάσματος θερμοκρασιών (5-28 °C) και αλατότητας (3 ‰ σε θαλάσσιο νερό πλήρους αντοχής). Έτσι συχνάζουν σε παράκτια ύδατα και εμφανίζονται σε εκβολές ποταμών και υφάλμυρες λιμνοθάλασσες. Μερικές φορές επιχειρούν να μεταναστεύσουν σε γλυκά ύδατα. Υπάρχει μόνο μία εποχή αναπαραγωγής ετησίως, η οποία πραγματοποιείται το χειμώνα στον πληθυσμό της Μεσογείου (Δεκέμβριος-Μάρτιος) και μέχρι τον Ιούνιο στους πληθυσμούς του Ατλαντικού. Τα λαβράκια δεν είναι ιδιαίτερα ευαίσθητα σε χαμηλές θερμοκρασίες, κι έτσι ορισμένα

ψάρια μπορεί να υπερβούν το χειμώνα σε παράκτιες λιμνοθάλασσες αντί να επιστρέψουν στην ανοιχτή θάλασσα. Είναι αρπακτικά ζώα και το εύρος της διατροφής τους περιλαμβάνει μικρά ψάρια, γαρίδες, καβούρια και σουπιές.



Κύκλος παραγωγής εκτρεφόμενων ψαριών του είδους *Dicentrarchus labrax* σε ιχθυοκαλλιέργεια
 Πηγή: <http://www.fao.org/fi/figis/culturespecies/data/assets/images/dicentrarchus/labrax-int-prdcycle.jpg>

Παρόλο που τα λαβράκια εκτρέφονται σε λίμνες και λιμνοθάλασσες θαλάσσιου ύδατος, ο μεγάλος όγκος της παραγωγής τους προέρχεται από τις ιχθυοκαλλιέργειες, εκκολαπτήρια ψαριών δηλαδή με εγκατεστημένους σταθμούς εκτροφής του είδους με σκοπό την αναπαραγωγή του.

1.1.3 Σύστημα καλλιέργειας

Πιο συγκεκριμένα, προκειμένου να εξασφαλιστεί μια αξιόπιστη και επαρκής προμήθεια αυτών ψαριών καλής ποιότητας, τα περισσότερα εκκολαπτήρια έχουν δημιουργήσει τις δικές τους μονάδες γενετικού υλικού, όπου τα εκτρεφόμενα ψάρια διαφορετικών ηλικιακών ομάδων διατηρούνται μακροπρόθεσμα. Τα άτομα που χρησιμοποιούνται ως γονείς μπορεί να προέρχονται είτε από άλλη ιχθυοκαλλιέργεια – εκκολαπτήριο, είτε από τη φύση. Η βέλτιστη ηλικία για τα θηλυκά μητρικά ψάρια είναι μεταξύ 5 και 8 ετών, ενώ για τα αρσενικά αυτή η περίοδος μειώνεται σε 2 με 4 χρόνια. Η διαχείριση της ιχθυοκαλλιέργειας περιλαμβάνει φυσική ωρίμανση, επαγωγή ωορρηξίας με χειρισμό φωτοπεριόδου

ή ορμονικές θεραπείες, γονιμοποίηση σε δεξαμενές αναπαραγωγής και επώαση σε σύστημα κυκλοφορίας ανοιχτού νερού.

1.1.4 Αναπαραγωγή σε ιχθυοκαλλιέργεια

Κατά την έναρξη της αναπαραγωγής είναι απαραίτητο να μετακινηθούν επιλεγμένες παρτίδες εκτρεφόμενων ψαριών από τις εγκαταστάσεις μακροχρόνιας φύλαξης τους στις δεξαμενές αναπαραγωγής, όπου μπορούν να υποστούν καλύτερη επεξεργασία και η απόδοσή τους να μπορεί να παρακολουθείται εύκολα. Η αναλογία αρσενικών:θηλυκών στις δεξαμενές ωοτοκίας διατηρείται στο 2:1. Ενώ τα ώριμα αρσενικά επιλέγονται όταν απελευθερώνουν σπέρμα αυθόρμητα ή κατά την έκδυση του σώματος, το στάδιο ωρίμανσης του θηλυκού πρέπει να εξακριβώνεται με την εξαγωγή ωοκυττάρων από την ωοθήκη με τη χρήση καθετήρα: μόνο θηλυκά με ωοκύτταρα στο τελικό **vitellogenic** στάδιο, δηλαδή επιλέγονται ωοκύτταρα με διάμετρο μεγαλύτερη από τα 650 μm.

1.1.5 Επεξεργασία φωτοπεριόδου της ιχθυοκαλλιέργειας

Όταν απαιτούνται γονιμοποιημένα αυγά εκτός της φυσικής περιόδου αναπαραγωγής τους, η ωρίμανσή τους εκτός εποχής επιτυγχάνεται με την προώθηση της γαμετογένεσης με χειρισμό της φωτοπεριόδου και της θερμοκρασίας. Η διαχείριση του εκκολαπτηρίου κάθε φορά χειρίζεται αυτές τις συνθήκες για τις περιόδους παραγωγής αυγών σύμφωνα με τις ανάγκες της στο εμπόριο ή / και τις απαιτήσεις της ιχθυοκαλλιέργειας.

1.1.6 Ορμονική θεραπεία και επεξεργασία

Η ορμονική θεραπεία χρησιμοποιείται για να ενεργοποιήσει την τελευταία φάση ωρίμανσης αυγών. Η ανθρώπινη χοριακή γοναδοτροπίνη (HCG) χρησιμοποιείται σε δόση 800-1000 IU ανά kg/bw, χορηγούμενη σε δύο ενέσεις στους ραχιαίους μυς, σε διάστημα 6 ωρών.

1.1.7 Εφαρμογές ιχθυοκαλλιέργειας λαβρακιού στο εμπόριο

Μία από τις μεγαλύτερες επιτυχίες στις ευρωπαϊκές υδατοκαλλιέργειες ήταν η μεσογειακή βιομηχανία του είδους λαβράκι, η οποία σε λιγότερο από 15 χρόνια αυξήθηκε από μερικές χιλιάδες τόνους σε 57.000 τόνους σήμερα, έχοντας κορυφωθεί σε περίπου 71.000 τόνους το 2000. Όταν άρχισε η καλλιέργεια λαβρακιού στην αγορά στα τέλη της δεκαετίας του 1980 και στις αρχές της δεκαετίας του 1990, η καλλιεργημένη ποιότητα θεωρήθηκε ότι συμπληρώνει αυτή του αγρίου τύπου και οι τιμές ήταν πολύ υψηλές. Οι τιμές του αγρίου τύπου μπορεί να έχουν υποτιμηθεί αρχικά, καθώς οι όγκοι παραγωγής από την ιχθυοκαλλιέργεια συνέχισαν να αναπτύσσονται, αλλά σήμερα υπάρχει σαφής διάκριση στην αγορά μεταξύ αγρίου και εκτρεφόμενου προϊόντος, ενώ οι τιμές για άγρια λαβράκι είναι αρκετές φορές υψηλότερες από αυτές των εκτρεφόμενων ψαριών.

Σε σύγκριση με πολλά άλλα είδη εκτρεφόμενων ψαριών, όπως ο σολομός ή η πέστροφα, μέχρι στιγμής το λαβράκι έχει διατεθεί κυρίως στο εμπόριο

ολόκληρο και φρέσκο. Σε κάθε περίπτωση, η ανάπτυξη προϊόντων στο λαβράκι ήταν πολύ περιορισμένη. Ένας σημαντικός λόγος είναι ο συντηρητικός καταναλωτισμός των μεσογειακών πληθυσμών, οι οποίοι συνηθίζουν να βλέπουν το σύνολο των ψαριών όταν πωλούνται στο λιανικό εμπόριο, παρά το γεγονός ότι τα ψάρια θα ήταν βέβαια καλύτερα αν είχαν εκτραφεί και αναπαραχθεί στην πηγή.

Έχει ήδη ξεκινήσει κάποια ανάπτυξη προϊόντων, τόσο μεταξύ των μεγαλύτερων ελληνικών παραγωγών όσο και από εξειδικευμένους Ιταλούς μεταποιητές ψαριών που εισάγουν ελληνικό προϊόν και στη συνέχεια το συσκευάζουν σε συσκευασία τροποποιημένης ατμόσφαιρας (MAP), προσδίδοντας στο προϊόν μακρύτερη διάρκεια ζωής. Ωστόσο, είναι απαραίτητη η ανάπτυξη περισσότερων προϊόντων, αν και οι επιπρόσθετες ποσότητες λαβρακιού απορροφούνται στις σημερινές αγορές.

Πηγή: http://www.fao.org/fishery/culturedspecies/Dicentrarchus_labrax/en#tcN80050

1.2 Τσιπούρα (*Sparus aurata*)

1.2.1 Προέλευση του είδους - ιστορική και γεωγραφική προέλευση

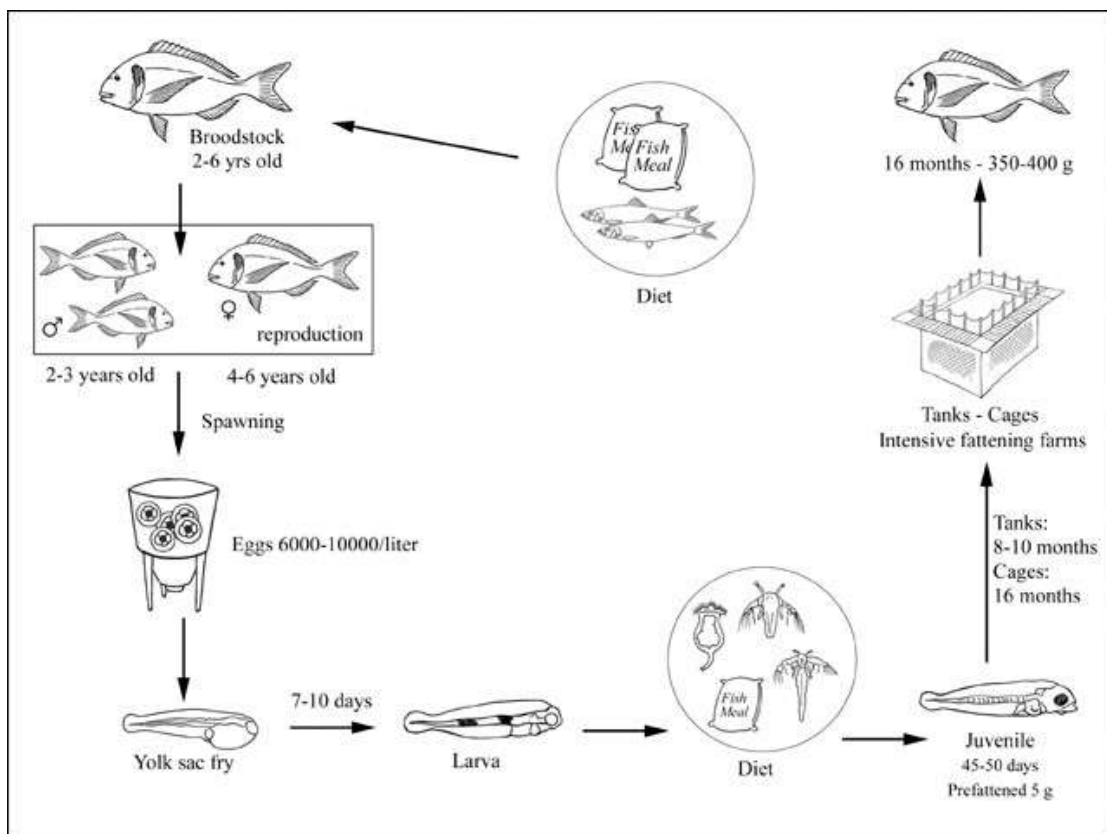
Παραδοσιακά, η τσιπούρα καλλιεργούνταν εκτενώς σε παράκτιες λιμνοθάλασσες και λίμνες αλμυρού νερού, έως ότου αναπτύχθηκαν συστήματα εντατικής εκτροφής της κατά τη διάρκεια της δεκαετίας του 1980. Η ιταλική «vallicoltura» ή η αιγυπτιακή «hosh» είναι εκτεταμένα συστήματα εκτροφής ψαριών που λειτουργούν σαν φυσικές παγίδες ιχθύων, εκμεταλλεζόμενες τη φυσική τροφική μετανάστευση νεαρών ψαριών από τη θάλασσα σε παράκτιες λιμνοθάλασσες. Τα ψάρια Gilthead (είδος τσιπούρας) είναι κατάλληλα για μεγάλη υδατοκαλλιέργεια στη Μεσόγειο, λόγω της καλής τιμής αγοράς, του υψηλού ποσοστού επιβίωσης και των διατροφικών συνηθειών, που είναι σχετικά χαμηλά στην τροφική αλυσίδα.

Η τεχνητή αναπαραγωγή της τσιπούρας σε συστήματα ιχθυοκαλλιέργειας επιτεύχθηκε με επιτυχία στην Ιταλία το 1981-1982 και η παραγωγή μεγάλης κλίμακας νεαρών ψαριών τσιπούρας επιτεύχθηκε οριστικά το 1988-1989 στην Ισπανία, την Ιταλία και την Ελλάδα. Η παραγωγή και η εκτροφή αυτού του είδους είναι μια από τις επιτυχίες της επιχείρησης υδατοκαλλιέργειας. Αυτό το είδος έδειξε πολύ γρήγορα μεγάλη προσαρμοστικότητα στις συνθήκες εντατικής εκτροφής, τόσο στις λίμνες όσο και στα τεχνητά συστήματα ιχθυοκαλλιέργειας, ενώ η ετήσια παραγωγή του αυξήθηκε μέχρι το 2000, όταν έφθασε στο ανώτατο όριο των 87.000 τόνων.

1.2.2 Βιολογία της τσιπούρας

Το είδος *Sparus aurata* είναι ευρέως γνωστό στη Μεσόγειο Θάλασσα και σπάνια στη Μαύρη Θάλασσα. Λόγω του ευρέος φάσματος συνθηκών θερμοκρασίας και αλατότητας, το είδος απαντάται τόσο σε θαλάσσια όσο και σε υφάλμυρα ύδατα, όπως οι παράκτιες λιμνοθάλασσες και οι εκβολές των ποταμών,

ιδίως κατά τα αρχικά στάδια του κύκλου ζωής του. Γεννημένοι στην ανοιχτή θάλασσα κατά τους μήνες Οκτώβριο – Δεκέμβριο, τα νεαρά ψάρια μεταναστεύουν συνήθως νωρίς την άνοιξη προς τα προστατευόμενα παράκτια ύδατα, όπου μπορούν να βρουν άφθονες πηγές για τροφή και πιο ήπιες θερμοκρασίες. Η τσιπούρα είναι πολύ ευαίσθητη στις χαμηλές θερμοκρασίες (χαμηλότερο και θανατηφόρο όριο είναι 4°C), και για αυτό το λόγο στα τέλη του φθινοπώρου επιστρέφει στην ανοιχτή θάλασσα, όπου εκτρέφονται τα ενήλικα πλέον ψάρια. Τα νεαρά ψάρια παραμένουν σε σχετικά αβαθείς περιοχές (έως 30μ.), ενώ τα ενήλικα μπορούν να φτάσουν σε βαθύτερα νερά, γενικά όχι πάνω από 50μ. Αυτό το είδος είναι αρχικά ερμαφρόδιτο. Η αναπαραγωγική ωρίμανση αναπτύσσεται στα αρσενικά σε ηλικία 2 ετών (20-30 cm) και στα θηλυκά σε 2-3 χρόνια (33-40 cm). Τα θηλυκά είναι γεννήτορες που μπορούν να παράγουν 20.000-80.000 αυγά κάθε μέρα για περίοδο έως και 4 μηνών. Στην ιχθυοκαλλιέργεια, η αντιστροφή του φύλου εξαρτάται από κοινωνικούς και ορμονικούς παράγοντες.



Κύκλος παραγωγής εκτρεφόμενων ψαριών του είδους *Sparus aurata* σε ιχθυοκαλλιέργεια

Πηγή: <http://www.fao.org/fi/figis/culturespecies/data/assets/images/sparus/aurata-prdcycle2.jpg>

1.2.3 Συστήματα εκτροφής τσιπούρας σε ιχθυοκαλλιέργειες

Συνήθως, κάθε εκκολαπτήριο έχει τη δική του μονάδα αναπαραγωγής, όπου τα εκτρεφόμενα ψάρια διαφόρων ηλικιακών ομάδων, από 1 έτους αρσενικά έως 5 ετών θηλυκά, διατηρούνται υπό μακροπρόθεσμες συνθήκες αποθέματος.

Τα εκτρεφόμενα ψάρια γονείς μπορούν να προέλθουν είτε από μια ιχθυοκαλλιέργεια, είτε από την φύση.

Στην αρχή της αναπαραγωγής, επιλεγμένες παρτίδες εκτρεφόμενων ψαριών μεταφέρονται από τη μακρινή τοποθεσία τους στις δεξαμενές αναπαραγωγής. Ο έλεγχος του ποσοστού διακύμανσης του φύλου στις δεξαμενές αυτές είναι ένας πολύ σημαντικός παράγοντας για την τσιπούρα και πρέπει να ληφθούν οι απαραίτητες προφυλάξεις επειδή η αντιστροφή του φύλου είναι κοινωνικά καθορισμένη. Η παρουσία νεαρών αρσενικών στο τέλος της περιόδου ωοτοκίας, για παράδειγμα, αυξάνει τον αριθμό των ηλικιωμένων ψαριών που γίνονται θηλυκά. Από την άλλη πλευρά, η εμφάνιση ηλικιωμένων θηλυκών μειώνει την αντιστροφή του φύλου σε νεαρότερα ψάρια.

Το γενετικό υλικό της τσιπούρας μπορεί να επηρεαστεί από περιβαλλοντικό χειρισμό, προκειμένου να επεκταθεί ή να τροποποιηθεί ο χρόνος αναπαραγωγής. Τα ψάρια αποθηκεύονται σε δεξαμενές εξοπλισμένες με σύστημα θέρμανσης/ψύξης του νερού και έλεγχο από ηλεκτρονικά συστήματα της θερμοκρασίας και της έντασης του φωτός. Η αναπαραγωγική ωρίμανση επιτυγχάνεται με την έκθεση του εκτρεφόμενου πληθυσμού σε συνθήκες φωτοπεριόδου και θερμοκρασίας νερού που συμβαίνουν κατά τη διάρκεια της φυσικής περιόδου ωοτοκίας του (όπως δηλαδή στα αγρίου τύπου ψάρια).

Υπάρχουν δύο βασικά συστήματα της εκτροφής των προνυμφών της τσιπούρας που ονομάζονται μικρής και μεγάλης κλίμακας. Το σύστημα εκτροφής μικρής κλίμακας (<10 m³) χαρακτηρίζεται από τον μέγιστο έλεγχο των περιβαλλοντικών παραμέτρων και σχεδιάζεται για να παράγει ένα μεγάλο αριθμό νεαρών ψαριών (150-250 ψάρια/λίτρο). Η τεχνική μεγάλης κλίμακας (~200 m³) προσομοιώνει ένα φυσικό οικοσύστημα. Αυτή η τεχνική εγγυάται πολύ καλύτερη ποιότητα νεαρής τσιπούρας από το σύστημα μικρής κλίμακας, αλλά παράγει πολύ λιγότερα ψάρια (το μέγιστο 10 ψάρια/λίτρο).

1.2.4 Εφαρμογές ιχθυοκαλλιέργειας της τσιπούρας στο εμπόριο

Όπως συμβαίνει στην περίπτωση γενικά της ευρωπαϊκής τσιπούρας, η εκτροφή νεαρών ειδών τσιπούρας ειδικά στην περιοχή της Μεσογείου υφίσταται μετασχηματισμό από την ύπαρξη βιομηχανίας υψηλών ορίων και χαμηλών όγκων σε ένα σύστημα με χαμηλά όρια και υψηλούς όγκους. Η ταχεία ανάπτυξη της παραγωγής σε ιχθυοκαλλιέργειες οδήγησε σε πτώση των τιμών. Οι τιμές των πρώην εκτρεφόμενων ειδών μειώθηκε κατά περίπου 60% μεταξύ του 1990 και του 2000 και εξακολουθούν να μειώνονται μέχρι και σήμερα.

Παρόλο που οι τιμές των πρώην εκτρεφόμενων πληθυσμών τσιπούρας άρχισαν να μειώνονται μεταξύ του 1990 και του 1995, εξακολουθούσαν να είναι στο επίκεντρο του ενδιαφέροντος πολλών επενδυτών και την εξασφάλιζαν ένα εύλογο όφελος για τους αγρότες μέχρι το 1998. Ωστόσο, κατά τα έτη 2000-2003 μέχρι και σήμερα η τιμή της εκτρεφόμενης τσιπούρας έχει καταρρεύσει. Τώρα η

ευρωπαϊκή τιμή αγοράς κυμαίνεται γύρω στα 5,50 ευρώ/kg, για το μέγεθος των 350gr. Σε αυτό το επίπεδο είναι πολύ δύσκολο για τους ιχθυοτρόφους να βγάλουν ένα ικανοποιητικό κέρδος.

Έτσι, αυτή τη στιγμή, οι συνθήκες της αγοράς φαίνονται πολύ μακριά από εκείνες που εμφανίστηκαν κατά το πρώτο μισό της δεκαετίας του 1990, αλλά υπάρχουν μερικές στρατηγικές στο εμπόριο για την κερδοφόρα εκτροφή της τσιπούρας. Ένα από αυτά είναι οι οικονομίες κλίμακας (καλλιέργεια πολλών ψαριών για τη μείωση του μοναδιαίου κόστους παραγωγής). Εναλλακτικά, τα μικρά συστήματα παραγωγής μπορούν να αυξήσουν την αξία του προϊόντος με την παραγωγή μικρών ποσοτήτων ψαριών υψηλότερης ποιότητας (π.χ. οργανικά ψάρια) ή με την παραγωγή μη συμβατικών μεγεθών ψαριών, τα οποία μπορούν επίσης να υποστούν επεξεργασία.

Πηγή: http://www.fao.org/fishery/culturedspecies/Sparus_aurata/en

1.3 Ιχθυοκαλλιέργεια στην Ελλάδα – λαβράκι και τσιπούρα

Η θαλάσσια ιχθυοκαλλιέργεια αποτελεί την κύρια δραστηριότητα υδατοκαλλιέργειας της Ελλάδας και μετράει πάνω από 30 χρόνια εμπειρίας. Από το 1981 που δημιουργήθηκαν οι πρώτες πειραματικές μονάδες ο κλάδος έφτασε το 2015 να παράγει 110.000 τόνους τσιπούρας (*Sparus aurata*) και λαβρακιού (*Dicentrarchus labrax*) κατέχοντας μια από τις πρώτες θέσεις παγκοσμίως στη μεσογειακή ιχθυοκαλλιέργεια. Αυτά τα δύο βασικά είδη αντιστοιχούν στο 98% του όγκου παραγωγής.

Η βιολογική ιχθυοκαλλιέργεια (εκτρεφόμενα είδη) αποτελεί ένα μικρό ποσοστό του συνολικού όγκου παραγωγής της χώρας. Το 2015 εκτιμάται πως η εκτροφή βιολογικής τσιπούρας και λαβρακιού κυμάνθηκε στους 1.000 τόνους αντιπροσωπεύοντας σχεδόν το 1% της συνολικής παραγωγής. Ωστόσο πρέπει να σημειωθεί πως από τους 1.000 περίπου τόνους, μόνο οι 330,3 τόνοι διατέθηκαν στην αγορά ως βιολογικό ψάρι (0,03 της συνολικής παραγωγής ιχθυοκαλλιέργειας), ενώ οι υπόλοιποι διατέθηκαν ως συμβατικό. Σε σχέση με το 2014 παρατηρείται συνολικά μείωση 13,3% στις πωλήσεις τους (12,4% στην τσιπούρα και 14,6% στο λαβράκι).

Πηγή: https://www.fgm.com.gr/uploads/file/FGM_2016.pdf

1.4 Ανάλυση γενετικής δομής εκτρεφόμενων πληθυσμών με μικροδορυφόρους

Η γενετική δομή των εκτρεφόμενων ζώων είναι ιδιαίτερα σημαντική, επειδή είναι αυτή που καθορίζει πρωταρχικά τις παραγωγικές δυνατότητές τους. Είναι πλέον γενικά αποδεκτό, ότι όσοι ασχολούνται με υδατοκαλλιέργειες πρέπει να εισάγουν τις βασικές έννοιες της γενετικής βελτίωσης στη διαχείριση της παραγωγικής μονάδας τους. Οι διαχειριστές πρέπει να γνωρίζουν ότι η παραγωγικότητα εξαρτάται, σε μεγάλο βαθμό, από την πρακτική διασταυρώσεων και από το πώς γίνεται η γενετική διαχείριση του πληθυσμού. Σήμερα, στις

ιχθυοκαλλιέργειες (του εξωτερικού, αλλά και στην Ελλάδα) εφαρμόζονται διάφορες μέθοδοι γενετικής, με σκοπό να βελτιώσουν ποικίλες βιολογικές παραμέτρους και να επηρεάσουν θετικά το τελικό αποτέλεσμα, είτε αυτό προσμετράται σε ποιοτικά προϊόντα, είτε σε οικονομική απόδοση. Μια μέθοδος γενετικής ανάλυσης αποτελεί η χρήση μικροδορυφορικού DNA.

Από το 1986 περίπου, η προσοχή των ερευνητών στράφηκε στον αριθμό των επαναλαμβανόμενων αντιγράφων σε μια ακολουθία του πυρηνικού DNA. Αυτές οι αλληλουχίες (δορυφορικές) μπορούν να καταταχθούν με βάση το μέγεθος των επαναλήψεων τους σε μινι- και μικροδορυφορικές. Οι πρώτες χαρακτηρίζονται από επαναλαμβανόμενες ακολουθίες DNA σε μήκος 6-100 ζευγάρια βάσεων, οι οποίες επαναλαμβάνονται διαδοχικά από δύο μέχρι εκατοντάδες φορές. Οι δεύτερες αποτελούνται από μονάδες μήκους από 1 έως 5 ζευγάρια βάσεων, που επαναλαμβάνονται από 5-100 φορές. Αυτοί οι γενετικοί δείκτες είναι ουδέτεροι στη δράση της φυσικής επιλογής, ακολουθούν συνεπικρατή Μεντελικό τρόπο κληρονομίσης, έχουν μεγάλο αριθμό αλληλομόρφων και υψηλό βαθμό ετεροζυγωτίας. Ανάλογα με τους στόχους της έρευνας, μπορεί να επιλεγούν μικροδορυφορικοί δείκτες που να παρουσιάζουν υψηλό ή χαμηλό βαθμό πολυμορφισμού. Η μεγάλη γενετική ποικιλότητα του μικροδορυφορικού DNA, καθώς και οι μικρές ποσότητες ιστού που απαιτούνται λόγω της ενίσχυσης με τη βοήθεια της PCR, το κατέστησαν ικανό να αντικαταστήσει, σε μεγάλο βαθμό, τη μέθοδο των αλλοενζύμων και τις αναλύσεις του μιτοχονδριακού DNA. Χρησιμοποιώντας πολλαπλό PCR και αυτόματα μηχανήματα είναι δυνατό να αναλυθεί η γενετική σύσταση μέχρι και 40.000 δειγμάτων ετησίως.

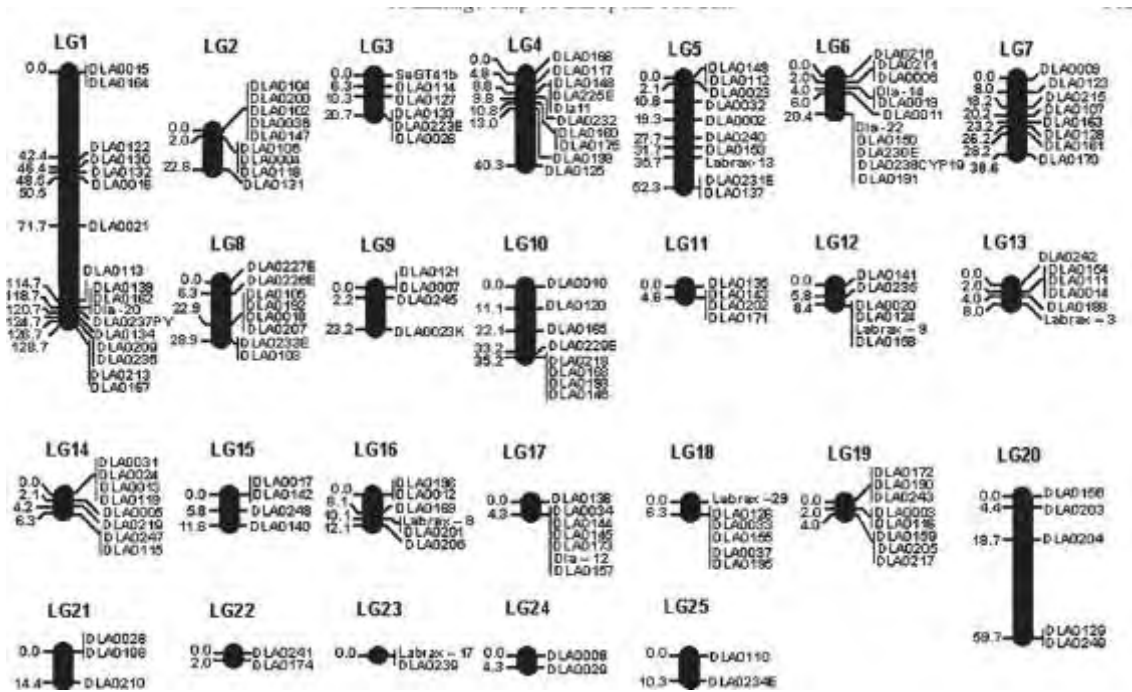
2 Δεδομένα που χρησιμοποιήθηκαν

Στην παρούσα εργασία χρησιμοποιούνται κυρίως ως δεδομένα των αναλύσεων πληροφορίες για γενετικούς τόπους στο είδος λαβράκι και στο είδος τσιπούρα, με τη χρήση μικροδορυφόρων. Οι μικροδορυφόροι είναι αμφιδρομες επαναλήψεις γενετικών στοιχείων (2 – 6 ζεύγη βάσεων) οι οποίες εκφράζουν τις διαφορές των αλληλομόρφων με το διαφορετικό αριθμό επαναλήψεων σε κάθε γενετικό τόπο (Webster and Reichart 2005).

Οι μικροδορυφορικοί δείκτες επιλέγονται ως μοριακοί δείκτες σε αναλύσεις προέλευσης και εύρεσης συγγένειας ειδών των ζώων ειδικά όταν απουσιάζουν από τα δεδομένα οι πληροφορίες γενεαλογίας ανάμεσα στα άτομα που είναι απαραίτητο να εξεταστούν, ενώ ταυτόχρονα παρουσιάζουν περισσότερα πλεονεκτήματα έναντι άλλων σχετικών μεθόδων και αναλύσεων.

Η σημασία της χρήσης μικροδορυφόρων αποδείχθηκε μέσω των ακόλουθων εφαρμογών που έχει αυτή η τεχνική: (i) χαρακτηρισμός και διαφοροποίηση αγρίου τύπου και εκτρεφόμενων πληθυσμών, (ii) ανακατασκευή

γενεαλογίας και εκτίμηση του δραστικού μεγέθους σε έναν εκτρεφόμενο πληθυσμό, (iii) ικανότητα ανακατασκευής γενεαλογίας σε διαφορετικές καταστάσεις και (iv) τον προσδιορισμό των γενετικών σχέσεων ελλείψει δεδομένων από τους γονείς.



Μικροδορυφόροι στο λαβράκι οργανωμένοι σε ομάδες σύνδεσης (*linkage group*) όπως εντοπίζονται στο γονιδίωμα

2.1 Πλεονεκτήματα των μικροδορυφόρων

Αρχικά, βασικότερο πλεονέκτημα των μικροδορυφόρων σε αναλύσεις που εξετάζονται πολλοί γενετικοί τόποι αποτελεί το γεγονός ότι είναι συνεπικρατείς δείκτες, δηλαδή παρουσιάζουν σαφή διάκριση των φαινοτυπικών τους χαρακτηριστικών ανάμεσα σε ετερόζυγα και ομόζυγα άτομα. Αυτό επιτρέπει τη δυνατότητα πιο ακριβή προσδιορισμού των γονοτύπων αλλά και πιο ακριβείς συγκρίσεις ανάμεσα στα άτομα.

Επιπλέον, οι μικροδορυφορικοί τόποι διαθέτουν υψηλά ποσοστά μεταλλάξεων και ως εκ τούτου σε ένα γενετικό τόπο εμφανίζονται περισσότερα του ενός αλληλόμορφα. Αυτό έχει αποτέλεσμα σε εξαιρετικά ισχυρές αναλύσεις συγγένειας, διότι άτομα που δεν έχουν συγγένεια μεταξύ τους, είναι αδύνατο να μοιράζονται κοινά αλληλόμορφα μεταξύ τους.

Τέλος, επειδή οι αναλύσεις των μικροδορυφόρων βασίζονται στην αλυσιδωτή αντίδραση πολυμεράσης (PCR), απαιτούνται πολύ μικρές ποσότητες DNA ή ακόμα μπορεί να χρησιμοποιηθεί αποικοδομημένο DNA. Το γεγονός αυτό επιτρέπει τη χρήση DNA από μη παραδοσιακές και συνήθεις πηγές (Webster and Reichart 2005).

Από την άλλη πλευρά, για να αντισταθμίζονται τα πλεονεκτήματα, οι μικροδορυφορικοί τόποι έχουν και ορισμένα μειονεκτήματα.

2.2 Μειονεκτήματα των μικροδορυφόρων

Αρχικά, οι εκκινητές που χρησιμοποιούνται για την ενίσχυση των μικροδορυφόρων είναι ιδιαίτερα ειδο-ειδικοί, με αποτέλεσμα να απαιτείται να απομονώνονται από το εκάστοτε είδος που μελετάται εκείνη τη στιγμή, διαδικασία επίπονη ειδικά για τις οικογένειες ειδών με μικρή συχνότητα μικροδορυφόρων. Αυτό το πρόβλημα βέβαια αρχίζει και λύνεται διότι αυξάνεται ο αριθμός των ειδών που διαθέτουν μικροδορυφόρους, αλλά και είναι πλέον διαθέσιμες ορισμένες τεχνικές που διευκολύνουν τη διαδικασία απομόνωσης μικροδορυφόρων.

Ένα επιπλέον μειονέκτημα που συναντά η χρήση μικροδορυφόρων είναι ότι βασίζονται σε αναλύσεις PCR, πράγμα που σημαίνει πως μεταλλάξεις σε περιοχές του εκκινητή μπορεί να μην ενισχύσουν μηδενικά αλληλόμορφα (null alleles). Αυτό μπορεί να δημιουργήσει προβλήματα στην μελέτη γονικής προέλευσης διότι ένας γονέας με τον απόγονό του που μοιράζονται ένα μηδενικό αλληλόμορφο, θα εμφανιστούν ως αναντιστοιχία συγγένειας. Ωστόσο είναι πλέον διαθέσιμες οι τεχνικές για την εκτίμηση της συχνότητας των μηδενικών αλληλομόρφων, αλλά δεν είναι πολύ ευαίσθητες (Webster and Reichart 2005).

2.3 Μηδενικά αλληλόμορφα σε μικροδορυφορικές αναλύσεις

Στα δεδομένα που χρησιμοποιήσαμε είχαμε σε κάποιους πληθυσμούς μηδενικά αλληλόμορφα στις αλληλομορφικές αλληλουχίες ορισμένων ατόμων. Πιθανές αιτίες αυτών των μηδενικών αλληλομόρφων είναι ότι οι μικροδορυφόροι ανιχνεύονται μετά από αντίδραση PCR στις περιοχές εκκινητών και συγκεκριμένα χαρακτηρίζονται από δι-, τρι-, τετρανουκλεοτιδικά μοτίβα σε αυτές. Κάθε εκκινητής έχει αλληλουχία συμπληρωματική προς την αλληλουχία που ενισχύεται με PCR, ωστόσο ορισμένα αλληλόμορφα στο ίδιο είδος αλλά και σε διαφορετικά είδη μπορεί να διαφέρουν σε διαφορετικό βαθμό στις περιοχές εκκινητών. Αυτό έχει ως συνέπεια να μην είναι επιτυχής η ενίσχυση με PCR χρησιμοποιώντας αυτούς τους εκκινητές. Επομένως, στη συγκεκριμένη περίπτωση παρουσιάζεται απόκλιση της αλληλουχίας του εκκινητή ή και των δυο εκκινητών της PCR με την αλληλουχία που ενισχύεται (πχ μεταλλάξεις).

Επίσης τα μηδενικά αλληλόμορφα μπορεί να δημιουργηθούν από την διαφορετική ενίσχυση ποικίλων μεγεθών αλληλομόρφων. Πιο συγκεκριμένα, λόγω της ανταγωνιστικής φύσης της αντίδρασης PCR, τα μικρά τμήματα ενισχύονται πιο αποτελεσματικά από τα μεγαλύτερου μεγέθους, έτσι ώστε μόνο το μικρότερο από τα δυο αλληλόμορφα να ανιχνεύεται στα ετερόζυγα άτομα. Αυτά τα μηδενικά αλληλόμορφα ονομάζονται «μερικώς μηδενικά» επειδή μπορεί να γίνουν ορατά στην ανίχνευση με PCR αν για παράδειγμα, φορτώσουμε περισσότερη ποσότητα δείγματος.

Μια τρίτη πηγή μηδενικών αλληλομόρφων είναι η αποτυχία της αντίδρασης PCR εξαιτίας τη ποιότητας του δείγματος του αποικοδομημένου DNA ή της χαμηλής ποιότητας του DNA του δείγματος. Αυτό δημιουργεί προβλήματα διότι σε μερικές περιπτώσεις ένας ή περισσότεροι γενετικοί τόποι (ή αλληλόμορφα) δεν ανιχνεύονται, τη στιγμή που άλλοι ανιχνεύονται με περισσότερη ευκολία, ακολουθώντας την ίδια διαδικασία παρασκευής δείγματος DNA για PCR. Όταν το τμήμα DNA σε ένα γενετικό τόπο είναι φτωχό σε ορισμένα δείγματα (ή άτομα πληθυσμού) και όχι σε κάποια άλλα, τότε τα πρώτα εμφανίζονται ως «ομόζυγα» και όχι «ετερόζυγα» για τα μηδενικά αλληλόμορφα (Dakin and Avise 2004).

Όλες αυτές οι πιθανές αιτίες παρουσίας μηδενικών αλληλομόρφων που αναφέρθηκαν παραπάνω είναι πολύ πιθανό να οδηγήσουν σε επιζήμια λάθη όπως ψευδείς πληροφορίες για την πατρική και μητρική προέλευση των ατόμων που εξετάζονται σε έναν πληθυσμό. Αυτές οι συνέπειες μπορεί να αποφευχθούν με την χρήση στρατηγικών που εξετάζουν την πιθανή παρουσία μηδενικών αλληλομόρφων.

2.4 Δοκιμές για χρήση βέλτιστων μικροδορυφόρων

Καθίσταται πλέον δυνατή η χρήση εκκινητών (μικροδορυφορικές αλληλουχίες) που απομονώθηκαν από διαφορετικά είδη οργανισμών. Η πιθανότητα ωστόσο ένα συγκεκριμένο ζεύγος εκκινητών να λειτουργήσει σε ένα άλλο είδος είναι αρκετά μικρή, αν και ορισμένοι μικροδορυφόροι φέρουν υψηλά συντηρημένες περιοχές κοντά στους εκκινητές (primers). Λεπτομερείς συγκρίσεις και αναλύσεις έδειξαν ότι η πιθανότητα αυτή αυξάνεται σημαντικά όταν το είδος από το οποίο θα απομονώσουμε τους εκκινητές είναι αρκετά συγγενικό με αυτό που εξετάζουμε.

Τέλος, πιο προφανής λύση για την εύρεση όσο το δυνατόν καλύτερου μικροδορυφόρου είναι η απομόνωσή του από το γονιδίωμα του ίδιου του οργανισμού που μελετάται. Φυσικά η επιτυχία των αποτελεσμάτων στην ενίσχυση του τμήματος μέσω PCR κρίνεται πρωτίστως από τη βελτιστοποίηση των συνθηκών στην προετοιμασία για την πραγματοποίηση της αντίδρασης PCR, έτσι ώστε η αντίδραση να δώσει μια μπάρα για τα ομόζυγα δείγματα και δυο για τα ετερόζυγα, στο τζελ ηλεκτροφόρησης.

3 Βιοπληροφορική ανάλυση δεδομένων

3.1 Συγγενικές σχέσεις μελών ενός πληθυσμού

Η γνώση των συγγενικών σχέσεων μεταξύ των μελών των ζωικών πληθυσμών έχει ευρεία εφαρμογή στην έρευνα για την οικολογία και την εξέλιξη επιτρέποντας την έρευνα των συστημάτων ζευγαρώματος, της αναγνώρισης των πατρότητας/μητρότητας, της εύρεσης των συγγενών καθώς και της διαχείρισης πληθυσμών των απειλούμενων ειδών. Ωστόσο η εκτίμηση συγγένειας μεταξύ

μελών ενός πληθυσμού είναι δύσκολη όταν απουσιάζουν λεπτομερείς πληροφορίες γενεαλογίας του.

Η σύγκριση γονοτύπων των διαφόρων ατόμων και η ταξινόμησή τους σε κατηγορίες συγγενικών σχέσεων, όπως πρώτου βαθμού, δεύτερου βαθμού συγγένειας κλπ., παρουσιάζει ορισμένες δυσκολίες. Αυτό οφείλεται στο γεγονός ότι η συγγένεια βάση γενετικών πληροφοριών και δεδομένων είναι μια συνεχής παράμετρος που καθορίζεται από την αναλογία του γονιδιώματος που μοιράζονται μεταξύ τους δύο άτομα από έναν κοινό πρόγονο και ιδιαίτερα αν αυτό συγκεντρώνεται από έναν περιορισμένο αριθμό δεικτών.

Επίσης ο διαχωρισμός των ζευγών χρωμοσωμάτων κατά τη διάρκεια της πρώτης κυτταρικής μειωτικής διαίρεσης καθώς και ο ανασυνδιασμός των χρωμοσωμάτων είναι διαδικασίες που οδηγούν σε μεγάλη διαφοροποίηση της ποσότητας του γονιδιώματος που είναι κοινή μεταξύ δύο συγγενών ίδια γενιάς, με εξαίρεση τη σχέση γονείς - απόγονοι και ανάμεσα σε μονοζυγωτικά δίδυμα. Για παράδειγμα, ενώ τα αδέρφια με δύο κοινούς γονείς (fullsib dyads) μοιράζονται κατά μέσο όρο το 50% του γονιδιώματός τους, κάποια μπορεί να μοιράζονται πολύ λιγότερο ή περισσότερο. Η διακύμανση αυτή εξαρτάται από τον αριθμό των χρωμοσωμάτων και τα ποσοστά ανασυνδιασμού τους (Städle and Vigilant 2016).

3.2 Εκτίμηση πατρότητας

Η σχέση γονέας – απόγονος μπορεί να καθοριστεί με σχετικά μεγαλύτερη σιγουριά από άλλες σχέσεις συγγένειας, με εξαίρεση βέβαια τις περιπτώσεις μεταλλάξεων και λαθών κατά την αντιγραφή, διότι γονέας και απόγονος πρέπει να έχουν κοινό τουλάχιστον ένα αλληλόμορφο σε κάθε γενετικό τόπο. Η άμεση σύγκριση των γονοτύπων του απογόνου, της μητέρας και του πιθανού πατέρα, αν διαφέρουν ικανοποιητικά, μπορεί άμεσα να αποκαλύψει σχέσεις γονέων μέσα στον πληθυσμό που μελετάται, μόνο εάν όλοι οι υποψήφιοι γονείς είναι γνωστοί στα δείγματα (Städle and Vigilant 2016).

Η ανάλυση πατρότητας γίνεται αξιoσημείωτα πιο δύσκολη σε καταστάσεις όπου ούτε ένας γονέας είναι γνωστός με παρατήρηση. Ουσιαστικά, μπορεί να εφαρμοστεί η ίδια αρχή των κοινών αλληλομόρφων τους και ο αποκλεισμός ατόμων που δεν τα μοιράζονται, αλλά η προέλευση γίνεται πολύ πιο περίπλοκη, καθώς οι ταυτότητες των μητρικών και των πατρικών αλληλομόρφων στους απογόνους είναι άγνωστες. Αρκετές διαφορετικές προσεγγίσεις έχουν επινοηθεί για να προσδιορίσουν την προέλευση εάν είναι λίγες ή καθόλου γνωστές οι σχέσεις γονέων-απογόνων ή δεν μπορούν να αποκλεισθούν αρκετοί τύποι.

Τα γενετικά δέντρα που απεικονίζουν γενεαλογικές σχέσεις μεταξύ των ατόμων ενός πληθυσμού, είναι ιδιαίτερα σημαντικά για διάφορους ερευνητικούς τομείς. Οι μοριακοί δείκτες συμβάλουν στην εξαγωγή συμπεράσματος της γενεαλογίας άγριων ειδών οργανισμών, όπου οι πληροφορίες συγγένειας είναι

δύσκολο να συλληθθούν με παρατήρηση. Για αυτό το λόγο τα δεδομένα που χρησιμοποιούν μοριακούς δείκτες αναλύονται στατιστικά χρησιμοποιώντας πολλές φορές Μεντελικές μεθόδους (O. R. Jones and Wang 2010).

3.3 Ανάλυση ετεροζυγωτίας και στατιστικών στοιχείων

Οι μικροδορυφόροι, τα τελευταία χρόνια, έχουν εξελιχθεί στους πιο δημοφιλείς γενετικούς δείκτες. Αρχικά, οι μεγάλες σαρώσεις γονιδιωμάτων που χρησιμοποιούν μικροδορυφόρους περιορίζονταν σε μελέτες χαρτογράφησης. Πιο πρόσφατα, με την εμφάνιση πλήρως αλληλουχημένων γονιδιωμάτων, η ταυτοποίηση μικροδορυφόρων σε μεγάλη κλίμακα παρουσιάζεται ως ένα σημαντικό εργαλείο για το διαχωρισμό των εξελικτικών δυνάμεων, όπως η επιλογή και η υβριδοποίηση σε φυσικούς πληθυσμούς.

Ενώ οι εργαστηριακές μέθοδοι έχουν αναπτυχθεί πλήρως για να συμβαδίζουν με αυτές τις απαιτήσεις της εξέλιξης, η ανάπτυξη εργαλείων ανάλυσης εξακολουθεί να υστερεί. Το μεγαλύτερο μέρος των προγραμμάτων λογισμικού της γενετικής πληθυσμών που χρησιμοποιείται ευρέως για τον υπολογισμό των βασικών στατιστικών στοιχείων, όπως η ετεροζυγωτία ή οι τιμές F_{ST} , σχεδιάστηκε αρχικά για την ανάλυση σχετικά μικρού αριθμού γενετικών τόπων μόνο. Ως εκ τούτου, τα προγράμματα αυτά συχνά χρησιμοποιούν τη δική τους ειδική μορφή εισόδου των δεδομένων, η οποία μερικές φορές απαιτεί σημαντικές αλλαγές και αναδιαμορφώσεις. Επί πλέον, ο αριθμός των τόπων και των ατόμων είναι συχνά περιορισμένος (DIERINGER 2003).

3.4 Ταυτοποίηση γενετικής δομής και ανακατασκευή γενεαλογίας πληθυσμών

Σε είδη για τα οποία ο πληθυσμός αναμένεται να περιέχει κυρίως ομάδες αδελφών με έναν κοινό γονέα (halfsib) ή και με του δύο γονείς κοινούς (fullsib), η ανακατασκευή της γενεαλογικής δομής είναι ένα ισχυρό εργαλείο για τον εντοπισμό των συγγενικών ατόμων. Αυτή η προσέγγιση είναι πιο ακριβής από την αξιολόγηση των δυάδων ατόμων, διότι συγκρίνει τις σχέσεις μεταξύ όλων των γονότυπων ταυτόχρονα.

Η επιτυχία της ανακατασκευής της γενεαλογίας βελτιώνεται γενικά με την αύξηση του αριθμού των ατόμων ανά οικογένεια με fullsib ή halfsib άτομα, αν και η fullsib σχέση μεταξύ των αδερφών μπορεί να προσδιοριστεί με μεγάλη ακρίβεια ακόμα και για ομάδες οικογένειας με τέσσερα μέλη – άτομα αλλά μπορεί να μειωθεί με έναν αυξανόμενο αριθμό οικογενειών. Τέτοιες αναλύσεις είναι εξαιρετικά ακριβείς, αλλά καθίστανται λιγότερο επιτυχείς αν συμπεριληφθούν και οι δυάδες ατόμων με χαμηλότερο βαθμό συγγένειας. Για παράδειγμα, η συμπερίληψη ατόμων 2^{ou} βαθμού συγγένειας μειώνει τη δύναμη και την ακρίβεια της ανάλυσης. Η ανάλυση των πληθυσμών με πολύπλοκες δομές συγγένειας, όπως μπορεί να προκύψει όταν και τα δύο φύλα είναι πολυγαμικά, μπορεί να οδηγήσει σε απαγορευτικά μεγάλους χρόνους για τις ανάγκες ανάλυσης και μη συμβατότητα.

Μια τέτοια πολυπλοκότητα, συμπεριλαμβανομένης της παρουσίας διαφόρων κατηγοριών στενών και μακρινών συγγενών, μπορεί να είναι παρούσα σε πληθυσμιακές ομάδες που χαρακτηρίζονται από ανάμεικτα συστήματα ζευγαρώματος, μικρά μεγέθη απορριμμάτων, μικρή διάρκεια ζωής, επικαλυπτόμενες γενιές ή μετανάστευση. Ως εκ τούτου, απαιτούνται προσεγγίσεις που αναγνωρίζουν διαφορετικούς τύπους συγγενών σε ομάδες με σύνθετες συγγενικές σχέσεις (Städle and Vigilant 2016).

4 Προγράμματα Βιοπληροφορικής στην γενετική ανάλυση πατρότητας

Υπάρχουν πολλά προγράμματα υπολογιστών που διατίθενται για γενεαλογική ανάλυση, αλλά το μεγαλύτερο μέρος του λογισμικού τους είναι μη προσαρμόσιμο, τόσο ως προς τις υποθέσεις όσο και ως προς τις απαιτήσεις των χαρακτηριστικών των δεδομένων. Οι περισσότερες μέθοδοι σχετίζονται με μονογαμικά διπλοειδή είδη χρησιμοποιώντας συνεπικρατείς (co-dominant) δείκτες χωρίς σφάλμα στη γονοτύπηση. Επιπλέον, οι συνηθέστερες μέθοδοι που χρησιμοποιούνται βασίζονται σε συγκρίσεις ανά ζεύγη αντί για προσέγγιση πιθανότητας πλήρους γενεαλογίας (full-pedigree likelihood), η οποία εξετάζει την πιθανότητα ολόκληρης της γενετικής δομής του εξεταζόμενου πληθυσμού και επιτρέπει την ταυτόχρονη ανάλυση πατρότητας και ομαδοποίηση ατόμων σε αδέρφια, δηλαδή οικογένειες (O. R. Jones and Wang 2010).

Τα προγράμματα που βρίσκουν εφαρμογή μεταξύ άλλων και οι παραπάνω δυνατότητες είναι:

4.1 Colony

Είναι ένα πρόγραμμα βιοπληροφορικής που εφαρμόζει μεθόδους πιθανοτήτων πλήρους γενεαλογίας του πληθυσμού χρησιμοποιώντας αλληλομορφικά δεδομένα από πολλούς γενετικούς τόπους. Το Colony μπορεί να χρησιμοποιηθεί τόσο για διπλοειδείς όσο και για απλοειδείς οργανισμούς. Το πρόγραμμα είναι διαθέσιμο σε έκδοση Microsoft Windows, το οποίο περιλαμβάνει ένα γραφικό περιβάλλον χρήστη και μια έκδοση Macintosh, η οποία διαθέτει περιβάλλον για το χρήστη με την τεχνική R (O. R. Jones and Wang 2010).

Το Colony μπορεί να χρησιμοποιηθεί, μεταξύ άλλων, για την εκτίμηση των σχέσεων συγγένειας, κατηγοριοποιώντας τα άτομα σε οικογένειες με έναν κοινό γονέα (halfsib family) ή και με τους δυο γονείς κοινούς (fullsib family), εντοπίζοντας έτσι τους κλώνους ή τα διπλότυπα άτομα. Επίσης μπορεί να προσδιορίσει την πατρότητα ανακατασκευάζοντας τη γενεαλογία, προτείνοντας συστήματα ζευγαρώματος (πολυγαμικό / μονογαμικό ποσοστό αναπαραγωγής σε ένα πληθυσμό) και να επανεκτιμήσει το σφάλμα γονοτύπησης. Με μια μικρή τροποποίηση των δεδομένων, μπορεί επίσης να εφαρμοστεί σε πιο πολύπλοκους οργανισμούς. Μπορεί να χρησιμοποιήσει δεδομένα από κυρίαρχους (dominant) ή και συγκυρίαρχους (co-dominant) δείκτες με ή χωρίς σφάλματα γονοτύπησης. Η

έκδοση Windows GUI του Colony μπορεί επίσης να χρησιμοποιηθεί για να προσομοιώσει δεδομένα από γονοτύπους ατόμων ενός πληθυσμού με συγκεκριμένη γενετική δομή (γονική προέλευση και ομαδοποίηση οικογενειών), τα οποία μπορούν να χρησιμοποιηθούν για τον έλεγχο της ακρίβειας διάφορων μεθόδων ανακατασκευής γενεαλογίας και της επάρκειας των πληροφοριών των μοριακών δεικτών που χρησιμοποιούνται (Wang 2016).

Συνοπτικά, η μέθοδος του προγράμματος Colony θεωρεί ένα σύνολο ατόμων (πληθυσμός) που υποδιαιρείται σε 3 υπό-ομάδες: απόγονοι (OFS), υποψήφια αρσενικά άτομα (CMS) και υποψήφια θηλυκά άτομα (CFS). Οι απόγονοι (OFS) είναι απαραίτητοι για την πραγματοποίηση της ανάλυσης, ενώ τα υποψήφια αρσενικά (CMS) και τα υποψήφια θηλυκά άτομα (CFS) είναι προαιρετικά. Τα άτομα της ομάδας OFS ομαδοποιούνται στις πατρικές οικογένειες K1 και στις μητρικές K2 (όπου τα K1 και K2 είναι άγνωστα) και τα άτομα των CMS και CFS, αν είναι γνωστά, αντιστοιχίζονται στις οικογένειες K1 και K2 ανάλογα με την πατρότητα και μητρότητα αντίστοιχα. Θεωρείται ότι τα άτομα των απογόνων (OFS) είναι: πανομοιότυπα (ή μέλη ενός κλώνου), ή μέλη οικογένειας fullsib (που μοιράζονται και τους δύο γονείς), είτε halvesib (που μοιράζονται μόνο έναν από τους δύο γονείς) ή άσχετα μεταξύ τους (δεν έχουν συγγένεια). Οι υποψήφιοι γονείς (CMS και CFS) θεωρούνται άσχετοι μεταξύ τους και μπορεί είναι γονείς των ατόμων που ανήκουν στην ομάδα απογόνων (OFS). Οι δείκτες θεωρείται ότι βρίσκονται σε ισορροπία σύνδεσης.

Η παρέκκλιση από αυτές τις υποθέσεις μπορεί να μειώσει την αποτελεσματικότητα της ανάλυσης, αλλά θα μπορούσε να αντισταθμιστεί με τη χρήση περισσότερων πληροφοριακών δεικτών (Wang 2004). Για παράδειγμα, οι πληροφορίες σχετικά με το φύλο και την ηλικία των ατόμων του δείγματος του πληθυσμού ενδέχεται να μην είναι διαθέσιμες. Σε αυτή την περίπτωση, κάθε άτομο επιτρέπεται να αντιστοιχηθεί και στις 3 υπό-ομάδες (OFS, CMS και CFS) και τα αποτελέσματα για την ανάλυση συγγένειας και την ανακατασκευή γενεαλογίας εξακολουθούν να συνάγονται ικανοποιητικά σε ορισμένες περιπτώσεις (Wang & Santure 2009). Παρομοίως, η παρουσία ιστορικού συγγένειας (όπως 2^{ου} και 3^{ου} βαθμού συγγένειας που θεωρείται ότι είτε απουσιάζουν, είτε δεν σχετίζονται με τη μέθοδο) θα μπορούσε να μειώσει την ακρίβεια της ανάλυσης.

Ωστόσο, η ακρίβεια βελτιώνεται γρήγορα με την παροχή περισσότερων διαθέσιμων πληροφοριών των μοριακών δεικτών. Το τρέχον μοντέλο παρουσιάζει αποκλίσεις από την ισορροπία Hardy – Weinberg. Όταν είναι επιθυμητό, μπορεί να υπολογιστεί και να εκτιμηθεί η συγγένεια μαζί με τη δομή των συγγενικών σχέσεων, δηλαδή της γενεαλογίας (Wang 2016).

4.2 Kinalyzer

Ένα πρόγραμμα λογισμικού στατιστικής ανάλυσης είναι το Kinalyzer, το οποίο χρησιμοποιείται για ανακατασκευή γενεαλογιών ομάδων ατόμων ενός

πληθυσμού, χωρίς πληροφορίες συγγένειας μεταξύ τους. Χρησιμοποιεί δεδομένα από μοριακούς δείκτες όπως μικροδορυφόρους. Το Kinalyzer χρησιμοποιεί έναν νέο αλγόριθμο για την ανακατασκευή γενεαλογίας των απογόνων των διπλοειδών οργανισμών που βασίζεται στην συνδυαστική βελτιστοποίηση. Χρησιμοποιεί μια προσέγγιση «2-Allele Minimum Set Cover» που βασίζεται σε Μεντελικές μεθόδους κληρονομιάς και βρίσκει τον μικρότερο αριθμό ομάδων αδελφών που εμπεριέχονται όλα τα άτομα του δείγματος του πληθυσμού. Σε αντίθεση με τις μεθόδους πιθανότητας ανακατασκευής, το Kinalyzer δεν απαιτεί πληροφορίες σχετικά με τις συχνότητες των αλληλομόρφων του πληθυσμού και δεν έχει ως δεδομένα τα σύστημα ζευγαρώματος του είδους. Το Kinalyzer είναι διαθέσιμο ως πρόγραμμα μέσω διαδικτύου [(ASHLEY et al. 2009), (<http://kinalyzer.cs.uic.edu/about.php>)]

4.3 Pedigree

Το Pedigree είναι ένα διαδικτυακό πρόγραμμα ανασυγκρότησης γενεαλογιών που χρησιμοποιεί τις πιθανότητες ανά ζεύγη ατόμων για την δημιουργία ομάδων συγγενικών ατόμων μεταξύ τους και τις αξιολογεί με βάση ένα ισοσταθμισμένο αποτέλεσμα (score) [(A. G. Jones et al. 2010), (<http://herbinger.biology.dal.ca:5080/Pedigree/>)]

4.4 Άλλα προγράμματα βιοπληροφορικής για συγγενικές σχέσεις μελών πληθυσμών

Εκτός από τα προγράμματα λογισμικού που αναφέρθηκαν παραπάνω και χρησιμοποιήθηκαν στην παρούσα ερευνητική εργασία για την ανάλυση πατρότητας και ανακατασκευή γενεαλογίας, διατίθενται και ορισμένα άλλα προγράμματα. Αυτά μπορούν να κατηγοριοποιηθούν σε τρεις κατηγορίες ανάλογα την πολυπλοκότητά τους, το περιβάλλον χρήστη που προσφέρουν, τα δεδομένα τα οποία μπορούν να επεξεργαστούν και την απαιτητική τους λειτουργία.

Ορισμένα από αυτά είναι:

- FAP 3.6, CERVUS 3.0, PASOS 1.0, PATRI, MASTERBAYES, GERUD 2.0, PARENTAGE 1.0, COLONY 2.0, PEDIGREE 2.2
- PEDAPP 1.0, TWO-SEX PATERNITY, NEWPAT 5, FAMOZ, PAE, PAPA 2.0, MATESOFT
- FAMSPHERE 0.4, PROBMAX 3.0, WHICHPARENTS 1.0, KINSHIP 1.3.1, PARENTE, PRT, ML-RELATE, KINALYZER, FAMILYFINDER, KINGROUP 2

4.5 Microsatellite Analyzer (MSA)

Για την στατιστική ανάλυση της δεδομένων ενός πληθυσμού και προσδιορισμού του ποσοστού ετεροζυγωτίας κάθε γενετικού τόπου χρησιμοποιούνται προγράμματα όπως το MSA, που αναλύεται στην παρούσα

εργασία και αποτελεί εκτελέσιμο πρόγραμμα με γλώσσα εντολών. Χαρακτηριστικά του MSA είναι:

- Απλή μορφή εισόδου
- Ικανότητα ανάλυσης μεγάλου συνόλου δεδομένων
- Παρουσίαση – εξαγωγή συνοπτικών στατιστικών στοιχείων σε ανοιχτού τύπου αρχείο, που επιτρέπει περαιτέρω αξιοποίηση των αποτελεσμάτων σε μεταγενέστερη ανάλυση ή ως αρχεία εισόδου (input files) σε άλλα προγράμματα λογισμικού
- Παροχή τρόπων απαρίθμησης των ειδικών προβλημάτων που αντιμετωπίζει η χρήση των συγγενικών σχέσεων ανάμεσα στα άτομα (DIERINGER 2003)

4.6 Άλλα προγράμματα στατιστικής ανάλυσης δεδομένων πληθυσμών **FSTAT**

Από ένα σύνολο δεδομένων από κωδικοποιητικούς ή απλοειδείς γενετικούς δείκτες, το FSTAT υπολογίζει τον αριθμό ατόμων ανά δείγμα και τόπο, τη συχνότητα αλληλομόρφων ανά δείγμα και συνολικά, τον παρατηρούμενο και προσδοκώμενο αριθμό κάθε γονότυπου ανά δείγμα και τόπο, την ακριβή γονιδιωματική ποικιλομορφία ανά δείγμα και τόπο, τον αριθμό αλληλομόρφων των δειγμάτων ανά τόπο και ανά δείγμα, καθώς και συνολικά και τέλος, την στατιστική μεταβλητή Fis ανά τόπο και δείγμα, καθώς και δοκιμασία για το αν είναι σημαντικά θετική ή αρνητική (σημαντικό έλλειμμα και περίσσεια ετερόζυγων αντίστοιχα).

Arlequin

Ο στόχος του Arlequin είναι να παρέχει στον μέσο ερευνητή της γενετικής πληθυσμών ένα αρκετά μεγάλο σύνολο βασικών μεθόδων και στατιστικών δοκιμών, προκειμένου να αντλήσει πληροφορίες σχετικά με τα γενετικά και δημογραφικά χαρακτηριστικά μιας συλλογής πληθυσμιακών δειγμάτων. Θεωρήθηκε σημαντικό να διερευνώνται τα δεδομένα, να πραγματοποιείται ανάλυση πολλές φορές στο ίδιο σύνολο δεδομένων από διαφορετικές προοπτικές, με διαφορετικές επιλεγμένες επιλογές. Οι στατιστικές δοκιμές που εφαρμόστηκαν στο Arlequin έχουν επιλεγεί έτσι ώστε να ελαχιστοποιούν τις κρυφές υποθέσεις και να είναι όσο το δυνατόν ισχυρότερες. Έτσι, συχνά λαμβάνουν τη μορφή είτε δοκιμών μετάλλαξης είτε ακριβών δοκιμών, με ορισμένες εξαιρέσεις. Τέλος, το Arlequin θα μπορούσε να χειριστεί τα γενετικά δεδομένα με πολλές διαφορετικές μορφές και να προσπαθήσει να εκτελέσει τους ίδιους τύπους αναλύσεων ανεξάρτητα από τη μορφή των δεδομένων. Επειδή το Arlequin διαθέτει ένα πλούσιο σύνολο χαρακτηριστικών και πολλές επιλογές, σημαίνει ότι ο χρήστης πρέπει να αφιερώσει λίγο χρόνο για την εκμάθησή του.

DiveRsity

Επιτρέπει τον υπολογισμό των στατιστικών κατανομής της γενετικής ποικιλότητας, των στατιστικών γενετικής διαφοροποίησης και της πληροφόρησης του τόπου για τον προσδιορισμό της προέλευσης των προγόνων. Παρέχει επίσης στους χρήστες τη δυνατότητα να υπολογίζουν με διαστήματα εμπιστοσύνης 95% τόσο σε τόπους, όσο και για συγκρίσεις ζευγών πληθυσμών. Επίσης υπολογίζει τις στατιστικές μεταβλητές (F-statistics). Παρέχονται διάφορα χαρακτηριστικά σχεδίασης, καθώς και Chi-square τεστ γενετικής ετερογένειας. Παρέχεται λειτουργικότητα για τον υπολογισμό διαφόρων παραμέτρων ποικιλομορφίας, είναι δυνατή για σύνολα δεδομένων SNP που προέρχονται από RAD-seq που περιέχουν χιλιάδες τόπους δείκτη.

GenoDive

Το GenoDive είναι ένα φιλικό προς το χρήστη πρόγραμμα για την εκτέλεση γενετικών αναλύσεων πληθυσμού. Διαθέτει ένα περιβάλλον εργασίας χωρίς ομαδοποιήσεις, πίσω από το οποίο βρίσκονται ισχυρά στατιστικά εργαλεία. Το πρόγραμμα διαθέτει πολλές αναλύσεις που δεν είναι διαθέσιμες σε άλλα προγράμματα, όπως ομαδοποιήσεις των ατόμων με βάση το AMOVA, τυποποιημένους συντελεστές διαφοροποίησης πληθυσμού, δοκιμές αναπαραγωγής κλώνων και AMOVA για δεδομένα πολυπλοειδικών ειδών. Το GenoDive μπορεί να χειριστεί γενετικά δεδομένα καθώς και γενετική απόσταση από τη μητέρα και οικολογικά δεδομένα, επιτρέποντας το συνδυασμό δεδομένων από διάφορες πηγές σε μια ενιαία ανάλυση.

4.7 Structure

Τα προγράμματα που χρησιμοποιούνται για την ανακατασκευή γενεαλογίας σε ένα δείγμα πληθυσμού είναι το Structure το οποίο χρησιμοποιεί τον αλγόριθμο Bayesian με σκοπό να ταυτοποιήσει τη γενετική δομή του πληθυσμού που μελετάται. Αποτελεί το πιο διαδεδομένο εργαλείο για την ανάλυση της γενετικής πληθυσμών. Το Structure προσφέρει τη δυνατότητα ταυτοποίησης υποομάδων σε ολόκληρο το δείγμα του πληθυσμού που μελετάται, εξετάζοντας διακύμανση της συχνότητας αλληλομόρφων των δεδομένων καθώς και την κατηγοριοποίηση όλων των ατόμων του πληθυσμού που εξετάζεται σε αυτές τις υποομάδες βάση ανάλυσης πιθανοτήτων.

Διαπιστώθηκε ωστόσο ότι στις περισσότερες περιπτώσεις ο εκτιμώμενος «λογάριθμος πιθανοτήτων των δεδομένων, log probability of data, LPD» δεν παρέχει μια σωστή εκτίμηση του αριθμού των ομάδων (clusters, K). Εντούτοις, χρησιμοποιώντας μια στατιστική μεταβλητή ΔK με βάση το ρυθμό μεταβολής του «log probability of data» μεταξύ διαδοχικών clusters K, διαπιστώθηκε ότι το Structure ανιχνεύει με ακρίβεια το ανώτατο ιεραρχικά επίπεδο γενετικής δομής.

Όπως είναι αναμενόμενο, τα αποτελέσματα που εξάγονται με τη χρήση του προγράμματος εξαρτώνται από τον τύπο γενετικού δείκτη που χρησιμοποιείται (AFLP ή μικροδορυφόροι), τον αριθμό των γενετικών τόπων, το μέγεθος του

πληθυσμού του δείγματος και τον αριθμό των ατόμων του κάθε δείγματος (Evanno, Regnaut, and Goudet 2005).

II. ΥΛΙΚΑ – ΜΕΘΟΔΟΙ

1. Δεδομένα που χρησιμοποιήθηκαν

Στην παρούσα εργασία αξιοποιούνται προς ανάλυση γονιδιακές δεξαμενές από πληθυσμό εκτρεφόμενων πληθυσμών από λαβράκι και τσιπούρα. Τα δεδομένα είναι από μικροδορυφόρους για 10 γενετικούς τόπους στο λαβράκι (συνολικό πλήθος ατόμων 565) και 9 στην τσιπούρα (συνολικό πλήθος ατόμων 803). Τα δεδομένα ήταν οργανωμένα σε πίνακα Excel με την πρώτη κάθετη στήλη να αναφέρεται στο χαρακτηριστικό ID κάθε ατόμου της δεξαμενής, και οι αλληλομορφικές τους αλληλουχίες χωρισμένες σε στήλες ανά γενετικό τόπο με δυο αλληλόμορφα για κάθε γενετικό τόπο.

Πιο συγκεκριμένα, οι γονιδιακές δεξαμενές είναι οργανωμένες ως εξής:

Λαβράκι	
Γονιδιακές δεξαμενές	Μέγεθος δεξαμενής
Γ11	54
Γ12	53
Δ19	41
Δ23	49
Δ25	86
Π1	72
Φ2	118
Φ3	92

Τσιπούρα	
Γονιδιακές δεξαμενές	Μέγεθος δεξαμενής
Δ20	65
Δ21	67
Δ22	80
Δ24	79
Δ26	72
Φ1	102
Φ4	142
Φ5	91
Φ6	105

Οι μοριακοί δείκτες που χρησιμοποιήθηκαν για τα δυο αλληλόμορφα των 10 γενετικών τόπων για το λαβράκι και των 9 γενετικών τόπων για την τσιπούρα είναι οι παρακάτω:

Λαβράκι: *DLA0018, DLA0039, DLA-12, DLA0013, LABRAX17, DLA0040, DLA0081, DLA0016, DLA261E, DLA0011*

Τσιπούρα: *BD06H, SAGT41B, SAIMBB26, SAU97INRA, FD78H, SAI12, SAI14, SAI21, SAUK140INRA*

2 Συγγενικές σχέσεις μελών ενός πληθυσμού – Εκτίμηση πατρότητας

Για την ομαδοποίηση των ατόμων κάθε δεξαμενής – πληθυσμού που αναλύθηκε σε οικογένειες, χρησιμοποιήθηκαν τα παρακάτω προγράμματα στατιστικής ανάλυσης της βιοπληροφορικής:

2.1 Colony

2.1.1 Παράμετροι που εισάγονται

Πριν την έναρξη του προγράμματος απαιτείται να καθοριστούν ορισμένοι παράμετροι που αφορούν το σύνολο των ατόμων του πληθυσμού που μελετάται. Στις περισσότερες περιπτώσεις, οι προεπιλεγμένες τιμές των παραμέτρων είναι ιδανικές για τα συνήθη δείγματα πληθυσμών προς ανάλυση. Οι παράμετροι που απαιτούνται είναι οι εξής:

- 1 *Σύστημα ζευγαρώματος (Mating System I, II)*: μονογαμία ή πολυγαμία, ύπαρξη ή μη ομομιξίας, ύπαρξη ή μη κλώνων
- 2 *Είδος πληθυσμού (Species)*: απλοειδής ή διπλοειδής, μόνοικα ή δίοικα
- 3 *Μέγεθος ανάλυσης (Length of Run)*: χαμηλό, μεσαίο, μεγάλο ή πολύ μεγάλο
- 4 *Μέθοδος ανάλυσης (Analysis Method)*: πλήρους πιθανότητας, πιθανότητα συγκριτικά ανά ζεύγη ή συνδυασμός των δύο

Στη συνέχεια τοποθετούνται τα δεδομένα των ατόμων του πληθυσμού που μελετάται. Τα δεδομένα που απαιτούνται να εισαχθούν εφόσον είναι γνωστά και διαθέσιμα, πριν την έναρξη του προγράμματος Colony, είναι:

- Αριθμός των γενετικών τόπων που εξετάζονται (*Markers*)
- Είδος των μοριακών δεικτών και ποσοστό λάθους γονοτύπων σε αυτούς (*Marker Types and Error Rates*)
- Συχνότητα των αλληλομόρφων μόνο όταν είναι γνωστή (*Allele Frequency*)
- Αριθμός των απογόνων (αριθμός ατόμων στον εξεταζόμενο πληθυσμό) (*Number of Offspring*)
- Γονότυποι ατόμων οργανωμένοι σε αλληλομορφικά δεδομένα μικροδορυφόρων ανά γενετικό τόπο, το αρχείο φορτώνεται ως κείμενο στη μορφή “.txt” (*Load Genotypes*)
- Γονότυποι αρσενικών και θηλυκών ατόμων (*Male Genotypes, Female Genotypes*)
- Πατρική και μητρική προέλευση των ατόμων (*Known PaternalSibs, Known MaternalSibs*)
- Άτομα που εξαιρούνται από υποψήφιοι γονείς, είτε αρσενικά είτε θηλυκά (*Excluded Paternity, Excluded Maternity*)
- Άτομα που αποκλείεται να έχουν κοινό πατέρα και/ή μητέρα (*Excluded PaternalSibs, Excluded MaternalSibs*)

Πριν την έναρξη του προγράμματος για ανάλυση των δεδομένων της παρούσας εργασίας, τέθηκαν οι κατάλληλες παράμετροι για τα συγκεκριμένα δεδομένα. Πιο συγκεκριμένα για τις παραμέτρους που αφορούν το σύστημα αναπαραγωγής του υπό μελέτη είδους:

Mating Systems: Female Polygamy, Male Polygamy, With Inbreeding, Without Clone.

Στη συνέχεια εισήχθησαν τα αρχεία με τους τύπους των δεικτών και τα ποσοστά λάθους, καθώς και το αρχείο με τους γονοτύπους των ατόμων της δεξαμενής που μελετάται κάθε φορά. Έπειτα, παραλείπονται τις καρτέλες με απαιτούμενα τους μητρικούς, πατρικούς γονοτύπους και τις καρτέλες που αφορούν στοιχεία πατρότητας και μητρότητας, διότι δεν είναι γνωστά. Τέλος αποθηκεύονται οι παράμετροι για τη δεδομένη ανάλυση κάθε φορά και το πρόγραμμα είναι έτοιμο για έναρξη.

2.1.2 Αποτελέσματα με τη χρήση του Colony

Τα αποτελέσματα της ανάλυσης που προκύπτει από το πρόγραμμα Colony περιλαμβάνουν κυρίως:

- Εκτίμηση συγγένειας fullsib ή halfsib οικογενειών μεταξύ των ατόμων των απογόνων (OFS)
- Εκτίμηση πατρότητας (εάν είναι διαθέσιμα τα άτομα CMS) και μητρότητας (εάν είναι διαθέσιμα τα CFS)
- Ανάλυση γονοτύπου σε κάθε γενετικό τόπο για κάθε απόγονο
- Ανάλυση γονοτύπου σε κάθε γενετικό τόπο για κάθε γονέα, ανεξάρτητα από το αν έχει αντιστοιχηθεί ή όχι σε υποψήφιο στους CFS ή CMS
- Πιθανά σφάλματα γονότυπου σε κάθε τόπο κάθε απογόνου
- Πιθανά σφάλματα γονότυπου σε κάθε τόπο κάθε υποψήφιου γονέα
- Εκτίμηση συχνότητας αλληλομόρφων λαμβάνοντας υπόψη τις πιθανές συγγενικές σχέσεις
- Εκτίμηση του ποσοστού σφαλμάτων των γονοτύπων σε κάθε γενετικό τόπο
- Εκτίμηση του πραγματικού μεγέθους του πληθυσμού από την εκτιμώμενη συχνότητα των αδελφών (Wang 2016)

Στη συγκεκριμένη ανάλυση που μελετάται στην παρούσα εργασία εξάγονται τα εξής αποτελέσματα για κάθε γονιδιακή δεξαμενή:

- Ομαδοποίηση ατόμων της γονιδιακής δεξαμενής ανά οικογένειες με υπολογισμένες πιθανότητες για κάθε οικογένεια
- Δυάδες ατόμων που έχουν κοινούς και τους δύο γονείς (fullsib dyad) με τις πιθανότητές τους. Στο σημείο αυτό αξίζει να αναφερθεί ότι πιο αξιόπιστη πιθανότητα για να αποτελέσει συμπέρασμα για τις περαιτέρω αναλύσεις είναι η πιθανότητα για fullsib dyad να είναι μεγαλύτερη από 0,05
- Οργάνωση όλων των αλληλομόρφων ανά γενετικό τόπο με τις αντίστοιχες πιθανότητες εμφάνισής τους σε ολόκληρη τη δεξαμενή
- Γράφημα για απεικόνιση των fullsib dyad

2.2 Kinalyzer

2.2.1 Η μέθοδος που χρησιμοποιείται

Οι συνδυαστικές προσεγγίσεις κατασκευάζουν πιθανές ομάδες αδελφών χρησιμοποιώντας μόνο Μεντελικές μεθόδους και αναζητούν την πιο φειδωλή λύση, όπως ο μικρότερος γονέων. Η μέθοδος αυτή χρησιμοποιεί μια διερευνητική προσέγγιση (και όχι καθιερωμένες μεθόδους υπολογιστικής βελτιστοποίησης) για να βρει μια άμεση βέλτιστη λύση όταν δεν είναι εγγυημένη η διερεύνηση της συνολικής βέλτιστης λύσης.

Το Kinalyzer χρησιμοποιεί συνδυαστική προσέγγιση που βασίζεται σε έναν απλό κανόνα για την κληρονομία αλληλομόρφων σε διπλοειδείς οργανισμούς: ένας απόγονος κληρονομεί ένα αλληλόμορφο από κάθε γονέα του για κάθε γενετικό τόπο. Αυτός ο κανόνας της Μεντελικής κληρονομικότητας εισάγει έναν απαραίτητο περιορισμό στις ομάδες απογόνων με δυο κοινούς γονείς, απουσία σφαλμάτων ή μεταλλάξεων των γονοτύπων: την ιδιότητα «2-Allele Minimum». Έτσι, υπάρχει μια αντιστοίχιση των αλληλομόρφων όλων των ατόμων ενός τόπου, σε ομάδα γονέων (μητέρα ή πατέρα), έτσι ώστε ο αριθμός διακριτών αλληλομόρφων που αντιστοιχούν σε κάθε γονέα σε αυτόν τον τόπο να μην υπερβαίνει τα δύο.

Το Kinalyzer επίσης διαθέτει μια μέθοδο συναινετική, ονομάζεται «consensus», που ανασυγκροτεί τις οικογένειες των απογόνων, χρησιμοποιώντας υποσύνολα των τόπων και αποδίδει τη συνένωση αυτών των διαφορετικών λύσεων.

Πηγή: <http://kinalyzer.cs.uic.edu/about.php>

Επιλέγοντας το Kinalyzer ως πρόγραμμα βιοπληροφορικής για ανακατασκευή γενεαλογιών, προτιμάται ταυτόχρονα η μέθοδος «2-Allele» που εφαρμόζεται, διότι πραγματοποιεί τον ελάχιστο αριθμό υποθέσεων και εκτελείται σε ένα ευρύ φάσμα παραμέτρων των δεδομένων. Επομένως, είναι γενικά μια καλή μέθοδος, ειδικά όταν τα δείγματα αφορούν λίγους γενετικούς τόπους ή η ποικιλία των αλληλομόρφων (ανά γενετικό τόπο) είναι χαμηλή (ASHLEY et al. 2009).

2.2.2 Παράμετροι που απαιτούνται

Το KINALYZER είναι ένα διαδικτυακό πρόγραμμα που απαιτεί μόνο ένα εισαγόμενο αρχείο με τα άτομα και τους γονότυπους για ανάλυση. Οι μορφές των εισαγόμενων αρχείων που απαιτούνται είναι ίδιες με ορισμένα άλλα προγράμματα βιοπληροφορικής όπως τα Kinship, Kingroup, Gerud ή Cervus με εξαίρεση ότι δεν απαιτούνται να εισαχθούν οι συχνότητες αλληλομόρφων των ατόμων του πληθυσμού για την εκτέλεση του προγράμματος. Η παράδοση των αποτελεσμάτων της ανάλυσης με το Kinalyzer γίνεται με την είσοδο του χρήστη στο ηλεκτρονικό του ταχυδρομείο (σύνδεση με προσωπικό email).

Τα αρχεία που θα εισαχθούν για ανάλυση, δηλαδή οι γονότυποι των ατόμων, πρέπει να είναι στη μορφή “.csv” του Excel. Τα δεδομένα που λείπουν (μηδενικά αλληλόμορφα) πρέπει να αποτυπώνονται με την τιμή -1. Οι στήλες πρέπει να αντιστοιχούν στην ταυτότητα των ατόμων (όνομα απογόνου, δηλαδή ID) και στο όνομα των γενετικών τόπων. Το αρχείο μπορεί να περιέχει επιπλέον στήλες που δεν χρησιμοποιούνται από τον Kinalyzer (δηλ. φύλο, τοποθεσία, κλπ). Δεν υπάρχει όριο στο μέγεθος του δείγματος του πληθυσμού ή στον αριθμό των τόπων. Ο χρόνος για την ανάλυση των δεδομένων εξαρτάται από τον αριθμό των εργασιών που επεξεργάζεται ο server εκείνη τη στιγμή.

2.2.3 Αποτελέσματα που εξάγονται

Τα αποτελέσματα της ανάλυσης χρησιμοποιώντας το Kinalyzer εξάγονται σε ένα αρχείο στο οποίο αποτυπώνονται τα άτομα χωρισμένα σε οικογένειες fullsib, δηλαδή σε οικογένειες στις οποίες οι απόγονοι έχουν κοινούς και τους δυο γονείς τους. Κάθε μια από αυτές τις οικογένειες θα αναφέρει τα άτομα με το όνομα ταυτότητάς τους (ή τον αριθμό ID) που αναγράφεται για το καθένα στο αρχείο “.csv” που εισάγεται πριν την έναρξη της ανάλυσης.

2.3 Pedigree

2.3.1 Μέθοδος που χρησιμοποιεί

Στο Pedigree χρησιμοποιείται η προσέγγιση Markov Chain Monte Carlo (MCMC) για την εύρεση της ομάδας με την υψηλότερη βαθμολογία (score). Τα άτομα μπορούν να χωριστούν σε ομάδες αδελφών που έχουν ένα κοινό γονέα ή σε ομάδες αδελφών που έχουν δύο κοινούς γονείς (halfsib και fullsib families αντίστοιχα). Οι ομάδες fullsib μπορούν να ενταχθούν μέσα σε halfsib οικογένειες, αλλά αυτή η διάταξη μπορεί να καταστεί δυνατή μόνο στην περίπτωση που ο ένας γονέας πραγματοποιεί πολλαπλά ζευγαρώματα.

2.3.2 Αποτελέσματα ανάλυσης

Τα αποτελέσματα της ανάλυσης παρουσιάζονται μόνο σε πίνακες με μορφοποίηση HTML (διαδικτυακή παρουσίαση), γεγονός που μπορεί να κάνει κάπως δύσκολη την εξαγωγή τους. Η σημασία οποιασδήποτε ομάδας μπορεί να εκτιμηθεί με τη σύγκριση των αποτελεσμάτων της συνεκτικότητας και της διαφορετικότητας ορισμένων ομάδων. Αυτές οι σχέσεις υπολογίζονται με βάση έναν λογάριθμο πιθανοτήτων συγγένειας ανάμεσα στα άτομα.

Η τρέχουσα έκδοση του Pedigree δεν ενσωματώνει έλεγχο σφάλματος ή ακρίβειας. Το πρόγραμμα μπορεί επίσης να ανασυγκροτήσει τους γονικούς γονότυπους των ατόμων. Με αυτή τη λειτουργία απαριθμούνται όλοι οι πιθανοί γονότυποι γονέων που θα μπορούσαν να δημιουργήσουν τους απογόνους μιας συγκεκριμένης ομάδας σε κάθε γενετικό τόπο.

3 Ανάλυση ετεροζυγωτίας και στατιστικών στοιχείων

3.1 Microsatellite Analyzer (MSA)

Για να ξεπεραστούν ορισμένοι περιορισμοί στον υπολογισμό ετεροζυγωτίας γενετικών τόπων πληθυσμών (οι μέθοδοι που αναπτύχθηκαν στην αρχή επέτρεπαν την ανάλυση μικρού αριθμού γενετικών τόπων) αναπτύχθηκε το πρόγραμμα Microsatellite Analyzer. Είναι ένα εκτελέσιμο πρόγραμμα που δέχεται εντολές για την έναρξη της ανάλυσης με τα δεδομένα και τις εντολές ως παραμέτρους που θα τεθούν από το χρήστη.

3.1.1 Ανάλυση δεδομένων με το MSA

Η παρατηρούμενη και αναμενόμενη ετεροζυγωτία (*Observed/Expected Heterozygosity*), η διακύμανση του μεγέθους του προϊόντος της αλυσιδωτής αντίδρασης πολυμεράσης (*PCR*), η διακύμανση του αριθμού επανάληψης (*Repeat Number*), η κατανομή της συχνότητας των αλληλομόρφων (*Allele Frequency*), ο αριθμός αλληλομόρφων (*Number of Alleles*) και το εύρος του μεγέθους των αλληλομόρφων είναι συνοπτικά τα στατιστικά στοιχεία που παρέχονται από το MSA. Επιπλέον, υπολογίζεται ο αναμενόμενος αριθμός αλληλομόρφων κάτω από το μοντέλο IAM (*infinite allele model*) και SMM (*stepwise mutation model*).

Επιπρόσθετα, με τη χρήση του MSA είναι δυνατή η εξαγωγή εννέα διαφορετικών μετρήσεων γενετικής απόστασης, είτε μεταξύ πληθυσμών είτε ατόμων. Μέσα σε αυτές υπολογίζονται οι μετρήσεις F_{ST} , F_{IS} και F_{IT} , των οποίων οι σημαντικές τιμές καθορίζονται από την ανταλλαγή των γονότυπων (ή αλληλομόρφων) μεταξύ των ομάδων ή πληθυσμού.

3.1.2 Εισαγωγή παραμέτρων

Στο MSA τέθηκαν στο περιβάλλον εργασίας, οι παρακάτω παράμετροι με τη μορφή εντολών για την έναρξη της ανάλυσης και την εξαγωγή των αναγκαίων για την παρούσα εργασία αποτελεσμάτων:

- *Distance settings*: All distances ON, Dfs ON, Dad ON, Dkf ON
- *Fst, Fit, Fis*: Distances calculated ON, Fst calculated pairwise ON
- *Heterozygosity range settings*: Calculate Heterozygosityrange ON

με σκοπό να υπολογιστούν οι απαραίτητες τιμές για τα F-statistics που απαιτούνται για τον υπολογισμό των γενετικών αποστάσεων.

3.1.3 Εξαγωγή δεδομένων

Το MSA εξήγαγε για την ανάλυση των δεδομένων της παρούσας εργασίας τα παρακάτω αποτελέσματα μετά την ανάλυση σε κάθε είδος:

- Τη συχνότητα που εμφανίζεται κάθε διαφορετικό αλληλόμορφο κάθε γενετικού τόπου για όλες τις δεξαμενές, ξεχωριστά για το λαβράκι και την τσιπούρα.
- Τις τιμές για τις στατιστικές μεταβλητές Fst-values, P-values, Global Fst για κάθε δεξαμενή χωριστά.

- Τα ποσοστά παρατηρούμενης (Het obs) και προσδοκώμενης (Het exp) ετεροζυγωτίας που παρατηρούνται ή αναμένονται αντίστοιχα, μετά την ανάλυση των δεδομένων που τέθηκαν. Τα ποσοστά αυτά παρουσιάζονται για κάθε γενετικό τόπο της κάθε δεξαμενής.
- Τον ελάχιστο ή μέγιστο αριθμό αλληλομόρφων που εντοπίζεται για κάθε γενετικό τόπο σε κάθε δεξαμενή.

Το MSA εξάγει επίσης ως αποτελέσματα αρχεία εισόδου για GENEPOP, MSVAR, STRUCTURE, ARLEQUIN και MIGRATE. Οι γενετικές αποστάσεις εξάγονται ως αρχεία PHYLIP.

3.1.4 Έλεγχος σφαλμάτων.

Τα αποτελέσματα που εξάγονται από το MSA εξετάστηκαν σε βάθος συγκριτικά με τα αποτελέσματα άλλων πακέτων λογισμικού, συμπεριλαμβανομένων των MICROSAT, FSTAT, GENETIX και ARLEQUIN. Σε αυτές τις συγκρίσεις δεν παρατηρήθηκαν αναντιστοιχίες. Μόνο οι τιμές P-value ανά ζεύγη στον πληθυσμό παρουσίαζαν διαφορά μεταξύ FSTAT και MSA. Αρχικά αποτελέσματα υποθέτουν ότι το FSTAT μεταβάλλεται σε όλους τους διαθέσιμους γονότυπους, ενώ το MSA και το GENETIX μεταβάλλονται μόνο μεταξύ δύο ομάδων των οποίων οι F-statistics ανά ζεύγη είναι καθορισμένες.

4 Ταυτοποίηση γενετικής δομής και ανακατασκευή γενεαλογίας πληθυσμών

Η ταυτοποίηση γενετικά όμοιων ομάδων ατόμων είναι μακροχρόνιο ζήτημα στη γενετική πληθυσμών. Ένας πρόσφατος αλγόριθμος τύπου Bayesian επιτρέπει την αναγνώριση τέτοιων ομάδων. Ωστόσο, δεν έχει εξεταστεί η ικανότητα αυτού του αλγορίθμου να ανιχνεύει τον πραγματικό αριθμό ομάδων (clusters, K) σε ένα δείγμα ατόμων στο οποίο δεν υπάρχουν όμοια πρότυπα διασποράς μεταξύ των πληθυσμών.

4.1 Structure software

4.1.1 Η μέθοδος που χρησιμοποιεί

Το Structure χρησιμοποιεί μια συστηματική προσέγγιση που βασίζεται στην ομαδοποίηση κατά Bayesian εφαρμόζοντας την εκτίμηση MCMC (Markov Chain Monte Carlo). Η διαδικασία ξεκινά με την τυχαία κατηγοριοποίηση των ατόμων σε ένα προκαθορισμένο αριθμό ομάδων, clusters, όπου στη συνέχεια εκτιμάται η διακύμανση των αλληλομορφικών συχνοτήτων σε κάθε ομάδα, με αποτέλεσμα τα άτομα να κατηγοριοποιούνται εκ νέου με βάση αυτές τις συχνότητες που εντοπίστηκαν. Αυτή η διαδικασία επαναλαμβάνεται πολλές φορές, τυπικά περίπου στις 100.000 επαναλήψεις, και οδηγεί προοδευτικά στη σύγκλιση αξιόπιστων εκτιμήσεων των αλληλομορφικών συχνοτήτων και των συχνοτήτων εμφάνισης των ατόμων στον πληθυσμό που εξετάζεται (Porrás-Hurtado et al. 2013).

4.1.2 Παράμετροι που είναι απαραίτητο να οριστούν

- Αριθμός ατόμων υπό μελέτη πληθυσμού (*Number of individuals*)
- Αν πρόκειται για απλοειδή ή διπλοειδή δεδομένα (*Ploidy of data*)
- Αριθμός γενετικών τόπων (*Number of loci*)
- Τιμή για τα μηδενικά δεδομένα (*Missing data value*)
- Το μοντέλο γενεαλογίας (*Admixture/ Non admixture model*)
- Το μέγεθος των κύκλων που θα πραγματοποιήσει η ανάλυση με το Structure (*Length of Burnin Period*)
- Αριθμός των επαναλήψεων του αλγόριθμου MCMC ανά κύκλο (*Number of MCMC Reps after Burnin*) (Pritchard 2010)

Για την ανακατασκευή γενεαλογίας των δεδομένων της παρούσας εργασίας ορίστηκαν οι παρακάτω παράμετροι:

- *Number of individuals*: σύνολο ατόμων από όλες τις γονιδιακές δεξαμενές του κάθε είδους ξεχωριστά, δηλαδή 565 άτομα για το είδος λαβράκι και 803 για το είδος τσιπούρα
- *Ploidy of data*: 2 και για τα δύο είδη
- *Number of loci*: 10 για το είδος λαβράκι και 9 για το είδος τσιπούρα
- *Missing data value*: 0 και για τα δύο είδη
- *Format of input data set*: Individual ID for each individual

Επιπλέον, επιλέγοντας το Parameter Set για το κάθε Project, δηλαδή για κάθε είδος, ορίστηκαν τα ακόλουθα:

- *Length of burning period*: 1000
- *Number of MCMC Reps after Burnin*: 25000

4.1.3 Αρχείο δεδομένων (data file – input file)

Το σύνολο των δεδομένων είναι διατεταγμένο ως ένας πίνακας σε ένα μόνο αρχείο, στο οποίο τα δεδομένα για τα άτομα είναι σε σειρές, και οι γενετικοί τόποι είναι σε στήλες. Ο χρήστης μπορεί να κάνει διάφορες επιλογές σχετικά με τη μορφή των δεδομένων, ενώ τα περισσότερα από αυτά τα δεδομένα είναι προαιρετικά (εκτός από τους γονότυπους).

Για ένα διπλοειδές οργανισμό, δεδομένα για κάθε άτομο μπορούν να αποθηκευτούν:

- είτε ως δυο διαδοχικές σειρές, όπου κάθε γενετικός τόπος βρίσκεται σε μία στήλη
- είτε σε μία σειρά, όπου κάθε γενετικός τόπος βρίσκεται σε δύο διαδοχικές στήλες.

Τα στοιχεία του αρχείου εισόδου εάν υπάρχουν, εμφανίζονται με την ακόλουθη σειρά, ωστόσο τα περισσότερα είναι προαιρετικά (όπως υποδεικνύεται) και μπορούν να διαγραφούν εντελώς. Ο χρήστης καθορίζει ποια δεδομένα

εμφανίζονται, είτε στο μπροστινό άκρο (όταν το πρόγραμμα εκτελείται από τη γραμμή εντολών), είτε σε ένα ξεχωριστό αρχείο, `mainprograms`. Παράλληλα, ο χρήστης καθορίζει τον αριθμό των ατόμων και τον αριθμό των γενετικών τόπων (Pritchard 2010).

4.1.4 Επιλογή μοντέλου γενεαλογίας

Υπάρχουν τέσσερα κύρια μοντέλα για την καταγωγή των ατόμων:

- το μοντέλο της μη πρόσμιξης (*no admixture model*): είναι διακριτή η προέλευση των ατόμων από τον έναν ή τον άλλο πληθυσμό
- το μοντέλο της πρόσμιξης (*admixture model*): κάθε άτομο αντλεί κάποιο ποσοστό του γονιδιώματός του από κάθε πληθυσμό K
- το μοντέλο σύνδεσης (*linkage model*): όπως το μοντέλο πρόσμιξης, αλλά οι συνδεδεμένοι γενετικοί τόποι είναι πιθανότερο να προέρχονται από τον ίδιο πληθυσμό
- τα μοντέλα με πληροφοριακές αρχές (*using prior population information*): επιτρέπουν στο Structure να χρησιμοποιεί πληροφορίες σχετικά με τις τοποθεσίες προέλευσης των ατόμων του δείγματος είτε για να βοηθήσει την ομαδοποίηση με αδύναμα δεδομένα, είτε για να ανιχνεύσει τα άτομα που προήλθαν από άλλους πληθυσμούς ή για να προκαθορίσει ορισμένους πληθυσμούς (Pritchard 2010).

4.1.5 Μοντέλα για την συχνότητα αλληλομόρφων

Υπάρχουν δύο βασικά μοντέλα για τις συχνότητες αλληλομόρφων:

- Ένα μοντέλο υποθέτει ότι οι συχνότητες αλληλομόρφων σε κάθε πληθυσμό που είναι ανεξάρτητες, αντλούνται από μια κατανομή η οποία καθορίζεται από μια παράμετρο που ονομάζεται λ . Συνήθως ορίζουμε $\lambda = 1$. Αυτή είναι η προεπιλεγμένη ρύθμιση.
- Το δεύτερο μοντέλο αφορά τις συσχετισμένες συχνότητες αλληλομόρφων. Σε αυτό το μοντέλο θεωρείται ότι οι συχνότητες στους διάφορους πληθυσμούς είναι πιθανό να είναι παρόμοιες (πιθανώς λόγω μετανάστευσης ατόμων ή κοινής καταγωγής). Το ανεξάρτητο μοντέλο λειτουργεί καλά για πολλά σύνολα δεδομένων.

Σε γενικές γραμμές, το πρώτο μοντέλο υποθέτει ότι περιμένουμε συχνότητες αλληλομόρφων σε διαφορετικούς πληθυσμούς να είναι αρκετά διαφορετικές μεταξύ τους. Αντιθέτως, το μοντέλο συσχετισμένων συχνοτήτων λέει ότι μπορεί να είναι στην πραγματικότητα αρκετά παρόμοιες οι συχνότητες αλληλομόρφων σε διαφορετικούς πληθυσμούς. Αυτό συχνά βελτιώνει την ομαδοποίηση για στενά συνδεδεμένους πληθυσμούς, αλλά μπορεί να αυξήσει τον κίνδυνο της λανθασμένης εκτίμησης του K . Εάν ένας πληθυσμός αποκλίνει αρκετά από τους άλλους, το συσχετισμένο μοντέλο μπορεί μερικές φορές να επιτύχει καλύτερα συμπεράσματα εάν απομακρυνθεί αυτός ο πληθυσμός από το σύνολο των δεδομένων (Pritchard 2010).

4.1.6 Επισκόπηση του Structure software

Το πρόγραμμα οργανώνει την ανάλυση δεδομένων σε "projects". Κάθε project συνδέεται με ένα μόνο αρχείο δεδομένων. Κατά τη δημιουργία ενός project, ο χρήστης παρέχει επίσης πληροφορίες που καθορίζουν τον τρόπο ανάγνωσης του αρχείου δεδομένων (αριθμός τόπων, αριθμός ατόμων κλπ). Αυτά είναι τα χαρακτηριστικά του αρχείου δεδομένων και είναι πάντα τα ίδια στο πλαίσιο του κάθε project.

Κάθε project περιέχει επίσης ένα ή περισσότερα "σύνολα παραμέτρων". Αυτά επιτρέπουν στο χρήστη να καθορίσει τις λεπτομέρειες του αλγόριθμου MCMC, συμπεριλαμβανομένου του αριθμού των επαναλήψεων, του μήκους ανάλυσης καθώς και τον προσδιορισμό του μοντέλου ανάλυσης (π.χ., να επιτρέπεται ή όχι το ενδεχόμενο πρόσμιξης, τα μοντέλα των συχνοτήτων των αλληλομόρφων κλπ). Ο χρήστης μπορεί στη συνέχεια να εκτελέσει το Structure software σε επιλεγμένες τιμές του K , για ένα συγκεκριμένο σύνολο παραμέτρων.

Το πρόγραμμα μπορεί στη συνέχεια να εκτελείται χρησιμοποιώντας αυτές τις τιμές παραμέτρων. Τέλος, το Structure αποθηκεύει διάφορες περιλήψεις των αποτελεσμάτων, συμπεριλαμβανομένου ενός αριθμού γραφικών παραστάσεων (Pritchard 2010).

III. ΑΠΟΤΕΛΕΣΜΑΤΑ

1 Συγγενικές σχέσεις μελών ενός πληθυσμού

Με βάση τα υπάρχοντα δεδομένα για τους πληθυσμούς του λαβρακιού και της τσιπούρας (γονιδιακές δεξαμενές) πραγματοποιείται ανάλυση των αλληλομορφικών τους δεδομένων και συγκεκριμένα με βάση αυτά η κατηγοριοποίηση των ατόμων κάθε δεξαμενής σε οικογένειες με βάση κυρίως τις αλληλομορφικές πιθανότητες, αλλά και άλλους παραμέτρους. Για το σκοπό αυτό χρησιμοποιήθηκαν τα προγράμματα Colony, Pedigree και Kinalyzer.

Σύμφωνα με τις οδηγίες για την οργάνωση των εισερχόμενων αρχείων δεδομένων για κάθε πρόγραμμα, πραγματοποιήθηκαν οι κατάλληλες διαμορφώσεις στα αρχεία των δεδομένων που προϋπήρχαν.

1.1 Ανάλυση συγγένειες με τα προγράμματα Colony, Pedigree και Kinalyzer

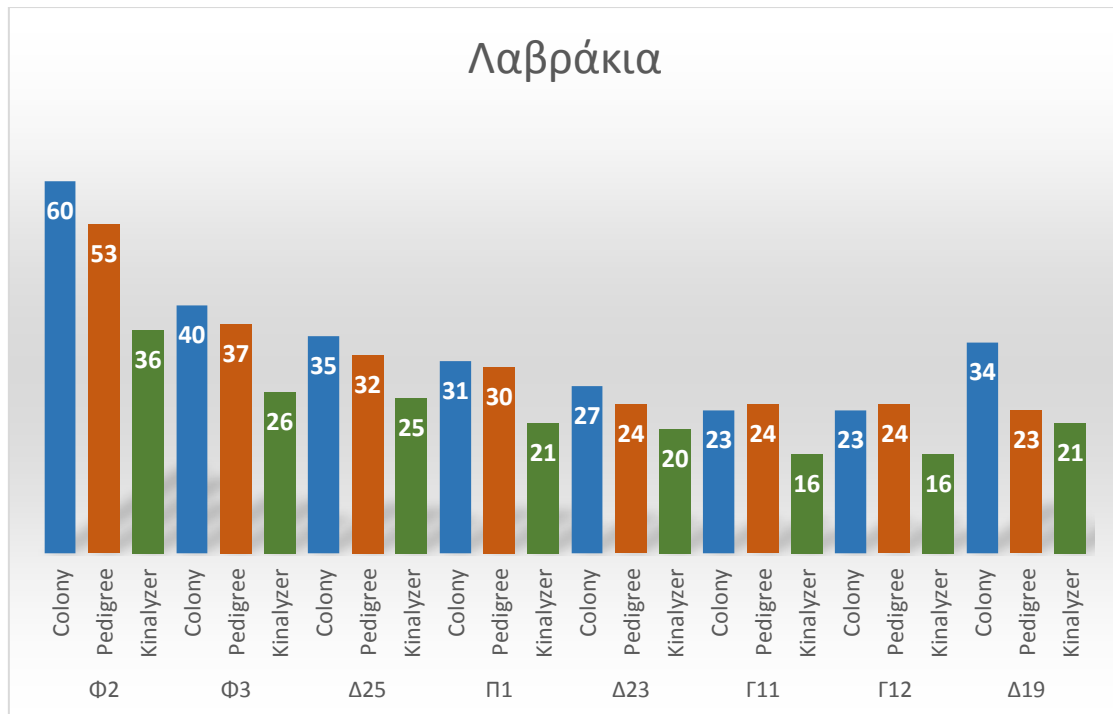
Στο Colony ορίστηκαν οι παράμετροι που αναφέρονται στην Ενότητα II και Παράγραφο 2.1. πριν την έναρξη του προγράμματος ανάλυσης των δεδομένων κάθε γονιδιακής δεξαμενής.

Τα προγράμματα Pedigree και Kinalyzer είναι διαδικτυακά και δεν απαιτούν ορισμό κάποιων παραμέτρων, με εξαίρεση το Pedigree που είναι απαραίτητη η εισαγωγή ενός αρχείου ελέγχου (control file) πριν την έναρξη του προγράμματος. Και τα δυο αυτά τα προγράμματα εξάγουν μόνο την ομαδοποίηση ατόμων της δεξαμενής με βάση τους γονοτύπους τους (αλληλομορφικές συχνότητες), χωρίς να υπολογίζουν τις πιθανότητες για κάθε οικογένεια.

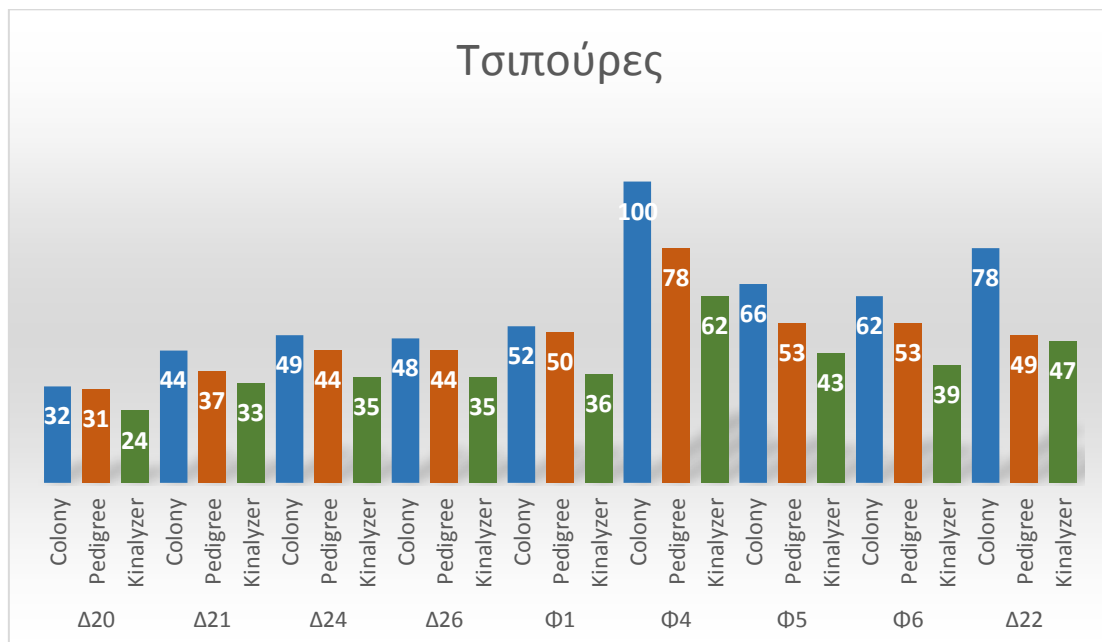
1.2 Σύγκριση των παραπάνω προγραμμάτων

Συγκεντρώνοντας τα παραπάνω αποτελέσματα από την ομαδοποίηση σε οικογένειες των ατόμων κάθε δεξαμενής και συγκρίνοντας την ομοιότητά τους σε άτομα, διαπιστώθηκε πως οι περισσότερες κοινές οικογένειες εμφανίζονται ανάμεσα στα προγράμματα Colony και Pedigree. Επομένως, θεωρήθηκαν ως οι πιο πιθανές οικογένειες για τα άτομα των γονιδιακών δεξαμενών.

Παρακάτω παρουσιάζονται τα διαγράμματα από τα οποία φαίνεται η ομοιότητα αυτή:



Ομαδοποιήσεις γονιδιακών δεξαμενών λαβρακιού με τα προγράμματα Colony, Pedigree, Kinalyzer



Ομαδοποιήσεις γονιδιακών δεξαμενών τσιπούρας με τα προγράμματα Colony, Pedigree, Kinalyzer

Επίσης, στον παρακάτω πίνακα παρουσιάζονται οι κοινές οικογένειες που διαπιστώθηκαν, ανάμεσα στα προγράμματα Colony και Pedigree:

Λαβράκια	
Γ11	8
Γ12	6
Δ19	11
Δ23	11
Δ25	
Π1	
Φ2	
Φ3	

Τσιπούρες	
Δ20	15
Δ21	12
Δ22	2
Δ24	20
Δ26	21
Φ1	25
Φ4	24
Φ5	21
Φ6	17

Στο λαβράκι στις οικογένειες Δ25, Π1, Φ2 και Φ3 δεν έγινε σύγκριση μεταξύ των προγραμμάτων διότι υπήρξε διαδικτυακό πρόβλημα στο σύστημα του προγράμματος Pedigree.

2 Γενετικές αποστάσεις μεταξύ γονιδιακών δεξαμενών – ανάλυση ετεροζυγωτίας

Αφού πραγματοποιήθηκε η ομαδοποίηση των δεδομένων των ατόμων για κάθε υπό-πληθυσμό ή αλλιώς γονιδιακή δεξαμενή, πρέπει να εξεταστεί ο βαθμός συσχέτισης μεταξύ των δεξαμενών, δηλαδή να υπολογιστεί η μεταξύ τους γενετική απόσταση. Για αυτό το σκοπό χρησιμοποιήθηκε το πρόγραμμα MSA με χρήση γλώσσας εντολών. Τα δεδομένα είναι οργανωμένα με όλες τις αλληλομορφικές αλληλουχίες των ατόμων από όλες τις δεξαμενές σε ένα συνολικό αρχείο Excel αλλά διαχωρισμένες ανά δεξαμενή με το χαρακτηριστικό "Pop" για να συγκρίνει τις γενετικές αποστάσεις μεταξύ των δεξαμενών.

Το MSA συγκρίνει κάθε μια δεξαμενή (population, Pop) με όλες τις υπόλοιπες του είδους αλλά και με τον εαυτό της και εντοπίζει την απόκλιση ή απόσταση μεταξύ τους χρησιμοποιώντας τη μεταβλητή Fst:

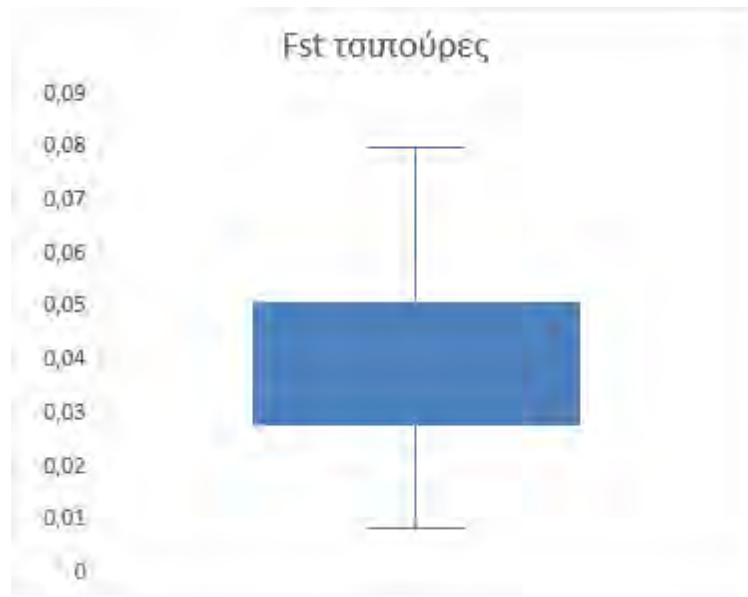
Λαβράκια	Γ11	Γ12	Δ19	Δ23	Δ25	Π1	Φ3
Γ11		0,002	0,022	0,022	0,029	0,019	0,029
Γ12	0,002		0,015	0,016	0,021	0,016	0,023
Δ19	0,022	0,015		0,004	0,019	0,016	0,016
Δ23	0,022	0,016	0,004		0,022	0,015	0,018
Δ25	0,029	0,021	0,019	0,022		0,034	0,041
Π1	0,019	0,016	0,016	0,015	0,034		0,003
Φ3	0,029	0,023	0,016	0,018	0,041	0,003	

Τσιπούρα	Δ20	Δ21	Δ22	Δ24	Δ26	Φ1	Φ4	Φ5	Φ6
Δ20		0,034	0,079	0,013	0,035	0,017	0,045	0,047	0,031
Δ21	0,034		0,052	0,017	0,008	0,038	0,033	0,024	0,036
Δ22	0,079	0,052		0,076	0,072	0,076	0,042	0,051	0,071
Δ24	0,013	0,017	0,076		0,011	0,026	0,048	0,043	0,042
Δ26	0,035	0,008	0,072	0,011		0,051	0,047	0,042	0,051
Φ1	0,017	0,038	0,076	0,026	0,051		0,045	0,047	0,015
Φ4	0,045	0,033	0,042	0,048	0,047	0,045		0,015	0,036
Φ5	0,047	0,024	0,051	0,043	0,042	0,047	0,015		0,038
Φ6	0,031	0,036	0,071	0,042	0,051	0,015	0,036	0,038	

Τα αποτελέσματα από τη σύγκριση αυτή φαίνονται στα παρακάτω γραφήματα:

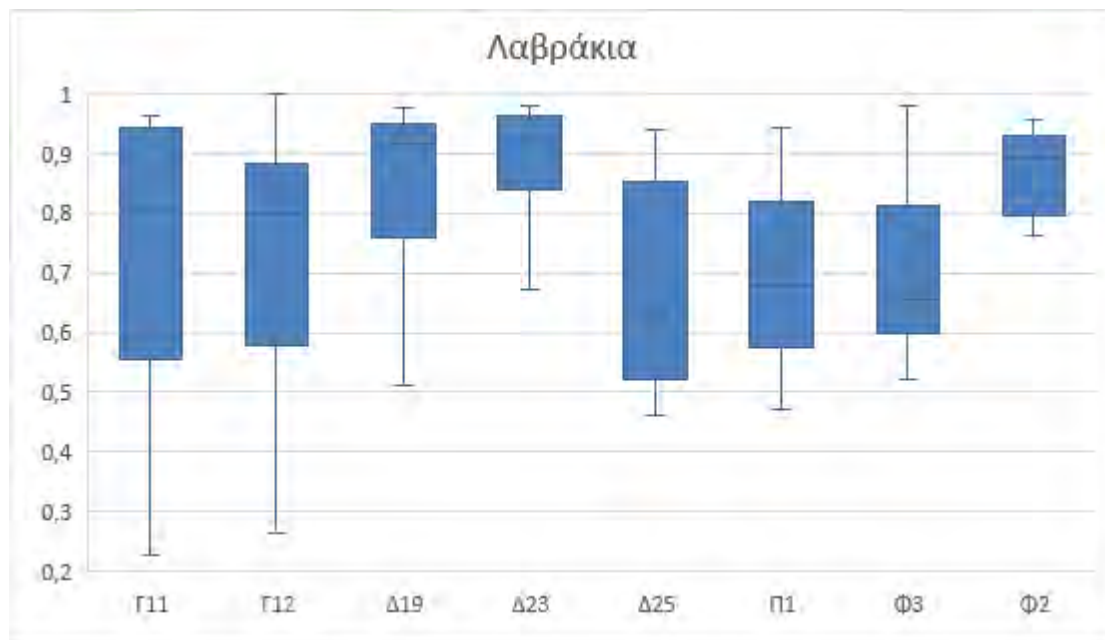


Υπολογισμός γενετικών αποστάσεων μεταξύ γονιδιακών δεξαμενών λαβρακιού και εύρος τιμών τους

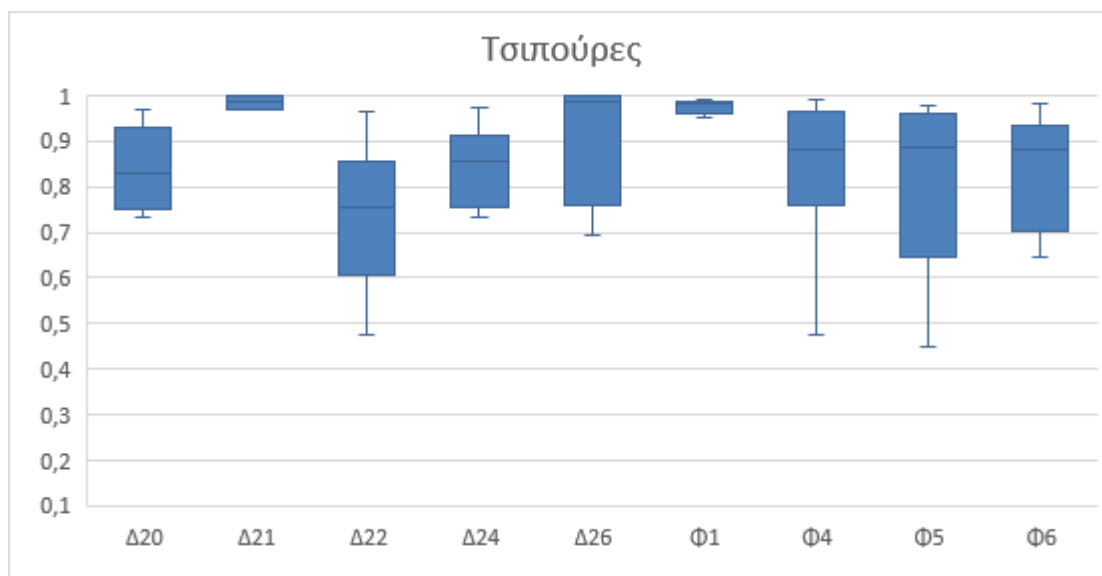


Υπολογισμός γενετικών αποστάσεων μεταξύ γονιδιακών δεξαμενών τσιπούρας και εύρος τιμών τους

Στη συνέχεια κρίνεται αναγκαίος ο υπολογισμός της μέσης τιμής του ποσοστού ετεροζυγωτίας που παρατηρείται στα δεδομένα ανάλυσης για κάθε γονιδιακή δεξαμενή. Τα αποτελέσματα αυτά παρουσιάζονται ανά γονιδιακή δεξαμενή σε ένα συνολικό αρχείο Excel από το οποίο υπολογίστηκε με συνάρτηση η μέση τιμή για κάθε δεξαμενή. Τα αποτελέσματα του υπολογισμού με τα ποσοστά παρατηρούμενης ετεροζυγωτίας φαίνονται στα παρακάτω γραφήματα:

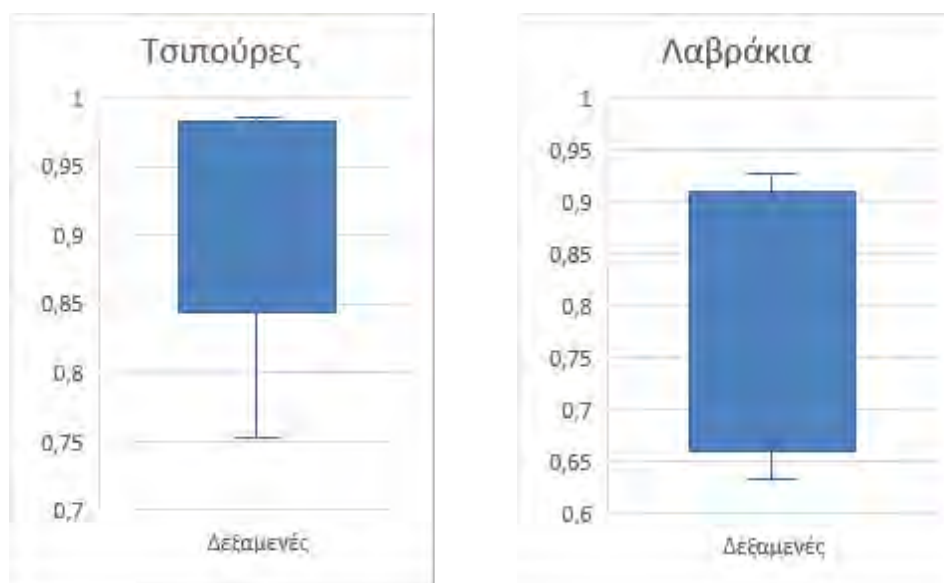


Εύρος τιμών ποσοστού ετεροζυγωτίας για κάθε γονιδιακή δεξαμενή στα λαβράκια



Εύρος τιμών ποσοστού ετεροζυγωτίας για κάθε γονιδιακή δεξαμενή στις τσιπούρες

Έπειτα, υπολογίστηκε και μια τιμή ποσοστού ετεροζυγωτίας για κάθε είδος συνολικά, ξεχωριστά για το λαβράκι και για την τσιπούρα, τα οποία ποσοστά αποτυπώνονται στο παρακάτω γράφημα:



Εύρος τιμών ποσοστού ετεροζυγωτίας από όλες τις γονιδιακές δεξαμενές του κάθε είδους

Στους παρακάτω πίνακες φαίνονται τα μεγέθη των γονιδιακών δεξαμενών με τη μέτρηση του λόγου αλληλόμορφα/γενετικό τόπο:

Δεξαμενές λαβράκι	αλληλόμορφα	άτομα	αλληλ/άτομο
Δ19	97	41	2,37
Δ23	101	49	2,06
Δ25	108	86	1,26
Φ3	84	92	0,91

Γ11	82	54	1,52
Γ12	76	53	1,43
Π1	84	72	1,17
Φ2	112	118	0,95

Δεξαμενές τσιπούρα	αλληλόμορφα	άτομα	αλληλ/άτομο
Δ20	136	65	2,09
Δ21	146	67	2,18
Δ22	150	80	1,88
Δ24	143	79	1,81
Δ26	151	72	2,10
Φ1	133	102	1,30
Φ4	162	142	1,14
Φ5	157	91	1,73
Φ6	153	105	1,46

3 Ταυτοποίηση γενετικής δομής – αρχικοί πληθυσμοί γονιδιακών δεξαμενών

Επόμενο βήμα αποτελεί ο καθορισμός του πιθανού αριθμού των αρχικών πληθυσμών και στα δύο είδη. Για αυτό το σκοπό χρησιμοποιήθηκε το πρόγραμμα Structure στο οποίο ορίζεται ένα σύνολο υποομάδων από τις οποίες αποτελείται ο πληθυσμός του κάθε είδους.

Ορίστηκε ο επιθυμητός αριθμός K που αντιστοιχεί στις υποομάδες, clusters, από τα οποία θα αποτελείται ο πληθυσμός κάθε είδους, όπως επίσης και οι φορές που το πρόγραμμα θα τρέξει την προγραμματισμένη ομαδοποίηση με βάση το K . Για αυτό το λόγο θεωρήθηκε αρχικά μια υποτιθέμενη ομαδοποίηση σε δυο υποομάδες και στη συνέχεια τέθηκε για κάθε είδος (λαβράκι και τσιπούρα ξεχωριστά) το K να είναι: $K=2 - K=n+1$, όπου n ο αριθμός των ξεχωριστών γονιδιακών δεξαμενών. Άρα $K=2-9$ για το λαβράκι και $K=2-10$ για την τσιπούρα, για την έναρξη του προγράμματος Structure. Οι κύκλοι για κάθε ανάλυση (για κάθε είδος) ορίστηκαν στους 20.

Μετά τον τερματισμό της ανάλυσης με τις παραπάνω παραμέτρους, εξήχθησαν ως αποτελέσματα η πιθανότητα για κάθε άτομο να ανήκει σε μια ομάδα από τις K που τέθηκαν καθώς και η τιμή Ln of Probability of Data (LnPD) για κάθε άτομο του πληθυσμού που αναλύθηκε κάθε φορά, δηλαδή μια τιμή LnPD για κάθε κύκλο για κάθε K .

Σύμφωνα όμως με αναλύσεις του G. Evanno και των συνεργατών του S. Regnaut, J. Goudet διαπιστώθηκε ότι το ΔK , μια επιπρόσθετη μεταβλητή που σχετίζεται με έναν δεύτερης τάξης ρυθμό αλλαγής των τιμών LnPD σε σχέση με τον αριθμό των υποομάδων – clusters, είναι καλός παράγοντας πρόβλεψης του

πραγματικού αριθμού υποομάδων, του πραγματικού K . Πιο αναλυτικά με βάση τα δεδομένα της παρούσας εργασίας υπολογίστηκαν:

$$\Delta K = m(|L''(K)|) / s[L(K)]$$

όπου $|L''(K)| = |L'(K+1) - L'(K)|$, $L'(K) = L(K) - L(K-1)$, $s[L(K)]$ η τυπική απόκλιση των τιμών L_nPD για κάθε K ξεχωριστά, δηλαδή το μέτρο της διασποράς των τιμών σε σχέση με την τιμή του μέσου όρου (αριθμητικού μέσου) και $L(K) =$ ο μέσος όρος των τιμών L_nPD των 20 κύκλων για κάθε K ξεχωριστά, που ορίστηκε για κάθε είδος. Άρα ο αρχικός τύπος επεκτείνεται σε:

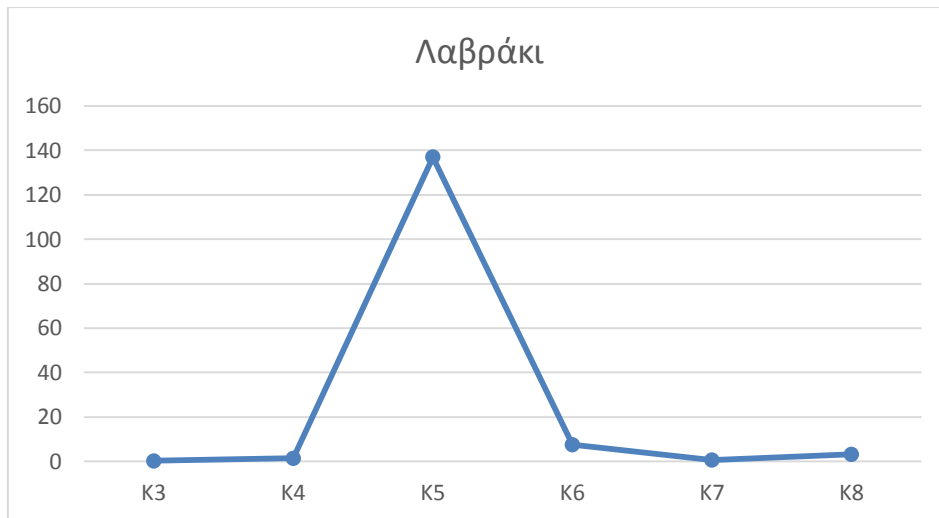
$$\Delta K = m(|L(K+1) - 2L(K) + L(K-1)|) / s[L(K)]$$

Έτσι με βάση τους παραπάνω τύπους εξάγονται ως αποτελέσματα οι ακόλουθες τιμές ΔK για το λαβράκι και την τσιπούρα:

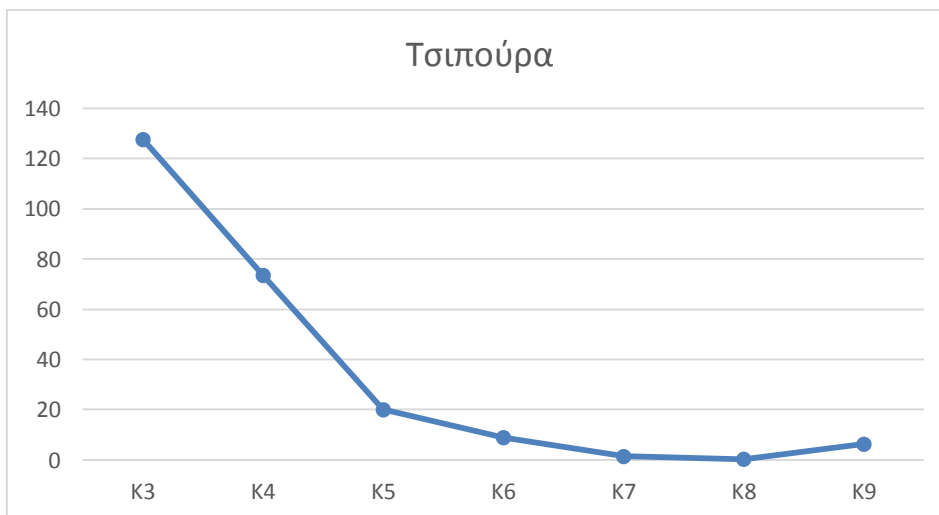
Λαβράκι	L(K)	s[L(K)]	L'(K)	L''(K)	ΔK
K2	-19367,01	1,08			
K3	-19104,82	269,69	262,19	65,54	0,24
K4	-18908,17	358,75	196,65	511,74	1,43
K5	-18199,78	4,16	708,39	570,89	137,08
K6	-18062,28	4,65	137,50	35,13	7,56
K7	-17959,91	37,25	102,37	25,29	0,68
K8	-17882,83	20,02	77,08	64,49	3,22
K9	-17870,24	29,46	12,59		

Τσιπούρα	L(K)	s[L(K)]	L'(K)	L''(K)	Δ(K)
K2	-34178,02	1,29			
K3	-33292,49	2,70	885,53	345,33	127,67
K4	-32752,29	4,52	540,20	332,65	73,54
K5	-32544,74	3,61	207,55	72,32	20,01
K6	-32409,51	4,24	135,23	37,73	8,90
K7	-32312,01	32,51	97,50	44,82	1,38
K8	-32259,33	68,65	52,68	20,38	0,30
K9	-32227,03	21,38	32,30	134,36	6,28
K10	-32329,09	46,28	-102,06		

Διαχωρίστηκε το $m(|L'(K)|)$ από το $s[L(K)]$ επειδή βρέθηκε μία σαφή και γενική τάση προς αύξηση της διακύμανσης $L(K)$ ανάμεσα στους 20 κύκλους και όσο αυξανόταν η τιμή του K . Διαπιστώθηκε ότι η μέγιστη τιμή της κατανομής του ΔK που αντιστοιχεί στην κορυφή του γραφήματος, πρέπει να αντιστοιχεί στην πραγματική τιμή K . Με βάση αυτά έχουμε τα παρακάτω γραφήματα με τις κορυφές τους για το λαβράκι και την τσιπούρα:



Κατανομή της μεταβλητής ΔK στο λαβράκι για την εύρεση του πραγματικού K που αντιστοιχεί στην κορυφή του γραφήματος



Κατανομή της μεταβλητής ΔK στην τσιπούρα για την εύρεση του πραγματικού K που αντιστοιχεί στην κορυφή του γραφήματος

Επομένως για το είδος λαβράκι η πραγματική τιμή K είναι K=5 και για την τσιπούρα K=3. Αυτές οι τιμές δηλώνουν τους αντίστοιχους αρχικούς πληθυσμούς από τους οποίους προήλθαν τα άτομα των γονιδιακών δεξαμενών που εξετάστηκαν και αντιστοιχούν περισσότερο στην πραγματικότητα.

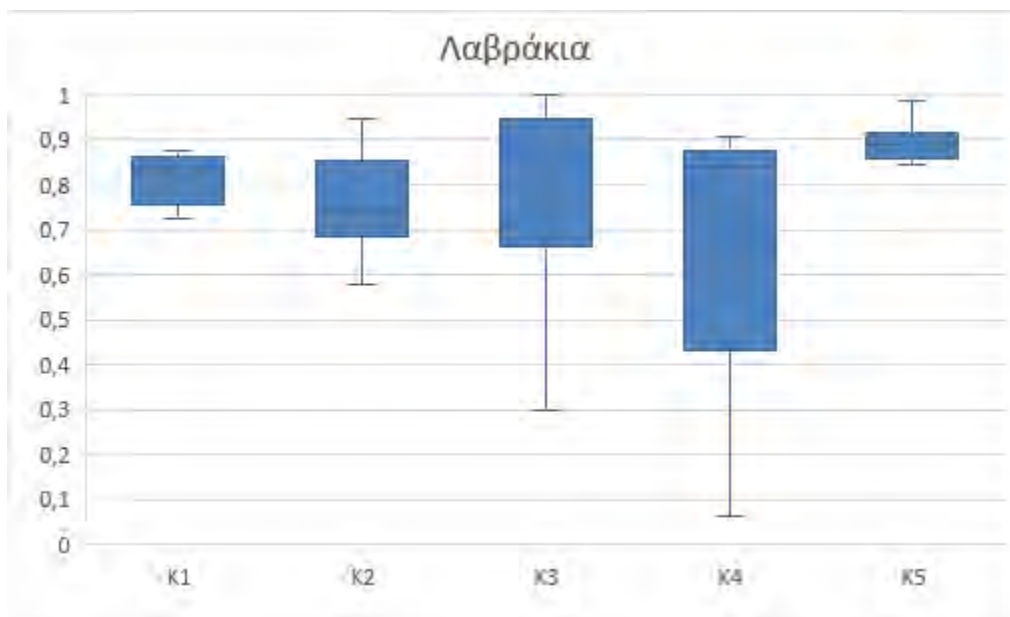
Στη συνέχεια και αφού βρέθηκαν οι πραγματικές τιμές K, ξαναχρησιμοποιήθηκε το Structure με τις παραμέτρους που αναφέρθηκαν και παραπάνω με τη διαφορά ότι η ορισμένη τιμή K τέθηκε K=5 για το λαβράκι και K=3 για την τσιπούρα. Με αυτή την ανάλυση εξάγονται πλέον ως αποτελέσματα η ομαδοποίηση των γονιδιακών δεξαμενών του κάθε είδους σε συγκεκριμένους αρχικούς πληθυσμούς (πραγματικό K) που ορίστηκε (5 αρχικοί πληθυσμοί για το λαβράκι και 3 για την τσιπούρα), αλλά και είναι δυνατός ο εντοπισμός των ατόμων που ανήκουν σε κάθε έναν αρχικό πληθυσμό – υποομάδα – cluster, σε κάθε K. Σε συνδυασμό με τα αλληλομορφικά δεδομένα από μικροδορυφόρους των

γονιδιακών δεξαμενών των υπό μελέτη ειδών, είναι δυνατή η εύρεση και των αλληλομορφικών αλληλουχιών των ατόμων που ανήκουν σε κάθε cluster. Άρα γίνονται πλέον γνωστοί με πληροφορίες αλληλομόρφων οι αρχικοί πληθυσμοί από τους οποίους προήλθαν τα άτομα που μελετώνται.

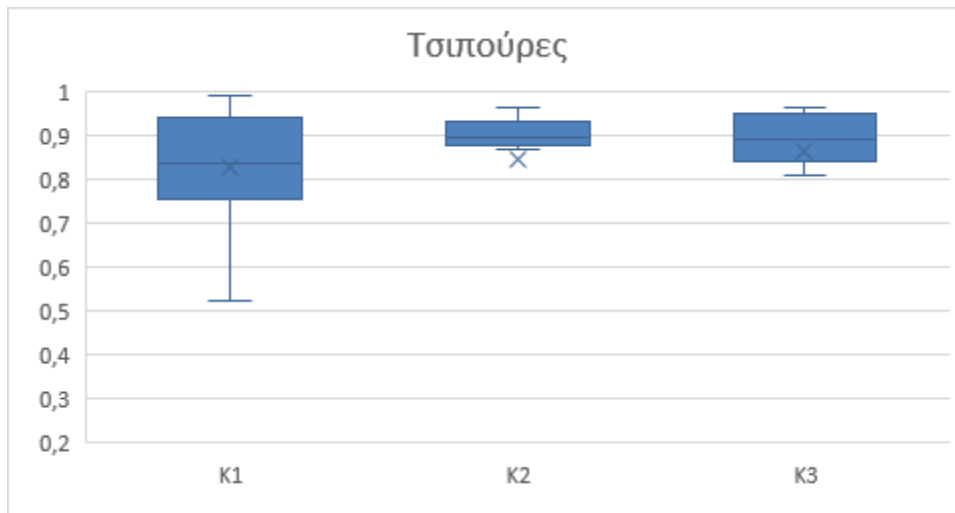
4 Ανακατασκευή γενεαλογίας – γενετικές αποστάσεις αρχικών πληθυσμών

Χρησιμοποιώντας τα αποτελέσματα της ανάλυσης του προγράμματος Structure για την εύρεση των αρχικών πληθυσμών, καθίσταται δυνατός ο καθορισμός τους από αλληλομορφικά δεδομένα. Με αυτό τον τρόπο και εκτελώντας το πρόγραμμα MSA με δεδομένα όμως αυτή τη φορά τα άτομα των 5 υποομάδων για το λαβράκι και των 3 υποομάδων για την τσιπούρα, υπολογίζεται η μεταξύ των αρχικών πληθυσμών γενετική απόσταση.

Αρχικά όπως προαναφέρθηκε εξάγονται ως αποτελέσματα η παρατηρούμενη και προσδοκώμενη ετεροζυγωτία ανάμεσα στους αρχικούς πληθυσμούς που εξετάζονται κάθε φορά. Έτσι υπολογίζοντας τη μέση τιμή της παρατηρούμενης ετεροζυγωτίας για κάθε γενετικό τόπο κάθε είδους εξάγονται τα ποσοστά ετεροζυγωτίας στα παρακάτω γραφήματα και κατά πόσο αποκλίνουν από τη μέση τιμή:



Εύρος τιμών ποσοστού ετεροζυγωτίας για κάθε αρχικό πληθυσμό K στα λαβράκια



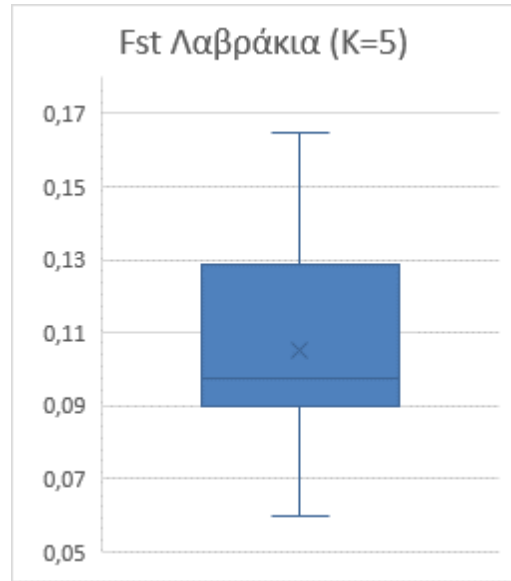
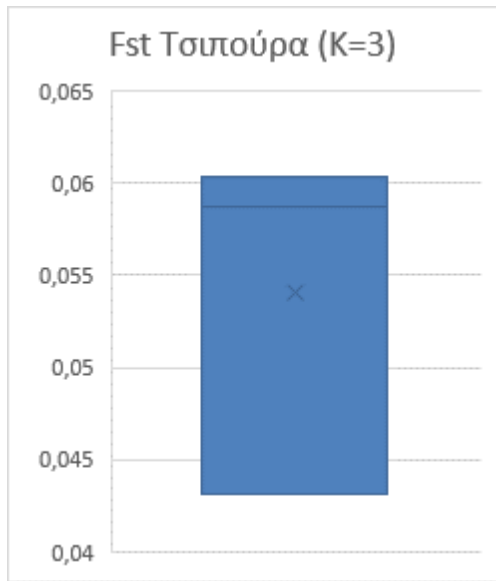
Εύρος τιμών ποσοστού ετεροζυγωτίας για κάθε αρχικό πληθυσμό K στις τσιπούρες

Στη συνέχεια χρησιμοποιώντας τις τιμές F_{st} που εξάγονται ως αποτέλεσμα της ανάλυσης του MSA, υπολογίζεται η γενετική απόσταση μεταξύ των υπό μελέτη πληθυσμών. Η σύγκριση απεικονίζεται στον παρακάτω πίνακα για κάθε είδος:

Λαβράκια ($K=5$)	K1	K2	K3	K4	K5
K1		0,098	0,081	0,102	0,060
K2	0,098		0,130	0,129	0,093
K3	0,081	0,130		0,165	0,097
K4	0,102	0,129	0,165		0,096
K5	0,060	0,092	0,097	0,096	

Τσιπούρα ($K=3$)	K1	K2	K3
K1		0,059	0,043
K2	0,059		0,060
K3	0,043	0,060	

Τα αποτελέσματα της παραπάνω σύγκρισης αποτυπώνονται στα παρακάτω γραφήματα στα οποία φαίνεται η τυπική τους απόκλιση με βάση τη σύγκριση:



Εύρος τιμών των γενετικών αποστάσεων των αρχικών πληθυσμών (K) του κάθε είδους

IV. ΣΥΜΠΕΡΑΣΜΑΤΑ – ΣΥΖΗΤΗΣΗ

Η παραπάνω μελέτη πραγματοποιήθηκε για την καλύτερη οργάνωση και διαχείριση των εκτρεφόμενων πληθυσμών του λαβρακιού και της τσιπούρας που μελετήθηκαν. Για το σκοπό αυτό μέσα από τις αναλύσεις εξήχθησαν τα παρακάτω συμπεράσματα μέσα από τις συγκρίσεις των γονιδιακών δεξαμενών του κάθε είδους αξιοποιώντας συγκριτικούς δείκτες που παρουσιάζονται παρακάτω.

1 Σύγκριση των προγραμμάτων Colony, Pedigree και Kinalyzer

Από τη σύγκριση των οικογενειών που δημιουργούν τα δύο αυτά προγράμματα εύρεσης συγγενικών σχέσεων ανάμεσα σε μέλη ενός πληθυσμού, παρατηρείται μια ομοιότητα σε σχέση με τις οικογένειες που εξάγει το Kinalyzer. Η ομοιότητα αυτή προκύπτει εξετάζοντας τα άτομα αυτών των οικογενειών και διαπιστώνοντας το πλήθος των κοινών ατόμων. Η διαφορά με το πρόγραμμα Kinalyzer με τα άλλα δύο προκύπτει από τη διαφορετική προσέγγιση που χρησιμοποιεί και τον διαφορετικό αλγόριθμο για την ανάλυση των δεδομένων. Επιπλέον, σε αντίθεση με τις μεθόδους πιθανότητας ανακατασκευής του Colony και Pedigree, το Kinalyzer δεν απαιτεί πληροφορίες σχετικά με τις συχνότητες των αλληλομόρφων του πληθυσμού και δεν έχει ως δεδομένα τα σύστημα ζευγαρώματος του είδους.

2 Δείκτης πολυμορφίας πληθυσμών

Σημαντικό δείκτη πολυμορφίας των γενετικών τόπων που μελετώνται σε κάθε είδος και για κάθε πληθυσμό αποτελεί η εξέταση που λόγω αλληλόμορφα/άτομο καθώς και η μέση τιμή της ετεροζυγωτίας για κάθε τόπο κάθε γονιδιακής δεξαμενής. Ο λόγος αυτός δείχνει το κατά πόσο σχετίζεται το μέγεθος του πληθυσμού με το πλήθος των αλληλομόρφων για του γενετικούς τόπους και επομένως πόσο πολυμορφικοί είναι αυτοί οι πληθυσμοί σε συνδυασμό με την τιμή της ετεροζυγωτίας για κάθε γενετικό τόπο. Παρατηρούνται στην παρούσα εργασία τα εξής:

Λαβράκι	Δεξαμενή	Άτομα	Αλληλ/άτομο	Mean Hobs
max	Δ19	41	2,37	0,85
min	Φ3	84	0,91	0,70

Τσιπούρες	Δεξαμενή	Άτομα	Αλληλ/άτομο	Mean Hobs
max	Δ21	67	2,18	0,87
min	Φ4	142	1,14	0,73

Η πιο πολυμορφική γονιδιακή δεξαμενή στα λαβράκια δεν είναι απαραίτητα και η μεγαλύτερη όπως ακριβώς και στις τσιπούρες. Αντιθέτως, παρατηρείται ότι η σχεδόν διπλάσιου μεγέθους δεξαμενή έχει λιγότερα αλληλόμορφα σε κάθε γενετικό τόπο ανά άτομο.

Βιβλιογραφία

- ASHLEY, M. V., I. C. CABALLERO, W. CHAOVALITWONGSE, B. DASGUPTA, P. GOVINDAN, S. I. SHEIKH, and T. Y. BERGER-WOLF. 2009. "KINALYZER, a Computer Program for Reconstructing Sibling Groups." *Molecular Ecology Resources* 9 (4): 1127–31. <https://doi.org/10.1111/j.1755-0998.2009.02562.x>.
- Dakin, E. E., and J. C. Avise. 2004. "Microsatellite Null Alleles in Parentage Analysis." *Heredity* 93 (5): 504–9. <https://doi.org/10.1038/sj.hdy.6800545>.
- DIERINGER, D. SCHLÖTTERER C. 2003. "Microsatellite Analyser (MSA): A Platform Independent Analysis Tool for Large Microsatellite Data Sets." *Molecular Ecology Notes*. <https://doi.org/10.1046/j.1471-8286>.
- Evanno, G., S. Regnaut, and J. Goudet. 2005. "Detecting the Number of Clusters of Individuals Using the Software STRUCTURE: A Simulation Study." *Molecular Ecology* 14 (8): 2611–20. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>.
- Jones, Adam G., Clayton M. Small, Kimberly A. Paczolt, and Nicholas L. Ratterman. 2010. "A Practical Guide to Methods of Parentage Analysis." *Molecular Ecology Resources* 10 (1): 6–30. <https://doi.org/10.1111/j.1755-0998.2009.02778.x>.
- Jones, Owen R., and Jinliang Wang. 2010. "COLONY: A Program for Parentage and Sibship Inference from Multilocus Genotype Data." *Molecular Ecology Resources* 10 (3): 551–55. <https://doi.org/10.1111/j.1755-0998.2009.02787.x>.
- Porras-Hurtado, Liliana, Yarimar Ruiz, Carla Santos, Christopher Phillips, Angel Carracedo, and Maria V. Lareu. 2013. "An Overview of STRUCTURE: Applications, Parameter Settings, and Supporting Software." *Frontiers in Genetics* 4 (MAY): 1–13. <https://doi.org/10.3389/fgene.2013.00098>.
- Pritchard, Jonathan K. 2010. "Documentation for Structure Software : Version 2 . 3" 6 (3): 321–26. <https://doi.org/10.1002/spe.4380060305>.
- Städele, Veronika, and Linda Vigilant. 2016. "Strategies for Determining Kinship in Wild Populations Using Genetic Data." *Ecology and Evolution* 6 (17): 6107–20. <https://doi.org/10.1002/ece3.2346>.
- Wang, Jinliang. 2016. "User ' s Guide for Software COLONY" 2: 1–72.
- Webster, Michael S., and Letitia Reichart. 2005. "Use of Microsatellites for Parentage and Kinship Analyses in Animals." *Methods in Enzymology* 395: 222–38. [https://doi.org/10.1016/S0076-6879\(05\)95014-3](https://doi.org/10.1016/S0076-6879(05)95014-3).