



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΚΡΗΤΗΣ
UNIVERSITY
OF CRETE

Πανεπιστήμιο Κρήτης
τμήμα Βιολογίας

Διδακτορική διατριβή

Ανάπτυξη και εφαρμογή γενετικών εργαλείων με στόχο τη μελλοντική βελτίωση της παραγωγής σε δύο νέα είδη της Μεσογειακής ιχθυοκαλλιέργειας, τον κρανιό (*Argyrosomus regius*) και το μαγιάτικο (*Seriola dumerili*)

Υποψήφιος διδάκτορας
Νούσσιας Ορέστης

Επιβλέπων καθηγητής : **Παυλίδης Μιχαήλ**

Επιβλέπων ερευνητής : **Τσιγγενόπουλος Κωνσταντίνος**

Μάιος 2021

Περιεχόμενα

Περίληψη	6
Abstract	6
1. Εισαγωγή	7
1.1 Βιολογία	7
1.2 Ο ρόλος της ιχθυοκαλλιέργειας στο σήμερα	8
1.3 Γενετική επιλογή σε Μεσογειακά είδη εκτρεφόμενων ιχθύων	9
1.4 Κρανιός	9
1.5 Μαγιάτικο	10
1.6 Γενετική βάση φαινοτυπικής παραλλακτικότητας	10
1.7 Αναλυτικές προσεγγίσεις	11
1.8 Πολυμορφικοί δείκτες με γονιδιωματικό εύρος κάλυψης	11
1.9 Έρευνα για QTL σε τελεόστεους ιχθείς	11
2. Γενετική περιγραφή, πληθυσμιακή δομή και ανάλυση συγγένειας αποθεμάτων γεννητόρων και απογόνων ως βάση για τον σχεδιασμό προγραμμάτων επιλογής	12
2.1 Εισαγωγή	12
2.1.1 Προγράμματα επιλογής, ένα πρώτο σχόλιο	13
2.1.2 Το ζήτημα της ενδογαμίας	14
2.1.3 Γενετική παρέκκλιση (genetic drift)	18
2.1.4 Ανάλυση σχέσεως γονέων-απογόνων	19
2.1.5 Μέθοδοι ανάλυσης	19
2.1.5.1 Αποκλεισμός (exclusion method)	20
2.1.5.2 Κατηγορική ανάλυση ανάθεσης (categorical allocations)	20
2.1.5.3 Μπαϋεσιανή μέθοδος	21
2.1.6 Ανακατασκευή των σχέσεων γονέων- απογόνων (sibship reconstruction and parentage assignment)	22
2.2 Δύο τρόποι σχεδιασμού επιλογής γεννητόρων, βασισμένες σε φαινοτυπικά δεδομένα: στην πρώτη περίπτωση το φύλο, και στη δεύτερη περίπτωση αποκλειστικά στην γενετική περιγραφή προερχόμενη από μικροδορυφορικούς γενετικούς δείκτες.	23
2.2.1 Περίπτωση 1: Λαμβάνοντας υπόψη τα δεδομένα για το φύλο και το βάρος των γονιών και των απογόνων, ως κριτήρια για την επιλογή γεννητόρων	23
2.2.2 Περίπτωση 2. Διαθέσιμα δεδομένα προερχόμενα αποκλειστικά από γενετικά δεδομένα (μικροδορυφορικοί γενετικοί δείκτες)	26
2.2.3 Σύγκριση των 2 περιπτώσεων/προσεγγίσεων για το μαγιάτικο	30
2.2.4 Γενετική περιγραφή, πληθυσμιακή δομή και ανάλυση συγγένειας σε 11 αποθέματα κρανιού, ως μία πληροφοριακή αφετηρία νέων προγραμμάτων αναπαραγωγής	33
2.2.6 Δείγματα	33
2.2.7 Εξαγωγή DNA, PCR, γονοτύπηση μικροδορυφορικών τόπων	33
	2

2.2.8 Πληθυσμιακή ανάλυση	35
2.2.8.1 Γενετική παραλλακτικότητα και ισορροπία Hardy-Weinberg (HWE)	35
2.2.8.2 Πληθυσμιακή δομή	36
2.2.9 Ανάλυση σχέσεων συγγένειας	37
2.3 Αποτελέσματα & συζήτηση	39
3. Ανάθεση γονέων, εκτίμηση της κληρονομησιμότητας, των γενετικών και φαινοτυπικών συσχετίσεων για αναπτυξιακούς χαρακτήρες στον κρανιό	48
3.1 Κληρονομησιμότητα και το ζωϊκό μοντέλο	48
3.1.1 Κληρονομησιμότητα	48
3.1.2 Γραμμικά μοντέλα και το ζωϊκό μοντέλο (<i>animal model</i>)	50
3.2 Ανάλυση απογόνων κρανιού ως προς την κληρονομησιμότητα των χαρακτήρων βάρος και μήκος, των γενετικών και φαινοτυπικών τους συσχετίσεων και του δραστικού μεγέθους πληθυσμού N_e	53
3.2.1 Εισαγωγή	53
3.2.2 Υλικά και μέθοδοι	54
3.2.2.1 Βιολογικό υλικό	54
3.2.2.2 Εξαγωγή DNA, συνθήκες PCR	55
3.2.2.3 Ανάλυση ανάθεσης γονέων	55
3.2.2.4 Ανάλυση δεδομένων	56
3.3 Αποτελέσματα	56
3.3.1 Ανάλυση ανάθεσης γονέων	56
3.3.2 Δραστικό μέγεθος πληθυσμού N_e	59
3.3.3 Εκτίμηση ετερογένειας των διακυμάνσεων για το βάρος και μήκος	60
3.3.4 Εκτίμηση κληρονομησιμότητας, γενετικών και φαινοτυπικών συσχετίσεων	60
3.3.5 Στατιστικά σημαντικές διαφορές στο βάρος και μήκος για τις κοινές οικογένειες των 2 παρτίδων	61
3.4 Συζήτηση	62
4. Γενετικός χάρτης σύνδεσης, ανάλυση ανίχνευσης QTL για τους χαρακτήρες βάρος και μήκος, και συγκριτική γονιδιωματική ανάλυση με 5 γονιδιώματα τελεόστων ιχθύων, στον κρανιό	66
4.1 Εισαγωγή	66
4.2 Γενετικός χάρτης σύνδεσης	69
4.2.1 Κατασκευή γενετικού χάρτη σύνδεσης	69
4.3 Ανάλυση QTL	70
4.3.1 Το πρόβλημα των αγνοούμενων δεδομένων στην ανάλυση QTL	71
4.3.2 Η προσέγγιση στην παρούσα μελέτη	73
4.5 Βιβλιοθήκη ddRAD για παραγωγή δεδομένων γονιδιωματικής εκπροσώπησης	73
4.6 Υλικά και μέθοδοι	74
4.6.1 Δείγματα, εξαγωγή DNA, PCR και ανάλυση ανάθεσης γονέων σε απογόνους	74

4.6.2 Κατασκευή βιβλιοθήκης ddRAD	74
4.6.3 Ποιοτικός έλεγχος και παραγωγή γονοτύπων μέσω του λογισμικού STACKS	75
4.6.4 Ποιοτικός έλεγχος με το Plink	75
4.6.5 Κατασκευή γενετικού χάρτη σύνδεσης	75
4.6.6 Ανάλυση QTL	76
4.6.7 Συγκριτική γονιδιωματική ανάλυση	76
4.6.8 Αποτελέσματα & συζήτηση	76
5.1 Ευρήματα της διατριβής	92
5.2 Συνεισφορά της διατριβής και χρηματοδότηση	93
Παράρτημα	94
Βιβλιογραφία	97

Ευχαριστίες

Πρώτα και κύρια, θα ήθελα να πω ένα ευχαριστώ, στους δύο επιβλέποντες της διδακτορικής διατριβής. Στον Μιχάλη Παυλίδη για την θετική του στάση και συνεισφορά κατά την διάρκεια της ερευνητικής δουλειάς και στον Κώστα Τσιγγενόπουλο για την καθοδήγηση του στην υλοποίηση της έρευνας σε αυτή την διατριβή, την εμπιστοσύνη του, την στήριξη του, την θετική του διάθεση και ενθουσιασμό για το υπο μελέτη αντικείμενο. Το τμήμα Βιολογίας του πανεπιστημίου Κρήτης και το ΕΛΚΕΘΕ Κρήτης είναι σημαντικές ερευνητικές υποδομές με παραγόμενη έρευνα υψηλών προδιαγραφών και σημαντικό αντίκτυπο στην κοινωνία και οικονομία, και είμαι περήφανος που διετέλεσα κομμάτι αυτής της αφήγησης. Ξεκινώντας από την γραμματεία, οφείλω ένα μεγάλο ευχαριστώ στην Ευφροσύνη Μπερβανάκη που από την αρχή μέχρι το τέλος ήταν πολύ ευγενική και εξυπηρετική. Τους τεχνικούς του εργαστηρίου του IMBBC του ΕΛΚΕΘΕ, εφαρμοσμένης γονιδιωματικής, Βάσω Τερζόγλου και Κατερίνα Οικονομάκη, με τις οποίες γνωριστήκαμε στενά και συνεργαστήκαμε τα χρόνια της διδακτορικής διατριβής. Οι δύο τους είναι πραγματικά σημαντικοί πυλώνες στήριξης του εργαστηρίου και του IMBBC συνολικά. Να ευχαριστήσω τον Στέλιο Δαριβιανάκη, ο οποίος με την τεχνική του υποστήριξη σε σειρά πειραμάτων αλλά και γενικότερης υλικοτεχνικής υποδομής έχει συμβάλλει τα μέγιστα στην ομαλή λειτουργία του εργαστηρίου και του ΕΛΚΕΘΕ Κρήτης συνολικότερα. Ένα ιδιαίτερο ευχαριστώ στον Δημήτρη Τσαπάρη, που περάσαμε πραγματικά πολλές στιγμές στο ΕΛΚΕΘΕ στην διεξαγωγή πειραμάτων, τον οποίο εκτιμώ και συμπαθώ ιδιαίτερα για την συγκρότηση και το ήθος του. Ένα ευχαριστώ στον Γιόν Κριστόφερσεν με τον οποίο συζητήσαμε πάμπολες φορές για την αλληλούχηση και τις διάφορες πτυχές της, και ιδιαίτερα για νέες τεχνολογίες, ήταν πολύ ευχάριστη και εποικοδομητική η αλληλεπίδραση μας. Να ευχαριστήσω τον Παναγιώτη Κασαπίδη για την φιλία και την βοήθεια του σε διάφορα ζητήματα καθώς επίσης την Τερέζα Μανουσάκη και τον Αλέξανδρο Τσακογιάννη για την απτή και κρίσιμη βοήθεια τους σε μία σειρά ζητημάτων πειραματικών και βιοπληροφορικών. Να ευχαριστήσω επίσης τους υπεύθυνους του υπολογιστικού cluster του ΕΛΚΕΘΕ Κρήτης, και ειδικά τον Στέλιο Νινιδάκη, με τον οποίο αλληλεπίδρασα, του οποίου η καλή διάθεση και διαθεσιμότητα είναι αξιοζήλευτα στοιχεία. Να ευχαριστήσω τον Δημήτρη Χατζηπλή, που πολύ συμπαθώ και εκτιμώ, για την βοήθεια του σε κομμάτια των αναλύσεων ποσοτικής γενετικής καθώς και την Σταυρούλα Οικονόμου για την συνεισφορά της στην ερευνητική δουλειά. Επίσης ένα ευχαριστώ στον Βασίλη Παπαδογιάννη και στην Νελίνα Αντζέλοβα για την βοήθεια και συνεισφορά τους. Θα ήθελα να ευχαριστήσω την γλυκιάτη Ευτυχία Μυρωνάκη, γραμματέα του IMBBC τον Μαρίνο Χουρδάκη και όλους τους εργαζόμενους στο ΕΛΚΕΘΕ Κρήτης ξεπερνώντας σειρά δυσκολιών και προσπαθώντας ως προς την επίλυση των καθημερινών προβλημάτων. Ευχαριστώ τον Νίκο Παπανδρουλάκη για τις προτάσεις διόρθωσης του κειμένου και την συνεργασία, καθώς και τους υπόλοιπους συνεργάτες του τμήματος των Aqualabs του ΕΛΚΕΘΕ. Ένα μεγάλο ευχαριστώ στον Γιώργο Κοτούλα, με τον οποίο επίσης συνδέθηκα πολύ στενά κατά την διάρκεια αυτής της διατριβής και με τον οποίο μοιραστήκαμε πάρα πολλές επιστημονικές, φιλοσοφικές, πολιτιστικές και κοινωνικές συζητήσεις. Ο Γιώργος αποτέλεσε συν τοις άλλοις, έναν άτυπο επιστημονικό «μέντορα» συμβάλλοντας σημαντικά στην θεωρητική μου ενδυνάμωση στην γενετική/γονιδιωματική και τροφοδότηση μου με σημαντικά βιβλία και άρθρα τα οποία με διαμόρφωσαν επιστημονικά. Τέλος ευχαριστώ ιδιαίτερα τα υπόλοιπα μέλη της επταμελούς επιτροπής, τον Χαράλαμπο Σπηλιανάκη για την θετική και εποικοδομητική του στάση, τον Νίκο Πουλακάκη, και ειδικά τους Χρήστο Δελιδάκη και Δημήτρη Χατζηπλή για την συμβολή τους στην βελτίωση του κειμένου, και στην διάνθιση κομματιών της συζήτησης του γενετικού χάρτη σύνδεσης.

Περίληψη

Η ανάπτυξη νέων γενετικών εργαλείων, όπως και η βέλτιστη χρήση των υπαρχόντων για την βελτίωση της παραγωγής στην ιχθυοκαλλιέργεια για τα είδη Κρανιό και Μαγιάτικο υπήρξε ο κεντρικός ερευνητικός στόχος στην παρούσα διδακτορική διατριβή. Ο εκσυγχρονισμός στον σχεδιασμό των προγραμμάτων αναπαραγωγής καθώς και η γενετική επιλογή υποβοηθούμενη από πολυμορφικούς δείκτες, είναι προϋποθέσεις οι οποίες στην υλοποίησή τους δίνουν σημαντικές δυνατότητες στην παραγωγική διαδικασία. Αναπτύξαμε δύο multiplex γενετικών τόπων για τον Κρανιό και το Μαγιάτικο, αντίστοιχα, τα οποία χρησιμοποιήσαμε για την εκτίμηση της γενετικής παραλλακτικότητας 11 αποθεμάτων αποτελούμενων από 946 ιχθύες σε συνθήκες εκτροφής, για την διερεύνηση των συνεπειών της ανάμιξης αποθεμάτων καθώς και της προκύπτουσας γενετικής ομαδοποίησης μέσω παραμέτρων πληθυσμιακής γενετικής. Προτείναμε μία μεθοδολογία εκτίμησης σχέσεων συγγένειας σε πολλαπλά αποθέματα, με βάση την οποία εκτιμήσαμε ατομικούς συντελεστές ενδογαμίας, όπως επίσης συγκρίναμε δύο μεθόδους επιλογής γεννητόρων από διασταυρώσεις 27 γονέων από ένα απόθεμα. Επιπρόσθετα, διερευνήσαμε την κληρονομησιμότητα για τους χαρακτήρες βάρος και μήκος στον Κρανιό, τις γενετικές και φαινοτυπικές συσχετίσεις βάρους και μήκους καθώς και την διακύμανση στην συνεισφορά των γεννητόρων από μία γέννα 800 απογόνων στην διαμόρφωση των F1 οικογενειών σε συνάρτηση με το δραστικό μέγεθος πληθυσμού. Η ανίχνευση στοιχείων συγκρότησης του γονιδιώματος ενός είδους μαζί με την περιγραφή της γενετικής βάσης σημαντικών χαρακτήρων μπορεί να συνεισφέρει στην διατήρηση του είδους και στις δυνατότητες παραγωγικής του αξιοποίησης. Ένας γενετικός χάρτης σύνδεσης για ένα είδος, αναδεικνύει στοιχεία εξέλιξης του γονιδιώματος, διευκολύνει την χαρτογράφηση χαρακτήρων μέσω της εύρεσης QTL, και βελτιώνει την συναρμολόγηση του γονιδιώματος. Χρησιμοποιήσαμε μία τεχνική μειωμένης γονιδιωματικής εκπροσώπησης μέσω της δημιουργίας βιβλιοθήκης ddRAD για έξι F1 οικογένειες στον Κρανιό σε ένα σύνολο 266 ιχθύων, για την κατασκευή γενετικού χάρτη υψηλής πυκνότητας αποτελούμενο από 4529 πολυμορφικούς SNP δείκτες. Διεξαγάγαμε ανάλυση QTL για το βάρος και το μήκος, η οποία ανίχνευσε ένα στατιστικό σημαντικά QTL στο linkage group 15. Ο γενετικός χάρτης χρησιμοποιήθηκε για την συγκριτική γονιδιωματική ανάλυση, μέσω εύρεσης ομολογιών, σε πέντε γονιώματα τελεόστεων ιχθύων και ανέδειξε έναν μία εξελικτικά συντηρημένη δομή του καρυοτύπου στον Κρανιό. Οι παρατηρούμενες συντενίες, συνέκλιναν με δημοσιευμένα δεδομένα για την κοντινή εξελικτικά απόσταση της οικογένειας Scianidae με την οικογένεια Sparidae και συγκεκριμένα με την τσιπούρα, με την οικογένεια Moronidae και το λαβράκι, ανίχνευσαν πιθανές χρωμοσωμικές αναδιατάξεις στον Κρανιό και ειδικά στο linkage group 17. Η διατριβή παρέχει στο σύνολο της, νέα γενετικά εργαλεία για τον Κρανιό και Μαγιάτικο για τον σχεδιασμό προγραμμάτων αναπαραγωγής, τον πρώτο γενετικό χάρτη σύνδεσης στον Κρανιό προερχόμενο από NGS δεδομένα, για πρώτη φορά ανιχνεύει την γενετική βάση των αναπτυξιακών χαρακτήρων βάρος και μήκος στον Κρανιό, παρουσιάζει για πρώτη φορά στοιχεία εξέλιξης του γονιδιώματος του Κρανιού, και παρέχει έναν υψηλής πυκνότητας γενετικό χάρτη ως ένα πολύτιμο εργαλείο για την έρευνα και την παραγωγή.

Abstract

The development of new genetic tools, as well as the optimal use of existing ones to improve production in fish farming for Meagre and Greater amberjack species has been the main research goal in this doctoral dissertation. Modernization in the design of breeding programs as well as genetic selection assisted by polymorphic markers, are conditions which in their

implementation give important possibilities in the production process. We developed two multiplex genetic loci for Meagre and Greater amberjack, respectively, which we used to estimate the genetic diversity of 11 stocks consisting of 946 farmed fish, to investigate the effects of stock mixing and the resulting genetic clustering via population genetics parameters. We proposed a methodology for estimating kinship relationships in multiple stocks, based on which we estimated individual inbreeding values, and we compared two methods of selecting broodstock from crosses of 27 parents from one stock. In addition, we investigated the heritability for the characters weight and length in Meagre, the genetic and phenotypic correlations of weight and length as well as the variation in the contribution of the offspring from a generation of 800 offspring to the formation of F1 families depending on effective population size. The detection of constituent elements of the genome of a species together with the description of the genetic base of important characters can contribute to the conservation of the species and the broadening of the possibilities of its productive utilization. A genetic linkage map for a species highlights genome evolution, facilitates trait mapping through QTL analysis, and improves genome assembly. We used a reduced genomic representation technique by creating a ddRAD library for six F1 families in the Meagre in a total of 266 fish, to construct a high density genetic map consisting of 4529 polymorphic SNP markers. We performed QTL analysis for weight and length, which detected a statistically significant QTL in linkage group 15. The genetic map was used for comparative genomic analysis, by finding homologs, in five genomes of teleost fish and showed a conserved evolutionarily karyotype structure in the Meagre. The observed synteny, converged with published data, on the evolutionary proximity of the Scianidae family with the Sparidae family and specifically with the seabream, with the Moronidae family and the seabass, detected possible chromosomal rearrangement in the Meagre and especially in linkage group 17. In its entirety the dissertation provided with new genetic tools for the Meagre and Greater amberjack for the design of breeding programs. Furthermore, the first genetic linkage map of the Meagre from NGS data was constructed which was used for the genetic appreciation of the growth-related characters weight and length in the Meagre, presents with elements of evolution of the Meagre genome, and provides a high-density genetic map as a valuable tool for research and production.

1. Εισαγωγή

1.1 Βιολογία

Η επιστήμη της βιολογίας, ως ερευνητικός κλάδος καθώς και ως σώμα ιδεών και γνώσεων ασκεί στον γράφοντα σημαντική επίδραση. Με την εξέλιξη των τεχνολογιών αλληλούχισης υψηλής απόδοσης νέας γενιάς, η βιολογία, αποκτά έναν περισσότερο “ολιστικό” χαρακτήρα όπου ο αναγωγικός τρόπος προσέγγισης αντικαθίσταται από την, όσο το δυνατόν, πολυδιάστατη βιολογική εικόνα των προς εξέταση ζητημάτων. Χαρακτηριστικό παράδειγμα αποτελεί το “human genome project” στο οποίο εφαρμόστηκε μία ολιστική προσέγγιση σε ένα βιολογικό ερώτημα που με την σειρά του άνοιξε νέους δρόμους στην απάντηση ερωτημάτων στον κλάδο της γενετικής. Είναι λοιπόν μια εξέλιξη της ίδιας της βιολογίας η οποία μας δίνει τη δυνατότητα να εξηγήσουμε φαινόμενα και διεργασίες που δεν μπορούσαν και δεν μπορούν να απαντηθούν με μονοσήμαντες προσεγγίσεις, αλλά ανιχνεύονται στις συσχετίσεις των επιπέδων οργάνωσης της βιολογικής πληροφορίας. Επιπρόσθετα, το ερευνητικό ενδιαφέρον, ακόμη και αν έλκει την καταγωγή του από την ανάγκη εξήγησης της

πραγματικότητας, οφείλει να προσπαθεί να βελτιώνει τους όρους, τις συνθήκες ζωής του ανθρώπου, καθώς και να διαφυλάσσει την βιοποικιλότητα της βιόσφαιρας συνολικά.

1.2 Ο ρόλος της ιχθυοκαλλιέργειας στο σήμερα

Είναι γεγονός ότι, η μεγάλη αύξηση του παγκόσμιου πληθυσμού σε συνδυασμό με την υπεραλίευση, τον αυξανόμενο ρυθμό εξαφάνισης ειδών της θαλάσσιας πανίδας, διαμορφώνουν μία κρίσιμη συνθήκη. Ο Παγκόσμιος Οργανισμός Τροφίμων και Γεωργίας (UN FAO) υπολογίζει ότι μέχρι το 2050 η παραγωγή τροφίμων πρέπει να έχει διπλασιαστεί ώστε να μπορούν να καλυφθούν οι διατροφικές ανάγκες του ανθρώπινου πληθυσμού. Λόγω της μείωσης των εκφορτώσεων της αλιείας (FAO 2012) φαίνεται πως δεν μπορεί να καλυφθεί η αύξηση της παραγωγής των αλιευτικών προϊόντων για ανθρώπινη κατανάλωση τα οποία έχουν υψηλή θρεπτική αξία και μεγάλη ζήτηση. Η υδατοκαλλιέργεια, ως ο ταχύτερα αναπτυσσόμενος κλάδος της πρωτογενούς παραγωγής, καλείται να καλύψει τις ελλείψεις της αγοράς καθώς το έλλειμμα σύμφωνα με τον FAO το 2030 θα προσεγγίζει τους 29 εκ. τόνους. Η απαίτηση για μείωση του κόστους παραγωγής της εκτροφής θαλασσινών ψαριών και η βελτίωση της ανταγωνιστικότητας τους ως προϊόντα είναι απαραίτητη προϋπόθεση ώστε η αύξηση της παραγωγής να καταστεί βιώσιμη.

1.3 Γενετική επιλογή σε Μεσογειακά είδη εκτρεφόμενων ιχθύων

Βασικό ρόλο στην επιθυμητή μείωση του κόστους αναμένεται να διαδραματίσει η γενετική επιλογή γεννητόρων ή και πληθυσμών που παρουσιάζουν ταχύτερη ανάπτυξη με στόχο την ταχύτερη αύξηση των ψαριών, την καλύτερη μετατρεψιμότητα της τροφής, την δυνατότητα να τραφούν με εναλλακτικές πρώτες ύλες, την ανθεκτικότητα σε ασθένειες κ.α.

Ωστόσο, στα Μεσογειακά είδη δεν υπάρχει μία εμπεριστατωμένη καταγραφή των γενετικών αποθεμάτων ακόμη και για βασικά είδη ιχθυοκαλλιέργειας όπως η τσιπούρα και το λαβράκι καθώς επίσης δεν υπάρχει μία πλατφόρμα γενετικής επιλογής βασισμένη σε μοριακούς δείκτες η οποία με ανταγωνιστικό τρόπο να βοηθά αποτελεσματικά στην αύξηση της παραγωγής και της ποιότητας των παραγόμενων προϊόντων. Τα τελευταία χρόνια έχει αναδειχθεί μία ερευνητική προσπάθεια εύρεσης και χρησιμοποίησης της γενετικής πληροφορίας προς όφελος της παραγωγής. Σειρές αποτελούμενες από πυκνούς γονοτυπημένους μονονουκλεοτιδικούς πολυμορφισμούς (Single Nucleotide Polymorphism), παρέχουν εκτενή πληροφορία για την διακύμανση των πολυμορφισμών του γονιδιώματος ενός συγκεκριμένου είδους. Τέτοια δεδομένα μπορούν να αξιοποιηθούν από μελέτες γενετικής αρχιτεκτονικής ή ποσοτικοποιημένων χαρακτηριστικών ώστε να αυξηθεί η ακρίβεια της επιλογής σε προγράμματα εκτροφής. Με δεδομένο ότι η φυσιολογία της ανάπτυξης στα ποικιλόθερμα σπονδυλωτά εμφανίζει δυναμικές και μπορεί να έχει αυξημένη πολυπλοκότητα λόγω αλληλεπιδράσεων με περιβαλλοντικές επιρροές, έχει παρατηρηθεί ισχυρή κληρονομικότητα για δείκτες μάζας σώματος και σωματικής κατάστασης (Kause et al., 2002).

Τα τελευταία χρόνια, αναδεικνύονται νέα είδη θαλάσσιας ιχθυοκαλλιέργειας στην Μεσόγειο. Ο κρανιός και το μαγιάτικο είναι δύο από τα σημαντικότερα. Λόγω της γρήγορης ανάπτυξης που εμφανίζουν και της δυνατότητας μεταποίησης τους ανταποκρίνονται στις ανάγκες της σύγχρονης διεθνοποιημένης αγοράς. Σαν είδη αποτελούν αντικείμενο μελέτης της ΕΕ και της Ελληνικής κοινότητας.

1.4 Κρανιός

Ο κρανιός, *Argyrosomus regius*, είναι ένα θαλάσσιο είδος που ανήκει στην οικογένεια Scianidae, ευρέως διαδεδομένο στη Μεσόγειο Θάλασσα και στον Ατλαντικό Ωκεανό από τη Νορβηγία προς το Κονγκό και από την Ισλανδία προς τον Κόλπο της Γουινέας (www.fishbase.org). Είναι ένα πελαγικό είδος με μια βενθοπελαγική τάση που μπορεί να αυξηθεί έως και 2 m σε μήκος και ζυγίζει περισσότερο από 50 kg. Το υψηλότερο ποσοστό αύξησης παρατηρείται να λαμβάνει χώρα κατά τη διάρκεια του καλοκαιριού και η δραστηριότητα, σίτιση μειώνονται σημαντικά όταν η θερμοκρασία της θάλασσας πέσει κάτω από 13-15°C (FAO, 2012). Είναι ένα εξαιρετικά γόνιμο είδος το οποίο ανέχεται ένα ευρύ φάσμα αλατότητας με το φιλέτο του να είναι ένα προϊόν υψηλής ποιότητας χάρη στα οργανοληπτικά χαρακτηριστικά του είδους (Poli et al., 2003). Έχει καθιερωθεί ως ένα νέο είδος για την ευρωπαϊκή υδατοκαλλιέργεια που μπορεί να συμβάλει στην διαφοροποίηση του κλάδου. Οι πρώτες δοκιμές με άγριους γεννήτορες διεξήχθησαν στη Νότια Γαλλία και η βιομηχανική παραγωγή έχει σταδιακά αυξηθεί. Η συνολική παραγωγή ήταν 23.440 τόνοι το

2016 (fao.org). Ο κρυνός είναι γονοχωριστικό είδος το οποίο είναι αναπαραγωγικά ώριμο στην ηλικία των δύο ετών για τα αρσενικά και τριών για τα θηλυκά Schiavone et al. (2011). Ο γοναδοσωματικός δείκτης (GSI), ο οποίος είναι ο υπολογισμός της μάζας των γονάδων ως αναλογία της συνολικής μάζας σώματος, έχει την μέγιστη τιμή του τον μήνα Μάρτιο για τα αρσενικά και τον Απρίλιο για τα θηλυκά, με το μήκος της πρώτης σεξουαλικής ωριμότητας των θηλυκών να ξεκινάει σε μεγαλύτερα μήκη από τα αρσενικά με 47 και 45 cm αντίστοιχα (Shabana et al., 2012). Κατά την αναπαραγωγική περίοδο η παραγωγή ωαρίων σε άγριους θηλυκούς γεννήτορες εκτιμήθηκε σε ~900,000 ωάρια kg^{-1} , για άτομα με βάρος 20 kg, ενώ η πυκνότητα των σπερματοζωαρίων σε άγριους αρσενικούς γεννήτορες ήταν από 18.9 ως 31.5×10^9 σπερματοζωάρια ml^{-1} με την μέση διάρκεια της κινητικότητας τους να κυμαίνεται από 0.78 ως 1.27 min (Mylonas et al., 2013).

1.5 Μαγιάτικο

Το μαγιάτικο, *Seriola dumerili*, ένα θαλάσσιο είδος της οικογένειας Carangidae, απαντάται κυρίως στην Μεσόγειο Θάλασσα, τον Ατλαντικό και Ειρηνικό Ωκεανό και τις Ινδικές ακτές. Είναι πελαγικό είδος, βρίσκεται σε βάθη από 20-70 μέτρα με το μέγιστο μήκος του να φτάνει τα 2m (Kozul et al., 2001). Οι μεγάλες δυνατότητες που δίνει στην υδατοκαλλιέργεια σχετίζονται με την παγκόσμια κατανομή του, την εύκολη προσαρμογή του σε συνθήκες αιχμαλωσίας, την ταχεία του ανάπτυξη (6kg σε 2.5 χρόνια), την εξαιρετική ποιότητα της σάρκας του και των δυνατοτήτων μεταποίησης. Η εκτροφή του ξεκίνησε την δεκαετία του '90 και η παραγωγή του παγκοσμίως είναι ακόμη αρκετά μικρή με την Ιαπωνία και Κορέα να βρίσκονται στις πρώτες θέση με παραγωγή 150,000 τόνους το 2017 (<http://assets.worldwildlife.org/>). Το μαγιάτικο είναι γονοχωριστικό είδος, και η αναλογία φύλου είναι 1: 1 (Lazzari and Barbera, 1989; Micale et al., 1999). Αναπαραγωγικά ώριμα αρσενικά και θηλυκά ανιχνεύονται στο ελάχιστο κανονικό μήκος 61 cm (2 έτη) και 80 cm (3 έτη), αντίστοιχα. Το διάμεσο μήκος που το 50% των ψαριών επιτυγχάνει αναπαραγωγική ωριμότητα είναι 1,09 m σε αρσενικά και 1,13 m σε θηλυκά (Marino et al., 1995). Η αναπαραγωγική περίοδος συμπίπτει με την αύξηση της θερμοκρασίας του νερού την άνοιξη, αν και η διάρκειά της εξαρτάται από τη γεωγραφική περιοχή Lazzari and Barbera (1988), (1989); Grau, (1992). Σε άτομα των ~25 kg ο μέσος αριθμός ωαρίων ανά κιλό ανήλθαν σε περίπου 1,200,000 με την αναπαραγωγική περίοδο να εκτείνεται από τα τέλη Απριλίου έως και τις αρχές Οκτώβρη (Jerez et al., 2006). Η διάρκεια κινητικότητας των σπερματοζωαρίων έχει δειχθεί να κυμαίνεται από 0,5 έως 5,2 min και η πυκνότητα τους μεταξύ 18 και 70×10^9 σπερματοζωάρια ml^{-1} (Fakriadis & Mylonas 2021).

1.6 Γενετική βάση φαινοτυπικής παραλλακτικότητας

Μια σημαντική πρόκληση στη βιολογία είναι η κατανόηση της γενετικής βάσης της φαινοτυπικής διακύμανσης χαρακτήρων. Η επιστημονική μέθοδος δείχνει ότι η φαινοτυπική διακύμανση προκαλείται από ποσοτικές γενετικές παραλλαγές που οφείλονται στον διαχωρισμό των αλληλόμορφων γονιδίων σε πολλαπλούς γενετικούς τόπους σχετιζόμενων με αυτούς τους χαρακτήρες (QTL), και στις επιδράσεις του περιβάλλοντος (Mackay 2001).

Η παραχθείσα έρευνα για αυτόν τον σύνθετο χάρτη γονότυπου-φαινοτύπου, συμπεριλαμβανομένων των περιβαλλοντικών επιπτώσεων, έχει αποδώσει σημαντική γνώση για την πρόβλεψη ασθενειών (Lehner 2013), υποστήριξη προγραμμάτων επιλογής/αναπαραγωγής (Dekkers 2012), καθώς επίσης και για τον εντοπισμό προσαρμοστικής παραλλακτικότητας σε φυσικούς πληθυσμούς (Savolainen et al., 2013). Η χαρτογράφηση τέτοιων συσχετίσεων μπορεί να χρησιμοποιηθεί για τον προσδιορισμό των γονιδιωματικών περιοχών που ελέγχουν φαινοτυπικές παραλλαγές χαρακτήρων, με τελικό στόχο τον εντοπισμό της αιτιώδους συνάφειας γονιδίων ή νουκλεοτιδικών μεταλλάξεων συνδεδεμένων με τον εν λόγω χαρακτήρα (Rockman 2012).

1.7 Αναλυτικές προσεγγίσεις

Για την συσχέτιση γονιδιωματικών περιοχών που συνδέονται με ποσοτικούς χαρακτήρες, μέθοδοι γενετικής σύνδεσης (linkage) και συσχέτισης (GWAS) με βάση την ανισοροπία σύνδεσης (linkage disequilibrium) έχουν αναπτυχθεί, καθώς επίσης και συνδυαστικές μέθοδοι των δύο. Το κοινό στοιχείο αυτών των μεθόδων είναι ότι λαμβάνουν υπόψη την φαινοτυπική παραλλακτικότητα χαρακτήρων μετρώντας την συσχέτιση τους με ένα γενετικό υπόβαθρο, με δείκτες που διαχωρίζονται με Μενδελιανό τρόπο. Από αυτές τις μεθόδους η ανάλυση βασιζόμενη στην γενετική σύνδεση (linkage) και η ανάλυση με βάση την ανισοροπία σύνδεσης (linkage disequilibrium/GWAS) είναι οι πιο διαδεδομένες. Μία αξιοσημείωτη διαφορά των δύο μεθόδων εντοπίζεται στο ότι η ανάλυση σύνδεσης είναι μεν πιο ακριβής στον εντοπισμό περιοχών QTL αλλά απαιτεί πολύ περισσότερους δείκτες για την επίτευξη ικανοποιητικής ακρίβειας (Ott et al., 2011). Μια ακόμη διαφορά είναι ότι οι μελέτες GWAS μπορούν να διεξαχθούν χρησιμοποιώντας δείγματα από άγριους ή μη πληθυσμούς με περιορισμένο αριθμό οικογενειών, ενώ οι μελέτες με βάση την γενετική σύνδεση απαιτούν το γενεαλογικό δέντρο των ατόμων προς εξέταση, που συνήθως προέρχεται από ελεγχόμενες διασταυρώσεις.

1.8 Πολυμορφικοί δείκτες με γονιδιωματικό εύρος κάλυψης

Ιστορικά, η έλλειψη πολυμορφικών μοριακών δεικτών είχε περιορίσει την εφαρμογή μεθόδων χαρτογράφησης QTL. Όμως, από την δεκαετία του 2000 και ύστερα νέες τεχνολογίες αλληλούχισης και ανάλυσης όπως η αλληλούχιση τμημάτων DNA προερχόμενων από την πέψη με περιοριστικά ένζυμα, Restriction Associated Dna (Baird et al., 2008), έδωσαν την δυνατότητα δημιουργίας δεικτών με γονιδιωματικό εύρος κάλυψης. Η ανάπτυξη νέων τεχνικών έδωσε με την σειρά της αυξημένες δυνατότητες στην στατιστική δύναμη και ακρίβεια της έρευνας για QTL καθώς επίσης και στην λεπτομερέστερη χαρτογράφηση της γενετικής βάσης φαινοτυπικών χαρακτήρων (Stranger et al., 2011; Visscher et al., 2012).

1.9 Έρευνα για QTL σε τελεόστεους ιχθείς

Η έρευνα για την συσχέτιση γονότυπων με φαινοτυπικούς χαρακτήρες απαντάται σε ένα εύρος οργανισμών μοντέλα (Kahraman et al., 2005), όπως *Saccharomyces cerevisiae*, *Drosophila melanogaster*, *Mus musculus* και *Danio rerio*. Οι δυνατότητες που προσδίδουν οι τεχνολογίες αλληλούχισης DNA νέας γενιάς καθώς και η μείωση του κόστους χρήσης τους επιτρέπουν πλέον την διεύρυνση της έρευνας σε οργανισμούς μη-μοντέλα (Ellegren 2014; Elshire et al., 2011). Οι τελεόστεοι ιχθύες αποτελούν μία ταξονομική ομάδα η οποία

περιλαμβάνει πολλούς οργανισμούς που δεν συγκαταλέγονται στους οργανισμούς μοντέλα στην βιολογία. Εμφανίζουν ιδιαίτερο ενδιαφέρον καθώς αποτελούν την μεγαλύτερη ταξινομική κατηγορία των σπονδυλωτών με περίπου 26.000 είδη, και αθροίζουν ποσοστιαία περίπου το 50% όλων των σπονδυλωτών ειδών παρουσιάζοντας μεγάλο εύρος μορφολογικής διαφοροποίησης. Λόγω αυτών των χαρακτηριστικών αλλά και λόγω του ότι αποτελούν τμήμα του πιο ραγδαίως αναπτυσσόμενου κλάδου της πρωτογενούς παραγωγής, αρκετές μελέτες διερευνώντας την ύπαρξη QTL έχουν διεκπαιρωθεί για μία σειρά είδη όπως *Dicentrarchus labrax* (λαβράκι), *Sparus aurata* (τσιπούρα), *Cyprinus carpio* (κυπρίνος), *Oreochromis niloticus* (τιλάπια), *Salmo salar* (σολωμός Ατλαντικού), κ.α (Hemmer-Hansen et al., 2014; Tong & Sun 2015) παράγοντας δεδομένα από μοριακούς δείκτες προερχόμενους από μονο-νουκλεοτιδικούς πολυμορφισμούς σχετιζόμενων με φαινοτυπικούς αναπτυξιακούς χαρακτήρες, ανθεκτικότητα σε ασθένειες, και αλληλεπίδραση με περιβαλλοντικές επιρροές. Μοριακοί δείκτες οι οποίοι χρησιμοποιήθηκαν για την γενετική επιλογή και τον σχεδιασμό προγραμμάτων αναπαραγωγής στα προαναφερθέντα είδη και όχι μόνο.

2. Γενετική περιγραφή, πληθυσμιακή δομή και ανάλυση συγγένειας αποθεμάτων γεννητόρων και απογόνων ως βάση για τον σχεδιασμό προγραμμάτων επιλογής

2.1 Εισαγωγή

Η γενετική ποικιλομορφία βοηθά στην παραγωγή ποικίλων σχημάτων, μεγεθών, συμπεριφορών, χρωμάτων και ικανοτήτων που κάνουν τα υδρόβια είδη πολύτιμα και ενδιαφέροντα. Επιτρέπει επίσης στα είδη να προσαρμόζονται σε νέα οικοσυστήματα, βιότοπους και περιβαλλοντικές συνθήκες. Η μεγαλύτερη γενετική ποικιλομορφία συνεπάγεται περισσότερα είδη και στελέχη οργανισμών, τα οποία οδηγούν σε μεγαλύτερες δυνατότητες προσαρμογής τους στις σημερινές συνθήκες κλιματικής αλλαγής και υπεραλίευσης. Τα είδη θαλάσσιων ιχθύων εμφανίζουν φυσικά υψηλά επίπεδα γενετικής ποικιλομορφίας (DeWoody & Avise 2000) που είναι πιθανώς ο κύριος μοχλός του ραγδαίου ρυθμού γενετικής τους βελτίωσης, σε σύγκριση με τα φυτά ή άλλα τάξα σπονδυλωτών (Gjedrem & Robinson 2014). Επιπλέον, η υψηλή γενετική παραλλακτικότητα αποτελεί αναγκαίο υπόβαθρο της γενετικής βελτίωσης, λόγω του ότι η διεύρυνση του φάσματος της βελτίωσης, για επιλεγμένα χαρακτηριστικά, στηρίζεται στην καταγραφή της γενετικής ποικιλομορφίας αυτών των χαρακτηριστικών στον βασικό πληθυσμό. Αντιθέτως, την γενετική απομόνωση του βασικού πληθυσμού η παραλλακτικότητα του θα οδηγηθεί σε μείωση. Η δυνατότητα αξιοποίησης του φάσματος της γενετικής ποικιλομορφίας στην αρχή ενός προγράμματος αναπαραγωγής αποδίδεται εν μέρει στην επιτυχία της γενετικής βελτίωσης (σολωμός Ατλαντικού, Gjæren & Gjerdde 1991), ενώ και άλλοι παράγοντες παίζουν επίσης ρόλο. Οι μελέτες γενετικής περιγραφής ή παραλλακτικότητας του αποθέματος γεννητόρων με βάση τους μοριακούς γενετικούς δείκτες είναι ένα κοινό εργαλείο για προγράμματα διατήρησης ή αναπαραγωγής, με τα είδη *Myxocyprinus asiaticus* (Liu D. et al 2018), *Hucho hucho* (Kucinski et al. 2015), *Oreochromis* sp. (Montoya-LópezAF et al. 2019) και *Lates calcarifer* (Loughnan, S.R et al. 2016) να είναι μερικά από τα πιο πρόσφατα παραδείγματα. Οι δείκτες πολυμορφικού DNA έχουν διευρύνει τα όρια της έρευνας παρέχοντας νέα γνώση σχετικά με τη γενετική δομή των πληθυσμών ιχθύων, την ένταση της φυσικής και σεξουαλικής επιλογής, τα επίπεδα αναπαραγωγής επιτυχίας καθώς και του

εμπλουτισμού των αναπαραγωγικών στρατηγικών (M.M. Ferguson et al. 1998). Οι μικροδορυφόροι ή οι απλές επαναλήψεις αλληλουχιών (SSRs) αποτελούνται από μικρά (1 ± 6 ζεύγη βάσεων) διαδοχικά μοτίβα νουκλεοτιδικών βάσεων που πλαισιώνονται από περιοχές μη επαναλαμβανόμενων μοναδικών αλληλουχιών (Tautz and Renz, 1984; Tautz, 1989). Τα χαρακτηριστικά των μικροδορυφορικών δεικτών, τους καθιστούν μία καλή επιλογή για την περιγραφή της γενετικής παραλλακτικότητας. Συγκεκριμένα ο συν-επικρατής τους χαρακτήρας (οι ετεροζυγώτες μπορούν να διακριθούν από τους ομοζυγώτες), οι ποικίλες θέσεις τους στο γονιδίωμα, ο υψηλός πολυμορφισμός, η διαφορετικότητα του επαναληπτικού τους μοτίβου, και η σχετική αφθονία τους στο γονιδίωμα, είναι τα κύρια πλεονεκτήματα των μικροδορυφορικών δεικτών. Τα μειονεκτήματα τους περιλαμβάνουν τα αυξημένα λάθη στην αντιγραφή τους λόγω των επαναλαμβανόμενων περιοχών (stutter bands), την ανομοιογένεια στα αλληλόμορφα τους (πχ null alleles), την ομοπλασία (εμφάνιση φυλογενετικής ανεξάρτητης ομοιότητας) και τον υψηλό αριθμό αλληλομόρφων σε συγκεκριμένες περιοχές (V. Mohindra et al. 2001). Τα μηδενικά αλληλόμορφα στους μικροδορυφορικούς τόπους εμφανίζονται με την ύπαρξη μετάλλαξης στην περιοχή πρόσδεσης του εκκινήτη με αποτέλεσμα την μη PCR επαύξηση του συγκεκριμένου αλληλόμορφου και την μη σωστή εκπροσώπηση των ετεροζυγωτικών αλληλομόρφων ή την παντελή έλλειψη κάποιων αλληλομόρφων από το δείγμα. Επιπλέον, οι μεταλλάξεις μετατόπισης (διαγραφή, εισαγωγή, διπλασιασμός) που έχουν εντοπιστεί σε πλευρικές περιοχές μικροδορυφορικών τόπων μπορεί να εισάγουν σφάλμα σε μελέτες πληθυσμιακής γενετικής (R. Zardoya et al. 1996). Οι δείκτες μικροδορυφορικού DNA είναι από τους “δόκιμους υπονήφιους” που εμπίπτουν στην εξελικτική θεωρία της μοριακής ουδετερότητας και έχειδειχθεί ότι αντανακλούν την γεωγραφική διαφοροποίηση απομονωμένων πληθυσμών, συγγενικών φυλογενετικά, ειδών και υποειδών (R. Zardoya et al. 1996). Μερικά από τα πρώτα είδη που διερευνήθηκαν χρησιμοποιώντας μικροδορυφορικούς δείκτες ήταν ο σολομός Ατλαντικού (Nielsen et al., 1994) και η ιριδίζουσα πέστροφα της λίμνης Οντάριο (O’Connell et al., 1997).

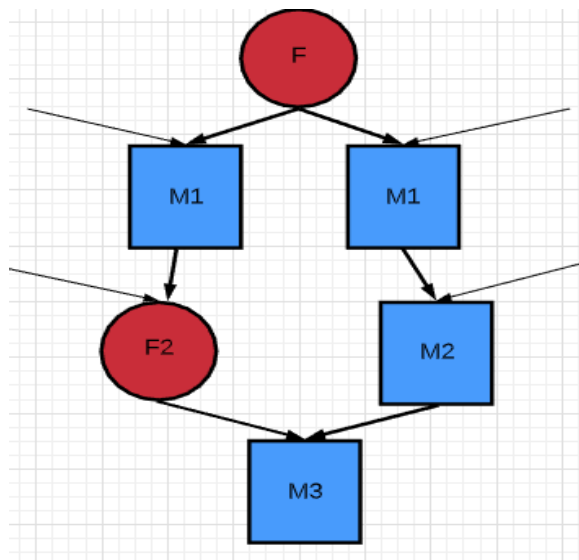
2.1.1 Προγράμματα επιλογής, ένα πρώτο σχόλιο

Η ιχθυοκαλλιέργεια σαν σύστημα παραγωγής έχει μία μακρά ιστορία. Παρόλα αυτά, η συστηματική και αποτελεσματική εφαρμογή προγραμμάτων αναπαραγωγής για την βελτίωση σημαντικών εμπορικών χαρακτήρων δεν άρχισε να εφαρμόζεται πριν τις αρχές της δεκαετίας του 1990 (Symonds et al. 2019). Τέτοια προγράμματα παρείχαν σημαντικό όφελος στην βιομηχανία της ιχθυοκαλλιέργειας τα οποία σήμερα αξιοποιούν τεχνολογίες όπως αυτή της εφαρμοσμένης γονιδιωματικής, της αλληλούχισης μέσω σύνθεσης (sequencing by synthesis) κ.α, για την αξιολόγηση της γενετικής ποικιλότητας των αποθεμάτων, σε φυσικό ή εμπορικό περιβάλλον, ως και την ανάπτυξη εργαλείων γενετικής επιλογής υποβοηθούμενης από μοριακούς δείκτες (*Marker Assisted Selection, MAS*) όπως οι μικροσυστοιχίες SNP (SNP-arrays ή SNP-chips). Η πολυπλοκότητα της αναπαραγωγικής βιολογίας των υδρόβιων ειδών αποτέλεσε έναν από τους βασικούς λόγους για την καθυστερημένη εφαρμογή τέτοιων προγραμμάτων επιλογής με εξαίρεση την AKVAFORSK που, από το 1971 έχει σχεδιάσει και υλοποιήσει μεγάλης κλίμακας προγράμματα για τα είδη της οικογένειας των σαλμονιδών

(Gjedrem 2005). Σε αυτή την κατεύθυνση, εφαρμόζουμε την ήδη υπάρχουσα γνώση σε συνδυασμό με τη βοήθεια λογισμικών για να επιλέξουμε γεννήτορες από αποθέματα γονιών και απογόνων, με την λογική της ευελιξίας και αξιοποίησης κάθε φορά, στο μέγιστο, των παρεχόμενων δεδομένων. Οι εμπορικές συνθήκες καθιστούν σχεδόν αδύνατη την εξολοκλήρου αντικατάσταση των γεννητόρων από κάθε νέα γενιά απογόνων. Αυτό συμβαίνει κυρίως λόγω της αυξημένης γονιμότητας και του χρόνου που απαιτείται για την σεξουαλική ωρίμανση, των ιχθύων μεγάλου μεγέθους, όπως του μαγιάτικου και του κρانيού. Οι μεγαλύτερης ηλικίας γεννήτορες παράγουν περισσότερα ωάρια/σπερματοζώαρια συνεπώς η παραγωγική τους δυνατότητα είναι αυξημένη συγκριτικά με τους γεννήτορες μικρότερης ηλικίας. Ωστόσο, για να διασφαλιστεί ένας βιώσιμος πληθυσμός μακροπρόθεσμα, η ενδογαμία (διασταυρώσεις μεταξύ συγγενικών ατόμων) πρέπει να είναι η μικρότερη δυνατή ώστε να αποφευχθεί ο εκφυλισμός και θεωρητικά η κατάρρευση των μελλοντικών γενιών λόγω της αύξησης της ομοζυγωτίας των υπολειπόμενων γονιδίων (που μπορεί π.χ. να σχετίζονται με ασθένειες κ.α) (Melfe 2011). Η αντικατάσταση του αποθέματος των γεννητόρων με νέους, προερχόμενους από επόμενες γενιές, παρέχει την δυνατότητα της μείωσης του εύρους της ενδογαμίας καθώς επίσης πιθανότατα και της ενίσχυσης της επιλογής των επιθυμητών χαρακτήρων.

2.1.2 Το ζήτημα της ενδογαμίας

Ο Sewall Wright το 1922 πρότεινε τον συντελεστή ενδογαμίας f (coefficient of inbreeding) ως την πιθανότητα που έχουν δύο αλληλόμορφα σε ένα τυχαίο γενετικό τόπο (ή γονίδιο) να προέρχονται από τον ίδιο πρόγονο (*Identical by Descent, IDB*). Αν και τα αλληλόμορφα μπορεί να είναι ομόλογα λόγω άλλων παραγόντων, αν και αυτό συμβαίνει αρκετά πιο σπάνια, όπως μεταλλάξεων, ανασυνδυασμού των χρωμοσωμάτων κατά την μείωση, στον υπολογισμό της ενδογαμίας κατά Wright ο ανασυνδυασμός ή οι πιθανές μεταλλάξεις δεν λαμβάνονται υπόψη και ο συντελεστής ενδογαμίας ισούται με την μαθηματική πιθανότητα της καταγωγής των αλληλομόρφων από έναν κοινό πρόγονο. Οι τιμές που μπορεί να λάβει ο f είναι από 0 ως 1 με την τιμή 0 να σηματοδοτεί μηδενική ενδογαμία και την τιμή 1 μέγιστη ενδογαμία (η πιθανότητα να είναι ομοζυγώτης, ο εξεταζόμενος για ενδογαμία απόγονος, με έναν πρόγονο για τον ίδιο γενετικό τόπο είναι 100% όταν ο συντελεστής ενδογαμίας παίρνει την τιμή 1).



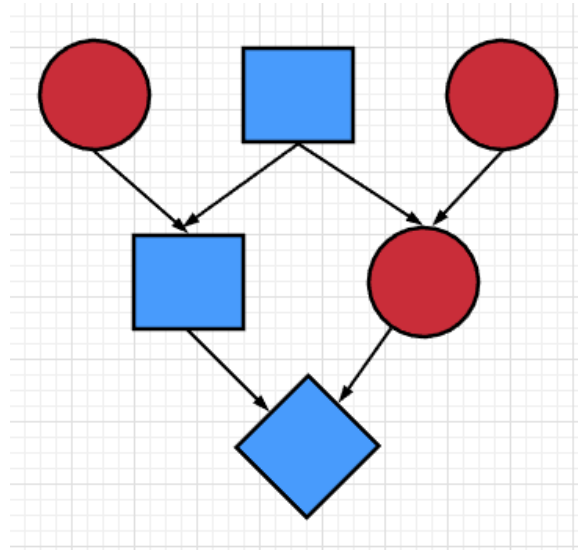
Σχήμα 1. Παράδειγμα 1, υπολογισμός ενδογαμίας. Με γαλάζιο τετράγωνο απεικονίζονται τα αρσενικά και με κόκκινο κύκλο τα θηλυκά. Καθώς επίσης οι αριθμοί 1,2,3 αντιστοιχούν στην πρώτη, δεύτερη και τρίτη γενιά αντίστοιχα.

Στο παραπάνω **Σχήμα 1** γενεαλογίας διπλοειδούς οργανισμού ο θηλυκός πρόγονος F είναι κοινός για όλους τους απογόνους $M1\alpha$, $M1\beta$, $M2$, $F2$, $M3$. Για να υπολογίσουμε τον συντελεστή ενδογαμίας του απογόνου $M3$ έχουμε : $M3 \text{ inbr. Coefficient} = (1/2)^5 = 1/32 = 3,1\%$ ή 0.031. Στο παράδειγμα, όλοι οι πρόγονοι του $M3$ παρουσιάζουν συγγένεια μεταξύ τους, αλλά μόνο οι $F2$, $M2$, έχουν διασταυρωθεί μεταξύ τους.

Ο υπολογισμός του συντελεστή ενδογαμίας f γίνεται με βάση την αρχή του διαχωρισμού των αλληλομόρφων γονιδίων κατά Mendel σε έναν διπλοειδή οργανισμό. Ο γενικός τύπος δίνεται από την εξίσωση (2.1.2.1).

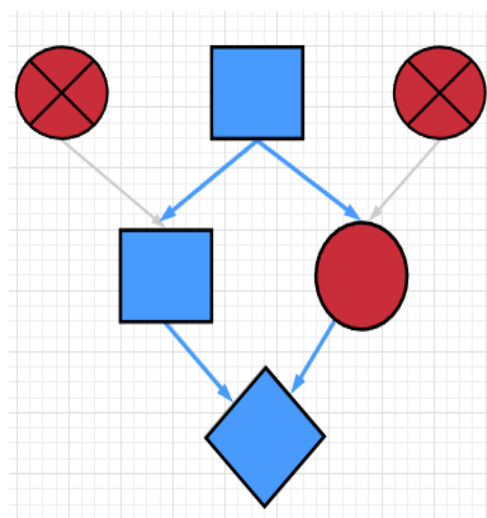
$$f = \left(\frac{1}{2}\right)^n \quad (2.1.2.1)$$

όπου n : ο αριθμός των ατόμων που συνεισέφεραν γονιδιακά στην διαμόρφωση του γονότυπου του εξεταζόμενου απογόνου. Έτσι, βλέπουμε στο παράδειγμα 2 που απεικονίζεται παρακάτω και αφορά την διασταύρωση μεταξύ ετεροθαλών αδελφών (half-sib mating, που μοιράζονται εδώ τον ίδιο αρσενικό γονέα), τον τρόπο υπολογισμού του συντελεστή ενδογαμίας.



Σχήμα 2. Παράδειγμα 2, διασταύρωση ετεροθαλών αδελφών

Στο **Σχήμα 2** έχουμε έναν αρσενικό πρόγονο ο οποίος διασταυρώθηκε με 2 θηλυκά άτομα και έδωσαν 2 απογόνους. Βλέπουμε ότι στην διαμόρφωση του γονότυπου του απογόνου που απεικονίζεται με τον γαλάζιο ρόμβο συνεισέφεραν : 1) οι γονείς του, ο αρσενικός συμβολίζεται με το τετράγωνο και ο θηλυκός με τον κύκλο, 2) ο κοινός αρσενικός πρόγονος.



Σχήμα 3. Παράδειγμα 3, η μοναδική “διαδρομή” για τον υπολογισμό της ενδογαμίας κατά Wright

Στο **Σχήμα 3** έχουμε μία μοναδική «ροή» του γενετικού υλικού για τον εξεταζόμενο απόγονο χωρίς κάποια εναλλακτική. Για τον υπολογισμό του συντελεστή ενδογαμίας από τον γενικό τύπο (1) για $n = 3$, $f = (1/2)^3$, λόγω του ότι 3 άτομα συνεισέφεραν στον γονότυπο του απογόνου με μία μοναδική διαδρομή. Αυτό συμβαίνει διότι οι δυο θηλυκοί πρόγονοι με τους οποίους διασταυρώθηκε ο κοινός αρσενικός πρόγονος δεν συνεισφέρουν και στους δυο γονείς του εξεταζόμενου απογόνου. Ο κάθε θηλυκός πρόγονος των ιδρυτών του γενεαλογικού δέντρου συνεισφέρει μόνο σε έναν από τους δύο γονείς. Συνεπώς δεν μπορούν να συνυπολογιστούν στην «ροή» του γενετικού υλικού για τον υπολογισμό του συντελεστή ενδογαμίας, καθώς στην περίπτωση της ενδογαμίας εξετάζονται μόνο οι κοινοί πρόγονοι. Έτσι ο συντελεστής ενδογαμίας ισούται με $f = 1/8$ ή 0,125 ή 12,5%. Εάν υποθέσουμε ότι υπήρχε ακόμη μία εναλλακτική ροή του γενετικού υλικού η οποία περιλάμβανε επίσης 3 άτομα, όπως στην περίπτωση διασταύρωσης ομοθαλών αδελφών θα είχαμε $f = (1/2)^3 + (1/2)^3 = 1/4$ ή 0.25 ή 25%.

Ελλείψει γενεαλογικού δέντρου, ο υπολογισμός της ενδογαμίας μπορεί να γίνει επίσης μέσω του N_e , τον δραστικό αριθμό αρσενικών και θηλυκών γεννητόρων (πχ σε έναν ιχθυογεννητικό σταθμό) ούτως ώστε να παραχθούν βιώσιμοι (viable) απόγονοι. Αν και ο πιο ακριβής τρόπος υπολογισμού του N_e γίνεται έμμεσα, με τις αλλαγές στις συχνότητες των αλληλομόρφων και την ανισορροπία σύνδεσης (δηλ. αν η συχνότητα της συσχέτισης των διαφορετικών αλληλομόρφων είναι μεγαλύτερη ή μικρότερη από την αναμενόμενη, **όταν** οι γενετικοί τόποι που βρίσκονται τα αλληλόμορφα αυτά είναι ανεξάρτητοι μεταξύ τους) μία σχετικά καλή εικόνα μπορεί να δοθεί από τον άμεσο υπολογισμό του N_e . Η σύνδεση του N_e με τον συντελεστή ενδογαμίας δίνεται από την εξίσωση (2.1.2.2) Wray and Thompson, (1990), με το f να είναι η μέση τιμή του συντελεστή ενδογαμίας για μία γενιά.

$$f = \frac{1}{(2 * N_e)} \quad (2.1.2.2)$$

Η εξίσωση (2) δείχνει ότι ο συντελεστής ενδογαμίας είναι αντιστρόφως ανάλογος με τον N_e όπως επίσης και την σημαντικότητα του N_e στην διαμορφούμενη ενδογαμία. Αξίζει επίσης να σημειωθεί ότι, αν δεν υπάρχει μία στοιχειώδης εικόνα των διασταυρώσεων και των απογόνων ο N_e δεν μπορεί να υπολογισθεί.

Η Οργάνωση Τροφίμων και Γεωργίας των Ηνωμένων Εθνών (FAO) στο τεχνικό της άρθρο 392 του Douglas Tave (<http://www.fao.org/3/x3840e/X3840E00.htm>) οριοθετεί και περιγράφει το ζήτημα της ενδογαμίας καθώς και της γενετικής παρέκκλισης.

Παρακάτω αναφέρονται δύο ποιοτικοί και ένας ποσοτικός λόγος ως μία πιο αναλυτική εξήγηση της εμφάνισης της κατάρρευσης ενός αποθέματος / στόκ λόγω ενδογαμίας.

1. Η αύξηση της ομοζυγωτίας των υπολειπόμενων αλληλομόρφων που σχετίζονται με παθογόνες ασθένειες είναι η πιο κοινή εξήγηση. Τα περισσότερα αλληλόμορφα που εκφράζουν προβληματικούς φαινοτύπους, είναι υπολειπόμενα. Η παραγωγή γαμετών στα ψάρια είναι ιδιαίτερα υψηλή και κυμαίνεται σε μεγέθη των δεκάδων ή και εκατοντάδων εκατομμυρίων ανά γέννα (Wootton & Smith 2014). Ακόμη και αν υποθέσουμε τον χαμηλότερο ρυθμό μετάλλαξης για μία νουκλεοτιδική βάση, είναι

μαθηματικά βέβαιο ότι σε κάποια από τα δεκάδες χιλιάδες γονίδια των παραχθέντων γαμετών θα προκληθούν μεταλλάξεις κατά την μείωση. Για κάθε 10^6 γαμέτες που παράγονται, αν υποθέσουμε ότι ένα μέσο γονιδίωμα τελεόστεων ιχθύων έχει 10.000 γονίδια (Glasauer & Neuhauss, 2014)) έχουμε την αντιγραφή 10^{10} γονιδίων. Πολλές από αυτές τις μεταλλάξεις οι οποίες δεν θα είναι θνησιγόνες και αφορούν υπολειπόμενα γονίδια, θα συσσωρευτούν στην νέα γενιά των απογόνων, οι οποίες θα εκφραστούν μέσω της ομοζυγωτίας. Οι μεταλλάξεις δεν συνδέονται πάντα με την απώλεια αρμοστικότητας (fitness) για ένα άτομο αλλά και με ευεργετικές φαινοτυπικές αλλαγές, κάποιες από τις οποίες θα μειώσουν την αρμοστικότητα μέσω φαινοτυπικών αλλαγών ως αποτέλεσμα της φυσικής επιλογής, μειώνοντας έτσι, την δυνατότητα επιβίωσης ή/και την αναπαραγωγική δυνατότητα ενός ατόμου σημαντικά. Η αποκάλυψη υπολειπόμενων αλληλομόρφων και η αύξηση της συχνότητας τους στις επόμενες γενιές, όντας σε ομοζυγωτία, ενισχύει την πιθανότητα του εκφυλισμού ενός πληθυσμού αναπαραγωγής.

2. Η «υπερκυριαρχία» (overdominance) εμφανίζεται όταν ένας ετερόζυγος φαινότυπος (πχ. Αα) παράγει έναν πιο επικρατή φαινότυπο (dominant) σε σύγκριση με τους δύο ομόζυγους (ΑΑ και αα). Η ενδογαμία μειώνει την ετεροζυγωτία συνεπώς μειώνει την εμφάνιση αυτού του φαινοτύπου.
3. Ένας ακόμη λόγος για την εμφάνιση της κατάρρευσης ενός πληθυσμού λόγω ενδογαμίας, δίνεται από την ποσοτική γενετική και την γενική εξίσωση της παραλλακτικότητας (2.1.2.1).

$$V_p = V_A + V_D + V_I + V_E \quad (2.1.2.1)$$

Για τους φαινότυπους που ελέγχονται από την αθροιστική δράση πολλών γονιδίων, η ενδογαμία δεν αναμένεται να παίξει σημαντικό ρόλο. Ωστόσο, οι φαινότυποι που εξαρτώνται **μόνο** από την αθροιστική δράση ομάδων γονιδίων είναι μειωθητικοί. Η κυριαρχική διακύμανση (παραλλακτικότητα) V_D , αφορά τις επιδράσεις των κυρίαρχων γονιδίων, που εκφράζουν ποσοτικούς φαινότυπους, ως το ποσοστό εκείνο της συνολικής γενετικής διακύμανσης (παραλλακτικότητας) και είναι ένας εκ των τριών παραμέτρων που διαμορφώνουν την συνολική γενετική διακύμανση. Οι άλλοι δύο παράμετροι, είναι το ποσοστό της αθροιστικής δράσης γονιδίων για ποσοτικά χαρακτηριστικά V_A και τέλος εκείνο που οφείλεται σε επιστατικές επιδράσεις V_I . Η V_D , όπως είναι αναμενόμενο μειώνεται με την μείωση της ετεροζυγωτίας, συνεπώς οι ποσοτικοί φαινότυποι που είναι στενότερα συνδεδεμένοι με αυτά τα επικρατή (κυριαρχικά) γονίδια θα μειωθούν επίσης. Όπου V_D είναι η γενετική διακύμανση που οφείλεται στην δράση επικρατών γονιδίων, V_A η γενετική διακύμανση που οφείλεται στην αθροιστική δράση γονιδίων, V_I η γενετική διακύμανση που οφείλεται στην επιστατικές δράσεις μεταξύ γονιδίων, V_E η διακύμανση που οφείλεται σε περιβαλλοντικές επιδράσεις, και V_p η συνολική φαινοτυπική διακύμανση.

Η ενδογαμία μπορεί να μειωθεί (Kincaid, 1977) με τους εξής τρόπους :

- την μεγαλύτερη δυνατή αύξηση του αριθμού των γεννητόρων, (αύξηση του N_e)
- την καταγραφή των γενεαλογιών και έλεγχο των συζεύξεων,

- την ομοιόμορφη συμμετοχή των οικογενειών στην δημιουργία της επόμενης γενιάς,
- την ανταλλαγή αποθεμάτων μεταξύ των ιχθυογεννητικών σταθμών,
- ένα σύστημα εκ περιτροπής διασταυρώσεων διαχωρισμένων ομάδων γεννητόρων με σκοπό την μείωση του ρυθμού ενδογαμίας

2.1.3 Γενετική παρέκκλιση (genetic drift)

Η γενετική παρέκκλιση γνωστή επίσης ως αλληλική παρέκκλιση ή επίδραση Sewall Wright (*Gould SJ 2002*) είναι η αλλαγή στην συχνότητα ενός αλληλόμορφου γονιδίου σε έναν πληθυσμό λόγω της τυχαίας “δειγματοληψίας” του (*Masel J 2011*). Ένας απόγονος μίας διασταύρωσης έχει στο γενετικό του υλικό ένα μέρος των αλληλομόρφων των γονέων του, και σε έναν πληθυσμό η συχνότητα ενός γονιδιακού αλληλόμορφου είναι το κλάσμα των αντιγράφων ενός γονιδίου που μοιράζονται μία λειτουργία. Αυτές οι γονιδιακές παραλλαγές, μπορεί να εκλείψουν ολοκληρωτικά από έναν πληθυσμό ως αποτέλεσμα της μείωσης της γενετικής ποικιλομορφίας. Η γενετική παρέκκλιση, έχει αυτή την επίδραση. Ο Μότου Κιμούρα, με την θεωρία της ουδετερότητας της μοριακής εξέλιξης, ανέδειξε την γενετική παρέκκλιση και την επίδραση της σε ουδέτερες μεταλλάξεις ως την βασική αιτία της διάδοσης μίας γενετικής αλλαγής σε έναν πληθυσμό (*Kimura 1968*).

Πολλοί παράγοντες μπορούν να επιδράσουν στην συχνότητα των γονιδιακών αλληλομόρφων κατά τα διάφορα στάδια της παραγωγής. Κάποιοι από αυτούς είναι: η επιλογή στον ιχθυογεννητικό σταθμό, η διαδικασία της “εξημέρωσης” καθώς και η γενετική παρέκκλιση. Οι αλλαγές στις γονιδιακές συχνότητες προκαλούμενες από την επιλογή και την εξημέρωση συνήθως παράγουν γενετικά βελτιωμένους απογόνους, οι οποίοι εμφανίζουν ταχύτερους ρυθμούς ανάπτυξης, έχουν αποδοτικότερη μετατρεψιμότητα τροφής σε βιομάζα, είναι πιο ανθεκτικοί σε ασθένειες, δέχονται την τροφή σε πελέτες πιο εύκολα κ.α.. Στον αντίποδα, οι αλλαγές στις γονιδιακές συχνότητες προκαλούμενες από την γενετική παρέκκλιση είναι τυχαίες, και αυτό μπορεί να σημαίνει ότι αντιβαίνουν στην όποια βελτίωση έχει επέλθει μέσω της τεχνητής επιλογής. Όταν η συχνότητα ενός γονιδιακού αλληλόμορφου μηδενιστεί, τότε εξαφανίζεται από την “γενετική δεξαμενή” του πληθυσμού, αυτό προφανώς συμβαίνει πιο συχνά σε αλληλόμορφα με χαμηλές συχνότητες και λιγότερο συχνά σε εκείνα με πιο υψηλές.

2.1.4 Ανάλυση σχέσεως γονέων-απογόνων

Η ανάλυση γονικών σχέσεων αποτελεί βασική πτυχή της μοριακής γενετικής και οικολογίας, με τις δεκαετίες του 1970 και του 1980 να φέρνουν στο προσκήνιο, την δυνατότητα των γενετικών δεδομένων να διαγνώσουν τις σχέσεις γονέων-απογόνων στη φύση (Ellstrand, 1984; Gowaty & Karlin, 1984). Από την δειλή της αρχή, χρησιμοποιώντας αλλοζυμικούς δείκτες, έχει ωριμάσει σε μια πρακτική που συνεχίζει να εξελίσσεται παράλληλα με την εξέλιξη των τεχνικών και γνωσιακών μέσων. Η θεμελιώδης ιδέα στην οποία βασίζεται είναι σχετικά απλή. Λαμβάνοντας υπόψη τους νόμους του Μέντελ, γνωρίζουμε ότι κάθε άτομο κληρονομεί γενετικό υλικό από τους γονείς του. Έτσι, ένας αριθμός από Μεντελιανούς τύπους, γονοτυπημένους σε ένα δείγμα ατόμων, μπορεί να διακρίνει τις σχέσεις γονέων-απογόνων από άλλα είδη σχέσεων (Thompson & Meagher, 1987), συνεπώς είναι σημαντικό να γνωρίζουμε την φύση με την οποία κληρονομούνται οι επιλεγμένοι δείκτες. Η πιο σημαντική πρόσφατη εξέλιξη στην εν λόγω ανάλυση είναι η εισαγωγή των τεχνολογιών αλληλούχησης νέας γενιάς. Η ύπαρξη γονότυπων, με χιλιάδες ή και εκατομμυρία γενετικούς τύπους σε σχεδόν οποιοδήποτε είδος μπορεί να την καταστήσει μία πολύ αποτελεσματική και πληροφοριακή ανάλυση (Glaubitz, Rhodes, & DeWoody, 2003). Από την άλλη μεριά, η χρησιμοποίηση μεγάλων συνόλων γενετικών δεδομένων απαιτεί και μία αναβαθμισμένη στάθμιση των διαφόρων παραμέτρων που προκύπτουν από τις ιδιότητες τους. Ένα από τα αξιοσημείωτα χαρακτηριστικά της εξελικτικής βιολογίας στον εικοστό αιώνα είναι ότι ένα μεγάλο μέρος της θεωρίας για την ανάλυση μοτίβων σε πληθυσμιακό επίπεδο, βασιζόμενη σε γενετικά δεδομένα, άνθισε πολύ πριν την εμφάνιση και την διαθεσιμότητα κατάλληλων μοριακών δεικτών. Οι πιο σημαντικές προσεγγίσεις σε αυτόν τον ερευνητικό τομέα αναπτύχθηκαν κατά τις δεκαετίες του 1970 και του 1980 (Meagher & Thompson, 1986). Οι ισοενζυμικοί δείκτες (ή αλλοένζυμα) έθεσαν τις βάσεις για την εξελικτική βιολογία και αποτέλεσαν τους πρώτους μοριακούς δείκτες. Όμως, λόγω της σχετικά μικρής τους παραλλακτικότητας (ειδικά στον άνθρωπο) δεν ήταν επαρκείς για την ανάλυση των σχέσεων γονέων-απογόνων με αξιόπιστο τρόπο (Hubby & Lewontin, 1966). Στα τέλη της δεκαετίας του 1980, έκανε την εμφάνιση του το πρώτο κύμα υπερμεταβλητών δεικτών (Tautz, 1989), των μικροδορυφορικών, όταν ήδη οι ερευνητές είχαν πρόσβαση σε μια αρκετά ισχυρή αναλυτική εργαλειοθήκη. Παρά την αυξανόμενη εμφάνιση προσεγγίσεων, η ανάλυση της ανάθεσης απογόνων σε υποψήφιους γονείς την εποχή των αλλοζυμικών δεικτών, ήταν σχεδόν αδύνατη δεδομένου του χαμηλού πολυμορφισμού των δεικτών (όπως προαναφέραμε), με τα λίγα επιτυχημένα παραδείγματα να αποτελούν εξαιρέσεις που επιβεβαιώνουν τον κανόνα (Ellstrand, 1984; Gowaty & Karlin, 1984; Hanken & Sherman, 1981, Meagher, 1986). Η σύγχρονη εποχή της ανάλυσης της σχέσης γονέων-απογόνων ξεκίνησε ουσιαστικά τη δεκαετία του 1990, όταν η δύναμη του κύματος των μικροδορυφορικών δεικτών ανέδειξε νέες διαδρομές στην έρευνα της μοριακής γενετικής και οικολογίας (Jarne & Lagoda, 1996; Powell, Machray, & Provan, 1996). Το γεγονός ότι οι μικροδορυφορικοί δείκτες χαρακτηρίζονταν από Μεντελιανή κληρονομικότητα και εμφάνιζαν πολλά αλληλόμορφα ανά τόπο επέτρεψε την πιο αποτελεσματική μελέτη των μοτίβων γονιών-απογόνων σε φυσικούς πληθυσμούς, με την εμφάνιση πολλών τέτοιων μελετών (Dow & Ashley, 1996; Jones & Avise, 1997a, 1997b; Kellogg, Markert, Stauffer, & Kocher, 1995; Morin, Wallis, Moore, & Woodruff, 1994).

2.1.5 Μέθοδοι ανάλυσης

Η θεωρία σχετικά με την ανάλυση της σχέσης γονέων-απογόνων δεν έχει αλλάξει ουσιαστικά την τελευταία δεκαετία και οι μοριακοί δείκτες που προέρχονται από τις τεχνολογίες αλληλούχισης νέας γενιάς (NGS), δεν άλλαξαν ριζικά την αποτελεσματικότητα των προσεγγίσεων που απαιτούνται για την επιτυχή έκβαση της ανάλυσης, ώστε οι μέθοδοι των τελευταίων δύο δεκαετιών (Jones & Ardren, 2003; Jones et al., 2010) να συνοψίζουν ακόμη και σήμερα τις βασικές προσεγγίσεις: αποκλεισμός, κατηγορική ανάθεση, Μπαϋεσιανή ανάλυση γονέων, ανακατασκευή των γονικών σχέσεων.

2.1.5.1 Αποκλεισμός (exclusion method)

Η πιο απλή και εύκολη προσέγγιση ενδεχομένως, είναι ο εντοπισμός ασυμβατότητας μεταξύ δύο ατόμων με τρόπο που να καταδεικνύει κατηγορηματικά ότι το ένα άτομο δεν θα μπορούσε να είναι ο γονέας του άλλου. Γνωρίζουμε ότι κάθε διπλοειδής απόγονος λαμβάνει μόνο ένα αλληλόμορφο ανά γενετικό τόπο από τη μητέρα του και ένα από τον πατέρα του. Έτσι, εάν ένας υποτιθέμενος απόγονος δεν μοιράζεται αλληλόμορφα σε κανέναν τόπο με έναν υποτιθέμενο γονέα, ο υποτιθέμενος γονέας μπορεί να αποκλειστεί από την ομάδα πιθανών γονέων (Chakraborty, Shaw, & Schull, 1974). Αυτό προϋποθέτει ότι δεν σημειώθηκαν μεταλλάξεις ή σφάλματα γονοτύπισης, μια υπόθεση της οποίας η εγκυρότητα ποικίλλει, ανάλογα με τον τύπο της υπό εξέταση γενετικής περιοχής. Ενώ οι μεταλλάξεις είναι σχετικά σπάνιες, ακόμη και στην πλειοψηφία των υπερμεταβλητών μικροδορυφορικών τόπων (Ellegren, 2000), τα σφάλματα γονοτύπισης είναι η κύρια αιτία στην μειούμενη χρήση της ως αποκλειστική μέθοδο ανάθεσης. Καθώς οι μελέτες αυξάνονται σε μέγεθος, όσον αφορά τόσο τον αριθμό δειγμάτων, όσο και τον αριθμό των μοριακών δεικτών, ένα σύνολο δεδομένων χωρίς λάθη είναι πρακτικά αδύνατο. Κατά συνέπεια, οι περισσότερες σύγχρονες εφαρμογές της μεθόδου του αποκλεισμού, λόγω της ύπαρξης σφαλμάτων, επιτρέπουν έναν ορισμένο αριθμό αναντιστοιχιών (mismatches) μεταξύ γονέων και απογόνων. Ενώ αυτή η προσέγγιση αποτρέπει τον πλήρη αποκλεισμό όλων των υποψηφίων γονέων, συμπεριλαμβανομένων των αληθινών, δεν παύει να εισαγάγει έναν υποκειμενικό παράγοντα στην μελέτη. Αν και η πιθανότητα ενός δεδομένου αριθμού ασυμβατότητας μεταξύ γονοτύπων γονέων και απογόνων μπορεί να εκτιμηθεί από τα ποσοστά σφάλματος γονότυπου (Chakraborty & Schull, 1976), αυτό το επιπλέον βήμα σπάνια υλοποιείται. Επιπλέον, η μέθοδος αποκλεισμού για διαλληλικά SNPs χρησιμοποιεί μόνο δεδομένα από ομόζυγους τόπους στους οποίους ένας υποτιθέμενος γονέας και ο απόγονος εμφανίζουν διαφορετικούς ομόζυγους γονότυπους π.χ AA και TT. Οι ετερόζυγοι τόποι σε γονείς ή απογόνους μπορεί να παρέχουν χρήσιμη πληροφορία (Kalinowski et al., 2007; Marshall et al., 1998), αλλά σε μια ανάλυση αποκλεισμού οι πληροφορίες αυτές απορρίπτονται, όσον αναφορά στα διαλληλικά SNP's. Σε γενικές γραμμές, η χρήση του αποκλεισμού ως μέθοδος για την ανάλυση σχέσεων γονέων-απογόνων εμφανίζει κάποια όρια τα οποία την καθιστούν επισφαλή κυρίως όσον αναφορά στα δεδομένα προερχόμενα από τις σύγχρονες τεχνολογίες νέας γενιάς. Βέβαια, η αξιολόγηση κάθε φορά της καταλληλότητας/ποιότητας των πρωτογενών δεδομένων, καθώς επίσης και της φύσης και χαρακτηριστικών του πληθυσμού

προς διερεύνηση είναι παράγοντες που μπορεί να επιτρέπουν την χρήση της μεθόδου με έναν αρκετά ασφαλή τρόπο.

2.1.5.2 Κατηγορική ανάλυση ανάθεσης (categorical allocations)

Μία συχνά χρησιμοποιούμενη προσέγγιση είναι η ανάθεση γονέων. Ιστορικά, η ανάθεση γονέων θα μπορούσε να χωριστεί σε δύο κατηγορίες: την κλασματική και κατηγορική κατανομή (Jones et al., 2010). Η κλασματική κατανομή αντιστοιχεί μερικούς απόγονους στους γονείς ως συνάρτηση των ύστερων πιθανοτήτων (posterior) (Nielsen et al., 2001), και έχει εξελιχθεί σε μια προσέγγιση γνωστή ως Μπαϋεσιανή γονική ανάλυση. Η κατηγορική ανάθεση, αναθέτει κάθε απόγονο εξ ολοκλήρου στον γονέα με τη μεγαλύτερη πιθανότητα, και αντιμετωπίζει την ανάλυση των γονέων ξεχωριστά από τις επακόλουθες εκτιμήσεις των μεταβλητών ενδιαφέροντος της πληθυσμιακής γενετικής. Η ανάθεση γονέων βασίζεται στον υπολογισμό και τη σύγκριση των πιθανοτήτων διαφορετικών υποθέσεων, σχετικά με τις σχέσεις μεταξύ υποτιθέμενων δυάδων γονέων-απογόνων, ή τριάδων, μητέρας-πατέρα-απογόνου. Γενικά, η μέθοδος της πιθανοφάνειας μελετά την πιθανότητα παρατήρησης των δεδομένων για μία δεδομένη υπόθεση. Σε αυτήν την περίπτωση, τα δεδομένα είναι οι γονότυποι και η υπόθεση είναι η προτεινόμενη σχέση μεταξύ των ατόμων. Οι απόλυτες πιθανοφάνειες, μπορούν να χρησιμοποιηθούν για τη σύγκριση εναλλακτικών υποθέσεων, δημιουργώντας τον λόγο (αναλογία) της πιθανοφάνειας για μια υπόθεση έναντι μιας δεύτερης (συχνά μηδενικής) υπόθεσης. Στην ανάθεση γονέων, αυτός ο λόγος περιλαμβάνει την υπόθεση ότι μια δυάδα ή τριάδα αντιπροσωπεύει ένα πραγματικό σύνολο γονέων-απογόνων έναντι της υπόθεσης ότι τα άτομα δεν σχετίζονται. Συνήθως, χρησιμοποιούνται οι λογάριθμοι των πιθανοτήτων και ο λόγος πιθανότητας μετατρέπεται σε μία τιμή LOD (*logarithm of odds*), που επί της ουσίας είναι το μέτρο της πιθανοφάνειας για ένα συγκεκριμένο αποτέλεσμα (Marshall et al., 1998). Μια θετική τιμή LOD δείχνει ότι η ύπαρξη γονικής σχέσης είναι πιο πιθανή, ενώ μια αρνητική τιμή LOD το αντίθετο, δεδομένων των γενετικών δεδομένων. Οι βαθμολογίες LOD, ενώ είναι χρήσιμες για την λύση της μέγιστης πιθανοφάνειας, έχουν το μειονέκτημα ότι δεν μπορεί να ερμηνευθεί η στατιστική τους σημαντικότητα λόγω της τοπικότητας τους. Οι Marshall et al. (1998) αναγνώρισαν ότι οι κρίσιμες τιμές για τις τιμές LOD θα μπορούσαν να προσδιοριστούν χρησιμοποιώντας την μέθοδο της προσομοίωσης. Η προσέγγιση αυτή, χρησιμοποιεί τη διαφορά Δ στη τιμή LOD μεταξύ του πιο πιθανού και του δεύτερου πιο πιθανού γονέα (ή της πρώτης τιμής LOD εάν μόνο ένας υποψήφιος έχει θετική τιμή) και προσομοιώνει πληθυσμούς γονέων και απογόνων για τον καθορισμό της κρίσιμης τιμής Δ που οδηγεί σε ένα επιθυμητό επίπεδο εμπιστοσύνης την ανάλυση. Αυτή η προσέγγιση, η οποία εξακολουθεί να χρησιμοποιείται ευρέως 20 χρόνια τώρα, ήταν και η πρώτη που ήλεγξε το πειραματικό σφάλμα στη γονική ανάλυση.

2.1.5.3 Μπαϋεσιανή μέθοδος

Η μέθοδος αυτή ξεκίνησε ως μια τεχνική κλασματικής (fractional) κατανομής των απογόνων στους γονείς ως συνάρτηση των ύστερων πιθανοτήτων (posterior) (Devlin κ.ά.,

1988). Όπου, ο υποτιθέμενος γονέας με την υψηλότερη ύστερη πιθανότητα αποδίδεται στο μεγαλύτερο μέρος του απογόνου, όπως επίσης και ο απόγονος κατανέμεται επίσης μερικώς σε οποιονδήποτε γονέα με μη-μηδενική πιθανότητα. Λόγω της βιολογικής της «παραδοξότητας» η κλασματική προσέγγιση δεν έτυχε ιδιαίτερης επιτυχίας αρχικά, καθώς ένας ενήλικος, φαινομενικά, δεν μπορεί να είναι 4% ο γονέας ενός απογόνου, παρόλο, που αυτό το αποτέλεσμα είναι δυνατό σε αυτού του τύπου την ανάλυση. Το γεγονός αυτό οδήγησε σε μία ευρεία υιοθέτηση της κατηγορικής ανάθεσης, παρόλο που η Μπαϋεσιανή μέθοδος έχει καλύτερες στατιστικές ιδιότητες για την εκτίμηση πολλών τιμών ενδιαφέροντος (Neff, Repka, & Gross, 2001). Ένα από τα πλεονεκτήματα της μεθόδου είναι ότι μπορεί να συνεκτιμήσει παραμέτρους ενδιαφέροντος παράλληλα με το μοτίβο σχέσεων γονέων-απογόνων. Ορισμένες από αυτές τις παραμέτρους σχετίζονται με την ίδια την ανάλυση της ανάθεσης, όπως το ποσοστό της δειγματοληψίας των υποψηφίων γονέων (Nielsen et al., 2001), η διακύμανση στην αναπαραγωγική επιτυχία ή το ποσοστό με το οποίο η γονική ηλικία μειώνεται ως συνάρτηση της απόστασης μεταξύ των πιθανών διασταυρώσεων (Hadfield et al., 2006). Το πλεονέκτημα αυτής της προσέγγισης για την εκτίμηση τέτοιων παραμέτρων είναι, ότι η αβεβαιότητα στην ανάλυση των σχέσεων γονέων-απογόνων, περιλαμβάνεται ως αβεβαιότητα στις τελικές εκτιμήσεις. Στην περίπτωση της κατηγορικής ανάθεσης, η αβεβαιότητα συνήθως απορρίπτεται στην επόμενη φάση της ανάλυσης, καθώς οι αναθέσεις αντιμετωπίζονται ως το σημείο αφετηρίας κατά την εκτίμηση παραμέτρων πληθυσμιακής γενετικής. Λόγω της φύσης της Μπαϋεσιανής μεθόδου, λεπτές διαφορές στις εκτιμήσεις των πιθανοτήτων επιλεγμένων ατόμων μπορούν να ενσωματωθούν με ένα φυσικό τρόπο, ενώ στην κατηγορική ανάλυση, τέτοιες πληροφορίες μπορούν να συμπεριληφθούν, είτε συμπεριλαμβάνοντας, είτε αφαιρώντας άτομα από τη λίστα των υποψηφίων γονέων. Το μεγαλύτερο μειονέκτημα της Μπαϋεσιανής μεθόδου είναι ότι κάθε ανάλυση απαιτεί προσεκτική επεξεργασία της πρότερης εξίσωσης πιθανότητας (*prior probability*) και μπορεί να επηρεαστεί λόγω των αποφάσεων που λαμβάνονται κατά τη διάρκεια αυτού του βήματος.

2.1.6 Ανακατασκευή των σχέσεων γονέων- απογόνων (sibship reconstruction and parentage assignment)

Η γνώση της κληρονομικότητας οδηγεί στο συμπέρασμα ότι ο πλήρης γονότυπος ενός γονέα μπορεί να ανακατασκευαστεί εάν είναι γνωστοί οι γονότυποι του άλλου γονέα και πολλών από τους απογόνους (Jones & Avise, 1997b). Δεδομένου ότι μια τέτοια πλήρης ανακατασκευή απαιτεί πολλούς απογόνους από μια οικογένεια, σε συνδυασμό με μια εκ των προτέρων γνώση ότι αυτοί οι απόγονοι είναι είτε ομοθαλή ή ετεροθαλή αδέρφια, την καθιστά χρήσιμη μόνο σε είδη για τα οποία μπορούν να συλλέγονται ταυτόχρονα και οι απόγονοι. Εάν υπάρχει ένας τέτοιος πίνακας απογόνων, τότε το πλήρες σύνολο πιθανών γονέων για τον πίνακα απογόνων μπορεί να οριστεί χρησιμοποιώντας έναν αλγόριθμο συνδυαστικής ή μέγιστης πιθανοφάνειας (Jones, 2001, 2005; Wang, 2004). Σε γενικές γραμμές, οι προσεγγίσεις μέγιστης πιθανοφάνειας είναι προτιμότερες, καθώς οι εξαντλητικοί συνδυαστικοί αλγόριθμοι είναι απαγορευτικά αργοί για περισσότερες από λίγες τοποθεσίες και απαιτούν ένα σύνολο δεδομένων υπερμεταβλητών τόπων, χωρίς σφάλματα. Η ανακατασκευή των αδελφικών σχέσεων εμφανίζει ομοιότητες με αυτή των γονέων-απογόνων και ορισμένοι αλγόριθμοι μπορούν να χρησιμοποιηθούν για την πραγματοποίηση μιας πολύ παρόμοιας ανάλυσης (Wang, 2004). Η ανακατασκευή των αδελφικών σχέσεων έχει το πλεονέκτημα ότι δεν απαιτεί εκ των προτέρων γνώση των σχέσεων μεταξύ των γονοτυπημένων ατόμων. Για ένα δείγμα που περιέχει υποτιθέμενα ομοθαλή και ετεροθαλή αδέρφια, χρησιμοποιούνται αλγόριθμοι ομαδοποίησης για να κατηγοριοποιήσουν τους

γονότυπους σε οικογένειες (Almudevar, 2003). Για τις τελευταίες εξελίξεις στην ανοικοδόμηση αδελφών και γενεαλογικών (συμπεριλαμβανομένων των απομακρυσμένων συγγενών), εκτός του πεδίου της ανάλυσης των γονέων καθαυτή, Staples, et al., (2014), Staples, et al., (2016) και Ko & Nielsen (2017).

2.2 Δύο τρόποι σχεδιασμού επιλογής γεννητόρων, βασισμένες σε φαινοτυπικά δεδομένα: στην πρώτη περίπτωση το φύλο, και στη δεύτερη περίπτωση αποκλειστικά στην γενετική περιγραφή προερχόμενη από μικροδορυφορικούς γενετικούς δείκτες.

Τα δύο παραδείγματα που εξετάζουμε προέρχονται από γονείς και απογόνους μαγιάτικου, *Seriola dumerili*, εντατικά εκτρεφόμενων σε μονάδες ιχθυοκαλλιέργειας, και αφορούν την επιλογή νέων υποψήφιων γεννητόρων προερχόμενων από μία διασταύρωση. Στους δύο τρόπους που προτείνουμε, έχει χρησιμοποιηθεί ένα πάνελ 13 μικροδορυφορικών τόπων για την ανάθεση των απογόνων στους γονείς (Πίνακας 1, Renshaw et al., 2006). Η διαδικασία της εξαγωγής DNA, PCR και γονοτύπισης περιγράφεται αναλυτικά στην ενότητα 2.2.7 με την διαφορά ότι η PCR για το μαγιάτικο έχει στον πρώτο κύκλο θερμοκρασία 59°C.

Πίνακας 1. Σύνοψη για τους 13 μικροδορυφορικούς τόπους που χρησιμοποιήθηκαν για τις αναλύσεις στο μαγιάτικο. Στην πρώτη στήλη δίνεται το όνομα του κάθε τόπου ενώ στην δεύτερη στήλη οι εμπρόσθιες και οι οπίσθιες αναστροφές τους αλληλουχίες συμβολίζονται με «F» και «R», αντίστοιχα. Ο σημασμένος φθορίζων εκκινητής (primer) δίνεται με το σύμβολο.

Μικροδορυφορικός τόπος	Primer αλληλουχία (5'-3')
Sdu29	F-CCTTGCCATACCGATGCCAG* R-GACTGCTCTGCCTGCTTGTG
Sdu31	F-CACATTTGGACGGATTCTTC* R-GCTGTTATCCTCCAGTGCT
Sdu32	F-CCTGTGAGAGCATTGGTAT** R-GTGCTTGTCTCTTCTGTTCAT
Sdu34	F-CCTTGTGTTGTATCTGCTGTAA*** R-GGAATAAACCTCGTCTGTCA
Sdu36	F-CTGTTATGAAGCAGTGAAGAGG* R-GGACCATCCTGCTCTGACA
Sdu37	F-CCTCTAATGGACTTCAGCG*** R-GGTTATTTTGAGAGCCGTC
Sdu39	F-AGTGGCTTCTGCTGCTGT** R-CGTGTGCGTGCTTGTA
Sdu40	F-CGATGCTTTCAACTCCGACACAC*** R-CCATCCTTCATCAGCAACAACATCC
Sdu41	F-AGCGTGACAGTTTATGG** R-GTCTGTTTACTGGTCGCA
Sdu46	F-GCAGTGTGAGCCATACATTAC*** R-CTACAGGACAAAAGCCATT
Sdu10	F-CCAAGTCCTCCTGCTACTACCAT* R-CCTTGTGGATGACCTGTTTG
Sdu 21	F-CTCAGGACAATGTTGGTAG* R-GCTAACAAGTTCACGACAT
Sdu 22	F-CATTCTCCAAGTATGTGACCTC** R-GCTCTATGCGAATACCTCCA

2.2.1 Περίπτωση 1: Λαμβάνοντας υπόψη τα δεδομένα για το φύλο και το βάρος των γονιών και των απογόνων, ως κριτήρια για την επιλογή γεννητόρων

Συγκεκριμένα, 14 αρσενικοί και 13 θηλυκοί γονείς διασταυρώθηκαν τυχαία και έδωσαν 117 απογόνους.

Οι στόχοι, τους οποίους έχουμε θέσει υποκειμενικά, είναι να γίνει επιλογή γεννητόρων από τους απογόνους, για τον περιορισμό της ενδογαμίας στις επόμενες γενιές, ώστε να πληρούνται οι εξής προϋποθέσεις :

- οι περισσότερες, αν όχι όλες οι οικογένειες πρέπει να εκπροσωπούνται γενετικά στην επόμενη γενιά,
- κάθε οικογένεια πρέπει να έχει ίσο αριθμό γεννητόρων, ώστε η γενετική παραλλακτικότητα να είναι η μέγιστη, καθώς επίσης και για την διατήρηση των συχνοτήτων των αλληλομόρφων,
- το βάρος συμπεριλαμβάνεται στις παραμέτρους επιλογής καθώς δίνονται οι τιμές των βαρών των απογόνων,
- το φύλο συμπεριλαμβάνεται στις παραμέτρους επιλογής καθώς δίνεται το φύλο των απογόνων,
- ο αριθμός των θηλυκών να είναι, κατά το δυνατόν, ίσος με αυτόν των αρσενικών γεννητόρων.

Οι παραδοχές που ακολουθούμε, οι οποίες έχουν τεθεί από εμάς, είναι οι παρακάτω :

1. Η απουσία συγγένειας μεταξύ των γονέων,
2. Οι απόγονοι που θα ανατεθούν, της F1 γενιάς, θα σχηματίσουν οικογένειες ανάμοιων μεγεθών λόγω της άνισης συνεισφοράς των θηλυκών και αρσενικών γεννητόρων,
3. Οι διασταυρώσεις μεταξύ των γεννητόρων να είναι τυχαίες

Στην παρούσα περίπτωση επιλογής γεννητόρων έχουμε την πληροφορία για το φύλο και το βάρος των απογόνων από τους οποίους θα επιλέξουμε τους υποψήφιους γεννήτορες. Η ύπαρξη της πληροφορίας για το φύλο μας δίνει μία επιπλέον δυνατότητα. Από την κάθε οικογένεια επιλέγονται γεννήτορες που ανήκουν σε ένα από τα δύο φύλα, αλλά όχι και στα δύο. Μέσω αυτής της επιλογής, η διασταύρωση μεταξύ ομοθαλών (*full-sibs*) αδελφών καθίσταται αδύνατη και συνεπώς μόνο ετεροθαλή (*half-sibs*) αδέλφια είναι δυνατόν να διασταυρωθούν μειώνοντας έτσι την ενδογαμία στην F2 γενιά.

Η ανάθεση των γονιών στους απογόνους είναι το πρώτο και απαραίτητο βήμα στην διαδικασία επιλογής γεννητόρων. Για την ανάθεση χρησιμοποιήθηκε το λογισμικό COLONY (Jones, O. and Wang, J. 2010).

Στην περίπτωση όπου θα επιλέξουμε γεννήτορες με βάση το φύλο και το βάρος, οκτώ (8) οικογένειες προκύπτουν με ανάμοια μεγέθη. Παρατηρείται επίσης ότι αυτές οι οικογένειες προέρχονται από ένα μικρό κλάσμα των συνολικών γονιών. Έξι αρσενικοί και 3 θηλυκοί γονείς συμμετέχουν από το σύνολο των 27 (14 αρσενικά και 13 θηλυκά άτομα). Για τους 91 απογόνους με βέβαιη ανάθεση βλέπουμε στον παρακάτω **Πίνακα 2** την συνεισφορά των γονιών.

Πίνακας 2. Διαμορφούμενες οικογένειες απογόνων μαγιάτικου μετά από μία διασταύρωση 27 (14 αρσενικών και 13 θηλυκών) γεννητόρων

	♂					
♀	A6	A10	A12	A22	A23	A27
Θ1	38	-	-	-	11	10
Θ3	-	8	18	6	-	-
Θ8	2	-	-	-	-	1

Από τις 8 οικογένειες, και για να διασφαλιστεί η μέγιστη γενετική παραλλακτικότητα στην επόμενη γενιά και η ελάχιστη ενδογαμία λαμβάνουμε υπόψη τους στόχους και τις παραμέτρους που θέσαμε παραπάνω.

Λόγω του γεγονότος ότι η μικρότερη οικογένεια με αριθμό απογόνων είναι μίας τάξης μεγέθους μικρότερη από αυτή των γονέων και > 1 , αριθμεί 6 άτομα (οικογένεια A22/Θ3), επιλέγουμε αυτό το νούμερο ως τον βέλτιστο αριθμό γεννητόρων από κάθε οικογένεια που έχουν ανατεθειμένους απογόνους περισσότερους των 6, με εξαίρεση τις οικογένειες που είχαν ανατεθειμένους λιγότερους απογόνους όπως η A27/Θ8. Έτσι θα διασφαλιστεί η μέγιστη εκπροσώπηση της κάθε οικογένειας στην επόμενη γενιά καθώς επίσης και η όσο το δυνατόν μεγαλύτερη γενετική παραλλακτικότητα στους απογόνους της F2 γενιάς (σύμφωνα με τα αποτελέσματα της ανάθεσης). Η επιλογή αυτή γίνεται απο εμάς, με βάση την κατανομή των απογόνων στις οικογένειες. Με βάση τα συγκεκριμένα δεδομένα δεν είναι εφικτό να επιλέξουμε 6 απογόνους από κάθε οικογένεια λόγω του ότι ο συνδυασμός του βέλτιστου βάρους, και της επιλογής ατόμων μόνο ενός εκ των δύο φύλων από κάθε οικογένεια θέτει τα όρια της επιλογής.

Συνολικά επιλέξαμε 15 θηλυκούς και 18 αρσενικούς απογόνους της F1 γενιάς για γεννήτορες. Το μέσο βάρος τους είναι 5.300 γραμ. με το μέσο βάρος των ανατεθειμένων απογόνων να είναι 5.441 γραμ. και του συνόλου των απογόνων 5.351 γραμ, **Πίνακας 3**.

Πίνακας 3: Σύνοψη των στοιχείων των προς επιλογή υποψήφιων γεννητόρων από την F1 γενιά

Οικογένεια	Επιλογή	Μέσο βάρος	Φύλο
A6/Θ8	2	5850	♀
A27/Θ8	1	5100	♂
A6/Θ1	6	6008	♂
A22/Θ3	6	4400	♂
A12/Θ3	5	5140	♀
A23/Θ1	4	5450	♂
A10/Θ3	3	5266	♀
A27/Θ1	6	6391	♀

2.2.2 Περίπτωση 2. Διαθέσιμα δεδομένα προερχόμενα αποκλειστικά από γενετικά δεδομένα (μικροδορυφορικοί γενετικοί δείκτες)

Οι απόγονοι στους οποίους ανατέθηκαν οι πιθανοί γονείς είναι 101 από το σύνολο των 106. Το πάνελ των μικροδορυφορικών τόπων που χρησιμοποιήθηκε είναι το ίδιο με την πρώτη περίπτωση (Πίνακας 1). Οι γονείς είναι οι ίδιοι με του πρώτου παραδείγματος, 27 στο σύνολο (14 αρσενικά και 13 θηλυκά άτομα). Οι γονείς που συνεισέφεραν στις 6 οικογένειες που διαμορφώθηκαν είναι 7. Επιπλέον, δεν είχαμε πληροφορία για το φύλο των απογόνων ή κάποιο άλλο χαρακτήρα και τα ψάρια ήταν μικρά χωρίς να έχουν διαφοροποιηθεί φυλετικά. Για την ανάθεση των γονιών στους απογόνους χρησιμοποιήθηκαν δύο λογισμικά. Το COLONY όπου έγινε η ανάθεση με την μέθοδο της μέγιστης πιθανοφάνειας και του VITASSIGN (Vandeputte et al. 2006). Το VITASSIGN χρησιμοποιήθηκε για την ανάθεση των μη-ανατεθειμένων απογόνων από την ανάλυση του COLONY, καθώς από την πρώτη ανάθεση έμειναν 46 απόγονοι εκτός της ανάλυσης. Το VITASSIGN, αναθέτει τους γονείς στους απογόνους αντιστοιχίζοντας τα αλληλόμορφα των μικροδορυφορικών τόπων των απογόνων με αυτά των γονιών με βάση την αρχή του Μέντελ, του διαχωρισμού των αλληλομόρφων γονιδίων, που στην προκειμένη περίπτωση είναι οι μικροδορυφορικοί τόποι. Λόγω της διαφορετικής προσέγγισης του VITASSIGN είναι δυνατόν να χρησιμοποιηθεί συμπληρωματικά ώστε να βελτιώσει το ποσοστό ανάθεσης καθώς χρησιμοποιεί μία εξατομικευμένη προσέγγιση, όπου η ανάθεση η μη ενός απογόνου σε έναν γονέα εξαρτάται από το όριο των αναντιστοιχιών των αλληλομόρφων που θα εισαχθεί από τον χρήστη. Στον παρακάτω Πίνακα 4 βλέπουμε τις διαμορφούμενες οικογένειες με βάση τα δύο λογισμικά.

Πίνακας 4: Διαμορφούμενες οικογένειες απογόνων μαγιάτικου μετά από μία διασταύρωση 14 αρσενικών και 13 θηλυκών γεννητόρων

	♂			
♀	A12	A16	A18	A20
Θ13	10	2	-	-
Θ14	58	22	-	-
Θ15	-	-	3	4

Σύμφωνα με τα προαναφερθέντα θεωρήσαμε τις παρακάτω παραδοχές για την επιλογή:

1. Ισορροπημένη εκπροσώπηση των οικογενειών στην επιλογή των γεννητόρων,
2. Η ετεροζυγωτία των προτεινόμενων απογόνων ως υποψήφιοι γεννήτορες.

Για κάθε αλληλόμορφο υπολογίστηκε η συχνότητα του σύμφωνα με την εξίσωση (2.2.2.1).

$$p = F(AA) + \frac{1}{2}F(AB) \quad (2.2.2.1)$$

όπου $F(AA)$ είναι η συχνότητα των ομόζυγων ατόμων ενός αλληλόμορφου και $F(AB)$ είναι η συχνότητα των ετερόζυγων ατόμων του ίδιου αλληλόμορφου, για έναν διπλοειδή μικροδορυφορικό τόπο σε όλους τους ανατεθειμένους απογόνους. Παρακάτω παραθέτουμε

τις συχνότητες των αλληλομόρφων για όλους τους τόπους, έχοντας συνολικά 73 αλληλόμορφα για 13 τόπους, **Πίνακας 5**.

Πίνακας 5: Σύνοψη των αλληλομόρφων, της συχνότητας και του αριθμού για το πάνελ των μικροδορυφορικών τόπων. Στην πρώτη στήλη αναφέρεται το όνομα του μικροδορυφορικού τόπου, στην δεύτερη στήλη ο αριθμός (N) των αλληλομόρφων σε κάθε τόπο και στην τρίτη στήλη η συχνότητα (Σ). Με πράσινο χρώμα είναι υπογραμμισμένα τα αλληλόμορφα τα οποία εμφανίζονται σε ομοζυγωτία τουλάχιστον μία φορά

sdu 10	N	Σ	sdu 21	N	Σ	sdu 22	N	Σ
297	48	0,2376	315	2	0,0099	318	128	0,6336
303	38	0,1881	323	47	0,2327	321	31	0,1534
306	9	0,0445	331	40	0,1980	327	2	0,0099
318	3	0,0014	335	1	0,0050	330	35	0,1732
324	15	0,0742	339	3	0,0149	348	6	0,0297
327	28	0,1386	343	61	0,3713			
330	6	0,0297	347	9	0,0446			
339	55	0,272	351	2	0,0099			
			367	1	0,0050			
			383	5	0,0248			
			387	1	0,0050			
			291	16	0,0792			

sdu 29	N	Σ	sdu 31	N	Σ	sdu 32	N	Σ
311	112	0,5742	086	76	0,3762	107	65	0,2821
329	6	0,0297	088	1	0,0049	135	13	0,0445
331	42	0,2178	094	50	0,2475	141	35	0,1732
333	8	0,0396	098	75	0,3712	143	88	0,4851
335	32	0,1584				157	1	0,0049

sdu 34	N	Σ	sdu 36	N	Σ	sdu 37	N	Σ
086	3	0,0148	206	6	0,0297	168	45	0,2227
088	50	0,2475	208	52	0,2574	174	67	0,3316
090	1	0,0049	210	43	0,2128	176	90	0,4455
092	75	0,3712	214	101	0,5			
096	6	0,0297						
102	62	0,3069						

sdu 39	N	Σ	sdu 40	N	Σ	sdu 41	N	Σ
146	1	0,0049	205	54	0,2673	004	36	0,1326
156	121	0,5990	217	37	0,1831	028	28	0,1425
160	1	0,0049	227	36	0,1782	104	64	0,3168
166	42	0,2079	249	75	0,3712	128	14	0,0779
178	37	0,1831				961	57	0,2118
						981	3	0,0187

sdu 46	N	Σ
211	2	0,0009
213	51	0,2547
219	5	0,0247
241	28	0,1386
243	111	0,5495
245	5	0,0247

Έχοντας υπολογίσει τις συχνότητες των αλληλομόρφων για όλους τους απογόνους στους οποίους έγινε ανάθεση, θα αξιολογήσουμε τον κάθε απόγονο ξεχωριστά στη βάση ενός κριτηρίου: του αριθμού των διαφορετικών αλληλομόρφων, διαιρεμένου με τον μέγιστο διαφορετικό αριθμό αλληλομόρφων δηλ. 26 αλληλόμορφα για τους 13 γενετικούς τύπους (στην περίπτωση που το άτομο είναι ετεροζυγο και για τους 13 τοπους), όπως φαίνεται στην εξίσωση (2.2.2.2). Όπου a είναι το μέτρο της ετεροζυγωτίας με βάση τους γονοτυπημένους σε κάθε άτομο τύπους.

$$a = \frac{(\text{αριθμός διαφορετικών αλληλομόρφων})}{26} \quad (2.2.2.2.)$$

Όσο πιο κοντά στη μονάδα είναι η τιμή, τόσο μεγαλύτερο το μέτρο της ετεροζυγωτία για τον κάθε απόγονο. Οι γεννήτορες με την μεγαλύτερη ετεροζυγωτία θα επιλεγθούν, λαμβάνοντας επίσης το κριτήριο που προαναφέρθηκε της βέλτιστης εκπροσώπησης της κάθε οικογένειας.

Ο στόχος είναι να επιλεγούν 3 ομάδες ιχθύων με μέγιστο των 25 ατόμων ανά δεξαμενή γεννητόρων. Επιλέγουμε τους 75 υποψήφιους γεννήτορες με την μεγαλύτερη ετεροζυγωτία.

Η ανάθεση των γονιών στους απογόνους μας έδωσε σημαντική ανομοιογένεια στις οικογένειες της F1 γενιάς. Η πλειοψηφία των ατόμων ανήκουν στην οικογένεια A12/Θ14. Αυτό πρέπει να ληφθεί υπόψη στην τοποθέτηση των γεννητόρων στις ομάδες. Τα 39 άτομα που ανήκουν σε αυτή την οικογένεια θα χωριστούν σε 3 ομάδες των 13 ατόμων. Στην τοποθέτηση των γεννητόρων στις ομάδες θα λάβουμε υπόψη το μέτρο της ετεροζυγωτίας. Θέλουμε η κάθε ομάδα να έχει παρόμοιο μέσο όρο, **Πίνακας 6**.

Πίνακας 6: Συνολική εικόνα των διαμορφούμενων τριών ομάδων με τους υποψήφιους προς επιλογή γεννήτορες

	Ομάδα 1			Ομάδα 2			Ομάδα 3		
1	16/14	E2-098	1	12/14	E2-078	1	16/14	E2-022	1
2	12/14	E2-060	0,9615	16/14	E2-037	1	12/14	E2-094	0,9615
3	16/14	E2-032	0,9615	16/14	E2-021	0,9615	16/14	E2-010	0,9615
4	12/14	E2-091	0,9615	12/13	E2-070	0,9615	12/14	E2-069	0,9615
5	12/14	E2-066	0,9615	12/14	E2-064	0,9615	12/14	E2-049	0,9615
6	16/13	E2-053	0,9615	16/14	E2-052	0,9615	16/14	E2-045	0,9615
7	20/15	E2-029	0,9615	12/14	E2-044	0,9615	12/14	E2-024	0,9615
8	16/14	E2-011	0,9615	16/14	E2-013	0,9615	16/14	E2-014	0,9615
9	16/14	E2-015	0,9615	16/14	E2-018	0,9615	12/14	E2-092	0,9231
10	12/14	E2-083	0,9231	12/14	E2-090	0,9231	12/14	E2-102	0,9231
11	12/13	E2-058	0,9231	16/14	E2-103	0,9231	12/13	E2-056	0,9231
12	16/14	E2-055	0,9231	12/14	E2-039	0,9231	16/14	E2-051	0,9231
13	12/14	E2-104	0,9231	12/14	E2-059	0,9231	12/14	E2-027	0,9231
14	12/14	E2-077	0,9231	12/14	E2-063	0,9231	12/14	E2-086	0,9231
15	12/14	E2-089	0,9231	12/14	E2-101	0,9231	16/14	E2-046	0,9231
16	16/14	E2-030	0,9231	12/14	E2-041	0,9231	12/14	E2-036	0,9231
17	12/14	E2-016	0,9231	16/13	E2-012	0,9231	12/14	E2-079	0,8846
18	12/14	E2-065	0,8846	12/14	E2-034	0,8846	12/14	E2-033	0,8846
19	12/14	E2-023	0,8846	12/14	E2-097	0,8846	18/15	E2-095	0,8846
20	12/14	E2-085	0,8846	16/14	E2-084	0,8846	16/14	E2-076	0,8846
21	12/14	E2-074	0,8846	12/14	E2-071	0,8846	20/15	E2-067	0,8846
22	16/14	E2-050	0,8846	12/13	E2-048	0,8846	16/14	E2-031	0,8846
23	20/15	E2-038	0,8846	12/14	E2-026	0,8846	12/14	E2-020	0,8846
24	12/14	E2-082	0,8462	12/13	E2-017	0,8846	12/14	E2-087	0,8462
25	18/15	E2-072	0,8462	20/15	E2-099	0,8462	12/13	E2-008	0,8846
		Μέσος όρος	0,8990		Μέσος όρος	0,9014		Μέσος όρος	0,8990

2.2.3 Σύγκριση των 2 περιπτώσεων/προσεγγίσεων για το μαγιάτικο

Με βάση τα δύο παραδείγματα επιλογής γεννητόρων, δείχνουμε την εκτίμηση του εύρους της ενδογαμίας στην επόμενη γενιά.

Για να εκτιμήσουμε το εύρος της ενδογαμίας στην επόμενη γενιά με βάση την επιλογή των γεννητόρων από τα δύο σετ δεδομένων, των μικροδορυφορικών τόπων αποκλειστικά και των δεδομένων που παρέχουν πληροφορία για το φύλο και το βάρος, ακολουθήσαμε τα παρακάτω βήματα :

1. Για το σετ δεδομένων των μικροδορυφορικών τόπων ανατέθηκε φύλο στους 75 γεννήτορες ακολουθώντας την αναλογία 1:1. Έτσι έχουμε 37 υποψήφιους «αρσενικούς» γεννήτορες και 38 «θηλυκούς». Για το σετ δεδομένων στο οποίο παρέχεται το φύλο και το βάρος των απογόνων, με βάση την επιλογή που έγινε είχαμε 17 αρσενικούς γεννήτορες και 16 θηλυκούς.
2. Ακολουθήθηκε ένα πλήρες σχήμα εικονικών διασταυρώσεων όπου όλοι οι αρσενικοί γεννήτορες διασταυρώθηκαν με όλους τους θηλυκούς.
3. Ο αριθμός των απογόνων καθορίστηκε με βάση την συνάρτηση randbetween του Excel όπου για κάθε οικογένεια ο ελάχιστος αριθμός απογόνων ήταν 2 και ο μέγιστος 15, με την συνάρτηση αυτή δίνεται μία τυχαία τιμή απογόνων στην κάθε οικογένεια στο καθορισμένο εύρος τιμών. Ο μέγιστος αριθμός απογόνων μίας οικογένειας καθορίστηκε από τον μέσο όρο που ήδη είχαμε στις οικογένειες των δύο σετ δεδομένων.
4. Στο σετ δεδομένων με μικροδορυφορικούς τόπους από την διασταύρωση των 37 «αρσενικών» γεννητόρων με τους 38 «θηλυκούς», είχαμε τον μέγιστο αριθμό οικογενειών, 1406 οικογένειες στην F2 γενιά, με σύνολο απογόνων 11.095. Στο σετ δεδομένων όπου παρέχονται το φύλο και το βάρος είχαμε τον μέγιστο αριθμό οικογενειών 270 και 2375 απογόνους στην F2 γενιά (τα δεδομένα δεν παρέχονται).
5. Το λογισμικό Pedigraph (John R. Garbe and Yang Da, 2008) από το πανεπιστήμιο της Μινεσότα χρησιμοποιήθηκε για την κατασκευή του γενεαλογικού δέντρου, τον υπολογισμό των συντελεστών ενδογαμίας και τον μέσο όρο τους.

Με βάση το γενεαλογικό δέντρο, το λογισμικό Pedigraph υπολόγισε τους συντελεστές ενδογαμίας ατομικά και συνολικά ως μέσο όρο για τους απογόνους της F2 γενιάς και στα δύο σετ δεδομένων. Τα αποτελέσματα παρατίθενται στον **Πίνακα 7**.

Πίνακας 7. Το εύρος των συντελεστών ενδογαμίας για τα δύο σετ δεδομένων με την προσομοίωση της F2 γενιάς. Το σετ δεδομένων 1 αντιστοιχεί στα δεδομένα όπου το φύλο και το βάρος δίνονται (τα γενετικά δεδομένα για τους μικροδορυφορικούς τόπους είναι κοινά μα χρησιμοποιούνται μόνο στην μία περίπτωση για την επιλογή γεννητόρων), ενώ το σετ δεδομένων 2 αντιστοιχεί στην περίπτωση που εξετάζουμε μόνο τα γενετικά δεδομένα μικροδορυφορικών τόπων για την επιλογή γεννητόρων. Στην στήλη «Κατηγορίες f», βλέπουμε, τους συντελεστές ενδογαμίας των απογόνων που προκύπτουν από την διασταύρωση ομοθαλών αδελφών $f = 0,25$, ετεροθαλών αδελφών $f = 0,125$ και όταν δεν υπάρχει καμία συγγένεια μεταξύ των γονιών $f = 0$. Οι τιμές αυτές είναι οι ελάχιστες τιμές. Στην στήλη με την ονομασία «Ποσοστό απογόνων» βλέπουμε το ποσοστό των απογόνων που προέκυψαν από την διασταύρωση των υποψήφιων γεννητόρων στα 2 σετ δεδομένων με βάση την στήλη «Κατηγορίες f», δηλαδή τους απογόνους, ομοθαλών αδελφών, ετεροθαλών αδελφών και μη συγγενών γεννητόρων. Στην στήλη «Μέσος όρος F2 γενιάς» βλέπουμε τον συνολικό μέσο όρο του συντελεστή ενδογαμίας για τα 2 σετ δεδομένων.

Δεδομένα	Κατηγορίες f	Ποσοστό απογόνων	Μέσος όρος F2 γενιάς
Σετ δεδομένων 1	0,25	<u>0,60%</u>	0.05552 ή 5.5%
	0,125	44%	
	0	55,40%	
Σετ δεδομένων 2	0,25	<u>37,85%</u>	0.14744 ή 14.7%
	0,125	43,08%	
	0	19,05%	

Η διασταύρωση ομοθαλών αδελφών είναι πρακτικά αδύνατη στο σετ δεδομένων 1. Η διασταύρωση ομοθαλών αδελφών δίνει συντελεστή ενδογαμίας 0,25, ενώ των ετεροθαλών 0,125. Στο σετ που χρησιμοποιούνται οι μικροδορυφορικοί γενετικοί δείκτες για την επιλογή, επειδή η επιλογή έγινε μόνο με βάση την μέγιστη ετερογυζωτία και το φύλο δεν ήταν γνωστό, δεν θα μπορούσε να μεθοδευτεί η μη διασταύρωση ομοθαλών αδελφών, γεγονός που αποτελεί και το μεγαλύτερο μειονέκτημα του. Ο συντελεστής ενδογαμίας για τα δεδομένα των μικροδορυφορικών τόπων διαμορφώνεται αρκετά υψηλός, με την προϋπόθεση ότι αυτοί οι γεννήτορες διασταυρώνονται μεταξύ τους αποκλειστικά και η αναλογία των φύλων είναι κοντά στο 1:1, οι απόγονοι της F2 γενιάς εμφανίζουν κατά μέσο όρο 14.7% ενδογαμία. Στις κατευθυντήριες γραμμές της, στο κείμενο «Δευτερεύουσες κατευθυντήριες γραμμές», «για την ανάπτυξη εθνικών γενετικών αποθεμάτων και διαχειριστικών σχεδίων των κτηνοτροφικών ζώων» (<http://www.fao.org/3/a-w9361e.pdf>) ο FAO θέτει τα αποδεκτά όρια της ενδογαμίας σε ένα μικρό πληθυσμό. Εισάγεται ή έννοια του ρυθμού ενδογαμίας ΔF για μία γενιά, όπου όταν δεν υπάρχει επιλογή ο ρυθμός της ενδογαμίας μπορεί να εκτιμηθεί με βάση την εξίσωση (2.2.3.1) (Wray and Thomson, 1990).

$$\Delta F = 18Nm + 18Nf \quad (2.2.3.1)$$

όπου Nm είναι ο αριθμός των αρσενικών γεννητόρων και Nf ο αριθμός των θηλυκών για μία γενιά. Στη δική μας περίπτωση όπου έχει γίνει επιλογή γεννητόρων, ο ΔF διαμορφώνεται υψηλότερος και η εξίσωση (5) δεν είναι η βέλτιστη. Η χρήση του GCI (δείκτης γενετικής διατήρησης, genetic conservation index) δίνει μία καλύτερη εικόνα, εξίσωση (2.2.3.2) (Alderson, 1990).

$$GCI = 1 / \sum Pi \quad (2.2.3.2)$$

όπου Pi είναι η αναλογία των γονιδίων τα οποία προέρχονται από έναν γεννήτορα i με το Σ να είναι το άθροισμα όλων των γεννητόρων, με την προϋπόθεση ότι το γενεαλογικό δέντρο όλων των διασταυρώσεων είναι διαθέσιμο. Εάν έχουν προηγηθεί πολλαπλές γενιές, οι οποίες έχουν καταγραφεί σε γενεαλογικά δέντρα μέχρι την αρχική, ο δείκτης GCI είναι ίδιος για όλους τους απογόνους και μπορεί να χρησιμοποιηθεί η κατά Wray and Thompson, (1990) εξίσωση (2.1.2.2).

Λόγω τού ότι, οι παράγοντες που εξισορροπούν την ενδογαμία, όπως

1. Η φυσική επιλογή
2. Ο ανασυνδυασμός χρωμοσωμάτων κατά τη μείωση
3. Η ανισορροπία σύνδεσης, και
4. Οι μεταλλάξεις

λειτουργούν στην χρονική κλίμακα μίας γενιάς (Wray and Thompson, 1990), ο ρυθμός ενδογαμίας μίας γενιάς, είναι ακριβέστερο μέτρο σε σχέση με τον ρυθμό ενδογαμίας ενός ημερολογιακού έτους. Στο κείμενο των κατευθυντήριων γραμμών, οριοθετείται ένα ελάχιστο 50 γεννητόρων το οποίο απαιτείται ώστε ο ρυθμός ενδογαμίας σε μία γενιά να είναι 1%. Το διάστημα μεταξύ 2 γενιών θεωρείται ως η χρονική μονάδα μέτρησης για ένα πληθυσμό και έχει να κάνει με τον χρόνο που απαιτείται για την αναπλήρωση των γεννητόρων του. Λόγω του γεγονότος ότι οι αρσενικοί και θηλυκοί γεννήτορες συνεισφέρουν το μισό γενετικό τους υλικό στον πληθυσμό (σε διπλοειδείς οργανισμούς), ο υπολογισμός του διαστήματος μεταξύ μίας γενιάς και της προηγούμενης ισούται με τον μέσο όρο ηλικίας των αρσενικών και θηλυκών γεννητόρων σε συνάρτηση με την συνεισφορά τους σε απογόνους στο πληθυσμό. Αν έχουμε για παράδειγμα, την γέννα απογόνων όταν οι αρσενικοί γεννήτορες είναι ηλικίας 1,5 έτους, και οι θηλυκοί γεννήτορες ήταν οι μισοί 2 ετών και οι άλλοι μισοί 1,7 ετών έχουμε τον χρόνο αναπλήρωσης των αρσενικών ως $Lm=1,5$ και των θηλυκών $Lf=0,5*2+0,5*1,7=1,85$, το διάστημα μεταξύ των γενιών είναι $L=2*(1,5+1,85)=6,7$. Με βάση το διάστημα μεταξύ των 2 γενιών, ο προτεινόμενος μέγιστος ρυθμός ενδογαμίας για μία γενιά υπολογίζεται ως $\Delta F \leq L$ άρα ο προτεινόμενος μέγιστος ΔF στο συγκεκριμένο παράδειγμα είναι 6.7%.

Για τα δύο παραδείγματα από τα διαφορετικά σετ δεδομένων που συγκρίναμε, πρέπει να σημειωθεί επίσης, ότι η ανάθεση φύλου στους γεννήτορες των μικροδορυφορικών δεδομένων, έγινε από εμάς, ώστε να μπορούμε να προσομοιώσουμε τις διασταυρώσεις και να

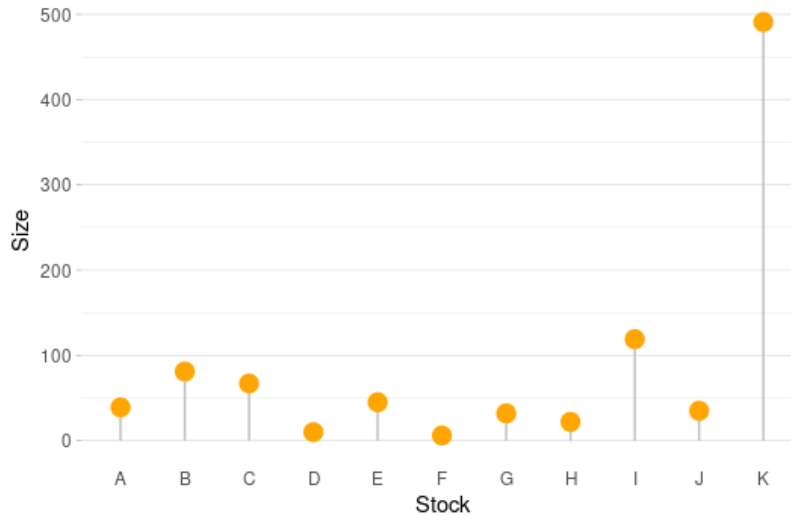
εξετάσουμε την ενδογαμία στην επόμενη γενιά, τηρώντας την αναλογία 1 αρσενικός γεννήτορας προς 1 θηλυκό γεννήτορα. Λόγω του ότι δεν μπορούμε να γνωρίζουμε το φύλο των υποψήφιων γεννητόρων στο δεύτερο σετ δεδομένων, υιοθετήσαμε μία βέλτιστη αναλογία, λόγω του ότι το φύλο δεν ήταν διαθέσιμο ως κριτήριο επιλογής. Με βάση αυτή την επισήμανση, αν η αναλογία είναι διαφορετική, μπορεί η ενδογαμία στην επόμενη γενιά να είναι μικρότερη ή μεγαλύτερη και αυτό θα εξαρτηθεί κυρίως από την ενδοοικογενειακή αναλογία των φύλων των υποψήφιων γεννητόρων, και τον αριθμό ομοθαλών αδελφών στους υποψήφιους γεννήτορες. Συγκεκριμένα βλέπουμε ότι στους υποψήφιους γεννήτορες, 61 άτομα ανήκουν σε 2 οικογένειες. 39 στην οικογένεια A12/Θ14 και 22 στην οικογένεια A16/Θ14. Σε μεγάλο βαθμό η διαμορφούμενη ενδογαμία στην επόμενη γενιά θα καθορισθεί από την αναλογία των φύλων στις 2 αυτές οικογένειες και τον αριθμό των πιθανών διασταυρώσεων μεταξύ ομοθαλών αδελφών αυτών των οικογενειών, η οποία όμως, ως μέσος όρος, δεν θα είναι ποτέ μικρότερη από την ενδογαμία του πρώτου σετ δεδομένων.

2.2.4 Γενετική περιγραφή, πληθυσμιακή δομή και ανάλυση συγγένειας σε 11 αποθέματα κρανιού, ως μία πληροφοριακή αφετηρία νέων προγραμμάτων αναπαραγωγής

Οι προηγούμενες δύο περιπτώσεις αφορούσαν την επιλογή γεννητόρων από απογόνους μίας διασταύρωσης. Αυτές οι δυο περιπτώσεις δεν επαρκούν για να περιγράψουν μίας μεγαλύτερης κλίμακας επιλογή η οποία περιπλέκεται από την ύπαρξη διαφορετικών αποθεμάτων που συνυπάρχουν και διαμορφώνουν την παραγωγική βάση μίας ή πολλών μονάδων ιχθυοκαλλιέργειών. Στην τρίτη περίπτωση αναλύεται και εξετάζεται η διαδικασία της επιλογής γεννητόρων από 11 αποθέματα, 8 εκ των οποίων αποτελούνται από γεννήτορες με τα υπόλοιπα 3 να είναι αποθέματα απογόνων τους.

2.2.6 Δείγματα

Από τον Σεπτέμβριο του 2017 έως τον Ιούνιο του 2020, 946 δείγματα από 11 αποθέματα κρανιού, 8 εκ των οποίων (A-H) από γεννήτορες και 3 (I-K) από απογόνους τους, έφτασαν στις εγκαταστάσεις του ΕΛΚΕΘΕ στην Κρήτη. Τα δείγματα προέρχονταν από μονάδες παραγωγής 3 εταιρειών που δραστηριοποιούνται στην περιοχή της Μεσογείου. Ο αριθμός των δειγμάτων στα 11 αποθέματα φαίνεται στο **Σχήμα 4**.



Σχήμα 4. Κατανομή μεγεθών των 11 αποθεμάτων κρανιού

2.2.7 Εξαγωγή DNA, PCR, γονοτύπηση μικροδορυφορικών τόπων

Έγινε εξαγωγή DNA από τα 946 δείγματα χρησιμοποιώντας ένα ευρέως διαδεδομένο πρωτόκολλο (Miller et al., 1988) και η ποιότητα και η ποσότητα του DNA αξιολογήθηκαν χρησιμοποιώντας το φασματοφωτόμετρο NanoDrop ND 1000 (Thermo Fisher Scientific, www.thermofisher.com). Χρησιμοποιήθηκε ένας συνδυασμός 12 μικροδορυφορικών τόπων (multiplex) (Πίνακας 8) που αναπτύχθηκαν για το *Sciaenops ocellatus*, ένα είδος φυλογενετικά συγγενικό με τον κρανιό. Οι συνθήκες για την PCR στον κρανιό βελτιστοποιήθηκαν (από τους Turner et al. 1998 και O'Malley et al. 2003). Οι αντιδράσεις έγιναν χρησιμοποιώντας το κιτ πολλαπλών PCR της Qiagen. Οι αντιδράσεις PCR πραγματοποιήθηκαν σε ένα μείγμα αντίδρασης 10 μl με συγκέντρωση 10 μmol/l για κάθε primer και 5 ng/μl DNA προτύπου. Το θερμικό προφίλ περιλάμβανε ένα στάδιο προ-αποδιάταξης στους 95°C για 15 λεπτά ακολουθούμενο από 30 κύκλους αποδιάταξης-ανόπτησης-επέκτασης στους 94°C για 30 δεύτερα, 57°C για 1,3 λεπτά, 72°C για 1 λεπτό και ένα τελικό βήμα επιμήκυνσης στους 60°C για 30 λεπτά. Τα ενισχυμένα τμήματα (amplicons) διαχωρίστηκαν με τριχοειδή ηλεκτροφόρηση σε ABI 3730 αλληλουχητή (Applied Biosystems, Foster City, CA) χρησιμοποιώντας τον τυποποιημένο δείκτη μεγέθους LIZ500. Το λογισμικό ανάλυσης μεγέθους θραυσμάτων STRand (<http://www.vgl.ucdavis.edu/informatics/STRand>) χρησιμοποιήθηκε για την γονοτύπηση των μικροδορυφορικών τόπων.

Πίνακας 8. Σύνοψη για τους 12 μικροδορυφορικούς τόπους που χρησιμοποιήθηκαν στον κρανίο. Στην πρώτη στήλη δίνεται το όνομα του κάθε τόπου ενώ στην δεύτερη στήλη ο τύπος της φθορίζουσας χρώσης. Στην τρίτη και τέταρτη στήλη δίνονται οι εμπρόσθιες και οι οπίσθιες ανάστροφες αλληλουχίες των εκκινητών.

Μικροδορυφορικοί τόποι	Χρώση	Εμπρόσθια (F) αλληλουχία εκκινητή (primer) 5'-3'	Ανάστροφη οπίσθια αλληλουχία (R) εκκινητή (primer) 5'-3'
Soc-11	FAM	GCCGAGTCACGAAGGAACAGAGAA	TGTCGTCTCATCTATCTCCATCTC
Soc-35	PET	TGTCCCATCAATCAAGCAGACTCT	CTCTACCTCACACTCCTCAAAGTT
Soc405	PET	AGCCTTTTGTITAGTTTCCCTCAT	GGGGTGTAGCAGAACCACAC
Soc-428	NET	GACATCGCATTTGTCTACAGAGTCG	AACTCCCAGTCATAATATCCCTTT
Soc431	VIC	GTGGTAGATGAAAACGTATAAAAAGGAG	GTTTCATATATATAGTGTACAGCTCCAGCTTC
Soc-44	VIC	GAGGGTGACGCTAACAGTTGA	CACAGCTCCACTCTGATATG
UBA005	NED	CATCAGGATTGGCAACTAGC	GTTTCCTCCAGGTTTATTCTTCATTGAC
UBA006	PET	AGCACACGTAATCACACACAGAT	GTTTCCACTAGTGCAAACGGTGGT
UBA042	NED	TTTCTGCCTGACTAGATGTTCTTTC	GATTGTTGCTGGTTTTTCCAAT
UBA050	FAM	GCACAACTGCATCCCTTAGAT	GTTTAGAAGTGAAGACTGCGGACTG
UBA054	FAM	CCTTGTGAGAACATTAATTTGGATG	GTTTCCAAACCCTGATAGATGGATAGTT
Cacmic14	FAM	TGTCCTCACTCCTCTTTTTCTTTC	GTTTAAGGCGCATCTCCAGTCTC

2.2.8 Πληθυσμιακή ανάλυση

2.2.8.1 Γενετική παραλλακτικότητα και ισορροπία Hardy-Weinberg (HWE)

Ο μέσος αριθμός αλληλομόρφων ανά τόπο (N_A), η παρατηρούμενη (H_{obs}) (Weir & Cockerharm 1984), η αναμενόμενη (H_{exp}) ετεροζυγωτία (Weir 1996 p.124) καθώς και ο δείκτης ποικιλομορφίας Shanon (Shannon, CE & W. Weaver, 1949) υπολογίστηκαν για κάθε στόκ και μικροδορυφορικό τόπο χρησιμοποιώντας τα λογισμικά Genetix 4.05.2 (Belkhir et al. 2004) και Genalex v. 6.5 (Peakall R, Smouse PE 2012). Η απόκλιση των συχνοτήτων των αλληλομόρφων από την ισορροπία Hardy-Weinberg (HWE) εκτιμήθηκε από τον συντελεστή F_{IS} του Wright (1969), σύμφωνα με τις εξισώσεις των Weir & Cockerham (1984). Η συμφωνία του F_{IS} με την αρχική υπόθεση της ύπαρξης ισορροπίας HW (η στατιστική σημαντικότητα του $F_{IS} > 0$, < 0 και $= 0$) εκτιμήθηκε μετά από 10.000 τυχαίες μεταθέσεις αλληλομόρφων και χρησιμοποιώντας την διαδικασία Bonferroni (Rice 1989) για την διόρθωση πολλαπλών δοκιμών και αποφυγή σφαλμάτων τύπου-1 (Rice 1989). Η παρουσία μηδενικών αλληλομόρφων ή άλλων σφαλμάτων προερχόμενων από την διαδικασία της γονοτύπισης εκτιμήθηκε για όλους τους τόπους χρησιμοποιώντας το λογισμικό Micro-checker v.2.2.3 (van Oosterhout et al. 2004).

2.2.8.2 Πληθυσμιακή δομή

Το λογισμικό Structure 2.3.2 (Falush et al., 2003) χρησιμοποιήθηκε για να συναγάγει την πιο πιθανή δομή του πληθυσμού βάσει των δεδομένων από τους μικροδορυφορικούς τόπους των 11 αποθεμάτων. Ο υπολογισμός έγινε με ένα μοντέλο μη-σύμμιξης (non-admixture), που προϋποθέτει ότι κάθε δείγμα/άτομο προέρχεται από έναν εκ των K πληθυσμών, χρησιμοποιώντας μια περίοδο burn-in για 250.000 δείγματα και 10^6 επακόλουθων επαναλήψεων MCMC (Markov Chain Monte Carlo) για κάθε τιμή K από 1-11 για 6 επαναλήψεις. Το μοντέλο προϋποθέτει ότι οι μοριακοί δείκτες είναι ασύνδετοι (σε ισορροπία γενετικής σύνδεσης) και σε ισορροπία HW. Στη συνέχεια, η μέθοδος ΔK (Evanno, Regnaut and Goudet, 2005) χρησιμοποιείται για την τελική επιλογή του αριθμού K υπολογίζοντας τη διακύμανση της πιθανοφάνειας (variation of rate of likelihood) μεταξύ K και $K+1$. Αυτές οι τιμές αντιστοιχούν στα τοπικά μέγιστα της συνάρτησης ΔK της καμπύλης του K . Η καλύτερη τιμή K υπολογίζεται από το διαδικτυακό πρόγραμμα STRUCTURE HARVESTER (Earl and vonHoldt, 2011). Η μέθοδος DAPC, χρησιμοποιήθηκε (discriminant analysis of principal components) ως μία δεύτερη αναλυτική μέθοδος διερεύνησης της πληθυσμιακής δομής των αποθεμάτων με το πακέτο adegenet (Jombart 2008) σε περιβάλλον R. Στην ανάλυση, πρώτα γίνεται μία βασική ανάλυση PCA (ανάλυση κύριων συνιστωσών) και έπειτα η DA (ανάλυση διακριτοποιητικών συστατικών). Στο κομμάτι της PCA έχουμε τον μετασχηματισμό των δεδομένων και την DA η οποία αναγνωρίζει τον αριθμό, των ομάδων/πληθυσμών, που εξηγεί με τον βέλτιστο τρόπο τα δεδομένα. Αυτή η διαδικασία, υλοποιεί έναν γραμμικό μετασχηματισμό των δεδομένων ψάχνοντας την μεγαλύτερη αλληλική διακύμανση μεταξύ των ομάδων και την μικρότερη αλληλική διακύμανση στο εσωτερικό μίας ομάδας/ενός πληθυσμού. Χρησιμοποιήσαμε την ανάλυση DAPC ως μία επιπλέον προσέγγιση, και ως μία επαλήθευση της ανάλυσης από το Structure, συγκρίνοντας τα αποτελέσματα των 2 μεθόδων.

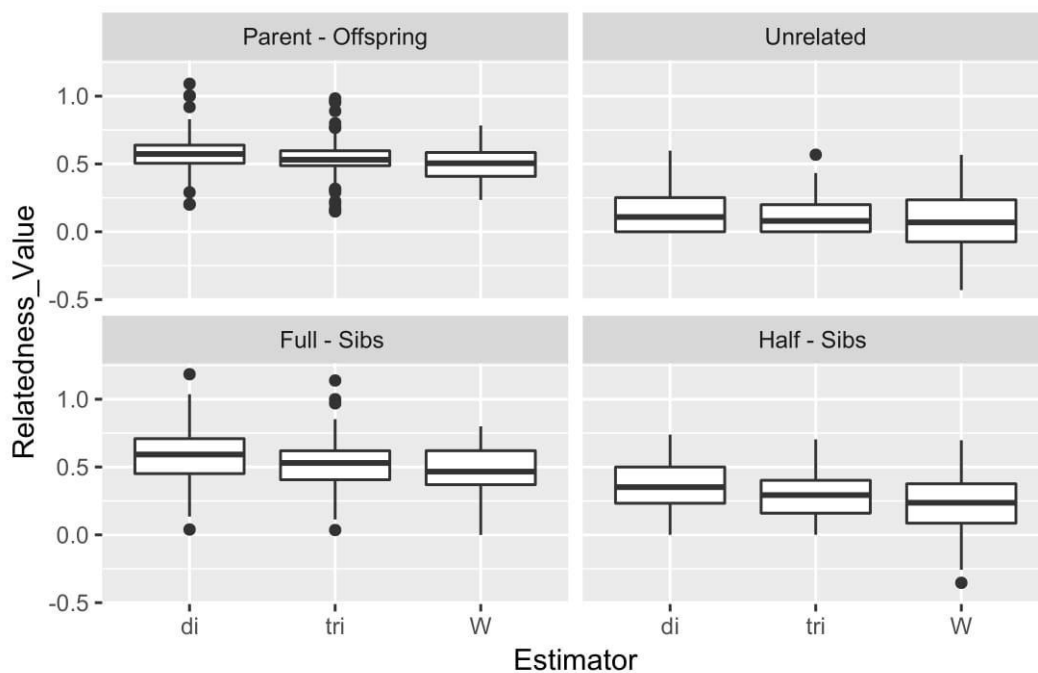
2.2.9 Ανάλυση σχέσεων συγγένειας

Το λογισμικό Coancestry (V1.0.1.9) εφαρμόζει 7 μεθόδους για την εκτίμηση της σχέσης μεταξύ ατόμων και 4 μεθόδους για την εκτίμηση ατομικών συντελεστών ενδογαμίας, χρησιμοποιώντας τον γονότυπο κάθε ατόμου για το σύνολο των γενετικών τόπων (Wang, J. 2011). Αρχικά έγινε μία ανάλυση χρησιμοποιώντας και τις 7 μεθόδους, και τα στοιχεία της στατιστικής της σύνοψης παρείχαν τις απαραίτητες πληροφορίες για τα επόμενα βήματα. Σύμφωνα με την στατιστική σύνοψη που εμφανίζεται στον **Πίνακα 9** επιλέχθηκαν οι 2 μέθοδοι που φαίνεται να εξηγούν με τον καλύτερο τρόπο τα δεδομένα. Συγκεκριμένα επιλέχθηκαν οι 2 μέθοδοι που εμφάνισαν τη μικρότερη διακύμανση και την καλύτερη γραμμική συσχέτιση μεταξύ τους. Αυτές είναι οι DyadML και TrioML των οποίων η γραμμική συσχέτιση είναι ίση με 0,976 και η διακύμανση 0,034 και 0,038 αντίστοιχα. Επιπλέον, πραγματοποιήθηκε μια προσομοίωση προς επαλήθευση αυτής της υπόθεσης, χρησιμοποιώντας το πακέτο “related” σε περιβάλλον R (Pew J, Muir P, Wang J, Frasier TR (in prep) related¹). Η προσομοίωση βασίστηκε σε όλους τους αρχικούς γονότυπους, από τους οποίους προσομοιώθηκαν, 400 ζεύγη ψαριών για καθέναν από τους 4 τύπους σχέσεων της επιλογής μας, PO, UR, FS, HS, γονέας-απόγονος, μη-συγγένεια, ομοθαλή αδέρφια και ετεροθαλή αδέρφια αντίστοιχα. Για το σκοπό αυτό επιλέχθηκαν οι δυο καλύτερες μέθοδοι μέγιστης πιθανοφάνειας (DyadML και TrioML) σύμφωνα με τη σύνοψη της προκαταρκτικής ανάλυσης, και μία ακόμη μέθοδος, η Wang, η οποία δεν λειτουργεί με βάση την μέγιστη πιθανοφάνεια, η μέθοδος Wang επιλέχθηκε για την σύγκριση με μία διαφορετική προσέγγιση. Τα αποτελέσματα απεικονίζονται στο **Σχήμα 5** όπου συγκρίνεται η απόδοση για κάθε τύπο σχέσης των 1600 προσομοιωμένων ζευγαριών για τις 4 κατηγορίες σχέσεων. Επιπλέον, πραγματοποιήθηκε μια εκτίμηση συσχέτισης των αποτελεσμάτων των μεθόδων DyadML, TrioML και Wang από τα αποτελέσματα προσομοίωσης με τις αναμενόμενες τιμές από τα πραγματικά αποτελέσματα. Η DyadML έδειξε την υψηλότερη συσχέτιση, με τιμή 0,7332469 ακολουθούμενη από την TrioML με 0,7281499 και την Wang με 0,6998539. Η συσχέτιση, των προσομοιωμένων τιμών κάθε μεθόδου με τα πραγματικά αποτελέσματα (των ίδιων μεθόδων) διερευνά την ενδεχόμενη λανθασμένη υποκειμενική επιλογή που μπορεί να προκύψει από το επιλεγμένο βάθος προσομοίωσης. Αυτό έγινε χρησιμοποιώντας τη συνάρτηση cor ως έχει: cor (relvals, simrel [, “στήλη των αποτελεσμάτων της μεθόδου”]). Δεδομένου ότι η απόδοση των μεθόδων είναι συγκρίσιμη με τα προσομοιωμένα και πραγματικά δεδομένα, ο επιλεγμένος αριθμός προσομοιωμένων ζευγών θεωρήθηκε αντιπροσωπευτικός του συνόλου των δεδομένων. Οι DyadML και TrioML εμφανίζουν τη μικρότερη διακύμανση σε όλους τους τύπους σχέσεων και είναι οι μέθοδοι που ταιριάζουν καλύτερα στα δεδομένα, ενώ η μέθοδος Wang δείχνει μεγαλύτερη διακύμανση με στατιστικά χαμηλότερο βαθμό εμπιστοσύνης των εκτιμήσεων, **Σχήμα 5**.

related: ένα πακέτο R για την ανάλυση δεδομένων συζευγμένης συγγένειας βασισμένη σε συνεπικρατείς μοριακούς δείκτες

Πίνακας 9. Στατιστική σύνοψη της απόδοσης των μεθόδων του λογισμικού Coancestry στην διερευνητική ανάλυση των 11 αποθεμάτων για επιλογή μεθόδου. Δίνεται ο συντελεστής γραμμικής συσχέτισης μεταξύ όλων των μεθόδων καθώς και η διακύμανση τους.

Συντελεστής συσχέτισης	TrioML	Wang	LynchLi	LynchRd	Ritland	QuellerGt	DyadML
TrioML	1,00000						
Wang	0,75285	1,00000					
LynchLi	0,76592	0,97722	1,00000				
LynchRd	0,87778	0,78018	0,78216	1,00000			
Ritland	0,63836	0,56091	0,58056	0,74478	1,00000		
QuellerGt	0,83700	0,87577	0,90519	0,88152	0,70389	1,00000	
DyadML	0,97694	0,77075	0,78289	0,90031	0,65864	0,85956	1,00000
Mean	0,13352	-0,01644	-0,01833	-0,00107	-0,00114	0,00271	0,14685
Variance	0,03442	0,11605	0,12933	0,05389	0,03443	0,08412	0,03852



Σχήμα 5. Σύνοψη αποτελεσμάτων των 3 μεθόδων για τους 4 τύπους προσομοιωμένων σχέσεων. di : DyadML, tri : TrioML, W : Wang.

2.3 Αποτελέσματα & συζήτηση

Οι τιμές της ετεροζυγωτίας σε κάθε απόθεμα για κάθε μικροδορυφορικό τόπο, παρουσιάζονται στον **Πίνακα 10**. Η αναμενόμενη ετεροζυγωτία H_{exp} (Weir 1996 p.124), εξίσωση (8), σε έναν τόπο είναι το ποσοστό των αλληλομόρφων που αναμένεται να είναι ετερόζυγα και παίρνει τιμές από 0-1. Εάν p_{lu} είναι η συχνότητα ενός u αλληλομόρφου σε έναν γενετικό τόπο l η αναμενόμενη ετεροζυγωτία ή gene diversity, δίνεται από την εξίσωση (2.3.1).

$$H_l = 1 - \sum_u p_{lu}^2 \quad (2.3.1)$$

Ενώ με την υπέρβαση πολλών τόπων m έχουμε την εξίσωση (2.3.1) να γίνεται

$$H = 1 - 1/m \left(\sum_l \sum_u p_{lu}^2 \right) \quad (2.3.2)$$

Η εξίσωση (8), υπολογίζεται από το άθροισμα των συχνοτήτων των αλληλομόρφων υψωμένων στο τετράγωνο, για ένα σύνολο γενετικών m τόπων, αφαιρούμενων από το άθροισμα όλων των επιμέρους αναμενόμενων ετεροζυγωτιών με την μονάδα διαιρώντας με τον αριθμό των τόπων. Η παρατηρούμενη ετεροζυγωτία H_{obs} υπολογίζεται από το άθροισμα των τιμών των συχνοτήτων των αλληλομόρφων σε έναν τόπο (2.3.3) και αντίστοιχα η συνολική H_{obs} είναι το άθροισμα για όλους τους τόπους διαιρεμένο με τον αριθμό των τόπων (2.3.4). Εάν το n_{luv} είναι ο παρατηρούμενος αριθμός των $A_u A_v$ ετεροζυγωτιών, με $u \neq v$, σε έναν γενετικό τόπο l , και πληθυσμό n , η παρατηρούμενη ετεροζυγωτία για αυτό τον τόπο είναι $H_{l,obs}$

$$H_{l,obs} = \sum_u \sum_{u \neq v} n_{luv} / n \quad (2.3.3)$$

ενώ για ένα σύνολο τόπων m η (9) γίνεται

$$H_{m,obs} = 1/m \sum_{l=1}^m H_{l,obs} \quad (2.3.4)$$

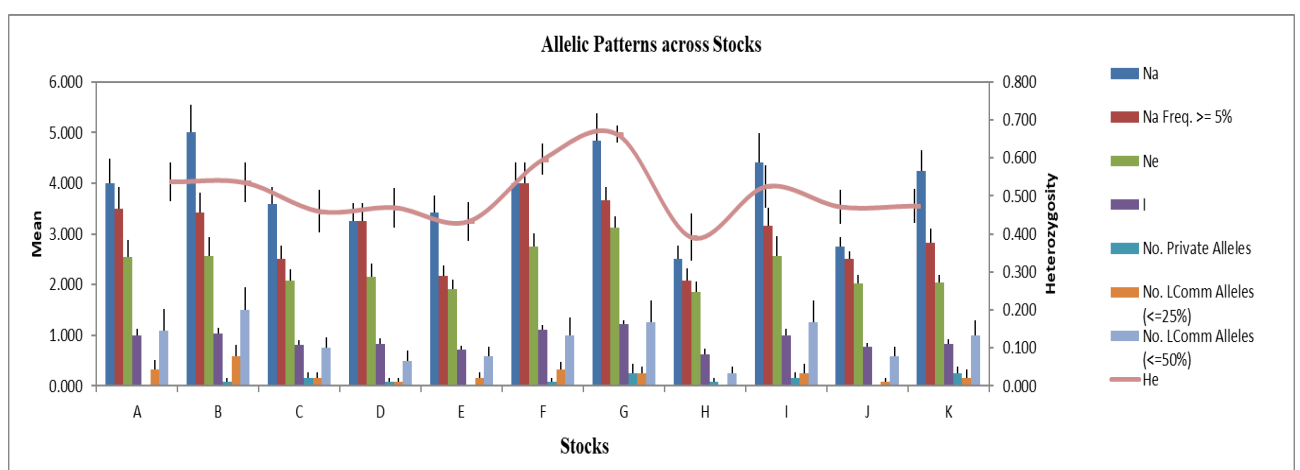
Οι τιμές είναι γενικά χαμηλές κυμαινόμενες από 0.4053 στο απόθεμα H, με την μέγιστη ετεροζυγωτία 0.8359 στο απόθεμα G.

Πίνακας 10. Παρουσίαση των τιμών ετεροζυγωτίας των μικροδορυφορικών αλληλομόρφων ανά τόπο και ανά απόθεμα (πληθυσμό). Στην πρώτη στήλη παρατίθενται οι μικροδορυφορικοί τόποι και στις στήλες 2 έως 11 οι αντίστοιχες τιμές για κάθε απόθεμα. Στις τελευταίες 2 σειρές δίνονται οι συνολικές τιμές της αναμενόμενης ετεροζυγωτίας (H_{exp}) και της παρατηρούμενης (H_{obs}) για όλους τους τόπους ανά απόθεμα.

Ετεροζυγωτία των αλληλομόρφων ανά τόπο και απόθεμα											
	Stock A	Stock B	Stock C	Stock D	Stock E	Stock F	Stock G	Stock H	Stock I	Stock J	Stock K
Casemic14	0,5302	0,5605	0,6533	0,585	0,6632	0,6389	0,6968	0,6994	0,7449	0,5351	0,5647
Soc-11	0,7436	0,7438	0,6691	0,685	0,559	0,7361	0,6245	0,657	0,784	0,6657	0,6857

Soc-35	0,3346	0,487	0,449	0,485	0,4886	0,6111	0,6841	0,0444	0,2363	0,602	0,533
Soc-405	0,473	0,5118	0,3247	0,34	0,022	0,4861	0,5488	0,3512	0,5913	0,571	0,4699
Soc-42	0,7988	0,8156	0,6883	0,78	0,4516	0,75	0,7983	0,5155	0,8144	0,6208	0,5858
Soc-431	0,6496	0,5871	0,1273	0,185	0,2526	0,7083	0,6587	0,5196	0,5203	0,2886	0,5424
Soc-44	0,4356	0,5021	0,4003	0,48	0,3968	0,5417	0,7046	0,2779	0,4614	0,4624	0,2359
Uba-005	0,5	0,4972	0,4639	0,18	0,4723	0,4861	0,6606	0,2975	0,4105	0,4412	0,5055
Uba-006	0,6052	0,4439	0,6479	0,585	0,3316	0,6944	0,6865	0	0,4194	0,3367	0,4436
Uba-042	0,429	0,3282	0,5711	0,445	0,4837	0,4861	0,6484	0,3967	0,5725	0,1327	0,1537
Uba-050	0,7449	0,7655	0,1898	0,51	0,664	0,7361	0,7422	0,5	0,5062	0,5029	0,5942
Uba-054	0,2041	0,178	0,3306	0,375	0,4109	0,2778	0,4956	0,4339	0,2366	0,4984	0,376
H_{exp}	0,5374	0,5351	0,4596	0,4696	0,433	0,5961	0,6624	0,3911	0,5248	0,4715	0,4742
H_{obs}	0,5791	0,608	0,5498	0,5167	0,4593	0,5972	0,8359	0,4053	0,5183	0,5429	0,4888

Επιπλέον, στο **Σχήμα 6** μπορούμε να δούμε ότι το μοτίβο της αναμενόμενης ετεροζυγωτίας, που δίνεται από διαφορετικούς δείκτες, δεν είναι το ίδιο μεταξύ των αποθεμάτων. Τα αλληλόμορφα που μοιράζονται στο ήμισυ ή στο ένα τέταρτο των αποθεμάτων και τα μοναδικά αλληλόμορφα κάθε αποθέματος διαφοροποιούνται. Αυτό, μπορεί εν μέρει να οφείλεται σε διαφορές μεγέθους μεταξύ των αποθεμάτων, και συνεπώς μια μεροληψία στη δειγματοληψία αλληλομόρφων, π.χ. στο απόθεμα F υπάρχουν μόνο 6 συμμετέχοντες, και στο απόθεμα K 491, καθώς επίσης και λόγω της διαφορετικής σύνθεσης των αποθεμάτων σε γεννήτορες και απογόνους. Το απόθεμα H δείχνει μειωμένη H_{exp} , όπως επίσης και κατά τη χρήση του δείκτη Weir για τον υπολογισμό της της ετεροζυγωτίας (H_{obs} και H_{exp}) στο GENETIX. Καθώς επίσης το ίδιο ισχύει σχετικά με το απόθεμα G.



Σχήμα 6. Σύγκριση της H_{exp} μεταξύ των αποθεμάτων. Na: αριθμός διαφορετικών αλληλομόρφων, Na (Freq $\geq 5\%$): αριθμός αλληλομόρφων με συχνότητα $\geq 5\%$, Ne: αποτελεσματικός αριθμός αλληλομόρφων = $1/(\sum \pi^2)$, I: δείκτης Shannon = $-\sum(\pi * \ln(\pi))$, No. Private Alleles: αριθμός αλληλομόρφων μοναδικών για έναν πληθυσμό, No. LComm Alleles ($\leq 25\%$): Αριθμός κοινών

αλληλομόρφων σε 25% ή λιγότερους πληθυσμούς, No. LComm Alleles ($\leq 50\%$): Αριθμός κοινών αλληλομόρφων σε 50% ή λιγότερους πληθυσμούς, He: Αναμενόμενη ετεροζυγωτία = $1 - \Sigma p_i^2$

Ο αριθμός των αλληλομόρφων ανά τόπο είναι επίσης χαμηλός. Στον **Πίνακα 11** βλέπουμε τον αριθμό των αλληλομόρφων των μικροδορυφορικών τόπων στα 11 αποθέματα, οι οποίες κατά μέσο όρο έχουν 3,58 αλληλόμορφα και εύρος τιμών από 2,5-5,0.

Οι αποκλίσεις από την ισοροπία HW στο σύνολο των δειγμάτων χαρακτηρίστηκαν από τον δείκτη F_{IS} με την διαδικασία που περιγράφεται στην ενότητα 2.2.8.1. Συγκεκριμένα ο δείκτης F_{IS} συνδέεται με την εξίσωση (2.3.5) των Weir & Cockerham (1984) ως εξής: το B είναι η διακύμανση των συχνοτήτων των αλληλομόρφων μεταξύ των ατόμων σε ένα πληθυσμό και C η διακύμανση των συχνοτήτων των αλληλομόρφων των γαμετών σε κάθε άτομο.

$$F_{IS} = 1 - C / (B + C) \quad (2.3.5)$$

Στον **Πίνακα 12** δίνονται οι τιμές του F_{IS} ανά απόθεμα. Ο δείκτης F_{IS} είναι χρήσιμος για την κατανόηση των εξελικτικών δυνάμεων που δρουν στους πληθυσμούς. Η περίσσεια ή η έλλειψη ετεροζυγωτίας σε έναν πληθυσμό, εμφανίζει διακύμανση η οποία σχετίζεται με το σύστημα ζευγαρώματος, και με το εάν και πόσο η ετεροζυγωτία προσδίδει κάποιο πλεονέκτημα επιλογής (Wright 1969). Όπως προαναφέρθηκε, η ετεροζυγωτία μπορεί να οφείλεται σε τεχνητούς (φαινόμενο Wahlund, null αλληλόμορφα, μη-επαρκής δειγματοληψία των αλληλομόρφων, ανεπαρκής ή αδύναμη αλληλική επικράτεια) ή βιολογικούς λόγους (ενδογαμία, ομογαμία, πλεονέκτημα επιλογής που συνδέεται με την ετεροζυγωτία). Οι αρνητικές τιμές του δείκτη F_{IS} σε έναν πληθυσμό όπου οι διασταυρώσεις γίνονται με τυχαίο τρόπο, και αν το δείγμα είναι αντιπροσωπευτικό, υποδηλώνουν τον αυξημένο βαθμό των διασταυρώσεων μεταξύ ατόμων λιγότερο συγγενικών μεταξύ τους, ενώ οι θετικές τιμές το αντίθετο. Στον **Πίνακα 12** βλέπουμε την απόκλιση από την ισοροπία HW με κόκκινο χρώμα. Τα αποθέματα A-H αποτελούνται από γεννήτορες, επομένως, να παρατηρείται μία αυξημένη ετεροζυγωτία, που ίσως αποδίδεται στο φαινόμενο Wahlund. Συμπληρωματικά, για τα αποθέματα των γεννητόρων, τα D και F είναι τα μικρότερα αριθμητικά, 6 και 10 άτομα αντίστοιχα, και η χρησιμότητά τους στην ανάλυση εντοπίζεται κυρίως στο κομμάτι της συγγένειας και στην συνεισφορά τους στην γενετική ομαδοποίηση των αποθεμάτων. Το απόθεμα K, το οποίο είναι και το μεγαλύτερο, απαρτίζοντας 491 απογόνους, φαίνεται να εμπίπτει στα στατιστικά σημαντικά όρια για την πλειοψηφία των γενετικών τόπων ενώ τα αποθέματα I και J εμφανίζουν παρόμοιες τιμές με κάποιους τόπους να δείχνουν υψηλότερη διακύμανση.

Πίνακας 11. Ο αριθμός των αλληλομόρφων των ανά τόπο και ανά απόθεμα

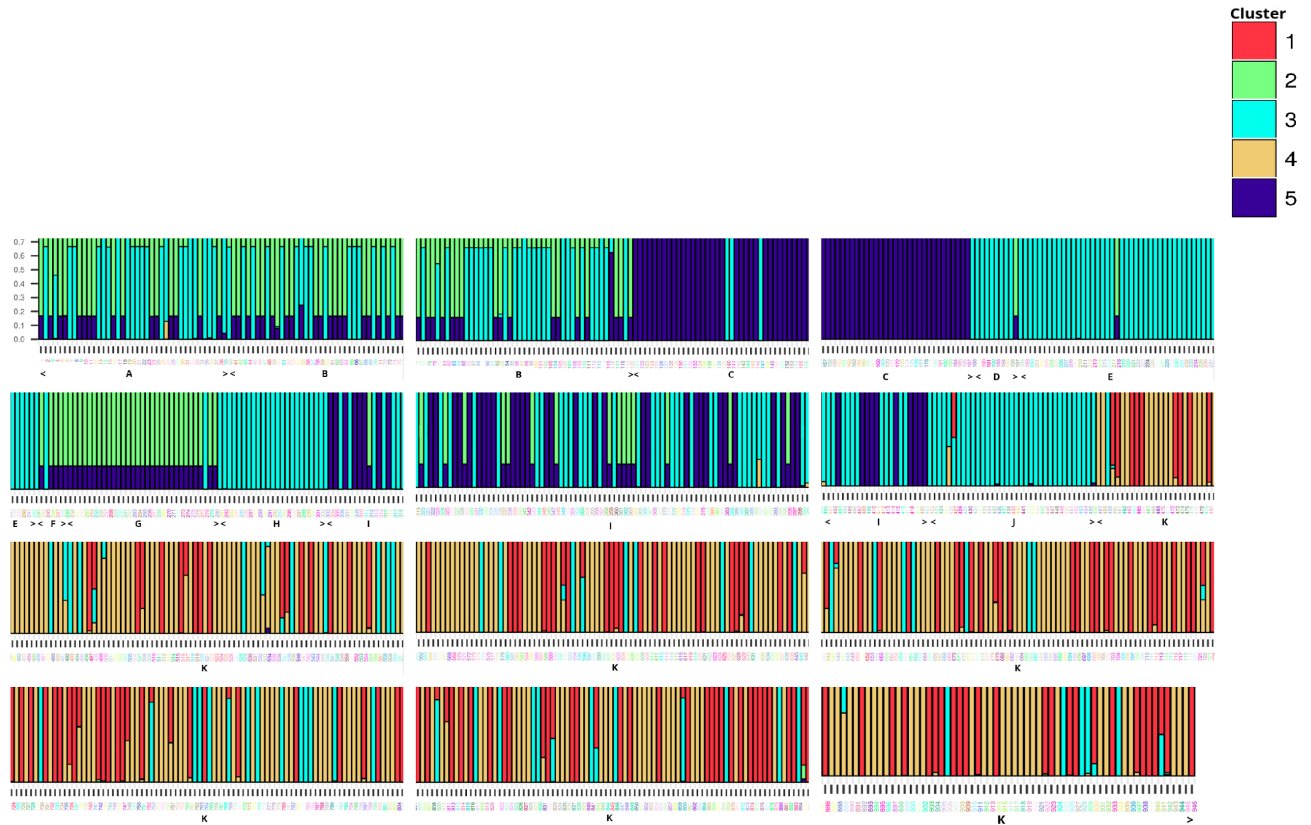
	Αριθμός αλληλομόρφων στα αποθέματα										
	Stock A	Stock B	Stock C	Stock D	Stock E	Stock F	Stock G	Stock H	Stock I	Stock J	Stock K
Cascmic14	4	5	4	3	5	4	5	4	5	3	5
Soc-11	5	7	4	5	6	5	8	4	6	4	7
Soc-35	5	5	4	3	3	5	6	2	4	3	4
Soc-405	3	3	3	3	2	3	3	2	4	3	3

Soc-42	7	9	6	6	3	6	8	3	10	3	6
Soc-431	4	4	3	3	3	4	4	3	3	3	4
Soc-44	2	5	5	2	3	3	4	3	4	3	5
Uba-005	2	3	2	2	3	2	3	2	3	2	3
Uba-006	5	6	4	3	4	6	6	1	4	2	5
Uba-042	3	5	3	3	3	3	4	2	3	2	3
Uba-050	6	6	3	4	4	5	5	2	4	3	4
Uba-054	2	2	2	2	2	2	2	2	3	2	2
Average	4,00	5,00	3,58	3,25	3,42	4,00	4,83	2,50	4,42	2,75	4,25

Πίνακας 12. Η διαμόρφωση των τιμών της παραμέτρου F_{IS} του Wright στα αποθέματα. Στην πρώτη στήλη αναγράφονται οι μικροδορυφορικοί τόποι και από την στήλη 2 έως την στήλη 12 οι αντίστοιχες τιμές για κάθε απόθεμα.

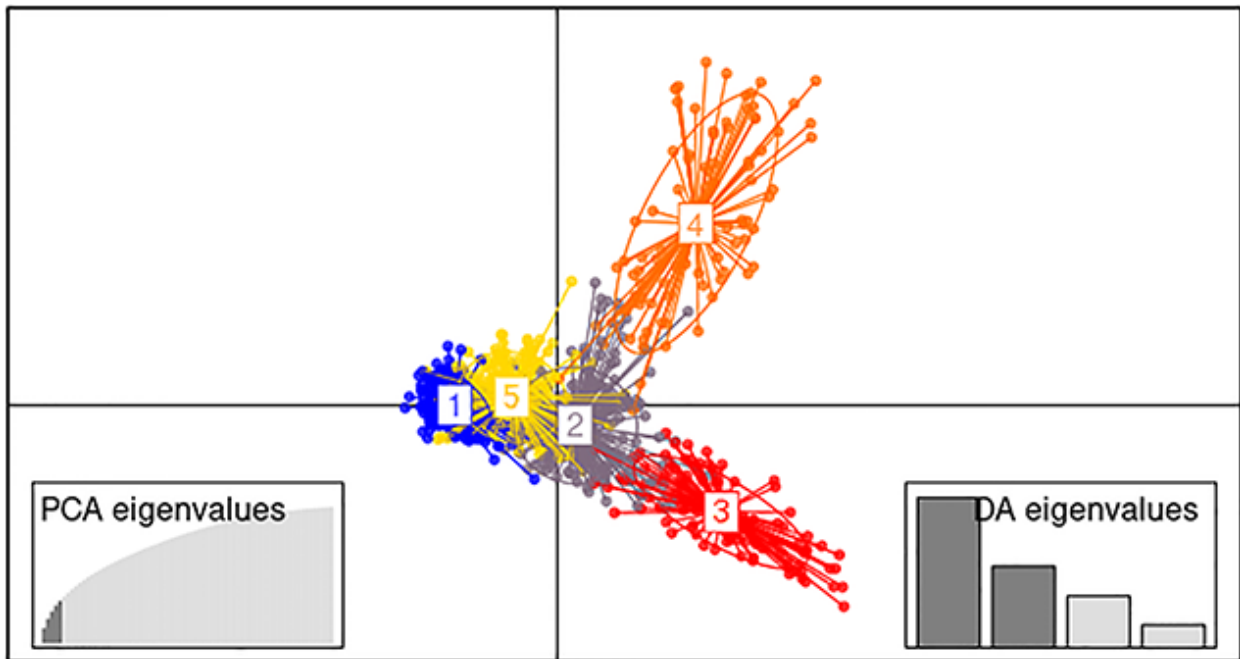
Οι τιμές F_{IS} στα αποθέματα											
	Stock A	Stock B	Stock C	Stock D	Stock E	Stock F	Stock G	Stock H	Stock I	Stock J	Stock K
Cascmic14	-0,051	-0,073	-0,135	-0,5	0,106	0,302	-0,151	-0,213	-0,145	0	-0,228
Soc-11	-0,091	-0,256	-0,175	0,459	0,017	0,184	-0,085	0,192	-0,046	-0,318	0,123
Soc-35	-0,137	-0,135	-0,356	-0,4	-0,08	0,268	-0,403	0	0,151	-0,22	-0,006
Soc-405	-0,017	-0,2	-0,234	0,455	0	0,394	-0,7	0,246	-0,062	-0,288	-0,002
Soc-42	-0,175	-0,144	-0,055	-0,233	-0,121	-0,25	-0,08	-0,035	0,137	-0,09	0,012
Soc-431	0,026	-0,214	-0,048	-0,029	0,043	-0,087	-0,362	-0,202	0,1	-0,174	0,011
Soc-44	-0,106	0,195	-0,111	-0,636	0,282	-0,471	-0,093	-0,122	0,457	-0,222	-0,08
Uba-005	0,141	-0,161	-0,248	1	-0,024	0,706	-0,31	0,106	-0,04	0,429	0,189
Uba-006	-0,131	-0,051	-0,052	-0,145	0,34	0,13	-0,168	-	-0,198	-0,089	-0,08
Uba-042	-0,003	-0,122	-0,458	-0,301	-0,508	0,062	-0,433	-0,355	-0,097	-0,062	-0,061
Uba-050	-0,158	-0,171	-0,093	-0,125	-0,228	-0,042	-0,291	0,294	0,124	-0,236	-0,308
Uba-054	0,134	-0,103	-0,257	0,757	-0,179	-0,111	0,07	-0,024	-0,142	-0,19	0,046
All	-0,064	-0,13	-0,188	-0,047	-0,049	0,089	-0,247	-0,013	0,017	-0,137	-0,029

Όσο αναφορά την γενετική δομή των αποθεμάτων, μέσω της διαδικασίας που αναφέρθηκε για το λογισμικό Structure, ο βέλτιστος αριθμός πληθυσμών K , που εμφανίζει την μεγαλύτερη πιθανότητα είχε την τιμή 5 σύμφωνα με το κριτήριο ΔK του Evanno. Στην απεικόνιση του **Σχήματος 7**, το οποίο είναι το γράφημα που δίνει το Structure με το πέρας της ανάλυσης, έχουμε την ανάθεση των ατόμων των αποθεμάτων στις 5 γενετικές ομάδες ή πληθυσμούς. Κάθε χρωματιστή στήλη αναπαριστά την πιθανότητα συμμετοχής κάθε ατόμου σε μία από τις 5 ομάδες ξεκινώντας από το πρώτο άτομο του αποθέματος (A) μέχρι το τελευταίο άτομο του αποθέματος (K). Η εικόνα είναι υψηλής ανάλυσης και η πληροφορία γίνεται ευδιάκριτη με την κατά βούληση μεγέθυνση του γραφήματος, **Σχήμα 7**. Αν και η ομαδοποίηση δεν είναι εμφανής με μία ματιά, μπορούμε να διακρίνουμε ότι μέχρι και το απόθεμα J τα άτομα ανατίθενται κυρίως στις ομάδες 2,3,5 ενώ το απόθεμα K ανατίθεται σχεδόν αποκλειστικά στις 1,4. Επιπρόσθετα η ομάδα 3 φαίνεται να έχει εμπεριέχει άτομα από όλα τα αποθέματα.



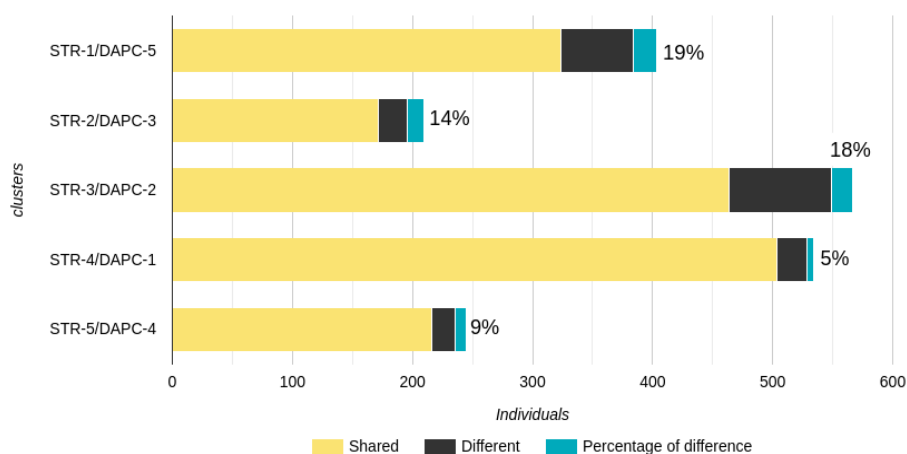
Σχήμα 7. Απεικόνιση της ανάθεσης των ατόμων των 11 αποθεμάτων στις 5 γενετικές ομαδοποιήσεις από το λογισμικό Structure. Στην κάθε στήλη, η οποία αναπαριστά ένα άτομο και την πιθανότητα της συμμετοχής του σε 1 από τους 5 πληθυσμούς, προσθέσαμε την ονομασία των ατόμων μέσω της γλώσσας R.

Η γενετική δομή των αποθεμάτων μέσω της ανάλυσης DAPC φαίνεται στο **Σχήμα 8**. Για την εύρεση του αριθμού K στην ανάλυση DAPC, χρησιμοποιήσαμε την συνάρτηση SNAPCLUST (Beugin et al. 2018), που εφαρμόζει την προσέγγιση ενός αλγορίθμου EM (expectation maximization) και διερευνεί την σύγκλιση των δεδομένων σε έναν συγκεκριμένο αριθμό K. Τα δεδομένα συνέκλιναν για $K = 5$, και η διαδικασία επαναλήφθηκε 3 φορές. Στην ανάλυση DAPC βλέπουμε ότι οι πληθυσμοί 4,3 έχουν γενετικά την μεγαλύτερη απόσταση μεταξύ τους, καθώς επίσης οι 1,2,5 εμφανίζουν μικρότερη γενετική απόσταση μεταξύ τους, **Σχήμα 8**. Τα αποτελέσματα του Structure και της DAPC συγκρίθηκαν ώστε να βρεθούν οι αντιστοιχίες μεταξύ των μεθόδων. Στο **Σχήμα 9**, βλέπουμε τις αντιστοιχίες μεταξύ των 2 μεθόδων. Οι διαφορές μεταξύ των 2 μεθόδων κυμάνθηκαν από 5% έως 19%.



Σχήμα 8. Ανάλυση DAPC (διακριτοποιητικών συστατικών και κύριων συνιστωσών) για την εύρεση της βέλτιστης γενετικής ομαδοποίησης των 11 αποθεμάτων. Η σχετική συνεισφορά της κάθε γενετικής ομάδας στην συνολική εικόνα έχει συνυπολογιστεί στους άξονες με βάση τα αποτελέσματα της PCA και της DA.

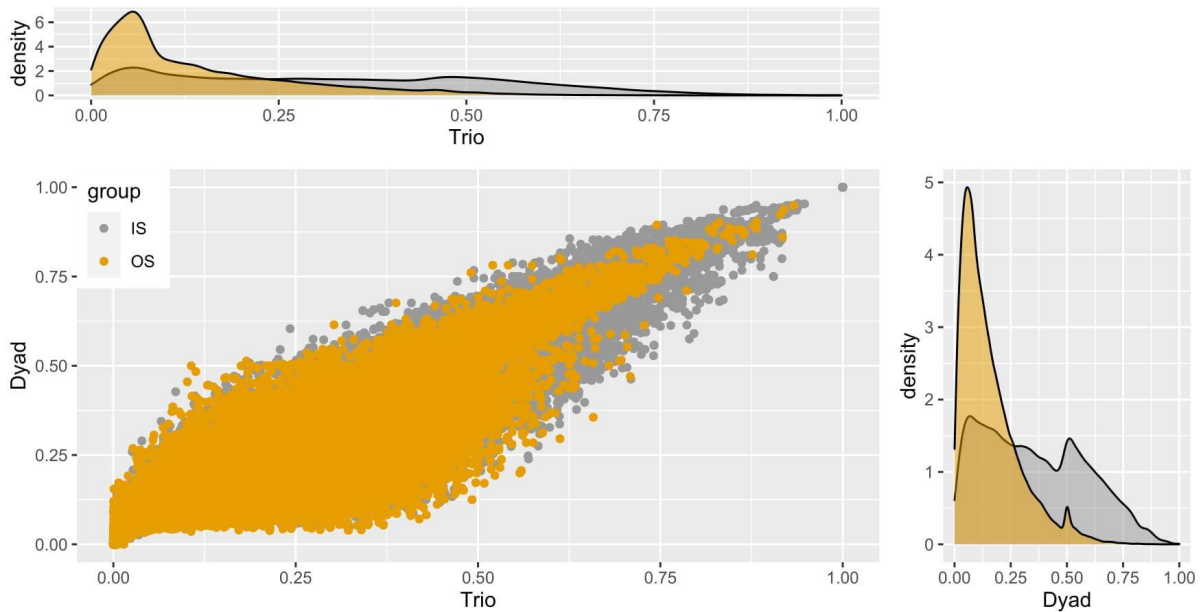
Για ευκολία, θα χρησιμοποιείται η ομαδοποίηση από την DAPC ως σημείο αναφοράς. Οι διαφορές μεταξύ των 2 μεθόδων εντοπίζονται κυρίως σε 2 ομάδες, στην 5 και 2. Αυτό είναι εμφανές από την τοπολογία της DAPC καθώς οι ομάδες 2,5 εμφανίζονται σε κοντινές γενετικές αποστάσεις με κάποιο βαθμό επικάλυψης ενώ οι 3,4 εμφανίζουν μεγαλύτερες γενετικές αποστάσεις συγκριτικά. Επιπλέον, στην ομάδα 5 από τα 60 διαφορετικά ανατεθειμένα άτομα, τα 48 ανήκουν στο απόθεμα Κ.



Σχήμα 9. Σύγκριση και αντιστοίχιση των γενετικών ομάδων από τις μεθόδους Structure και DAPC. Ο άξονας ψ δείχνει την αντιστοίχια των ομάδων και ο άξονας χ τον αριθμό των ατόμων που ανατέθηκαν σε αυτές

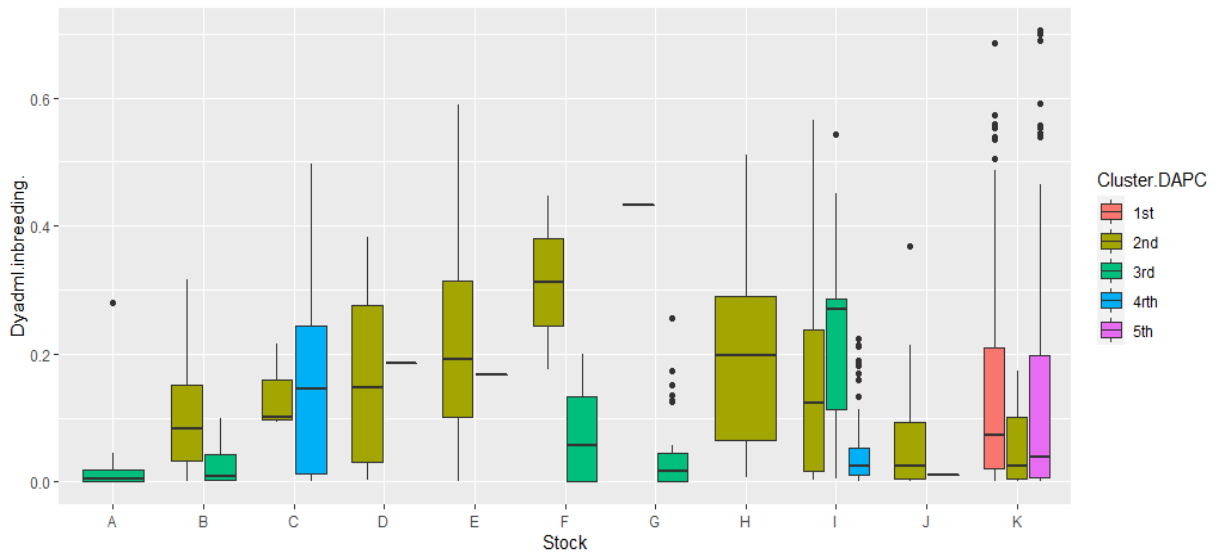
Αν και είναι γνωστό, ότι τα αποθέματα από το Α έως το Η αποτελούνται από γεννήτορες και τα αποθέματα Ι έως Κ από μικτούς πληθυσμούς απογόνων, διερευνήθηκε ο βαθμός της ενδογαμίας και η πιθανή σχέση όλων έναντι όλων. Με αυτόν τον τρόπο, αντικατοπτρίζεται η γενετική ακεραιότητα των αποθεμάτων, αποκλείεται η πιθανή γενετική ροή μεταξύ των αρχικών φυσικών πληθυσμών των γεννητόρων, καθώς και ελέγχονται τα αποτελέσματα από τις αναλύσεις Structure και DAPC. Επιπλέον, η ανάλυση όλων έναντι όλων δίνει την πρόσθετη δυνατότητα αναπροσαρμογής του υπάρχοντος συστήματος αναπαραγωγής. Έτσι, υλοποιείται και ο βασικός στόχος της ανάλυσης των συγγενικών σχέσεων των αποθεμάτων Ι, J, Κ με τα υπόλοιπα για την επιλογή γεννητόρων, έχοντας μία συνολική εικόνα των αποθεμάτων.

Στο **Σχήμα 10**, βλέπουμε τα αποτελέσματα της ανάλυσης συγγένειας με τις 2 επιλεγμένες μεθόδους, ομαδοποιώντας τις σχέσεις μέσα σε κάθε απόθεμα καθώς και μεταξύ των αποθεμάτων, IS και OS αντίστοιχα, με τις εκτιμήσεις σχέσεων μεταξύ των διαφορετικών αποθεμάτων (OS) να βρίσκονται στο επίκεντρο. Η συμφωνία των 2 μεθόδων, φαίνεται σε μεγάλο βαθμό στο κεντρικό γράφημα διασποράς όπου απεικονίζεται η γραμμική τους συσχέτιση. Βλέπουμε ότι η πλειοψηφία των εκτιμώμενων τιμών σχετικά με την εκτίμηση συγγένειας, κυμαίνεται μεταξύ 0,1 και 0,4 με την τιμή 0,25 να είναι έντονα παρούσα. Κάθε τιμή αντιπροσωπεύει μια πιθανή σχέση των 447,458 ζευγών που εξετάστηκαν ($946 * 946/2$). Η τιμή 0,125 αφορά τα πρώτα ξαδέφια, 0,25 για τα ετεροθαλή αδέλφια και για τις τιμές $\geq 0,5$ για τα ομοθαλή αδέλφια και γονείς-απογόνους. Η υψηλή συσχέτιση μεταξύ των δύο μεθόδων είναι αναμενόμενη, όπως αναφέρθηκε προηγουμένως. Βλέπουμε ότι οι τιμές με κίτρινο χρώμα (μεταξύ διαφορετικών αποθεμάτων) είναι κυρίως 0,1 που υποδηλώνει ότι δεν υπάρχει ισχυρή συγγενική σχέση μεταξύ των αποθεμάτων και, αντίθετα, οι τιμές με γκριζο χρώμα για το εσωτερικό των αποθεμάτων κυμαίνονται γύρω από το 0,5 όπως αναμενόταν που υποδηλώνει σχέσεων γονιών-απογόνων και/η ομοθαλών αδελφών. Αυτή η παρατήρηση χρησιμεύει ως μία επιπρόσθετη επαλήθευση της γενετικής απόστασης μεταξύ των αποθεμάτων και της πιθανής βιωσιμότητας τους ως βάση ενός νέου προγράμματος αναπαραγωγής.



Σχήμα 10. Κατανομή των συγγενικών σχέσεων μεταξύ των αποθεμάτων με τις μεθόδους DyadML και TrioML. Τα ζεύγη σχέσεων που απεικονίζονται είναι 447,458. Οι σχέσεις που απεικονίζονται με κίτρινο χρώμα (OS), αφορούν τις σχέσεις μεταξύ διαφορετικών αποθεμάτων, ενώ αυτές με το γκριζο (IS) χρώμα αφορούν τις σχέσεις στο εσωτερικό των αποθεμάτων. Το εύρος των τιμών αντιστοιχεί : 0,125 σε πρώτα ξαδέλφια, 0,25 σε ετεροθαλή αδέρφια, και από 0,5+ σε ομοθαλή αδέρφια/γονιός-απόγονος. Οι μηδενικές τιμές έχουν αφαιρεθεί για την βέλτιστη απεικόνιση των γραφημάτων

Δεδομένου ότι το Coancestry παρέχει εκτιμήσεις για τις ατομικές τιμές ενδογαμίας με βάση τις εκτιμήσεις των συγγενικών σχέσεων, κρίθηκε σκόπιμο να συσχετιστούν οι τιμές ενδογαμίας από την μέθοδο DyadML με τις 5 γενετικές ομαδοποιήσεις και με τα αποθέματα. Στο **Σχήμα 11**, παρέχονται οι εκτιμήσεις της DyadML για την ενδογαμία ανά άτομο για τα 11 αποθέματα. Μπορούμε να δούμε ότι τα αποθέματα με την μικρότερη ενδογαμία έχουν να κάνουν με την γενετική ομάδα 3 και 4 του DAPC, το οποίο είναι επίσης εμφανές στην ίδια την απεικόνιση του DAPC. Με το απόθεμα C να δείχνει κάποια αυξημένη διακύμανση, καθώς και υψηλές τιμές στις εκτιμήσεις ενδογαμίας του, στην ομάδα 4. Φαίνεται επίσης, ότι η ομάδα 2 του DAPC είναι η πιο συνδεδεμένη με αποθέματα που δείχνουν τις υψηλότερες τιμές ενδογαμίας ενώ έχει στη σύνθεσή της συγκεκριμένα αποθέματα: D, E, H, J που ανήκουν αποκλειστικά σε αυτή (σύμπλεγμα 2) και άτομα από άλλα αποθέματα, B, K, C, I που συνδέονται με αυτό. Το απόθεμα K φαίνεται να διαιρείται, όσον αφορά τις τιμές ενδογαμίας και τη σχέση μεταξύ των ομάδων 1,5 και 2, ομάδες που φαίνεται να έχουν κάποια γενετική εγγύτητα. Αυτό μπορεί να οφείλεται στο γεγονός ότι οι ιδρυτές του αποθέματος K μπορεί να προέρχονται από 3 διαφορετικούς πληθυσμούς λόγω του γεγονότος ότι για το απόθεμα K οι τιμές της ενδογαμίας είναι συγκριτικά χαμηλές.



Σχήμα 11. Γραφική απεικόνιση της συσχέτισης των τιμών ενδογαμίας των αποθεμάτων όπως υπολογίστηκαν με την μέθοδο DyadML στα αποθέματα, με τις 5 γενετικές ομάδες από τα κοινά αποτελέσματα Structure-DAPC, χρησιμοποιώντας την τοπολογία της DAPC ως αναφορά. Στον άξονα των ψ δίνονται οι εκτιμήσεις της ενδογαμίας, ενώ στον άξονα χ τα αποθέματα.

Με βάση τα αποτελέσματα, από τα αποθέματα I, J, K προτείνονται τα 100 άτομα με τις μικρότερες τιμές ενδογαμίας ως υποψήφιοι γεννήτορες.

3. Ανάθεση γονέων, εκτίμηση της κληρονομησιμότητας, των γενετικών και φαινοτυπικών συσχετίσεων για αναπτυξιακούς χαρακτήρες στον κρανίο

3.1 Κληρονομησιμότητα και το ζωϊκό μοντέλο

3.1.1 Κληρονομησιμότητα

Ένα γονίδιο που διαχωρίζεται σε έναν πληθυσμό μπορεί να επηρεάσει τον φαινότυπο ενός χαρακτηριστικού. Για ένα σύνθετο ή ποσοτικά κληρονομήσιμο χαρακτηριστικό, τα γονίδια που το καθορίζουν μπορεί να είναι πολυάριθμα και οι σχέσεις τους με το περιβάλλον μπορεί να είναι περίπλοκες. Η μελέτη της γενετικής βάσης ενός ποσοτικού χαρακτηριστικού είναι το ζητούμενο της ποσοτικής γενετικής. Ένα ποσοτικό χαρακτηριστικό με φαινοτυπική τιμή P , η οποία καθορίζεται από γενετικούς (G) και περιβαλλοντικούς παράγοντες (E), και η αλληλεπίδρασή τους ($G \times E$), εκφράζεται από την εξίσωση (3.1.1.1)

$$P = G + E + (G \times E) \quad (3.1.1.1)$$

Με την προϋπόθεση ότι όλοι οι όροι στην εξίσωση (3.1.1.1) είναι ανεξάρτητοι μεταξύ τους, χωρίζουμε την φαινοτυπική διακύμανση του χαρακτηριστικού στα αντίστοιχα τμήματα, γενετικό, περιβαλλοντικό, και γενετικό \times περιβάλλον (3.1.1.2)

$$V_P = V_G + V_E + V_{G \times E} \quad (3.1.1.2)$$

Η διακύμανση συμβολίζεται γενικά με V ή σ^2 . Η γενετική διακύμανση, V_G ή σ_G^2 , οφείλεται στις επιδράσεις όλων των γονιδίων που καθορίζουν το χαρακτηριστικό. Ας πάρουμε ως παράδειγμα, ένα γονίδιο με δύο αλληλόμορφα και τρεις γονότυπους AA , Aa και aa των οποίων οι συχνότητες και οι γονοτυπικές τιμές σε έναν πληθυσμό βρίσκονται σε ισορροπία Hardy-Weinberg. Οι γονοτυπικές τιμές αποτελούνται από το συνολικό μέσο όρο του χαρακτηριστικού (μ), την προστιθέμενη επίδραση (α) του γονιδίου λόγω της υποκατάστασης αλληλομόρφων από το A στο a ή το φαινόμενο κυριαρχίας (d) λόγω της αλληλεπίδρασης των διαφορετικών αλληλομόρφων A και a . Εάν δεν υπάρχει επικρατές αλληλόμορφο τότε, $d = 0$, ενώ αν το αλληλόμορφο A κυριαρχεί πάνω στο a , το $d > 0$, και αν το αλληλόμορφο a κυριαρχεί έναντι του A , το $d < 0$. Η κυριαρχία είναι απόλυτη εάν το d είναι ίσο με το $+a$ ή το $-a$, και υπάρχει υπερκυριαρχία αν το d είναι είτε $> +a$ ή $< -a$. Ο βαθμός της κυριαρχίας περιγράφεται από την αναλογία d/a . Ο λόγος για τον οποίο υφίσταται αυτός ο διαχωρισμός σε α και d αποδίδεται στην φύση της κληρονομικότητας σε έναν διπλοειδή (ή και πολυπλοειδή) δίοικο οργανισμό όπου έχουμε τυχαίες διασταυρώσεις. Κάθε γονέας συνεισφέρει ένα αλληλόμορφο για κάθε τόπο σε κάθε απόγονο, το κάθε αλληλόμορφο παρουσιάζει την προστιθέμενη επίδραση συνδυαζόμενο τυχαία με το δεύτερο αλληλόμορφο από τον δεύτερο γονέα, και η επίδραση κληρονομείται. Όμως, οι επιδράσεις που σχετίζονται με τα επικρατή αλληλόμορφα (d) είναι συνάρτηση μεταξύ των 2 αλληλομόρφων του ενός γονέα, που κατά την παραγωγή γαμετών διαχωρίζονται, και έτσι αυτές οι αλληλεπιδράσεις δεν κληρονομούνται αυτούσιες.

Ο λόγος της γενετικής διακύμανσης με την φαινοτυπική διακύμανση ορίζεται ως H^2 ή ευρεία κληρονομησιμότητα (*broad sense heritability*) και δίνεται από την εξίσωση (3.1.1.3)

$$H^2 = V_G / (V_G + V_E + V_{G \times E}) \quad (3.1.1.3)$$

Όπως αναφέρεται παραπάνω, η γενετική επίδραση ή διακύμανση μπορεί να διαχωριστεί σε αθροιστικές (A) και μη αθροιστικές (NA) επιδράσεις. Έτσι, από την εξίσωση (3.1.1.1) έχουμε

$$P = G + E + (G \times E) \rightarrow P = (A + NA) + E + (A \times E) + (NA \times E) \quad (3.1.1.4)$$

από την εξίσωση (3.1.1.2) έχουμε

$$V_P = V_G + V_E + V_{G \times E} \rightarrow V_P = (V_A + V_{NA}) + V_E + (V_{A \times E} + V_{NA \times E}) \quad (3.1.1.5)$$

Οι μη αθροιστικές επιδράσεις (ή η διακύμανση τους) ορίζονται ως το άθροισμα των κυρίαρχων και των επιστατικών επιδράσεων. Λόγω του ότι, οι αθροιστικές επιδράσεις μπορεί να κληρονομηθούν, από τους γονείς στον απόγονο, ενώ οι μη αθροιστικές όχι, όπως αναφέρθηκε, η αναλογία της διακύμανσης των αθροιστικών επιδράσεων με την συνολική φαινοτυπική διακύμανση ορίζεται ως κληρονομησιμότητα με την στενή έννοια h^2 (3.1.1.6).

$$h^2 = V_A / (V_A + V_{NA} + V_E + V_{A \times E} + V_{NA \times E}) \quad (3.1.1.6)$$

Οι επιστατικές επιδράσεις είναι πιθανό ότι διαδραματίζουν σημαντικό ρόλο στην έκφραση ποσοτικών χαρακτήρων. Όταν n γενετικοί τόποι συνδέονται με έναν χαρακτήρα, ο αριθμός των αθροιστικών επιδράσεων (α) είναι της τάξης μεγέθους του αριθμού n , ενώ οι διαλληλικές, τριαλληλικές ή πολυαλληλικές επιστατικές επιδράσεις, είναι της τάξης μεγέθους του n^2 , $n^3 \dots n^n$. Συνεπώς ακόμη και με μικρό εύρος επιστατικών επιδράσεων, η συνολική τους συνεισφορά μπορεί να έχει αξιοσημείωτο αντίκτυπο. Τα πέντε βασικά είδη επιστατικών επιδράσεων διαμορφώνονται ως εξής :

1. Υπολειπόμενη επίσταση: Η συγκάλυψη της έκφρασης υπολειπόμενων ή επικρατών γονιδίων σε έναν γενετικό τόπο από την δράση υπολειπόμενου γονιδίου σε άλλο γενετικό τόπο
2. Επικρατής επίσταση: Η συγκάλυψη της έκφρασης υπολειπόμενων ή επικρατών γονιδίων σε έναν γενετικό τόπο από την δράση επικρατούς γονιδίου σε άλλο γενετικό τόπο
3. Επικρατής ανασταλτική επίσταση: Η αναστολή της έκφρασης υπολειπόμενων ή επικρατών γονιδίων σε έναν γενετικό τόπο από την δράση επικρατούς γονιδίου σε άλλο γενετικό τόπο
4. Συνδυαστική επικρατής επίσταση: Η συγκάλυψη της έκφρασης επικρατών γονιδίων σε δύο γενετικούς τόπους από την δράση ενός επικρατούς γονιδίου σε έναν από τους δύο τόπους

5. Υπερεπικρατής επίσταση: Η ομοζυγωτία σε έναν τόπο επιφέρει ιδιότητες οι οποίες είναι πιο ισχυρές από τις δύο ομοζυγωτίες.

Οι μέθοδοι για την αξιολόγηση της ποσοτικής σημαντικότητας των επιστατικών επιδράσεων εμφανίζουν αρκετές αδυναμίες αλλά έχει σημειωθεί πρόοδος (Lee & Park 2007).

Με την κληρονομησιμότητα υπό στενή έννοια μπορεί να ποσοτικοποιηθεί ο βαθμός της μη-μεταβολής της φαινοτυπικής τιμής ενός ποσοτικού χαρακτηριστικού από τη μία γενιά στην επόμενη. Οι δύο τύποι κληρονομησιμότητας (3) και (6) χρησιμοποιούνται για να περιγράψουν τον βαθμό της γενετικής βάσης ενός χαρακτηριστικού, συμπεριλαμβανομένων των συνεισφορών όλων των συσχετιζόμενων γονιδίων (Lynch and Walsh 1998). Μία κεντρική πλευρά της h^2 είναι ότι συνδέεται με την εξίσωση της κληροδοτικής αξίας ενός ατόμου (3.1.1.7) όπου το S είναι η απόκριση ενός ατόμου στην επιλογή (διαφορικό της επιλογής), ή η αλλαγή στην μέση τιμή ενός φαινοτυπικού χαρακτήρα πριν την αναπαραγωγή. Οι κληροδοτικές τιμές είναι εκτιμήσεις της αθροιστικής δράσης των γονιδίων με βάση το ατομικό ζωικό πρότυπο.

$$\Delta\mu = h^2 S \quad (3.1.1.7)$$

Η στενή κληρονομησιμότητα h^2 μπορεί να εκφραστεί έτσι ως η απόδοση της απόκρισης $\Delta\mu$ στην επιλογή ενός φαινοτυπικού χαρακτήρα.

Οι δύο προσεγγίσεις για την κληρονομησιμότητα, χρησιμοποιούνται συνήθως για να περιγράψουν την συνεισφορά συγκεκριμένων γονιδίων και το κατά πόσο αυτά τα γονίδια μπορούν να ανιχνευθούν με μια μέθοδο όπως η γενετική χαρτογράφηση, η οποία περιγράφεται στο επόμενο κεφάλαιο. Στην πράξη, οι γενετικές αποκλίσεις μπορούν να εκτιμηθούν βάσει της ποσοτικής γενετικής θεωρίας που ιδρύθηκε από τον Cockerham (1954, 1963). Σύμφωνα με την θεωρία, ένα σύνολο γονέων διασταυρώνεται συνδυαστικά πολλαπλές φορές. Στη συνέχεια, και με βάση έναν συγκεκριμένο πειραματικό σχεδιασμό, συλλέγονται τα φαινοτυπικά δεδομένα και αναλύονται με στατιστικές μεθόδους, όπως την ανάλυση της διακύμανσης, ώστε να ληφθούν μία σειρά τιμών διακυμάνσεων που να αντανακλούν τον πειραματικό σχεδιασμό. Με βάση την ομοιότητα μεταξύ των συγγενών, οι εκτιμήσεις των διακυμάνσεων χρησιμοποιούνται για την εκτίμηση της προστιθέμενης γενετικής διακύμανσης, και της διακύμανσης της γενετικής κυριαρχίας, και ως εκ τούτου και των κληρονομησιμότητων.

3.1.2 Γραμμικά μοντέλα και το ζωικό μοντέλο (*animal model*)

Τα μικτά μοντέλα χρησιμοποιούνται στην βιολογία από τη δεκαετία του 1930. Αρχικά, το βάρος είχε δοθεί στην εκτίμηση των σταθερών επιδράσεων (fixed effects) όσο το δυνατόν ακριβέστερα λαμβάνοντας υπόψη τις τυχαίες επιδράσεις προερχόμενες από το πειραματικό πλαίσιο. Ο Yates (1940), εισήγαγε την μέθοδο του ισορροπημένου ημιτελούς μπλοκ (BIB balanced incomplete block). Υποθετικά, έχοντας άτομα προερχόμενα από διάφορες διασταυρώσεις και διατηρώντας τον αριθμό των απογόνων σταθερό, αν ασκήσουμε μία σειρά επιδράσεων σε κάθε άτομο θα έχουμε διαφορετικές αποκρίσεις οι οποίες αντιμετωπίζονται ως μη προβλέψιμες ή τυχαίες. Αν η σύγκριση των διάφορων διακυμάνσεων

στα άτομα κάθε διασταύρωσης, και συνολικά σε όλα τα άτομα, είναι το ζητούμενο, τότε προτείνεται από τον Yates η κατασκευή ενός σταθμισμένου μέσου όρου αυτών των δύο εκτιμήσεων. Όπου, με βάση μια ανάλυση της διακύμανσης, διαχωρίζεται πρώτα η παραλλακτικότητα σε δύο στρώματα, εντός και μεταξύ των μπλοκ, και η στάθμιση εξαρτάται από την παραλλαγή στα δύο στρώματα και τη διακύμανση της επίδρασης του πειράματος σε αυτά. Έτσι έχουμε τρία μέρη από τα οποία πηγάζει η πληροφορία για την εκτίμηση της διακύμανσης: τα δύο υπόλοιπα (*residuals*) του αθροίσματος των τετραγώνων στα δύο στρώματα και τη σύγκριση των πειραματικών αποτελεσμάτων στα δύο στρώματα. Ο Nelder (1968) επέκτεινε και γενίκευσε τη μέθοδο ώστε να επιτρέπει περισσότερες από δύο πηγές διακύμανσης, ή στρώματα, κάνοντας δυνατή την κατανομή των πειραματικών επιδράσεων σε σύνολα με διαφορετικές αποδόσεις. Αυτή η επέκταση του Nelder εξελίχθηκε εν τέλει στην πιο ολοκληρωμένη δουλειά του Henderson (Henderson 1985), την βέλτιστη γραμμική αμερόληπτη πρόβλεψη (BLUP best linear unbiased prediction).

Η εξίσωση (19) δίνει τον γενικό τύπο για αυτό που ονομάζεται μικτό μοντέλο ή ζωικό μοντέλο. Η εκτίμηση των γενετικών παραμέτρων που εμφανίζουν ενδιαφέρον στις αναλύσεις ποσοτικής γενετικής, όπως η στενή κληρονομησιμότητα, ή οι φαινοτυπικές και γενετικές συσχετίσεις χαρακτήρων δεν είναι το μόνο ζητούμενο. Σε ορισμένες περιπτώσεις, είναι πιθανό να απαιτείται η εκτίμηση των τυχαίων επιδράσεων όπως η κληροδοτική τιμή κάθε ατόμου, ή σταθερών επιδράσεων όπως οι περιβαλλοντικές επιδράσεις (στην ποσοτική γενετική τα ευρήματα που αφορούν τυχαίες επιδράσεις ονομάζονται προβλέψεις, ενώ αυτά που αφορούν σταθερές επιδράσεις, εκτιμήσεις). Η μέθοδος BLUP η οποία είναι μία γενική μέθοδος για την πρόβλεψη τυχαίων επιδράσεων προϋποθέτει ότι οι διακυμάνσεις των συστατικών προς διερεύνηση είναι γνωστές, ενώ η REML (*restricted maximum likelihood*) μέθοδος, υπό περιορισμό μέγιστη πιθανοφάνεια, βασίζεται στις προβλέψεις της BLUP μεθόδου για τις τυχαίες επιδράσεις ώστε να εκτιμήσει τις διακυμάνσεις των συστατικών. Ένα μοντέλο λέγεται μικτό όταν περιέχει τυχαίες (*u*) και σταθερές επιδράσεις (*b*). Από την (3.1.1.8) έχουμε

$$y = Xb + Zu + e \quad (3.1.1.8)$$

- $y = n \times 1$ διάνυσμα των παρατηρήσεων, n = αριθμός καταγραφών
- $b = p \times 1$ διάνυσμα των σταθερών επιδράσεων, p = αριθμός επιπέδων των σταθερών επιδράσεων (fixed effects)
- $u = q \times 1$ διάνυσμα των τυχαίων επιδράσεων του κάθε ατόμου, q = αριθμός των επιπέδων των τυχαίων επιδράσεων (random effects)
- $e = n \times 1$ διάνυσμα των τυχαίων υπολοίπων επιδράσεων, από το περιβάλλον, μη προστιθέμενες γενετικές επιδράσεις
- X = πίνακας της τάξης $n \times p$, ο οποίος συσχετίζει τις καταγραφές με τις σταθερές επιδράσεις
- Z = πίνακας της τάξης $n \times q$, ο οποίος συσχετίζει τις καταγραφές με τις τυχαίες επιδράσεις του κάθε ατόμου

Το X και το Z ονομάζονται πίνακες επίπτωσης ή incidence matrices και είναι πίνακες για τις επιδράσεις του b και του u αντίστοιχα των οποίων όλα τα στοιχεία είναι 0 ή 1, ανάλογα με το αν ο υπό μελέτη χαρακτήρας συνεισφέρει στον φαινότυπο του ατόμου. Για το u και το e , προϋπόθεση είναι να αποτελούν κανονικές κατανομές. Οι συναρτήσεις διακύμανσης για τα

y , u , e ($\text{var}(y)$, $\text{var}(u)$, $\text{var}(e)$), ισούνται αντίστοιχα με τους V , G , R πίνακες. Συγκεκριμένα, όπου G έχουμε τον πίνακα ($q \times q$) γενετικών διακυμάνσεων και συνδιακυμάνσεων για το διάνυσμα τυχαίων γενετικών επιδράσεων (u), και R είναι ο ($n \times n$) πίνακας διακυμάνσεων και συνδιακυμάνσεων για τις υπόλοιπες επιδράσεις (e). Ο πίνακας V των διακυμάνσεων και συνδιακυμάνσεων του διανύματος y δίνεται από την σχέση $V = V(y) = ZGZ' + R$. Όπου Z' ο ανάστροφος πίνακας του Z . Ο όρος (Z) αφορά την συνεισφορά των τυχαίων γενετικών επιδράσεων και ο (R) των υπολοίπων e . Η γενική παραδοχή είναι ότι τα σφάλματα του e δεν συσχετίζονται με την διακύμανση του R και ο πίνακας παίρνει την μορφή ενός διαγώνιου πίνακα : $R = \sigma_e^2 I$, όπου I είναι ο μοναδιαίος $n \times n$ πίνακας. Για ένα μικτό ζωικό μοντέλο γνωρίζουμε τα y , X , Z ενώ τα ζητούμενα είναι τα b , u , R , G . Η διαδικασία γίνεται σε δύο βήματα,

1. Εκτίμηση των b και u
Εάν οι διακυμάνσεις δεν είναι γνωστές εφαρμόζεται μία από τις μεθόδους, ANOVA, REML για την εκτίμηση τους.
2. Εκτίμηση των πινάκων διακυμάνσεων συνδιακυμάνσεων των G , R .

Αν μελετώνται ταυτόχρονα δύο ποσοτικοί χαρακτήρες έχουμε μία επιπρόσθετη εξίσωση όπως η (3.1.1.8) για τον δεύτερο χαρακτήρα. Για να έχουμε εκτιμήσεις των συσχετίσεων των δύο ή περισσότερων μεταβλητών (ποσοτικοί χαρακτήρες) γίνεται χρήση άλγεβρας πινάκων στο σύστημα των γραμμικών εξισώσεων των δύο (ή περισσότερων μεταβλητών). Η λύσεις του συστήματος των γραμμικών εξισώσεων αποτελούν και τις εκτιμήσεις των συσχετίσεων. Υπάρχουν διάφορες προσεγγίσεις στο τρόπο επίλυσης του συστήματος των εξισώσεων η αλγεβρική φύση όμως του ζητήματος αυτού ξεφεύγει από το πλαίσιο αυτής της διδακτορικής διατριβής, αλλά θα μπορούσε να αναζητηθεί στην αντίστοιχη στατιστική-μαθηματική βιβλιογραφία.

3.2 Ανάλυση απογόνων κρανιού ως προς την κληρονομησιμότητα των χαρακτηρισμών βάρους και μήκους, των γενετικών και φαινοτυπικών τους συσχετίσεων και του δραστικού μεγέθους πληθυσμού Ne

Σε αυτή την ενότητα, παρατίθεται το δημοσιευμένο άρθρο στο περιοδικό *Aquaculture*, Volume 518, 15 Μαρτίου 2020, 734663 και doi <https://doi.org/10.1016/j.aquaculture.2019.734663>

3.2.1 Εισαγωγή

Τις τελευταίες δεκαετίες, έχει σημειωθεί σημαντική ερευνητική προσπάθεια σε καθιερωμένα είδη στην υδατοκαλλιέργεια, όπως τιλάπια, κυπρίνο, γαρίδα και σολομό. Οι όποιες προσπάθειες βελτίωσης που αυξάνουν την απόδοση της παραγωγής και την ανθεκτικότητα σε ασθένειες, είναι απαραίτητες για τη μείωση του κόστους και την επίτευξη ενός ανταγωνιστικού προϊόντος. Ο κρανιός (*Argyrosomus regius*) είναι ένα είδος της οικογένειας Sciaenidae το οποίο εμφανίζει υψηλές δυνατότητες εντατικής καλλιέργειας. Τα χαρακτηριστικά του είδους που αποδεικνύουν τις δυνατότητες για την εκτροφή του σε μεγάλη κλίμακα είναι: η καλή αναλογία μετατρεψιμότητας της τροφής σε βιομάζα (FCR feed conversion ratio) και ο ρυθμός ανάπτυξης (Fountoulaki et al., 2017). Έχει ελκυστικό σχήμα σώματος, χαμηλό ποσοστό λίπος φιλέτου και καλή απόδοση κατά την επεξεργασία. Επιπρόσθετα οφέλη είναι η σχετικά εύκολη διαχείριση του γόνου, η παραγωγή ωαρίων καλής ποιότητας (Duncan et al., 2012, 2013; 2018; Mylonas et al., 2013, 2015; 2016; Fernandez-Palacios et al., 2014), και η εύκολη εκτροφή ιχθυοσυμμάτων, με την υποστήριξη μίας εδραιωμένης και τυποποιημένης βιομηχανίας ζωντανών τροφών και σχημάτων διατροφής (Duncan et al., 2013; Vallés and Estévez, 2013). Σήμερα έχει αναδειχθεί ως σημαντικό είδος για τη μεσογειακή υδατοκαλλιέργεια (Duncan et al., 2013) με ετήσια παραγωγή άνω των 23.000 τόνων (FAO, 2016). Η γεωγραφική του κατανομή είναι από την Μεσόγειο και τη Μαύρη Θάλασσα, κατά μήκος των ακτών του Ατλαντικού της Ευρώπης και της Αφρικής (Poli et al., 2003). Η εκτροφή του κρανιού ξεκίνησε τη δεκαετία του 1990 στη Γαλλία και επεκτάθηκε στο μεγαλύτερο μέρος της Μεσογείου. Το είδος χαρακτηρίζεται από ανοχή σε ευρύ φάσμα αλατότητας (5–39 ‰) καθώς και θερμοκρασίας (13–28 ° C). Η χρησιμότητα των μικροδορυφορικών δεικτών έχει δείχθει για την περιγραφή της γενεαλογίας σε είδη υδατοκαλλιέργειας (Herbinger et al., 1995; Moore et al., 1999), και τα αποτελέσματα μελετών δείχνουν ότι η γενεαλογία μικτών πληθυσμών μπορεί να καθοριστεί μέσω της χρήσης μικροδορυφορικών δεικτών (O'Reilly and Wright, 1995; Perez-Enriquez et al., 1999; Waldbieser and Wolters, 1999; Dong et al., 2008). Ένας εκ των πρωταρχικών στόχων αυτής της μελέτης είναι η επιτυχής ανάθεση των απογόνων στους υποτιθέμενους γονείς τους, και η διερεύνηση της γενετικής βάσης για το μήκος και το βάρος ως δύο σημαντικά χαρακτηριστικά για την παραγωγή. Έχοντας ακολουθήσει τις τυπικές βιομηχανικές πρακτικές διαχωρισμού των ιχθυδίων, δηλαδή τη διαλογή ιχθυδίων ίδιου μεγέθους, οι δύο παρτίδες οι οποίες ανατέθηκαν στους πιθανούς γονείς τους εμφάνισαν διαφορετικούς ρυθμούς ανάπτυξης. Εξετάζουμε τον πιθανό αντίκτυπο της διαφορικής ανάπτυξης μεταξύ αυτών των δύο παρτίδων στο σωματικό βάρος και το συνολικό μήκος. Οι διαφορές στο βάρος και το μήκος, καθώς και οι εκτιμήσεις κληρονομησιμότητας, γενετικών και φαινοτυπικών συσχετίσεων, εξετάστηκαν σε αυτό το κεφάλαιο ως μία προκαταρκτική ανάλυση που μπορεί ενδεχομένως να ρίξει φως στη μελλοντική μας έρευνα σχετικά με την αναγνώριση QTL για αυτούς τους χαρακτήρες στον κρανιό.

3.2.2 Υλικά και μέθοδοι

3.2.2.1 Βιολογικό υλικό

Τον Δεκέμβριο του 2012 έγινε δειγματοληψία πτερυγίων από δεκαεννέα γεννήτορες (Tank M7). Όλοι οι ιχθύες είχαν ηλικία άνω των 6 ετών. Σε όλους τους (6) θηλυκούς, και στους 7 από τους 13 αρσενικούς γεννήτορες έγιναν ενέσεις με GnRHα για την συγχρονισμένη απελευθέρωση ωαρίων και σπερματοζωαρίων τον Απρίλιο του 2013. Η πρώτη παρτίδα μεγαλύτερων ιχθυδίων τοποθετήθηκε σε κλωβό στις 2 Αυγούστου 2013 και η δεύτερη παρτίδα με μικρότερα ιχθύδια στις 14 Αυγούστου 2013, σε μια εμπορική εκμετάλλευση στα ανοικτά της ισπανικής ακτής στην κοινότητα της Βαλένθια. Οι δύο παρτίδες μπήκαν στους κλωβούς με διαφορά 12 ημερών. Στις 20 Ιανουαρίου 2016, πραγματοποιήθηκε δειγματοληψία των πτερυγίων από 400 ιχθύες (παρτίδα 1) στο εμπορικό μέγεθος (~ 2 κιλά), και το συνολικό μήκος και το βάρος μετρήθηκε για όλα τα άτομα στο εργοστάσιο επεξεργασίας της εταιρείας. Στις 5 Μαΐου 2016, πραγματοποιήθηκε η ίδια διαδικασία για άλλους 400 ιχθύες από τον θαλάσσιο κλωβό της δεύτερης παρτίδας που πραγματοποιήθηκε στον ίδιο εμπορικό χώρο εκμετάλλευσης (παρτίδα 2).

3.2.2.2 Εξαγωγή DNA, συνθήκες PCR

Έγινε εξαγωγή DNA από τα 800 δείγματα χρησιμοποιώντας ένα ευρέως διαδεδομένο πρωτόκολλο (Miller et al., 1988) και η ποιότητα και η ποσότητα του DNA αξιολογήθηκαν χρησιμοποιώντας το φασματοφωτόμετρο NanoDrop ND 1000 (Thermo Fisher Scientific, www.thermofisher.com). Χρησιμοποιήθηκε ένας συνδυασμός 10 μικροδορυφορικών τόπων (multiplex) (Soula et al., 2011) (Πίνακας 1). Οι συνθήκες για την PCR στον κρανιό βελτιστοποιήθηκαν (από τους Turner et al. 1998 και O'Malley et al. 2003). Οι αντιδράσεις έγιναν χρησιμοποιώντας με ένα ειδικό kit για τον ταυτόχρονο πολλαπλασιασμό πολλών γενετικών τόπων (multiplex kit) της Qiagen. Οι αντιδράσεις PCR πραγματοποιήθηκαν σε ένα μείγμα αντίδρασης 10 μl με συγκέντρωση 10 μmol/l για κάθε primer και 5 ng/μl DNA προτύπου. Το θερμικό προφίλ περιλάμβανε ένα στάδιο προ-αποδιάταξης στους 95°C για 15 λεπτά ακολουθούμενο από 30 κύκλους αποδιάταξης-ανόπτησης-επέκτασης στους 94°C για 30 δεύτερα, 57°C για 1,3 λεπτά, 72°C για 1 λεπτό και ένα τελικό βήμα επιμήκυνσης στους 60°C για 30 λεπτά. Τα ενισχυμένα τμήματα (amplicons) διαχωρίστηκαν με τριχοειδή ηλεκτροφόρηση σε ABI 3730 αλληλουχητή (Applied Biosystems, Foster City, CA) χρησιμοποιώντας τον τυποποιημένο δείκτη μεγέθους LIZ500. Το λογισμικό ανάλυσης μεγέθους θραυσμάτων STRand (<http://www.vgl.ucdavis.edu/informatics/STRand>) χρησιμοποιήθηκε για την γονοτύπηση των μικροδορυφορικών τόπων.

Πίνακας 1. Σύνοψη για τους 10 μικροδορυφορικούς τόπους που χρησιμοποιήθηκαν στον κρανιό. Στην πρώτη στήλη δίνεται το όνομα του κάθε τόπου, ενώ στην δεύτερη στήλη το μοτίβο της επανάληψης. Στην τρίτη στήλη ο τύπος της φθορίζουσας χρώσης που προστέθηκε στο 5' άκρο του εκκινητή R. Στην τέταρτη και πέμπτη στήλη δίνονται αντίστοιχα οι εμπρόσθιες και οι οπίσθιες ανάστροφες αλληλουχίες των εκκινητών.

Μικροδορυφορικοί τόποι	Μοτίβο	Χρώση	Εμπρόσθιες αλληλουχίες (5'–3')	Οπίσθιες ανάστροφες αλληλουχίες (5'–3')
Casmic 14	(CT) ₁₂	6-FAM	TGTCCTCACTCCTCTTTTTCTTTC	GTTTAAGGCGCATCTCCAGTCTC
UBA005	(CT) ₁₆	NED	CATCAGGATTGGCAACTAGC	GTTTCCTCCAGGTTTATTCTTCATTGAC
UBA006	(CA) ₂₂ (GT) ₂₆ /(GT)	PET	AGCACACGTAATCACACACAGAT	GTTCCACTAGTGCAAAACGGTGGT
UBA050	₉	6-FAM	GCACAACCTGCATCCCTTAGAT	GTTTAGAAGTGAAGACTGCGGACTG
UBA054	(CA) ₁₇	6-FAM	CCTTGTGAGAACAATTAATTTGGATG	GTTTCAAACCCTGATAGATGGATAGTT
SOC11	(GA) ₁₁	6-FAM	GCCGAGTCACGAAGGAACAGAGAA	TGTCGTCTCATCTATCTCCATCTC
SOC405	(CA) ₁₂	PET	AGGCTTTTGTTTAGTTTCCCTCAT GTGGTAGATGAAAACGTATAAAAGG	GGGGTGTAGCAGAACCACAC
SOC431	(TG) ₃₀	HEX	AG	GTTTCATATATATAGTGTACAGCTCCAGCT
SOC35	(CT) ₅ (GA) ₉	PET	GAGGGTGACGCTAACAGTTGA	CTCTACCTCACACTCCTCAAAGT
SOC428	(TG) ₃₈	NED	GACATCGCATTTGTCTACAGAGTC	AACTCCAGTCATAATATCCCTT

3.2.2.3 Ανάλυση ανάθεσης γονέων

Η ανάθεση των απογόνων στους γονείς τους έγινε μέσω του λογισμικού Vitassign (Vandeputte et al., 2006). Δεκατρείς αρσενικοί και έξι θηλυκοί γεννήτορες διασταυρώθηκαν και δεδομένου ότι δεν υπήρχαν προηγούμενες πληροφορίες σχετικά με την γενεαλογία, τα ομοθαλή αδέρφια, ετεροθαλή αδέρφια και μη-συγγένεια ήταν οι μόνες πιθανές σχέσεις συγγένειας για οποιοδήποτε ζευγάρι ιχθύων από τους απογόνους. Το Vitassign χρησιμοποιεί τη μέθοδο του αποκλεισμού που βασίζεται στον διαχωρισμό των αλληλομόρφων γονιδίων κατά Μέντελ ως μοναδική του υπόθεση. Στις αναλύσεις επιτράπηκε οι αναντιστοιχίες

μεταξύ γονέων και απογόνων τέθηκαν όπως στο κεφάλαιο 2, και όπου υπήρχαν διπλά ζεύγη γονέων για έναν απόγονο, ο απόγονος δεν συμπεριλαμβάνονταν στα αποτελέσματα.

3.2.2.4 Ανάλυση δεδομένων

Μετρήθηκε το βάρος (σε κιλά, kg) και το συνολικό μήκος (σε εκατοστά, cm), και στους 800 ιχθύες των δύο παρτίδων, που αποτελεί το σύνολο των φαινοτυπικών δεδομένων. Οι μέσες τιμές για το μήκος ήταν 61,92 cm για την παρτίδα 1 και 62,17 cm στην παρτίδα 2 και για το βάρος ήταν 2,289 kg για την παρτίδα 1 και 2,317 kg για την παρτίδα 2. Χρησιμοποιήθηκε το τεστ του Levene για την εξέταση της ετερογένειας της διακύμανσης για τα χαρακτηριστικά βάρος και μήκος των δύο παρτίδων. Οι διαφορές στο βάρος και το μήκος για κάθε κοινή οικογένεια των δύο παρτίδων, ελέγχθηκαν για στατιστικά σημαντικές διαφορές. Έγιναν δύο t-τέστ σε κάθε μία από τις 15 κοινές οικογένειες, για το βάρος και μήκος αντίστοιχα. Το όριο τιμής P για τη στατιστική σημαντικότητα ορίστηκε σε $P \leq 0,05$. Η τιμή P, εκτιμήθηκε για μια δοκιμή δύο ουρών (*two-tail test*), δηλαδή ο μέσος όρος θεωρείται στατιστικά σημαντικός εάν βρίσκεται στο άνω 2,5% ή κάτω 2,5% της κατανομής πιθανότητας, παίρνοντας μια τιμή $p < 0,05$ υποθέτοντας άνισες διακυμάνσεις με το SPSS v24.0 (IBM Corp.). Για κάθε κοινή οικογένεια στις δύο παρτίδες, οι απόγονοι σχηματίζουν δύο ξεχωριστές ομάδες άνισων μεγεθών. Τελευταία, υπολογίστηκε το δραστικό μέγεθος πληθυσμού N_e (Falconer και Mackay, 1996) χρησιμοποιώντας:

α) τον τύπο της άνισης συμβολής αρσενικών και θηλυκών (3.2.2.4.1)

$$N_p = 4N_m N_f / (N_m + N_f) \quad (3.2.2.4.1)$$

όπου N_m και N_f είναι ο αριθμός των αρσενικών και θηλυκών γονέων, αντίστοιχα, και β) τον τύπο που λαμβάνει υπόψη τη διακύμανση του οικογενειακού μεγέθους (3.2.2.4.2)

$$N_v = 8N / (V_m + V_f + 4) \quad (3.2.2.4.2)$$

όπου V_m και V_f είναι η διακύμανση της συνεισφοράς των αρσενικών και θηλυκών γεννητόρων στις οικογένειες, αντίστοιχα (Hill, 1979). Στον υπολογισμό του N_v , χρησιμοποιήθηκαν οι τιμές των V_m και V_f που υπολογίστηκαν από τα δεδομένα N_{v1} , και για το N_{v2} οι τιμές διακύμανσης θεωρούνται ίσες με το μ , μέσο μέγεθος οικογένειας, ώστε να θεωρούνται κατανομή κατά Poisson. Οι εκτιμήσεις της κληρονομησιμότητας, των γενετικών και φαινοτυπικών συσχετίσεων υπολογίστηκαν από τους 728 ανατεθειμένους απογόνους των δύο παρτίδων χρησιμοποιώντας ένα ζωικό μοντέλο με τη μέθοδο μέγιστης πιθανοφάνειας στο λογισμικό VCE4.0 (Kovach et al., 2002). Στο ζωικό μοντέλο, η παρτίδα ήταν η μοναδική σταθερή επίδραση, προκειμένου να ληφθούν υπόψη τυχόν διαφορές λόγω του χρόνου τοποθετήσεων στους κλωβούς, όπως επίσης και για την περίοδο ανάπτυξης. Πραγματοποιήθηκαν δύο επιπρόσθετες αναλύσεις για την εκτίμηση της κληρονομησιμότητας, των γενετικών και φαινοτυπικών συσχετίσεων για κάθε παρτίδα ξεχωριστά, χωρίς να συμπεριλαμβάνεται κάποια σταθερή επίδραση στην ανάλυση.

3.3 Αποτελέσματα

3.3.1 Ανάλυση ανάθεσης γονέων

Οι απόγονοι της παρτίδας 1 γονοτυπήθηκαν επιτυχώς και η ανάλυση ανάθεσης ήταν επιτυχής στο 87,5%, επιτρέποντας μόνο μια αναντιστοιχία. Συνολικά, 348 απόγονοι ανατέθηκαν στους υποτιθέμενους γονείς τους. Η κατανομή των οικογενειών ορίστηκε στις 18, από τις θεωρητικά 78 αναμενόμενες. Μόνο 5 από τους 6 θηλυκούς γεννήτορες αναγνωρίστηκαν ως πιθανοί γονείς των απογόνων. Τα θηλυκά 4F και 3F συμμετείχαν περισσότερο, και σε μικρότερο βαθμό τα θηλυκά 1F και 5F. Παρομοίως, 7 στους 13 αρσενικούς γεννήτορες αναγνωρίστηκαν ως πιθανοί γονείς. Ο αρσενικός γεννήτορας 6M φαίνεται να είναι υπεύθυνος για σχεδόν το ήμισυ των απογόνων με τους 2M, 1M και 5M (βλ. **Πίνακα 2**) να ακολουθούν.

Πίνακας 2. Διαμορφούμενες οικογένειες απογόνων κρανιού μετά από μία διασταύρωση 19 γεννητόρων στην παρτίδα 1.

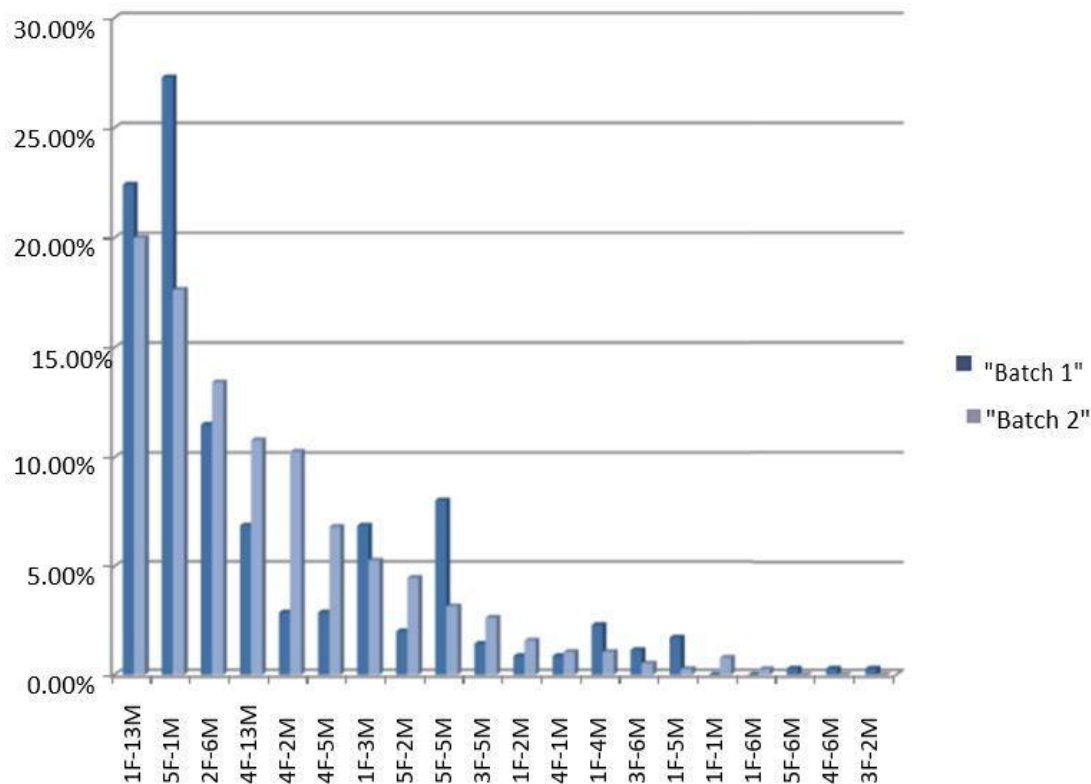
Παρτίδα 1 ♀	♂	1M	2M	3M	4M	5M	6M	13M	Sum
1F		2,87%	1,44%	1,15%	2,01%	2,87%	6,90%	0,29%	17,53%
2F							0,29%		0,29%
3F			2,41%			0,86%	6,90%		30,17%
4F		8,05%				1,72%	27,30%		37,07%
5F		0,29%	2,30%			0,86%	11,49%		14,94%
Sum		11,21%	26,15%	1,15%	2,01%	6,32%	52,87%	0,29%	

Για τους ιχθύες της παρτίδας 2, το τελικό ποσοστό ανάθεσης διαμορφώθηκε στο 95%, ή 380 απογόνους. Οι 17 οικογένειες της παρτίδας 2 ακολούθησαν παρόμοια κατανομή ειδικά σε σχέση με την συμμετοχή των αρσενικών γεννητόρων. (**Πίνακας 3**). Συνολικά, το ποσοστό ανάθεσης υπολογίζεται στο 91,1% και στις δύο παρτίδες.

Πίνακας 3. Διαμορφούμενες οικογένειες απογόνων κρανιού μετά από μία διασταύρωση 19 γεννητόρων στην παρτίδα 2.

Παρτίδα 2 ♀	♂	1M	2M	3M	4M	5M	6M	13M	Sum
1F		10,26%	2,63%	0,53%	4,47%	6,84%	10,79%		35,53%
2F									0
3F			20,00%			1,58%	5,26%		26,84%
4F		3,16%	0,79%			0,26%	17,63%	0,26%	22,11%
5F			1,05%			1,05%	13,42%		15,53%
Sum		13,42%	24,47%	0,53%	4,47%	9,74%	47,11%	0,26%	

Το στατιστικό τεστ χ^2 (Chi square test) για τις οικογένειες δείχνει ότι η κατανομή των οικογενειακών μεγεθών μεταξύ των παρτίδων δεν είναι ανεξάρτητη ($P = 0,044$), και κατά συνέπεια, η διακύμανση της οικογενειακής σύνθεσης μεταξύ των δύο παρτίδων δεν διαφέρει σημαντικά με τις δύο παρτίδες να είναι ομοιογενείς όσον αφορά το μέγεθος των οικογενειών (**Σχήμα 1**). Υπήρχαν εκπρόσωποι από όλες τις οικογένειες και στις δύο παρτίδες, δηλαδή και στα μεγάλα ψάρια διαλογής και στα μικρά. Και στις δύο παρτίδες, τρεις θηλυκοί γεννήτορες, οι 1F, 3F και 4F, συμμετείχαν περισσότερο (26,9, 28,4 και 29,3%, αντίστοιχα), και σε πολύ μικρότερη έκταση ο 5F (0,7%). Η 2F αναγνωρίζεται ως γονέας σε έναν μόνο απόγονο. Ομοίως, επτά στους 13 αρσενικούς γεννήτορες αναγνωρίστηκαν ως πιθανοί γονείς. Ο 6M φαίνεται να είναι υπεύθυνος για σχεδόν το ήμισυ των απογόνων (49,9%) ακολουθούμενος από τους 2M, 1M, 5M και 4M (25,3, 12,4, 8,1 και 3,3%, αντίστοιχα). Από τους έξι θηλυκούς γεννήτορες που έλαβαν ορμόνη GnRHα, μόνο ένας δεν έδωσε κανένα απόγονο (6F). Ενώ από τους επτά αρσενικούς που έλαβαν ορμόνη GnRHα, μόνο οι 7M και 11M δεν είχαν κάποιο απόγονο. Αντιθέτως, ο 1M στον οποίο δεν χορηγήθηκε η ορμόνη, συμμετείχε έντονα στις διαμορφούμενες οικογένειες με ποσοστό 11,2% και 13,4%, αντίστοιχα (**Πίνακες 2 και 3**).



Σχήμα 1. Διαφορική ανάθεση στις δύο παρτίδες. Όλοι οι ιχθύες απομακρύνθηκαν από τους κλωβούς τους έχοντας μέσο βάρος 2 κιλά και περίπου 100 ημέρες διαφοράς εξαλίευσης από τους κλωβούς. Οι οικογένειες απεικονίζονται στον άξονα X και το μέγεθος τους εκφρασμένο ως ποσοστό στον άξονα Y.

3.3.2 Δραστικό μέγεθος πληθυσμού N_e

Η διακύμανση ενός εξιδανικευμένου πληθυσμού, με άνισο αριθμό αρσενικών και θηλυκών γονέων, δίνεται από την εξίσωση $N_v = 8N / (V_m + V_f + 4)$, όπου $V_m = V_f = \mu$ (μέσο οικογενειακό μέγεθος) για τον N_{v2} . Στον πίνακα 4, δείχνουμε τους υπολογισμούς των N_{v1} , N_{v2} και N_p . Οι αναλογίες που λάβαμε N_{v1}/N , N_{v2}/N και N_p/N ήταν 0,02, 0,08 και 0,02, αντίστοιχα. Οι τιμές μας του N_e ήταν σημαντικά χαμηλότερες από το N (τους δεκαεννέα γεννήτορες), με διακύμανση $V_{family} = 2,351$ και το μέσο οικογενειακό μέγεθος $\mu = 42,6$.

Πίνακας 4. Υπολογισμός του δραστικού μεγέθους του πληθυσμού μετά από τρεις προσεγγίσεις, τη μέση συνεισφορά των θηλυκών (C_f) και των αρσενικών (C_m), καθώς και των διαφορών τους στα μεγέθη των οικογενειών (V_f και V_m για θηλυκά και αρσενικά, αντίστοιχα). Πρώτον, το δραστικό μέγεθος πληθυσμού εκτιμάται σύμφωνα με $N_p = (4N_m N_f) / (N_m + N_f)$. τότε το N_{v1} υπολογίζεται σύμφωνα με τη διακύμανση των οικογενειακών μεγεθών των κοινών οικογενειών στις δύο παρτίδες και το τελευταίο N_{v2} υπολογίζεται για τις κοινές οικογένειες και στις δύο παρτίδες με τις διακυμάνσεις στα οικογενειακά μεγέθη να είναι $V_f = V_m = \mu$. N αντιπροσωπεύει το μέγεθος του πληθυσμού απογραφής, το μ είναι το μέσο οικογενειακό μέγεθος, ενώ και τα N_{v1} και N_{v2} υπολογίζονται από την εξίσωση $N_v = 8N / (V_m + V_f + 4)$.

	N	N_p	N_{v1}	N_{v2}	C_f	C_m	V_f	V_m
Παρτίδα 1&2 ανατεθειμένοι απόγονοι	728	16,42	14,92	64,95	120,83	55	8364	10555
N_e/N	-	0,02	0,02	0,08	-	-	-	-

3.3.3 Εκτίμηση ετερογένειας των διακυμάνσεων για το βάρος και μήκος

Λόγω του γεγονότος ότι διεξήχθη ανάλυση για την εκτίμηση δύο μεταβλητών σε ένα συνδυασμένο σύνολο δεδομένων, έγινε εκτίμηση της υπόθεσης ότι οι διαφορές βάρους και μήκους που λαμβάνονται από δύο διαφορετικά υποσύνολα (παρτίδες) δεν είναι ίσες. Οι τιμές μήκους δεν έχουν ίσες διακυμάνσεις μεταξύ των δύο παρτίδων. Αυτό αναμένεται, λόγω της εφαρμογής της διαλογής η οποία γίνεται με βάση το μήκος αρχικά στα ιχθύδια. Αντιθέτως, οι τιμές βάρους δείχνουν ίσες διακυμάνσεις και πληρούνται τα κριτήρια ομοιογένειας. Διαπιστώθηκε ετερογένεια των διακυμάνσεων όσων αφορά το μήκος σε επίπεδο σημαντικότητας 0.05. Κατά συνέπεια, η προέλευση των ιχθύων (παρτίδα 1 ή 2) προσαρμόστηκε ως σταθερή επίδραση στην γενετική ανάλυση, προκειμένου να ληφθούν υπόψη οι διαφορές των διακυμάνσεων του μήκους στις δύο παρτίδες.

3.3.4 Εκτίμηση των συντελεστών κληρονομησιμότητας, και των γενετικών και φαινοτυπικών συσχετίσεων

Οι εκτιμήσεις του συντελεστή κληρονομησιμότητας για το σωματικό βάρος κυμαινόνταν από 0,64 έως 0,76 για τις δύο παρτίδες, με τους απογόνους της παρτίδας 1 να δείχνουν 12% υψηλότερες εκτιμήσεις από την παρτίδα 2 (Πίνακας 5). Οι εκτιμήσεις του συντελεστή κληρονομησιμότητας για το χαρακτηριστικό του μήκους κυμαινόνταν από 0,69 έως 0,72 για τις δύο παρτίδες, με τους απογόνους της παρτίδας 1 να δείχνουν 3% υψηλότερες εκτιμήσεις από τη παρτίδα 2. Οι εκτιμήσεις του συντελεστή κληρονομησιμότητας της κοινής ανάλυσης των δύο παρτίδων, ήταν 0,62 για το βάρος και 0,64 για το μήκος. Οι εκτιμήσεις της γενετικής συσχέτισης μεταξύ βάρους και μήκους ήταν 0,97 για την πρώτη παρτίδα, 0,94 για τη δεύτερη παρτίδα, και 0,96 και για τις δύο παρτίδες. Οι εκτιμήσεις της φαινοτυπικής συσχέτισης μεταξύ βάρους και μήκους ήταν 0,78 για την πρώτη παρτίδα, 0,80 για τη δεύτερη παρτίδα 2, και 0,82 και για τις δύο παρτίδες. Επιπλέον, οι ίδιοι γενετικοί παράμετροι (κληρονομησιμότητα, γενετικές και φαινοτυπικές συσχετίσεις) εκτιμήθηκαν σε μία ανάλυση 4 μεταβλητών, όπου οι μετρήσεις βάρους και μήκους θεωρήθηκαν ως διαφορετικοί χαρακτήρες σε κάθε μία από τις 2 διαφορετικές παρτίδες. Τα αποτελέσματα έδειξαν ότι οι εκτιμήσεις κληρονομησιμότητας μειώθηκαν δραστικά στην πρώτη παρτίδα (0,27 και 0,26 για το σωματικό βάρος και το μήκος, αντίστοιχα) και σε μικρότερη έκταση στην δεύτερη παρτίδα, (0,46 και 0,52 για σωματικό βάρος και μήκος, αντίστοιχα). Ωστόσο, οι γενετικές συσχετίσεις μεταξύ του βάρους και του μήκους ήταν υψηλές μεταξύ των δύο παρτίδων (0,90 και 0,88, αντίστοιχα). Επιπλέον, οι συσχετίσεις κατάταξης (Spearman) μεταξύ των εκτιμώμενων κληροδοτικών τιμών του βάρους και του μήκους στις δύο παρτίδες ήταν πολύ υψηλές (0,95 και για τα δύο χαρακτηριστικά).

Πίνακας 5. Εκτιμήσεις κληρονομησιμότητας, γενετικών και φαινοτυπικών συσχετίσεων για τις δύο παρτίδες καθώς και για κάθε παρτίδα ξεχωριστά. Οι εκτιμήσεις κληρονομησιμότητας εμφανίζονται με έντονη γραφή και τα τυπικά σφάλματα βρίσκονται σε παρένθεση.

Απόθεμα κρανιού	Βάρος	Μήκος	Γενετική συσχέτιση	Φαινοτυπική συσχέτιση
Παρτίδα 1	0,769 (0,14)	0,725 (0,14)	0,974 (0,01)	0,788 (0,09)
Παρτίδα 2	0,643 (0,13)	0,699 (0,14)	0,945 (0,02)	0,809 (0,06)
Παρτίδες 1&2	0,627 (0,12)	0,643 (0,12)	0,962 (0,01)	0,826 (0,04)

3.3.5 Στατιστικά σημαντικές διαφορές στο βάρος και μήκος για τις κοινές οικογένειες των 2 παρτίδων

Η ύπαρξη δύο παρτίδων απογόνων με διαφορετικό ρυθμό ανάπτυξης που προέρχονται από τους ίδιους γεννήτορες, εγείρει το ζήτημα των στατιστικά σημαντικών διαφορών στο σωματικό βάρος και το μήκος μεταξύ των παρτίδων. Για κάθε κοινή οικογένεια στις δύο παρτίδες, οι απόγονοι σχηματίζουν δύο ξεχωριστές ομάδες άνισων μεγεθών. Τρεις οικογένειες παρουσίασαν στατιστικά σημαντικές διαφορές στο βάρος, και δύο οικογένειες στο μήκος, δείχνοντας ότι παρόλο που σε κάθε παρτίδα δεν ήταν δυνατό να εντοπιστεί μία στατιστικά σημαντική διαφορά, δεν ισχύει το ίδιο για το εσωτερικό κάποιων οικογενειών. (Πίνακες 6 και 7). Όπως αναφέρθηκε συγκρίναμε τα μέλη 15 κοινών οικογενειών που προέρχονταν και από τις δύο παρτίδες. Το ζήτημα της σύγκρισης δειγμάτων με άνισα μεγέθη, όπως στην περίπτωση μας, μικρά και μεγάλα ψάρια, λύνεται με την προσέγγιση του two tail t-test, το οποίο δίνει την δυνατότητα της σύγκρισης δειγμάτων με ανόμοιο μέγεθος. Με αυτή την στατιστικά σημαντική φαινοτυπική ανομοιογένεια στο εσωτερικό των κοινών οικογενειών, δίνεται και μία έμμεση ιδέα για το ποιοί γονείς, πιθανά, να εμφανίζουν και πιο έντονη γενετική διακύμανση για τον εξεταζόμενο χαρακτήρα.

Πίνακας 6. t-test για τρεις κοινές οικογένειες μεταξύ των παρτίδων 1 και 2 που έδειξαν στατιστικά σημαντική διαφορά για το βάρος. Οι παρατηρήσεις αντιπροσωπεύουν τον αριθμό των απογόνων σε κάθε παρτίδα για μια οικογένεια.

		Βάρος					
Οικογένεια		3F/2M		4F/6M		1F/4M	
Παρτίδα		1	2	1	2	1	2
Μέσος		2,366	2,497	2,379	2,143	3,178	2,806
Παρατηρήσεις		77	76	95	67	8	17
P		0,04998		0,00045		0,01386	

Πίνακας 7. t-test για δύο κοινές οικογένειες που έδειξαν στατιστικά σημαντικές διαφορές για το μήκος.

		Μήκος			
Οικογένεια		4F-6M		1F-4M	
Παρτίδα		1	2	1	2
Μέσος		63,01	61,44	70,06	65,97
Παρατηρήσεις		95	67	8	17
P		0,01552		0,00448	

3.4 Συζήτηση

Σε αυτή τη μελέτη, αξιολογείται για πρώτη φορά στον κρανίο (*Argyrosomus regius*) η σχετική συμβολή κάθε γεννήτορα στο σχηματισμό των οικογενειών της F1 γενιάς. Τα ποσοστά ανάθεσης διαφέρουν ελαφρώς μεταξύ των δύο παρτίδων, αποτελούμενη κάθε μία από 400 ιχθύες, που εξετάστηκαν, κυμαινόμενες από 87,5% έως 95,0% και υπολογιζόμενο στο 91,1% και στις δύο παρτίδες (728 ιχθύες ανατέθηκαν επιτυχώς). Η χρήση γονότυπων αποτελούμενων από πολλαπλούς γενετικούς τόπους για να εξακριβωθεί η γενεαλογία είναι μια ευρέως χρησιμοποιούμενη μέθοδος (Milner and Brotherstone, 2000; Garant and Kruuk, 2005). Ένα μεγάλο μειονέκτημα της μεθόδου αποκλεισμού είναι ότι μία μόνο αναντιστοιχία αρκεί για να αποκλείσει έναν υποτιθέμενο γονέα (Jones and Ardren, 2003). Ένας κρίσιμος παράγοντας για την απόκτηση υψηλών επιπέδων μοναδικών αναθέσεων είναι η αποτελεσματικότητα ή ισχύς του γενετικών δεικτών που χρησιμοποιείται ως εργαλείο. Οι μικροδορυφορικοί δείκτες λόγω του πολυμορφισμού και της μεταβλητότητάς τους επιτρέπουν την αποτελεσματική ανάθεση γονέων (Perez-Enriquez et al., 1999; Waldbieser and Wolters, 1999; Boudry et al., 2002; Brown et al., 2005; Fessehaye et al. 2006, Herlin et al., 2007; Wang et al., 2008). Η ανάπτυξη ενός πάνελ δεικτών για την διεξαγωγή ανάλυσης ανάθεσης γονέων, είναι σημαντική για την εκτίμηση της αποτελεσματικότητας της ανάλυσης από τα αρχικά δεδομένα. Οι συχνότητες των αλληλομόρφων χρησιμοποιούνται για τον υπολογισμό των πιθανοτήτων αποκλεισμού και της ποσοτικοποίησης των μεμονωμένων δεικτών. Το Vitassign χρησιμοποιεί την μέθοδο της προσομοίωσης για να προσδιορίσει την ισχύ ανάθεσης ολόκληρου του συνόλου δεικτών, με την προϋπόθεση της ύπαρξης των γονότυπων των γονέων, χρησιμοποιώντας το σχέδιο ζευγαρώματος που δηλώνεται. Το αποτέλεσμα είναι το αναμενόμενο ποσοστό μοναδικών αναθέσεων ελλείψει σφαλμάτων γονοτύπισης. Το ποσοστό ανάθεσης (μ.ο. 91,1% και για τις δύο παρτίδες) είναι συγκρίσιμο με εκείνο που έχει ήδη δείχθει σε μέλη της οικογένειας Sciaenidae. Στο *Larimichthys crocea*, η ανάθεση χρησιμοποιώντας 6-7 μικροδορυφορικούς τόπους ήταν 99,6% (Liu et al., 2012), ενώ για απογόνους *Argyrosomus japonicus*, ήταν 84,7% χρησιμοποιώντας 14 μικροδορυφορικούς τόπους (Mirimin & Roodt-Wilding, 2015).

Ο θεωρητικός (μέγιστος) αριθμός των οικογενειών, για τις δύο παρτίδες, είναι 78 με τις οικογένειες των απογόνων να είναι 20. Στις δύο παρτίδες, η συμβολή των υποτιθέμενων γονέων δείχνει μικρή και μη-στατιστικά σημαντική διακύμανση. Η κατανομή των οικογενειών στον πληθυσμό των απογόνων της παρτίδας 1 δείχνει ότι ορισμένοι γονείς έχουν πιθανώς αυξημένη συνεισφορά σε σύγκριση με τους υπόλοιπους. Στο **Σχήμα 1**, δείχνεται ότι οι διασταυρώσεις 1F-1M, 1F-5M, 1F-6M, 4F-1M και 4F-6M δείχνουν ανομοιογένεια. Μεταξύ αυτών των γονέων, οι 1F και 3F από τους θηλυκούς γεννήτορες και οι 4M, 5M και 6M από τους αρσενικούς παίζουν έναν πιθανώς κυρίαρχο ρόλο με βάση τις παρατηρούμενες συνεισφορές και τις οικογένειες.

Συνολικά, οι πιο κυρίαρχοι θηλυκοί ήταν ο 4F με 37% στην παρτίδα 1 και ο 1F με 35,53% στην παρτίδα 2, καθώς επίσης ο αρσενικός 6M ήταν εξίσου κυρίαρχος και στις δύο παρτίδες με τους αρσενικούς 5M και 4M να δείχνουν κάποια διακύμανση. Μόνο τρεις εκ των γεννητόρων στους οποίους έγινε ένεση με GnRHα δεν έδωσαν κανένα απόγονο, ένας θηλυκός και δύο αρσενικοί. Αυτό μπορεί να οφείλεται σε ενδεχόμενο δειγματοληπτικό σφάλμα, καθώς άλλοι γεννήτορες, όπως οι 13M και 2F, φαίνεται να έχουν έναν απόγονο στις δύο παρτίδες που αναλύθηκαν. Παραδόξως, ένας μόνο αρσενικός στον οποίο δεν έγινε ένεση, ήταν υπεύθυνος για το 11 και το 13% των απογόνων στις δύο παρτίδες αντίστοιχα. Τα t-τέστ που πραγματοποιήθηκαν μεταξύ των κοινών οικογενειών των δύο παρτίδων έδειξαν ότι υπάρχει κάποια στατιστικά σημαντική διακύμανση στις οικογένειες. Τρεις οικογένειες εμφάνισαν τέτοιες στατιστικά σημαντικές διαφορές στο βάρος (πίνακες 6 και 7). Αυτές οι οικογένειες περιλαμβάνουν τους γεννήτορες 1F, 3F, 4F και 2M, 4M, 6M. Ωστόσο, η διακύμανση του μεγέθους και η ανάπτυξη γενικά, παρατηρείται πολύ συχνά μεταξύ των ιχθύων της ίδιας ηλικίας σε πολλά είδη. Αυτή η μεταβλητότητα στην ανάπτυξη μπορεί να είναι ένα σημαντικό εμπόδιο στη μακροπρόθεσμη βιωσιμότητα μιας εμπορικής εκμετάλλευσης στην ιχθυοκαλλιέργεια. Η διαλογή των ιχθυδίων με βάση το μέγεθος βελτιστοποιεί την παραγωγή περιορίζοντας τον κανιβαλισμό, με την μείωση της διακύμανσης του μεγέθους μεταξύ των συλλεχθέντων ιχθυδίων και βελτιώνοντας την αποτελεσματικότητα της μετατρεψιμότητας των ζωοτροφών σε βιομάζα (Loughnan et al., 2013). Η τοποθέτηση ιχθύων σε δεξαμενές οι οποίοι δεν έχουν υποστεί διαλογή μπορεί να οδηγήσει σε υπο-ανάπτυξη του αποθέματος και σε κακές διατροφικές συνθήκες (Ghozlan et al., 2018). Οι πρακτικές της επιλογής (δηλαδή τα προγράμματα αναπαραγωγής) θα μπορούσαν να χρησιμοποιηθούν για τη μείωση της διακύμανσης στην ανάπτυξη και τη βελτίωση της μακροπρόθεσμης ανάπτυξης, την ομοιομορφία των προϊόντων και τη μείωση του κόστους διαχείρισης.

Η παρούσα μελέτη είναι η πρώτη προσπάθεια να προσδιορίσουμε στον κρανίο, το κατά πόσο το μέγεθος και η μεταβλητότητα της ανάπτυξης έχει ισχυρή γενετική αιτίαση, η οποία με τη σειρά της μπορεί να παρέχει τις πρώτες ενδείξεις για τη υλοποίηση ενός προγράμματος επιλεκτικής αναπαραγωγής. Ωστόσο, δεδομένου ότι οι δύο παρτίδες περιέχουν άτομα που προέρχονται από μία γέννα, και συλλέχθηκαν από δύο κλουβιά σε δύο διαφορετικά χρονικά σημεία (το πρώτο τον Ιανουάριο του 2016 και το δεύτερο τον Μάιο του 2016), παρουσιάστηκαν δύο χρονικά σημεία στην συγκεντρωτική ανάλυση και των δύο παρτίδων ως σταθερή επίδραση. Όπως επισημάνθηκε, λόγω της διαδικασίας διαλογής, οι φαινοτυπικοί χαρακτήρες των δύο παρτίδων δεν μετρήθηκαν ταυτόχρονα, αλλά όταν τα άτομα έφθασαν στο εμπορικό τους βάρος. Λαμβάνοντας υπόψη αυτό το γεγονός και τον αριθμό των απογόνων, είναι πιθανό να υπάρχει μία υπερεκτίμηση των τιμών, όσον αφορά την κληρονομησιμότητα και τις εκτιμήσεις γενετικών και φαινοτυπικών συσχετίσεων βάρους και μήκους. Ωστόσο, ο υπολογισμός και η αξιολόγηση αυτών των τιμών δεν μπορεί να

επηρεαστεί δραματικά, παρέχοντας έτσι χρήσιμες ενδείξεις. Οι τιμές κληρονομησιμότητας των δύο παρτίδων συγκεντρωτικά για το βάρος και το μήκος (0,62 για το βάρος και 0,64 για το μήκος), δείχνουν ότι είναι ιδιαίτερα πιθανό τα συγκεκριμένα χαρακτηριστικά να συσχετίζονται στενά και ότι το περιβάλλον παίζει, δυνητικά, λιγότερο σημαντικό ρόλο. Αυτό ενισχύεται από το γεγονός της ανάπτυξης των ιχθύων ουσιαστικά στο ίδιο ή σε πολύ παρόμοιο περιβάλλον. Επιπλέον, η αναμενόμενη υψηλή γενετική συσχέτιση (96%) μεταξύ αυτών των δύο χαρακτήρων υποδηλώνει ότι υπάρχει μια σχέση μεταξύ βάρους και μήκους στον κρανίο, η οποία φαίνεται να ισχύει και για άλλα είδη (Chatziplis et al., 2007; Fernandes et al., 2017; Rutten et al., 2005). Επίσης, η χωριστή εκτίμηση των κληρονομησιμότητων για το βάρος και το μήκος στις δύο παρτίδες, σκιαγραφεί τις διαφορές ανάμεσα στις δύο παρτίδες. Η πρώτη παρτίδα εμφανίζει τον υψηλότερο ρυθμό ανάπτυξης, καθώς δείχνει υψηλότερες τιμές κληρονομησιμότητας βάρους και μήκους (0,76 και 0,72) σε σύγκριση με 0,64 και 0,69, αντίστοιχα, για τη δεύτερη παρτίδα 2. Ο πληθυσμός που είχε τον υψηλότερο ρυθμό ανάπτυξης δείχνει επίσης υψηλότερες τιμές κληρονομησιμότητας από το πληθυσμό με χαμηλότερο ρυθμό ανάπτυξης. Οι συνθήκες υπό τις οποίες γίνεται η επιλογή ενδέχεται να επηρεάσουν την απόκριση σε αυτήν την επιλογή. Τα εμπορικά και μη εμπορικά περιβάλλοντα είναι διαφορετικά. Ως εκ τούτου, το εύρος των γενετικών αλληλεπιδράσεων με το περιβάλλον μπορεί να διαφέρουν μεταξύ διαφορετικών περιβαλλόντων (Lynch και Walsh, 1998). Στην τιλάπια π.χ., δεν παρατηρήθηκαν σημαντικές διαφορές στις εκτιμήσεις κληρονομησιμότητας μεταξύ του εμπορικού και του πειραματικού περιβάλλοντος (Khaw et al., 2009) ενώ, διαφορές, αν και χαμηλές, παρατηρήθηκαν σε μια μελέτη για το λαβράκι (Dupont-Nivet et al., 2008). Όταν η διαλογή εφαρμόζεται σε αναλύσεις γενετικών παραμέτρων για 2 ή πολλαπλούς χαρακτήρες, μια στρατηγική είναι η δημιουργία κλάσεων για το χαρακτηριστικό στο οποίο βασίζεται η διαλογή. Ο διαχωρισμός σε ξεχωριστές μικρότερες ομάδες συσχετίζεται με μια υποτίμηση στις τιμές κληρονομησιμότητας (Blonk et al., 2010). Στην παρούσα μελέτη, ο διαχωρισμός σε 2 κλάσεις (παρτίδες) οδηγεί σε δραστική μείωση των εκτιμήσεων κληρονομησιμότητας. Ωστόσο, οι γενετικές συσχετίσεις μεταξύ των ίδιων χαρακτήρων σε διαφορετικές παρτίδες, καθώς και οι συσχετίσεις υψηλής κατάταξης (0,95) μεταξύ των εκτιμώμενων κληροδοτικών τιμών των ίδιων χαρακτήρων, δείχνουν ότι αναλύεται ο ίδιος χαρακτήρας. Από την άλλη πλευρά, ο συγγών παράγοντας του κλωβού (*confounding factor*), η ετερογένεια του μήκους στους απογόνους, έχουν δυνητικά συμβάλλει στο πρόβλημα, ενώ θα μπορούσαν επίσης να οδηγήσουν σε υπερεκτίμηση των τιμών των γενετικών παραμέτρων. Μία ενδεχόμενη λύση θα μπορούσε να είναι η αύξηση του μεγέθους του συνόλου δεδομένων. Τα σχετικά υψηλά τυπικά σφάλματα που παρατηρούνται (Πίνακας 5) σε σχέση με άλλες μελέτες (Gjedrem, 2000) υπογραμμίζουν αυτές τις παρατηρήσεις.

Ένα από τα προβλήματα στα προγράμματα επιλογής ιχθύων που βασίζονται στην μαζική γονιμοποίηση είναι η άνιση συνεισφορά των γαμετών και η μη τυχαία κατανομή των οικογενειακών μεγεθών. Δυστυχώς, η άνιση κατανομή στα οικογενειακά μεγέθη που προέρχεται από ένα προϋπάρχον και δεδομένο σχήμα γεννητόρων έχει αρνητικές επιπτώσεις στο δραστικό μέγεθος του πληθυσμού (N_e). Σε έναν εξιδανικευμένο πληθυσμό, κάθε γονέας έχει την ίδια πιθανότητα με τους άλλους στην συνεισφορά γονιδίων ή απογόνων, και αυτή η συνεισφορά κατανέμεται μεταξύ των γονέων με την επακόλουθη αποτύπωση στα οικογενειακά μεγέθη (Falconer και Mackay, 1996). Ωστόσο, στην πραγματικότητα, οι γονείς δεν έχουν ίσες πιθανότητες συνεισφοράς λόγω διαφορών στη γονιμότητα, στην επιβίωση των απογόνων κ.λπ. Η διακύμανση μεταξύ της συνεισφοράς των γονέων οδηγεί σε μεταβολή των μεγεθών των οικογενειών και αυτό οδηγεί σε μια επόμενη γενιά που πηγάζει από μικρότερο από τον επιθυμητό, αριθμό γονέων. Οι αναλογίες που ελήφθησαν (βλ. Πίνακα 2) δείχνουν ότι οι τιμές του N_e ήταν σημαντικά χαμηλότερες από το μέγεθος του πληθυσμού

αναφοράς που μπορεί να αποδοθεί στην αυξημένη διακύμανση των μεγεθών των οικογενειών (μέγεθος $V_{family} = 2351$ και $\mu = 42,6$), καθώς και στην μειωμένη συνεισφορά των αρσενικών σε σχέση με τους θηλυκούς γεννήτορες. Η παρούσα μελέτη είναι μια πρώτη ανάλυση της ποσοτικοποίησης του μεγέθους της γενετικής προέλευσης των διαφορών σε χαρακτηριστές ανάπτυξης στον κρανίο καθώς επίσης να παράσχει τις πρώτες ενδείξεις σχετικά με την αναζήτηση QTL.

Ο κρανιός είναι ένα σχετικά νέο είδος υδατοκαλλιέργειας με μεγάλες δυνατότητες στην ευρωπαϊκή υδατοκαλλιέργεια μεγάλης κλίμακας. Ο πρωταρχικός στόχος της μελέτης ήταν να περιγράψει, για πρώτη φορά, την κατανομή συνεισφερόμενων γονέων και την ανάθεση απογόνων στους γονείς τους σε ένα σύστημα παραγωγής βιομηχανικής κλίμακας. Ωστόσο, λόγω της διαφορετικής ανάπτυξης κατά τη διάρκεια των αρχικών σταδίων ανάπτυξης, έγινε διαλογή των ιχθύων σε δύο ομάδες, μια ομάδα μεγαλύτερων νεαρών που μεταφέρθηκε σε έναν κλωβό (παρτίδα 1), και μια ομάδα μικρότερων νεαρών που μεταφέρθηκαν σε δεύτερο κλωβό (παρτίδα 2). Το συνολικό μήκος και το βάρος μετρήθηκαν για όλα τα άτομα που γονοτυπήθηκαν με ένα πάνελ 10 μικροδορυφορικών τόπων ώστε να ανατεθούν στους πιθανούς γονείς τους. Το ποσοστό ανάθεσης ήταν υψηλό (87,5% για την παρτίδα 1 και 95% για την παρτίδα 2), με τον αριθμό των συνολικών οικογενειών να ανέρχεται στις 20. Οι πιθανοί γονείς των απογόνων περιείχαν περίπου τους μισούς γεννήτορες, πέντε θηλυκούς και επτά αρσενικούς. Μεταξύ των δύο κλωβών, παρουσιάστηκε διαφορά στην σύνθεση των κοινών οικογενειών. Οι δεκαπέντε κοινές οικογένειες στις οποίες ανατέθηκαν οι απόγονοι, αναλύθηκαν για στατιστικά σημαντικές διαφορές σχετικά με το σωματικό βάρος και το συνολικό μήκος του σώματος, διαφορές που παρατηρήθηκαν σε 3 οικογένειες και στις δύο παρτίδες. Εκτιμήθηκε η κληρονομικότητα για το σωματικό βάρος και το συνολικό μήκος του σώματος, καθώς και οι γενετικές και φαινοτυπικές συσχετίσεις των δύο χαρακτηριστικών. Η παρτίδα 1 έδειξε υψηλότερες εκτιμήσεις κληρονομησιμότητας από την παρτίδα 2 με τις γενετικές και φαινοτυπικές συσχετίσεις να είναι σχεδόν ίδιες και για τις δύο παρτίδες. Ορισμένοι γονείς συνεισέφεραν περισσότερους απογόνους και επέδειξαν κυριαρχία στην γονιμοποίηση. Ομοίως, τα χαρακτηριστικά ανάπτυξης του σωματικού βάρους και το συνολικό μήκος του σώματος των κυρίαρχων γονέων συσχετίζονται, υποθετικά, με τις στατιστικά σημαντικές διαφορές που παρατηρούνται σε αυτές τις τρεις οικογένειες.

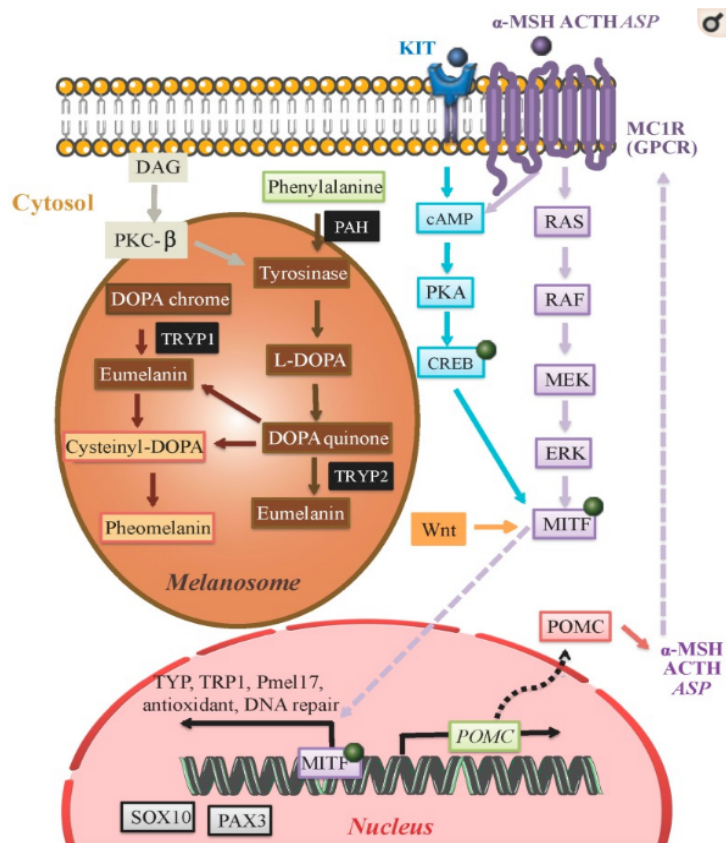
4. Γενετικός χάρτης σύνδεσης, ανάλυση ανίχνευσης QTL για τους χαρακτήρες βάρος και μήκος, και συγκριτική γονιδιοματική ανάλυση με 5 γονιδιώματα τελεόστεων ιχθύων, στον κρανίο

4.1 Εισαγωγή

Η ανάπτυξη είναι μια περίπλοκη διαδικασία όσο αναφορά την φυσιολογία της η οποία επηρεάζεται από γενετικούς και περιβαλλοντικούς παράγοντες. Ο ρυθμός ανάπτυξης είναι ο κύριος στόχος των προγραμμάτων αναπαραγωγής στην ιχθυοκαλλιέργεια και την κτηνοτροφία γενικότερα για δύο λόγους: (1) μείωση του κόστους παραγωγής και, (2) η μέτρια/υψηλή κληρονομησιμότητα που επιτρέπει σημαντικό γενετικό κέρδος μέσω της κλασικής επιλογής (Thodesen et al., 2011; Gjedrem, 2016; Bentsen et al., 2017). Όμως, η αναγνώριση σημαντικών γονιδιοματικών επιδράσεων στην ανάπτυξη μπορεί να είναι χρήσιμη για μία επιλογή υποβοηθούμενη από γονιδιοματικούς δείκτες (MAS, marker assisted selection), ειδικά στα είδη τα οποία εμφανίζουν μεγάλη περίοδο αναπαραγωγικής ωρίμανσης (Salem et al., 2012), όπως το καλκάνι (~3-4 χρόνια) (Janssen et al., 2017), για την εύρεση δύσκολα εντοπίσιμων γενετικών παραλλαγών που διαφορετικά θα μπορούσαν να χαθούν χρησιμοποιώντας την κλασική επιλογή (Ruane and Sonnino, 2007). Πολλά QTL (Γονιδιακοί τόποι ποσοτικών ιδιοτήτων) που σχετίζονται με την ανάπτυξη μπορεί να επηρεάζουν μικρό ως μέτριο ποσοστό της συνολικής φαινοτυπικής διακύμανσης (V_p) (~10-40%), που έχει εντοπιστεί στους ιχθείς, ποσοστά τα οποία δείχνουν την άνιση επίδραση των σχετιζόμενων γονιδίων στο γονιδίωμα. Οι Wringe et al. (2010), εντόπισαν 10 QTL για το σωματικό βάρος (BW) και 6 QTL για τον παράγοντα Fulton (το βάρος είναι ανάλογο του μήκους³) σε 9 οικογένειες πέστροφας (*Oncorhynchus mykiss*). Παρόμοια ευρήματα προέκυψαν στο είδος *Salvelinus alpinus*, για 10 QTL σχετιζόμενα με το BW και 9 QTL για τον παράγοντα Fulton, εξηγώντας ένα εύρος της φαινοτυπικής διακύμανσης από 10 έως 34% (Moghadam et al., 2007). Ομοίως στο ασιατικό λαβράκι (*Lates calcarifer*), 6 BW και BL (σωματικού μήκους) QTL σε 6 LG (linkage groups) για το 8-16% της V_p (Wang et al., 2015), για το *Carassius auratus* 8 BW-QTL για 10-13% της V_p (Liu et al., 2017), και για ένα είδος κυπρίνου *Hypophthalmichthys nobilis* 38 QTL σε έξι LG, για πέντε χαρακτήρες ανάπτυξης για το 15-38% της V_p (Fu et al., 2016). Μεταξύ των πλατύψαρων, 3 BW-QTL βρέθηκαν στο είδος γλώσσας *Cynoglossus semilaevis* και στο *Paralichthys olivaceous*, αντίστοιχα, εξηγώντας από 11% έως 26% της V_p (Song et al., 2012a, 2012b). Μέσω χαρτογράφησης υψηλής πυκνότητας εντοπίστηκαν πολλαπλά QTL για χαρακτήρες ανάπτυξης σε 12 LG σε μία οικογένεια καλκανιού (Wang et al., 2015b). Μέσω της ανάλυσης GWAS (γονιδιοματική ανάλυση συσχέτισης) έγινε δυνατή μια πιο πληροφοριακή γνώση σχετικά με αναπτυξιακούς χαρακτήρες στους ιχθείς (Peng et al., 2016; Li et al., 2017), καθώς επίσης αναδείχθηκαν συγκεκριμένες αλληλικές παραλλαγές οι οποίες μένει να επικυρωθούν περαιτέρω μέσω λειτουργικών προσεγγίσεων, συγκριτικής γονιδιοματικής, και εκτίμηση παραμέτρων πληθυσμιακής γενετικής (Fu et al., 2016; Tsai et al., 2016; Ashton et al., 2017; Liu et al., 2017). Παρά το γεγονός ότι η επιλογή βασιζόμενη σε γονιδιοματικούς δείκτες έχει εφαρμοστεί κυρίως σε «δύσκολους» χαρακτήρες στην ιχθυοκαλλιέργεια, όπως η ανθεκτικότητα σε ασθένειες (Ozaki et al., 2013), η πληροφορία που έχει συγκεντρωθεί σχετικά με την γενετική αρχιτεκτονική των χαρακτήρων που σχετίζονται με την ανάπτυξη πρέπει να ενθαρρύνει την εφαρμογή της MAS. Ωστόσο, η MAS απαιτεί την παρουσία δεικτών, ιδανικά, σε πολλές οικογένειες για την ουσιαστική εκπροσώπηση ενός είδους, κάτι

που δεν έχει υλοποιηθεί στον βέλτιστο βαθμό μέχρι σήμερα σε είδη υδατοκαλλιέργειας (Tsai et al., 2015).

Οι προσπάθειες για τον χαρακτηρισμό της γενετικής παραλλακτικότητας που επηρεάζει τους φαινοτυπικούς χαρακτήρες είχαν ιστορικά ανάμεικτα αποτελέσματα, με μεγάλο μέρος της κληρονομησιμότητας για σύνθετους χαρακτήρες να μην εξηγείται πάντα από γνωστές αιτιώδεις παραλλαγές (συχνά αναφέρεται ως το πρόβλημα της αγνοούμενης κληρονομησιμότητας, (missing heritability) (Benfey και Mitchell-Olds, 2008; Manolio et al., 2009). Ένα παράδειγμα το οποίο επισημαίνει το πρόβλημα και τις λύσεις σε αυτό, είναι το ύψος του ανθρώπινου σώματος (Yang et al., 2010, 2015). Μέχρι πρόσφατα, αυτός ο χαρακτήρας εμφάνιζε κληρονομησιμότητα 0,8, αλλά για την οποία μόνο το 0,45 εξηγείται σε μια από τις πιο ενδελεχείς μελέτες (Yang et al., 2010), και μόνο το 0,05 από όλες τις προηγούμενες μελέτες συνδυαστικά (Yang et al., 2010). Επιπρόσθετα, πραγματοποιήθηκε μια ακόμη μελέτη που έδειξε ότι η αγνοούμενη κληρονομησιμότητα είναι αμελητέα μετά από εξέταση της πιθανής υπερεκτίμησης της μεταξύ οικογενειών (η κληρονομησιμότητα ήταν πιο κοντά στο 0,6-0,7), και χρησιμοποιώντας μέγεθος δείγματος ($n = 44.126$), το οποίο είναι αρκετά μεγάλο για να αποτυπώσει την διακύμανση σπάνιων παραλλαγών (Yang et al., 2015). Σε μελέτες με μικρότερα μεγέθη δείγματος, οι αιτιώδεις παραλλαγές, οι οποίες αναγνωρίζονται επιτυχώς, συχνά εξηγούν μεγάλο ποσοστό της V_p στον διερευνούμενο χαρακτήρα, αλλά πιθανώς να μην είναι αντιπροσωπευτικές των υπόλοιπων αιτιωδών παραλλαγών, οι οποίες είναι πιθανό να έχουν μικρή αθροιστική επίδραση στον χαρακτήρα και κατά συνέπεια είναι δύσκολο να εντοπιστούν (Rockman, 2012). **Στο Σχήμα 1** βλέπουμε ένα παράδειγμα για έναν χαρακτήρα με πολυγονιδιακή βάση.



Σχήμα 1. Ένα παράδειγμα σύνθετου χαρακτήρα με πολυγονιδιακή βάση. Το χρώμα δέρματος, που καθορίζεται σε μεγάλο βαθμό από το μοριακό μονοπάτι της μελανίνης, μοιράζεται μεταξύ ενός ευρέος φάσματος ειδών (ιχθύες, τρωκτικά, άνθρωποι). Η ύπαρξη πολλών αλληλεπιδρώντων συστατικών απαιτούν πολυγονιδιακές μεθόδους ανάλυσης για την πλήρη κατανόηση αυτών των πολύπλοκων χαρακτήρων (D'Mello et al., 2016).

Ορισμένες από τις λύσεις για την αγνοούμενη κληρονομησιμότητα περιλαμβάνουν αυξημένο μέγεθος δείγματος και πυκνότητα δεικτών, χρησιμοποιώντας την διακύμανση μεταξύ των οικογενειών για τη διερεύνηση σπάνιων παραλλαγών χαμηλής συχνότητας, την βελτίωση της ακρίβειας των εκτιμήσεων κληρονομησιμότητας, και την χρήση πολυγονιδιακών μοντέλων ή διαγονιδιακών επιδράσεων για τη βελτίωση της ισχύς (Manolio et al., 2009; Yang et al., 2010; Rockman, 2012; Yang et al., 2015). Η αύξηση του μεγέθους δείγματος μπορεί να αυξήσει την ισχύ στην διερεύνηση αλληλομόρφων χαμηλής συχνότητας. Ωστόσο, όπως σημείωσε ο Gibson (2012), κάτω από ορισμένα μοντέλα (π.χ. όπου εκατοντάδες γενετικές παραλλαγές επηρεάζουν έναν χαρακτήρα) το μέγεθος του δείγματος, ενδεχομένως, θα πρέπει να είναι μεγαλύτερο από τον πληθυσμό ώστε να ανιχνεύσει τις σπάνιες αιτιώδεις παραλλαγές. Ακόμη, η αυξημένη πυκνότητα δεικτών θα μπορούσε να αυξήσει την ανίχνευση αιτιωδών παραλλαγών με την άυξηση της κάλυψης του γονιδιώματος. Η χρήση της φαινοτυπικής διακύμανσης μεταξύ των οικογενειών θα μπορούσε να βελτιώσει την ισχύ όταν οι αιτιώδεις παραλλαγές με χαμηλή συχνότητα στον ευρύτερο πληθυσμό έχουν μεγαλύτερη συχνότητα εντός της εστιακής οικογένειας. Τέλος, η εφαρμογή των πολυγονιδιακών μεθόδων μπορεί να βοηθήσει στον εντοπισμό των αιτιωδών αλληλομόρφων λαμβάνοντας υπόψη τις διαγονιδιακές επιδράσεις (Wellenreuther and Hansson, 2016). Ωστόσο, τα μεγέθη δείγματος που απαιτούνται για τις πολυγονιδιακές μεθόδους μπορεί να είναι μεγάλα (π.χ. χιλιάδων ατόμων). Η εφαρμογή αυτών των προσεγγίσεων πιθανότατα να

αποτελεί μέρος της λύσης της αντιμετώπισης του πρόβληματος της αγνοούμενης κληρονομησιμότητας.

4.2 Γενετικός χάρτης σύνδεσης

Ο ανασυνδυασμός και η ανισορροπία σύνδεσης είναι δύο αναπόσπαστα συνδεδεμένες όψεις των διεργασιών που οδηγούν τον σχηματισμό απλοτύπων, πάνω στους οποίους μπορεί να δράσει η φυσική επιλογή (Gessler και Xu 2000, Otto και Lenormand 2002, Posada et al., 2002). Η αξιολόγηση της έκτασης του ανασυνδυασμού και της σύνδεσης στο γονιδίωμα ενός οργανισμού είναι βασικά στοιχεία για την κατανόηση των δομικών, λειτουργικών και εξελικτικών χαρακτηριστικών του γονιδιώματος (Wang et al., 2009, Hohenlohe et al., 2011). Ο γενετικός χάρτης σύνδεσης δεν παρέχει μόνο τη σχετική σειρά των δεικτών, αλλά και ένα άμεσο μέτρο της έκτασης του ανασυνδυασμού και της ανισορροπίας σύνδεσης μεταξύ των χρωμοσωμάτων. Σε σεξουαλικά αναπαραγωγικούς οργανισμούς, επιτρέπει επίσης την αναγνώριση του ρόλου κάθε φύλου στη δημιουργία νέων απλοτύπων (Broman et al., 1998, Lenormand 2003). Από δομική γονιδιωματική άποψη, ένας χάρτης σύνδεσης παρέχει τα απαραίτητα δεδομένα για την ανάλυση της παρουσίας, της τοποθεσίας, και του σχετικού μεγέθους πιθανών χρωμοσωμικών αναδιατάξεων, π.χ αναστροφών (Ganal et al., 1992, Bansal et al., 2007). Ως εκ τούτου, οι γενετικοί χάρτες σύνδεσης μπορούν να διευκολύνουν την de novo συναρμολόγηση ενός γονιδιώματος επιτρέποντας την αναγνώριση πιθανών χμαιοικών κριωμάτων (scaffolds) (Rastas et al., 2013, Fierst 2015). Επιπλέον, επειδή το μήκος των απλοτύπων και η παρουσία τους, επηρεάζονται έντονα από την φυσική επιλογή, είναι δυνατές οι συγκρίσεις των επιπέδων σύνδεσης, που περιβάλλουν τους γενετικούς τόπους, σε πολλαπλούς πληθυσμούς. Αυτές οι συγκρίσεις ενδέχεται να παρέχουν στοιχεία για την κατανόηση της εξελικτικής ιστορίας του υπό εξέταση γενετικού τόπου (Birky και Walsh 1988, Kreitman and Hudson 1991). Εκτός από την παροχή πληροφορίας για τα γονιδιώματα και τους πληθυσμούς των μελετούμενων ειδών, χάρτες γενετικής σύνδεσης υψηλής πυκνότητας από οργανισμούς μη-μοντέλα μπορούν να αποτελέσουν ισχυρή βάση για την συγκριτική γονιδιωματική ανάλυση και την εύρεση συντενιών μεταξύ των ειδών, παρέχοντας σημαντικά στοιχεία για την κατανόηση της εξέλιξης του γονιδιώματος (Kulathinal et al., 2009, Lewin et al., 2009, Michalak de Jimenez et al., 2013). Ακόμη, αυτοί οι χάρτες επιτρέπουν την σύνδεση φαινότυπων με συγκεκριμένες περιοχές των γονότυπων (μέσω ανάλυσης χαρτογράφησης QTL ή γονιδιωματικής συσχέτισης (GWAS), και την κατανόηση της γενετικής βάσης σύνθετων φαινοτυπικών χαρακτήρων (Paterson et al., 1988, Flint and Mackay 2009, Goddard and Hayes 2009). Η ανάπτυξη τεχνολογιών αλληλούχισης υψηλής απόδοσης, που σχετίζονται με γονιδιωματικές αλληλουχίες αναγνώρισης από περιοριστικές ενδονουκλεάσες (RADtags, Miller et al., 2007), επέτρεψαν με οικονομικά αποδοτικό τρόπο την ανίχνευση αρκετών χιλιάδων νουκλεοτιδικών πολυμορφισμών, SNP's, ως γενετικούς δείκτες σε γονιδιώματα μη-μοντέλων οργανισμών. Σε σύγκριση με τους μικροδορυφορικούς δείκτες, τα SNP's έχουν τη δυνατότητα της απλοποίησης της δημιουργίας χαρτών γενετικής σύνδεσης λόγω του ότι μπορεί να διαμορφώνουν, δυνητικά, γονότυπους με μεγαλύτερη ακρίβεια και κάλυψη γονιδιώματος από τους μικροδορυφόρους (Kruglyak 1997, Slate et al., 2009).

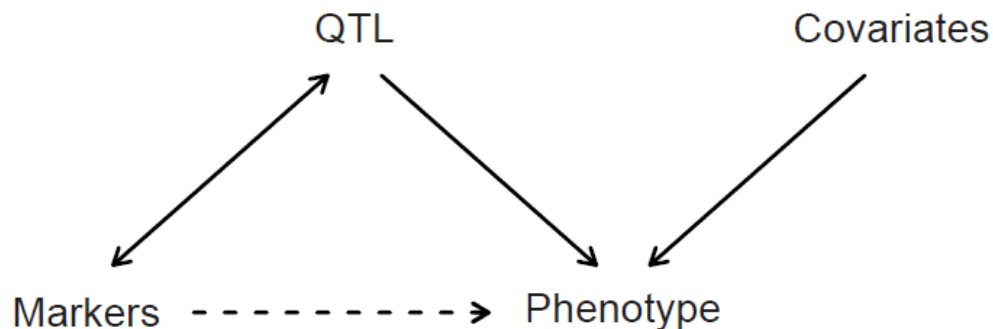
4.2.1 Κατασκευή γενετικού χάρτη σύνδεσης

Το λογισμικό Lep-MAP2 (Rastas et al., 2016) χρησιμοποιήθηκε για την κατασκευή των γενετικών χαρτών σύνδεσης σε αυτή τη μελέτη, το οποίο βασίζεται στην πρώτη έκδοση του Lep-MAP (Rastas et al., 2013). Είναι σε μεγάλο βαθμό αυτοματοποιημένο και απαιτεί

ελάχιστη αλληλεπίδραση με τον χρήστη ενώ μπορεί να αναλύσει ταυτόχρονα πολλαπλές οικογένειες και να χειριστεί όλους τους τύπους των δεδομένων από γενετικούς δείκτες (π.χ. SNPs, μικροδορυφόροι). Η ανάθεση των ομάδων σύνδεσης, (Linkage Groups) γίνεται με την επιλογή «Separate Chromosomes» και εκτελείται με ένα όριο LOD (logarithm of odds) καθορισμένο από τον χρήστη. Στη συνέχεια, οι μοναδικοί δείκτες προστίθενται στα LG χρησιμοποιώντας την επιλογή «JoinSingles» με όριο LOD καθορισμένο από τον χρήστη. Ορίζεται το επιτρεπτό όριο των ελλειματικών γονότυπων των γενετικών δεικτών και οι δείκτες που εμφανίζουν ελλείψεις μεγαλύτερες του ορίου αφαιρούνται από τα LG. Στη συνέχεια το λογισμικό φιλτράρει τους γενετικούς δείκτες συγκρίνοντας την κατανομή των γονότυπων των απογόνων και τις αναμενόμενες αναλογίες κατά Μέντελ μέσω μίας δοκιμής παραμόρφωσης του διαχωρισμού των αλληλομόρφων (segregation distortion test) χρησιμοποιώντας μία τιμή ανοχής (data tolerance threshold) που ως default είναι ίση με 0.01. Αυτό, σημαίνει ότι 1/100 δείκτες θα αφαιρεθούν τυχαία (χ^2 τεστ, $p < 0.01$) για την εκτίμηση της στατιστικής σημαντικότητας της δοκιμής «παραμόρφωσης». Η σειρά των δεικτών προσδιορίζεται επιτρέποντας διαφορετικές πιθανότητες ανασυνδυασμού στα δύο φύλα, και επιλέγονται οι βέλτιστες τιμές πιθανοτήτων εξ αυτών. Ο τελικός χάρτης περιλαμβάνει επίσης μια εκτίμηση για τα γονοτυπικά σφάλματα για κάθε δείκτη και οι παράμετροι σφάλματος αντιστοιχούν στο κρυφό μοντέλο Markov (HMM) το οποίο χρησιμοποιείται για τη μοντελοποίηση των ανασυνδυασμένων απλότυπων στο Lep-MAP2.

4.3 Ανάλυση QTL

Ο προσδιορισμός ενός QTL και η σχετική του θέση είναι εφικτά με την χαρτογράφηση του σε έναν γενετικό χάρτη. Αυτό γίνεται με την εκτίμηση μιας στατιστικά σημαντικής σχέσης μεταξύ του ποσοτικού φαινοτυπικού χαρακτήρα και ενός ή περισσότερων γενετικών δεικτών που βρίσκονται στον χάρτη, **Σχήμα 2**.



Σχήμα 2. Η στατιστική δομή του προβλήματος χαρτογράφησης ενός QTL. Το QTL και οι παράμετροι συνδιακύμανσης (covariates) ευθύνονται για την φαινοτυπική παραλλαγή (κατευθυνόμενα συνεχή βέλη). Οι γενετικοί δείκτες και το QTL συσχετίζονται μεταξύ τους λόγω γενετικής σύνδεσης (υποδεικνύεται από το αμφίδρομο συνεχές βέλος). Οι γενετικοί δείκτες δεν προκαλούν άμεσα τον φαινότυπο καθώς ορισμένοι δείκτες μπορεί να σχετίζονται με τον φαινότυπο μέσω της γενετικής σύνδεσης με το QTL (διακεκομμένο βέλος) και άλλοι όχι.

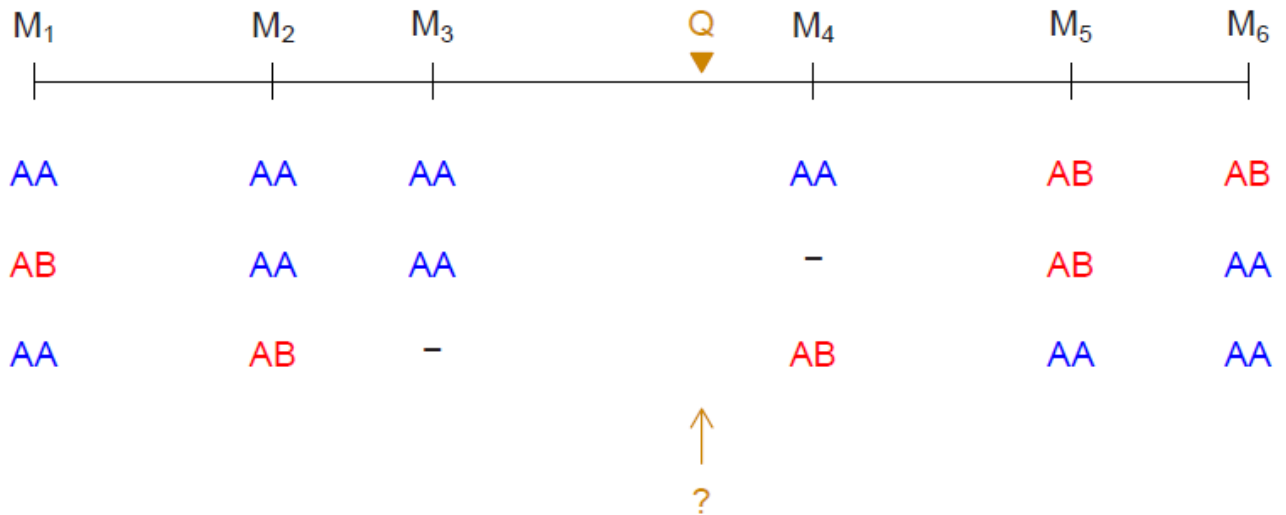
Όπως απεικονίζεται στο **Σχήμα 2**, ο φαινότυπος επηρεάζεται από τον γονότυπο για ένα QTL συνυπολογίζοντας τις πιθανές παραμέτρους συνδιακύμανσης, όπως το φύλο, ή οι περιβαλλοντικές επιδράσεις. Ωστόσο ο γονότυπος, δεν βρίσκεται στο QTL, παρά στους γενετικούς δείκτες. Οι γονότυποι των γενετικών δεικτών και το QTL σχετίζονται μέσω της

γενετικής σύνδεσης, η οποία οδηγεί σε συσχέτιση μεταξύ του φαινότυπου και των γονότυπων των γενετικών δεικτών. Αν θεωρητικά ήταν γνωστός ο γονότυπος κάθε ατόμου σε κάθε θέση στο γονιδίωμα, η χαρτογράφηση QTL θα ήταν απλώς η αναγνώριση του συνόλου των τόπων στο γονιδίωμα που παίζουν ρόλο στην παραγωγή του φαινότυπου, και του τρόπου με τον οποίο συνδυάζονται αυτές οι τοποθεσίες (μαζί με παραμέτρους συνδιακύμανσης) για την παραγωγή του φαινότυπου. Αυτό χαρακτηρίζεται ως το πρόβλημα της επιλογής μοντέλου. Ωστόσο, λόγω του ότι παρατηρούμε τους γονότυπους των ατόμων μόνο σε ένα υπο-σύνολο γενετικών δεικτών και την εξέταση των θέσεων μεταξύ αυτών των δεικτών ως πιθανές τοποθεσίες QTL, χρειάζεται να συμπεράνουμε τους γονότυπους σε παρεμβαλλόμενες τοποθεσίες από τους γονότυπους των γενετικών δεικτών. Αυτό είναι το πρόβλημα των αγνοούμενων δεδομένων. Υπάρχουν πολλές λύσεις στο πρόβλημα των αγνοούμενων δεδομένων, οι οποίες είναι όλες ικανοποιητικές όταν η πυκνότητα των γενετικών δεικτών στον γενετικό χάρτη είναι σχετικά μεγάλη. Όμως, το γενικό πρόβλημα της μεταβλητής επιλογής (variable selection) στην παλινδρόμηση (regression) είναι ένας ενεργός ερευνητικός τομέας για τον οποίο δεν υπάρχει μία γενικά αποδεκτή πρόταση.

4.3.1 Το πρόβλημα των αγνοούμενων δεδομένων στην ανάλυση QTL

Οι λύσεις για το πρόβλημα των αγνοούμενων δεδομένων βασίζονται σε μοντέλα ανασυνδυασμού. Αυτά τα μοντέλα βοηθούν να συνδεθούν, πιθανολογικά, μη παρατηρούμενοι και παρατηρούμενοι γονότυποι. Όλες οι προσεγγίσεις για την αντιμετώπιση των αγνοούμενων δεδομένων βασίζονται στον υπολογισμό των γονοτυπικών πιθανοτήτων για ένα QTL με βάση τα διαθέσιμα δεδομένα των γενετικών μοριακών δεικτών (SNPs εν προκειμένω). Για να γίνει αυτό, χρειάζεται ένα μοντέλο για τη διαδικασία του ανασυνδυασμού. Το πιο βολικό μοντέλο θεωρεί τις τοποθεσίες των ανασυνδυασμών κατά την μείωση να ακολουθούν την κατανομή Poisson. Άτυπα, αυτό σημαίνει ότι οι τοποθεσίες χρωμοσωμικών επιχιασμών είναι τυχαίες. Κάτω από το μοντέλο της μη-παρεμβολής, τα συμβάντα ανασυνδυασμού σε διακριτά διαστήματα θεωρούνται ανεξάρτητα μεταξύ τους. Ως αποτέλεσμα, οι γονότυποι των γενετικών δεικτών κατά μήκος ενός χρωμοσώματος σχηματίζουν μία αλυσίδα Markov (όπου περιγράφεται στοχαστικά μία σειρά πιθανών γεγονότων κατά την οποία ένα γεγονός εξαρτάται μόνο από την παρατήρηση του προηγούμενου γεγονότος, ώστε η παρούσα κατάσταση ενός συστήματος να είναι ανεξάρτητη γεγονότων και στις δύο κατευθύνσεις ($5' - 3'$) ή ($3' - 5'$). Με άλλα λόγια, σύμφωνα με τον γονότυπο σε έναν συγκεκριμένο τόπο, οι γονότυποι στις πλευρικές θέσεις στα αριστερά είναι ανεξάρτητοι από τους γονότυπους στις πλευρικές θέσεις στα δεξιά. Αυτό ακολουθείται από την προϋπόθεση ότι οι συχνότητες των γονότυπων σε έναν αυτοσωμικό τόπο ακολουθούν τον Μεντελιανό νόμο του ανεξάρτητου διαχωρισμού κατά τη μείωση.

Η ευκολία της υπόθεσης χωρίς παρεμβολές ανασυνδυασμών, απεικονίζεται καλύτερα με ένα παράδειγμα στο **Σχήμα 3**, όπου παρουσιάζονται υποθετικά γονοτυπικά δεδομένα για τρία διαφορετικά άτομα σε ένα σύνολο έξι γενετικών δεικτών κατά μήκος ενός χρωμοσώματος όπου κάθε σειρά αντιστοιχεί σε διαφορετικό άτομο.



Σχήμα 3. Επεξήγηση του προβλήματος της εξαγωγής αγνοούμενων γονοτυπικών δεδομένων. Σε κάθε σειρά είναι τα γονοτυπικά δεδομένα για έναν γενετικό δείκτη για κάθε ένα από τα τρία διαφορετικά άτομα. Οι παύλες συμβολίζουν τα αγνοούμενα δεδομένα. Αναζητείται η πιθανότητα κάθε ατόμου να έχει γονότυπο AA ή AB στο πιθανό QTL που υποδεικνύεται από το τρίγωνο.

Προς εύρεση είναι η πιθανότητα ότι ένα άτομο είναι AA ή AB στον γενετικό τόπο (που ίσως είναι πιθανό QTL) που υποδεικνύεται από το τρίγωνο, δεδομένων των διαθέσιμων δεδομένων των γενετικών δεικτών. Εάν δεν παρεμβάλλεται, θεωρητικά, κάποιος ανασυνδυασμός, λαμβάνονται υπόψη μόνο οι γονότυποι στους πλησιέστερους πλευρικούς γονοτυπημένους γενετικούς δείκτες. Για το πρώτο άτομο, λαμβάνονται υπόψη οι γονότυποι στους δείκτες M3 και M4. Εάν το r_{iQ} είναι ο λόγος ανασυνδυασμού μεταξύ ενός δείκτη M_i και του πιθανού QTL, και εάν το r_{ij} είναι ο λόγος ανασυνδυασμού μεταξύ των δεικτών M_i και M_j , τότε η πιθανότητα ότι το άτομο έχει γονότυπο AA στο υποθετικό QTL είναι $(1 - r_{3Q})(1 - r_{4Q}) / (1 - r_{34})$, και η πιθανότητα ότι είναι AB είναι $r_{3Q} r_{4Q} / (1 - r_{34})$. Το γεγονός του ανασυνδυασμού μεταξύ των M4 και M5 για το άτομο στη 1^η σειρά δεν παίζει ρόλο καθώς το μοντέλο θεωρεί ότι όλα τα συμβάντα ανασυνδυασμού είναι ανεξάρτητα μεταξύ τους. Αν, παρόλα αυτά τα συμβάντα ανασυνδυασμού δεν ήταν ανεξάρτητα, τότε η πιθανότητα να έχουν συμβεί 2 εξαρτημένοι ανασυνδυασμοί, ένας στους M4-M5 και ένας στους M3-M4 θα ήταν μικρότερη. Συνεπώς η πιθανότητα, να είναι όντως ο γονότυπος AA για το άτομο 1 στο πιθανό QTL θα ήταν μεγαλύτερη. Για το άτομο στη δεύτερη σειρά, στο οποίο λείπει ο γονότυπος για τον δείκτη M4, λαμβάνονται υπόψη οι M3 και M5. Η πιθανότητα ότι το άτομο 2 είναι AA στο πιθανό QTL είναι $(1 - r_{3Q}) r_{5Q} / r_{35}$, και η πιθανότητα ότι είναι AB είναι $r_{3Q} (1 - r_{5Q}) / r_{35}$. Για το άτομο 3, οι δείκτες M2 και M4 λαμβάνονται υπόψη. Η πιθανότητα ότι είναι AA είναι στο πιθανό QTL είναι $(1 - r_{3Q}) r_{5Q} / r_{35}$ και για AB $(1 - r_{2Q})(1 - r_{4Q}) / (1 - r_{24})$. Γενικά, μπορεί κανείς να χρησιμοποιήσει αλγόριθμους για κρυφά Markov μοντέλα (HMMs) για τέτοιου είδους υπολογισμούς, όπως επίσης είναι δυνατό να επιτραπεί η παρουσία γονοτυπικών σφαλμάτων για αγνοούμενα γονοτυπικά δεδομένα. Μία τελευταία σημείωση αναφορικά με την παρούσα υποενοότητα έχει να κάνει με την συνάρτηση χάρτη (map function). Η συνάρτηση αυτή, ανάλογα με το μοντέλο ανασυνδυασμού που έχει επιλεγεί, μετατρέπει τις αποστάσεις των LG από cM (centiMorgan μονάδα γενετικής σύνδεσης) σε λόγους ανασυνδυασμού (fractions of recombination). Στην περίπτωση του μοντέλου των ανεξάρτητων ανασυνδυασμών, οι γενετικές αποστάσεις αντιστοιχούν στους λόγους ανασυνδυασμού.

4.3.2 Η προσέγγιση στην παρούσα μελέτη

Η προσέγγιση για την ανάλυση QTL, που χρησιμοποιήθηκε σε αυτή την μελέτη, αναφορικά με πληθυσμούς που προκύπτουν από πολλούς γονείς μπορεί να χωριστεί σε δύο μέρη: υπολογισμός των πιθανών γονότυπων, με την χρήση πολυσημιακών γονότυπων προερχόμενων από SNPs, και σάρωση του γενετικού χάρτη σύνδεσης για την εκτίμηση της συσχέτισης μεταξύ γονότυπου και φαινότυπου, χρησιμοποιώντας τους γονότυπους με τις βέλτιστες τιμές πιθανοφάνειας. Το κρυφό μοντέλο Markov (Hidden Markov Model) βλ. Broman and Sen (2009), είναι μία από τις προσεγγίσεις για τον υπολογισμό των πιθανοτήτων των γονότυπων. Το HMM υλοποιείται στην `rqtl2` (Broman et al., 2018), το οποίο είναι το πακέτο στο περιβάλλον R που χρησιμοποιήσαμε, έχει τροποποιηθεί από την προηγούμενη έκδοση του `rqtl1` (Broman 2014) για την δυνατότητα ανάλυσης δεδομένων από πολλαπλούς γονείς. Ένα βασικό συστατικό του HMM είναι ο πίνακας μετάβασης. Οι μεταβάσεις αντιπροσωπεύουν γενετικούς τύπους όπου η καταγωγή των χρωμοσωμικών τμημάτων μετατίθενται μεταξύ των απλότυπων των γονέων. Το αποτέλεσμα του HMM είναι μια λίστα τρισδιάστατων συστοιχιών, μία ανά χρωμόσωμα, με αντίστοιχες διαστάσεις σε άτομα x γονότυπους x θέσεις γενετικών δεικτών. Τα στοιχεία της κάθε σειράς αντιπροσωπεύουν γονοτυπικές πιθανότητες που αντανακλούν τόσο την αβεβαιότητα της εκτίμησης των απλότυπων όσο και της ετεροζυγωτίας. Για να καθοριστεί η στατιστική σημασία του εντοπισμού ενός QTL, το λογισμικό υλοποιεί τεστ μετάθεσης (Churchill και Doerge 1994). Για πληθυσμούς με πολλαπλούς γονείς υλοποιείται η ανάλυση μέσω γραμμικού μικτού μοντέλου (βλ. Κεφάλαιο 4), όπου οι σειρές των απλοτυπικών ανακατασκευών των γονότυπων μετατίθενται με την προσέγγιση των Cheng και Palmer (2013).

4.5 Βιβλιοθήκη ddRAD για παραγωγή δεδομένων γονιδιωματικής εκπροσώπησης

Οι προσεγγίσεις μειωμένης εκπροσώπησης γονότυπων αποτελούν ένα ισχυρό εργαλείο για τη διεξαγωγή μελετών πληθυσμιακής γενετικής για οποιοδήποτε είδος. Μετά την εισαγωγή της αλληλούχισης DNA σχετιζόμενης με τόπους αναγνώρισης περιοριστικών ενδονουκλεασών (Baird et al., 2008), ένα ευρύ φάσμα σχετικών μεθοδολογιών που χρησιμοποιούν περιοριστικά ένζυμα έχουν εισαχθεί όπως η γονοτύπιση μέσω αλληλούχισης (genotyping by sequencing) (Elshire et al., 2011), ddRAD-seq (Peterson, Weber, Kay, Fisher, & Hoekstra, 2012), 2b-RAD (Wang, Meyer, McKay, & Matz, 2012), ezRAD (Toonen et al., 2013), quaddRAD (Franchini, Monné Parera, Kautt, & Meyer, 2017) και 2RAD/3RAD (Bayona - Vásquez et al., 2019). Οι προαναφερθείσες πλατφόρμες έχουν χρησιμοποιηθεί εκτενώς σε μελέτες για υδρόβιους οργανισμούς που εστιάζουν τόσο σε γενετικές πτυχές των πληθυσμών (Andrews, Good, Miller, Luikart, & Hohenlohe, 2016) και στη μελέτη χαρακτηριστικών ενδιαφέροντος για παραγωγικούς σκοπούς (Houston et al., 2020; You, Shan, & Shi, 2020). Η ddRAD-seq είναι μία από τις πιο συχνά χρησιμοποιούμενες μεθόδους της οικογένειας της μειωμένης-εκπροσώπησης που συνδυάζει την απλότητα και την αποδοτικότητα κόστους κατά την κατασκευή της βιβλιοθήκης (Peterson et al., 2012). Τα τελευταία χρόνια, η ddRAD χρησιμοποιήθηκε με επιτυχία σε μια πληθώρα μελετών εστιασμένων στην τιλάπια για την διερεύνηση την υποκείμενης γενετικής δομής των χαρακτήρων οικονομικού ενδιαφέροντος (Jiang et al., 2019; Li, Zhu, Gu, Lin, & Xia, 2019; Li et al., 2017; Palaiokostas et al., 2015; Taslima et al., 2020), για την εύρεση SNP σε επίπεδο είδους (Syaifudin et al., 2019), και για την αποκρυπτογράφηση της γενετικής

ποικιλότητας και πληθυσμιακής δομής άγριων και εκτρεφόμενων πληθυσμών (Kajungiro, Palaiokostas, et al., 2019; Moses et al., 2019).

4.6 Υλικά και μέθοδοι

4.6.1 Δείγματα, εξαγωγή DNA, PCR και ανάλυση ανάθεσης γονέων σε απογόνους

Τα δείγματα που χρησιμοποιήθηκαν σε αυτή τη μελέτη, ανήκουν στην παρτίδα 1 όπως περιγράφηκαν στο κεφάλαιο 3. Συγκεκριμένα επιλέχθηκαν 6 οικογένειες από την παρτίδα 1 οι οποίες εμφάνισαν υψηλή φαινοτυπική διακύμανση στο βάρος και στο μήκος. Οι οικογένειες αυτές είναι οι, 1F-6M, 4F-6M, 4F-1M, 3F-2M, 3F-6M, 1F-5M. Ο αριθμός των δειγμάτων ήταν 288, ενώ τα μοναδικά δείγματα ήταν 266 λόγω του ότι οι γονείς γονοτυπήθηκαν για SNPs εις τριπλούν για λόγους ασφάλειας. Η διαδικασία της εξαγωγής DNA, PCR, γονοτύπισης των μικροδορυφορικών δεικτών, καθώς και της ανάλυσης ανάθεσης γονέων είναι αυτή του κεφαλαίου 3.

4.6.2 Κατασκευή βιβλιοθήκης ddRAD

Τα 266 δείγματα μετά την εξαγωγή DNA υποβλήθηκαν σε επεξεργασία με RNase για την απομάκρυνση του υπολειμματικού RNA. Το γονιδιωματικό DNA εκλούστηκε σε 5 mmol/L Tris, pH 8,5, και φυλάχθηκε στους 4°C. Κάθε δείγμα ποσοτικοποιήθηκε με φασματοφωτομετρία (Nanodrop 1000-Thermo Fisher Scientific) και η ποιότητα εκτιμήθηκε με ηλεκτροφόρηση σε πήκτωμα αγαρόζης 0,7%. Για την κατασκευή της βιβλιοθήκης ddRAD, χρησιμοποιήσαμε το πρωτόκολλο που περιγράφεται από τους Manousaki et al. (2016), με κάποιες μικρές τροποποιήσεις. Σε κάθε ένα από τα 266 δείγματα DNA (7 γονείς εις τριπλούν και 259 απόγονοι · 15 ng DNA ανά δείγμα) έγινε ξεχωριστή πέψη, στον ίδιο χρόνο επώασης και συγχρονισμένα, από τα δύο περιοριστικά ένζυμα υψηλής πιστότητας (RE): SbfI (CCTGCA | τόπος αναγνώρισης GG) και NlaIII (CATG | τόπος αναγνώρισης C), New England Biolabs (NEB), Ηνωμένο Βασίλειο. Οι πέψεις επώαστηκαν στους 37°C για 90 λεπτά, χρησιμοποιώντας 20 U ενζύμου ανά μικρογραμμάριο DNA και 0.6 μl CutSmart Buffer (NEB), σε συνολικό όγκο αντίδρασης 6 μl. Οι αντιδράσεις ψύχθηκαν βραδέως σε θερμοκρασία δωματίου, και 3 μl μίγματος προσαρμογέων (adaptors) προστέθηκαν στο DNA μετά την πέψη, και επώαστηκαν σε θερμοκρασία δωματίου για 10 λεπτά. Το μείγμα προσαρμογέων περιείχε ατομικούς συνδυασμούς P1 (SbfI-συμβατός) και P2 (NlaIII-συμβατός) σε συγκεντρώσεις 6 και 96 nM, αντίστοιχα, σε 1x ρυθμιστικό διάλυμα αντίδρασης vo2 (NEB). Ο λόγος του προσαρμογέα P1 προς P2 (1:16) επιλέχθηκε ώστε να αντικατοπτρίζει τη σχετική αφθονία των θέσεων κοπής SbfI και NlaIII. Ο προσαρμογέας P1 και P2 περιελάμβανε μία αλληλουχία πέντε ή επτά βάσεων για την ταυτοποίηση του κάθε δείγματος μετά την αλληλούχιση. Οι απολινώσεις (ligations) πραγματοποιήθηκαν σε διάστημα 3 ωρών στους 22°C με προσθήκη επιπλέον 3 μl μίγματος απολίνωσης που περιλαμβάνει 4 mM rATP (Promega, UK) και 2000 μονάδες T4 λιγάσης (NEB) σε 1x CutSmart buffer (NEB). Τα απολινωμένα δείγματα ομαδοποιήθηκαν μαζί, καθαρίστηκαν με στήλη (MinElute PCR PurificationKit, Qiagen, UK), και εκλούστηκαν σε 70 μl EB ρυθμιστικό διάλυμα (Qiagen, UK). Η επιλογή μεγέθους πραγματοποιήθηκε με διαχωρισμό πηκτώματος αγαρόζης, διατηρώντας τα θραύσματα μεταξύ 450 και 750 bp. Μετά τον καθαρισμό του πηκτώματος (kit εκχύλισης πηκτωμάτων αγαρόζης MinElute, Qiagen, UK), τα εκλουσμένα επιλεγμένα θραύσματα DNA (68 μl σε ρυθμιστικό διάλυμα EB) ενισχύθηκαν

με PCR (15 κύκλοι PCR · 32 ξεχωριστές αντιδράσεις 12,5 μl, καθεμία με 1 μl DNA προτύπου) χρησιμοποιώντας πολυμεράση υψηλής πιστότητας Taq (Q5 Hot Start High-Fidelity DNA Polymerase, NEB). Οι αντιδράσεις PCR συνδυάστηκαν (400 μl συνολικά) και καθαρίστηκαν σε στήλη (κιτ καθαρισμού MinElute PCR). Το έκλουσμα των 57 μl, σε ρυθμιστικό διάλυμα EB, στη συνέχεια υποβλήθηκε σε περαιτέρω καθαρισμό επιλογής μεγέθους χρησιμοποιώντας ίσο όγκο μαγνητικών σφαιριδίων AMPure (Perkin-Elmer, UK) για μεγιστοποίηση της αφαίρεσης μικρών θραυσμάτων. Οι τελικές βιβλιοθήκες εκλούστηκαν σε 24 μl EB ρυθμιστικό διάλυμα. Τέλος, οι βιβλιοθήκες ddRAD αλληλουχήθηκαν σε μια λωρίδα HiSeq 2500 (2x125 bp).

4.6.3 Ποιοτικός έλεγχος και παραγωγή γονοτύπων μέσω του λογισμικού STACKS

Χρησιμοποιήσαμε λογισμικό FastQC v.0.11.5 για την διεξαγωγή του ποιοτικού ελέγχου των πρωτογενών δεδομένων αλληλούχισης (Andrews and Babraham Bioinformatics Group, 2010). Για να ανακτήσουμε τις αλληλουχίες (reads) που αντιστοιχούν σε κάθε άτομο, φιλτράραμε και διαχωρίσαμε τις αλληλουχίες χρησιμοποιώντας το πρόγραμμα Process radtags από το λογισμικό STACKS v.2.3 (Catchen, 2013). Σε αυτό το βήμα, η παράμετρος -q χρησιμοποιήθηκε για την απόρριψη αλληλουχιών χαμηλής ποιότητας (κάτω από 20) χρησιμοποιώντας τις βαθμολογίες Phred που παρέχονται από τα αρχεία FASTQ (Catchen, 2013). Για την ευθυγράμμιση των αλληλουχιών στο γονιδίωμα του κρυνίου (Papadogiannis et al., 2021 in preparation) χρησιμοποιήθηκε το λογισμικό BWA (Li and Durbin, 2009). Στην συνέχεια το πρόγραμμα gstacks χρησιμοποιήθηκε για την δημιουργία των ddRAD τόπων ενσωματώνοντας τις αλληλουχίες ζευγών (paired reads) που έχουν ευθυγραμμιστεί στο γονιδίωμα αναφοράς και ταξινομηθεί, χρησιμοποιώντας έναν αλγόριθμο συρόμενου παραθύρου (sliding window algorithm). Τέλος το πρόγραμμα populations από το stacks χρησιμοποιήθηκε για το τελικό φιλτράρισμα των δεδομένων ανάλογα με το εύρος της κάλυψης των ddRAD τόπων στο σύνολο των δειγμάτων, την maf (minor allele frequency) η οποία είναι η συχνότητα του δεύτερου πιο κοινού (most shared) αλληλόμορφου. Σε αυτό το βήμα χρησιμοποιήθηκε η παράμετρος -r 70, (για την επιλογή τόπων που αντιπροσωπεύουν το 70% των δειγμάτων) καθώς και -maf 0.01.

4.6.4 Ποιοτικός έλεγχος με το Plink

Δύο επιπλέον βήματα ελέγχου των ddRAD τόπων από το STACKS υλοποιήθηκαν με το Plink1.9 (Purcell et al., 2007). Σε αυτό το βήμα χρησιμοποιήθηκε η παράμετρος -geno 0.3, που είναι το ανάλογο του -r 70 του STACKS ώστε να επιβεβαιωθεί η κάλυψη των τόπων στα δείγματα, καθώς επίσης και η παράμετρος -mendel για την εύρεση λαθών γονοτύπισης μεταξύ γονέων και απογόνων που δεν εξηγούνται μέσω της κληρονομικότητας κατά Μέντελ. Τα επιτρεπτά λάθη ανά τόπο ορίστηκαν σε 5.

4.6.5 Κατασκευή γενετικού χάρτη σύνδεσης

Για την κατασκευή γενετικού χάρτη χρησιμοποιήθηκε το λογισμικό LepMap2 (Rastas et al., 2016). Τα SNPs που πέρασαν το Data Tolerance threshold (0.001), ήταν 4529 από τα 5988 και κατανεμήθηκαν στα Linkage groups χρησιμοποιώντας LOD score 4.80. Η επιλογή του έγινε ώστε ο παραγόμενος αριθμός των LG να είναι ίσος με τον αριθμό των χρωμοσωμάτων που είχε οριστεί σε n=24. Χρησιμοποιήθηκαν οι εντολές όπως περιγράφηκαν από τους

Oikonomou et al. 2021 (in submission). Τέλος, η απεικόνιση του έγινε με το LinkageMapView (Lisa A Ouellette et al., 2018).

4.6.6 Ανάλυση QTL

Ο γενετικός χάρτης σε συνδυασμό με τα φαινοτυπικά χαρακτηριστικά χρησιμοποιήθηκαν για την ανίχνευση γενετικών περιοχών που σχετίζονται με τον μελετώμενο φαινότυπο. Η ανάλυση πραγματοποιήθηκε με το λογισμικό *rqtl2* (Broman et al., 2018) χρησιμοποιώντας δυο μοντέλα με και χωρίς την επίδραση του πολυγονιδιακού παράγοντα. Ενώ ως όριο για την δοκιμή μετάθεσης χρησιμοποιήθηκε το 1000 με το επίπεδο σημαντικότητας να είναι 0,05. Τέλος, τα όρια για το μοντέλο χωρίς την επίδραση του πολυγονιδιακού παράγοντα ήταν 4,77 και 4,59 για το σωματικό βάρος και για το μήκος αντίστοιχα. Ενώ για το μοντέλο με την επίδραση του πολυγονιδιακού παράγοντα ήταν 4,76 και 4,58 το σωματικό βάρος και για το μήκος, αντίστοιχα. Η διαδικασία που ακολουθήθηκε καθώς και οι εντολές αναφέρονται στο παράρτημα εκτενώς.

4.6.7 Συγκριτική γονιδιωματική ανάλυση

Οι χαρτογραφημένοι RAD τόποι του κρανίου χρησιμοποιήθηκαν σε συγκριτική ανάλυση με γονιδιώματα από τους ακόλουθους σχετικά κοντινούς φυλεγενετικά, τελεόστεους ιχθύες : medaka (*Oryzias latipes*, Ensembl 100), τιλάπια του Νείλου (*Oreochromis niloticus*, ncbi GCA_000188235.2), λαβράκι (*Dicentrarchus labrax*, dicLab v1.0c <http://seabass.mpipz.de>), τσιπούρα (Ensembl 100), yellow croaker (*Lamirichthys crocea*, Ensembl 100). Οι νουκλεοτιδικές αλληλουχίες από τον γενετικό χάρτη του κρανίου οι οποίες συνέθεσαν τα 24 Linkage groups χρησιμοποιήθηκαν σε BLASTN ανάλυση εύρεσης ομοιότητας, πρὸς εύρεση ομόλογων περιοχών στα αντίστοιχα 5 προαναφερθέντα γονιδιώματα. Το όριο της τιμής e ορίστηκε σε 10^{-9} και τα αποτελέσματα με περισσότερες από 10 αντιστοιχίες (hits) ή περισσότερα από 10 HSP (high scoring segment pairs) στην πρώτη σε βαθμολογία αντιστοιχία, αποκλείστηκαν για την εξάλειψη επαναλαμβανόμενων περιοχών στο γονιδίωμα. Μόνο η πρώτη αντιστοιχία σε βαθμολογία e ανά αλληλουχία, διατηρήθηκε, και θεωρήθηκε ομόλογη με το γονιδίωμα στόχο. Η αντιστοιχία των Linkage groups του κρανίου με τα χρωμοσώματα των γονιδιωμάτων στόχων των 6 ιχθύων, συνήχθη από αυτή την διαδικασία. Εάν η πλειοψηφία των αλληλουχιών του κρανίου από ένα LG εύρισκε ομόλογες περιοχές σε ένα από τα χρωμοσώματα στόχο, τότε το συγκεκριμένο LG του κρανίου και το συγκεκριμένο χρωμόσωμα στόχος θεωρούνταν ομόλογα. Η ίδια ακριβώς διαδικασία πραγματοποιήθηκε για τις κωδικοποιούσες περιοχές. Οι κωδικοποιούσες αλληλουχίες των 5 γονιδιωμάτων, χρησιμοποιήθηκαν σε ένα δεύτερο γύρο ανάλυσης BLASTN, με τις ίδιες παραμέτρους, ώστε να διερευνηθεί η ομολογία του γενετικού χάρτη αποκλειστικά με κωδικοποιούσες περιοχές. Οι συνδέσεις μεταξύ των RAD γενετικών τόπων του κρανίου και των 6 γονιδιωμάτων οπτικοποιήθηκαν με το λογισμικό Circos (Krzywinski et al. 2009).

4.6.8 Αποτελέσματα & συζήτηση

Από τις 688×10^6 πρωτογενείς BAM καταγραφές, από την ευθυγράμμιση των αναγνώσεων (reads) με το γονιδίωμα, πέρασαν τον έλεγχο οι 577×10^6 ή το 87% από τις οποίες οι 281×10^6 ήταν reads του συμπληρωματικού κλώνου με 3'-5' κατεύθυνση. Το 4% ή 26×10^6 δεν ευθυγραμμίστηκαν επαρκώς με το γονιδίωμα και αποκλείστηκαν από την ανάλυση. Οι 51×10^6 ή το 7.8% δεν ευθυγραμμίστηκαν μερικώς και ως εκ τούτου αποκλείστηκαν, καθώς και 3×10^6 ή το 0.4% δεν βρήκε κάποια αντιστοιχία με το γονιδίωμα. Οι τόποι RAD που

προέκυψαν από το gstacks ήταν 316425 με 288×10^6 εμπρόσθιες αλληλουχίες και 247×10^6 paired-end αλληλουχίες, με μέσο μήκος βάσεων 353.1 (sd 106.7). Η κάλυψη για κάθε τόπο ήταν 43.3x.

Το populations με τις ρυθμίσεις για το φιλτράρισμα των τόπων που είναι κοινοί στο 70% των ατόμων καθώς και για maf 0.01 απέκλεισε 294.953 από τους 316.425 τόπους καταλήγοντας σε 21.472 τόπους, οι οποίοι είχαν 9.710 SNP's. Ο μέσος όρος των βάσεων, ή ο μέσος γονότυπος, για κάθε RAD τόπο αποτελούνταν από 414.67 ζεύγη βάσεων (sd 0.73).

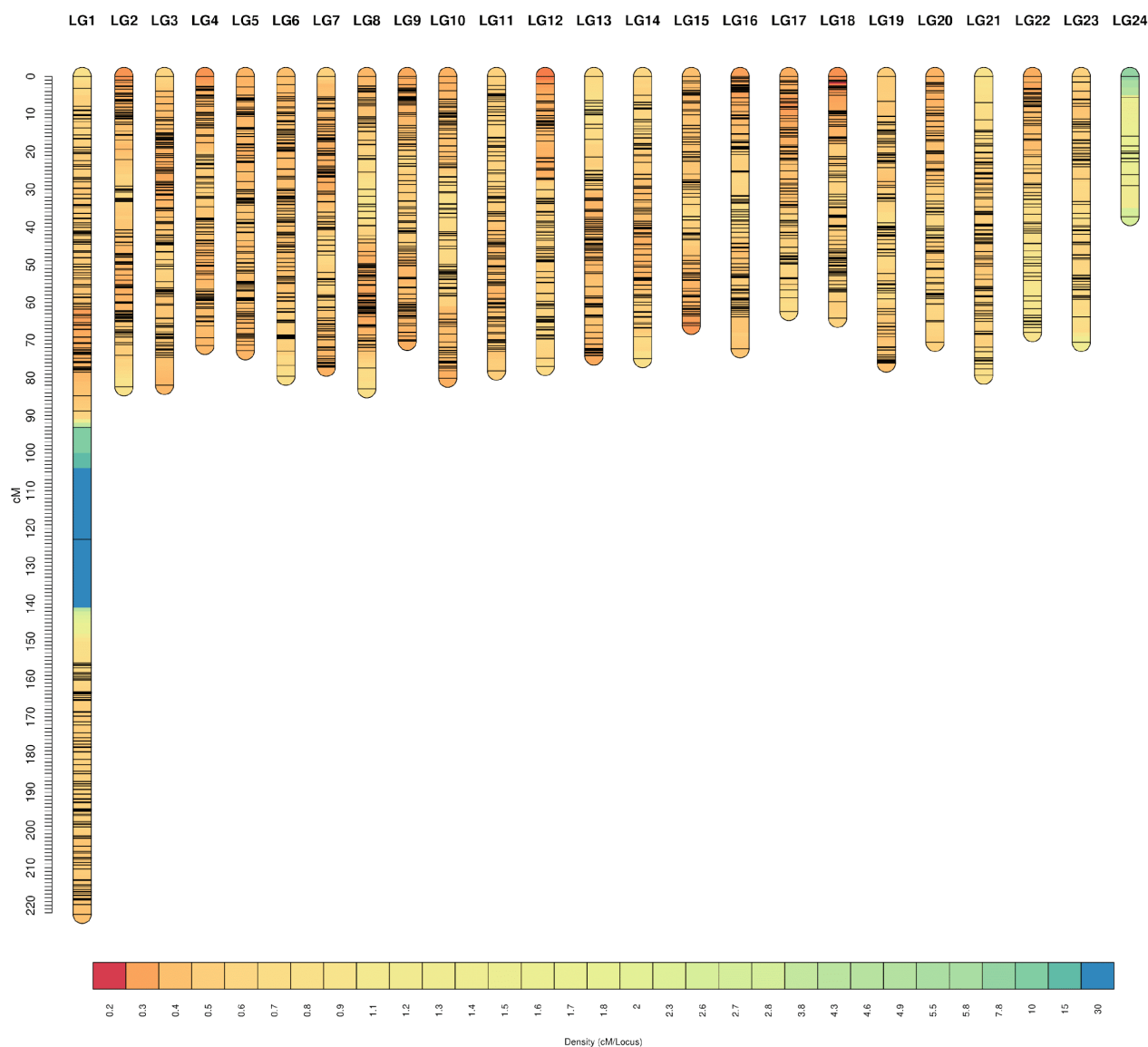
Από τους 9710 RAD δείκτες, μετά τον έλεγχο στο plink για το επιτρεπτό όριο λαθών κατά Μέντελ, 5988 παρέμειναν. Από τους 5988 τόπους οι 4529 συμμετείχαν στα διαμορφούμενα 24 Linkage groups του γενετικού χάρτη σύνδεσης, του οποίου η σύνθεση δίνεται στον πίνακα 1. Το μικρότερο LG αποτελείται από 20 RAD τόπους (LG24) και το μεγαλύτερο από 396, (LG1). Ενώ η μέση απόσταση στον χάρτη ήταν 0.4 cM και η ολική 1158.356 cM.

Πίνακας 1. Η σύνοψη του γενετικού χάρτη σύνδεσης για τον κρανίο, βασισμένος σε 6 οικογένειες από 3 θηλυκούς και 4 αρσενικούς γεννήτορες. Στην πρώτη στήλη παρατίθεται ο αριθμός των Linkage group, στην δεύτερη ο αριθμός των RAD τόπων ανά LG και στην τρίτη το μήκος σε μονάδες cM ανά LG. Στον χάρτη συμμετείχαν συνολικά 4529 RAD τόποι.

Σύνοψη του γενετικού χάρτη σύνδεσης		
LG	Αριθμός δεικτών	Μήκος (cM)
1	396	222.415
2	218	82.417
3	221	82.003
4	208	71.361
5	218	72.767
6	207	79.57
7	215	77.159
8	208	82.911
9	205	70.251
10	199	80.137
11	191	78.158
12	189	77.025
13	194	74.192
14	182	74.915
15	182	66.145
16	182	72.286
17	171	62.38
18	172	64.156
19	166	76.145
20	165	70.594
21	154	79.321
22	138	68.02
23	128	70.623

Ο γενετικός χάρτης, όπως αναφέρθηκε και στην εισαγωγή, βασίζεται στους λόγους χρωμοσωμικών ανασυνδυασμών γονίων και απογόνων. Εάν οι αποστάσεις μεταξύ των γενετικών δεικτών στον χάρτη είναι μεγάλες, π.χ > 50 cM τότε ενδεχομένως η πυκνότητα των δεδομένων να μην είναι επαρκής για μία αντικειμενική εκπροσώπηση του γονιδιώματος. Όσο μεγαλύτερη κάλυψη υπάρχει μεταξύ των γονότυπων των απογόνων και των γονέων, ιδανικά όλο το γονιδίωμα, οι λόγοι των χρωμοσωμικών ανασυνδυασμών εμφανίζουν μεγαλύτερη ακρίβεια, και ως εκ τούτου ο γενετικός χάρτης δεν εμφανίζει διαστήματα. Ο χάρτης που παρουσιάζεται εδώ, δεν εμφανίζει μεγάλα διαστήματα μεταξύ των Linkage groups, όμως, εμφανίζει 2 διαστήματα των ~ 30 cM στο LG1, το οποίο είναι επίσης αρκετά μεγαλύτερο των υπόλοιπων **Σχήμα 4**. Αξίζει να σημειωθεί ότι η κατανομή των τόπων στον χάρτη είναι ομοιόμορφη, με τα περισσότερα LG να περιλαμβάνουν ~ 200 δείκτες.

Σχήμα 4. Ο γενετικός χάρτης σύνδεσης όπως οπτικοποιήθηκε με το πακέτο LinkageMapView (Lisa A Ouellette et al., 2018) σε περιβάλλον R. Το ιδεόγραμμα των Linkage groups απεικονίζεται ώστε η πυκνότητα των περιοχών του χάρτη να συμβολίζεται με το χρωματικό εύρος ερυθρό-γαλάζιο. Με το γαλάζιο χρώμα να είναι 30 cM και το ερυθρό 0.2 cM ως μέγιστες και ελάχιστες αποστάσεις μεταξύ των τόπων στον χάρτη



Λόγω του ότι ο γενετικός χάρτης συμπεριλάμβανε τόπους που ήταν κοινί στο 70% των γονέων και απογόνων, είναι πιθανό ένα υποσύνολο των τόπων με μεγαλύτερη κάλυψη (90%) να βελτιώνει την εικόνα του χάρτη, λόγω του ότι θα εμφάνιζε μεγαλύτερη ανάλυση και ακρίβεια στους λόγους των χρωμοσωμικών ανασυνδυασμών. Βάση αυτής της υπόθεσης, ένας γενετικός χάρτης σύνδεσης δημιουργήθηκε για ένα εύρος κοινών τόπων 90% (παράρτημα) για τον οποίο όμως δεν προέκυψαν κάποια στατιστικά σημαντικά αποτελέσματα στην ανίχνευση QTL. Συνεπώς ο απεικονιζόμενος χάρτης στο **Σχήμα 4**, είναι αυτός ο οποίος επιλέχθηκε τελικώς. Λόγω του ότι το εύρος κάλυψης των τόπων ddRAD δεν φαίνεται να εξηγεί στην ολότητα του την τελική διαμόρφωση του γενετικού χάρτη είναι πολύ πιθανό ένας συνδυασμός παραγόντων να οδήγησαν σε αυτή την τελική διαμόρφωση.

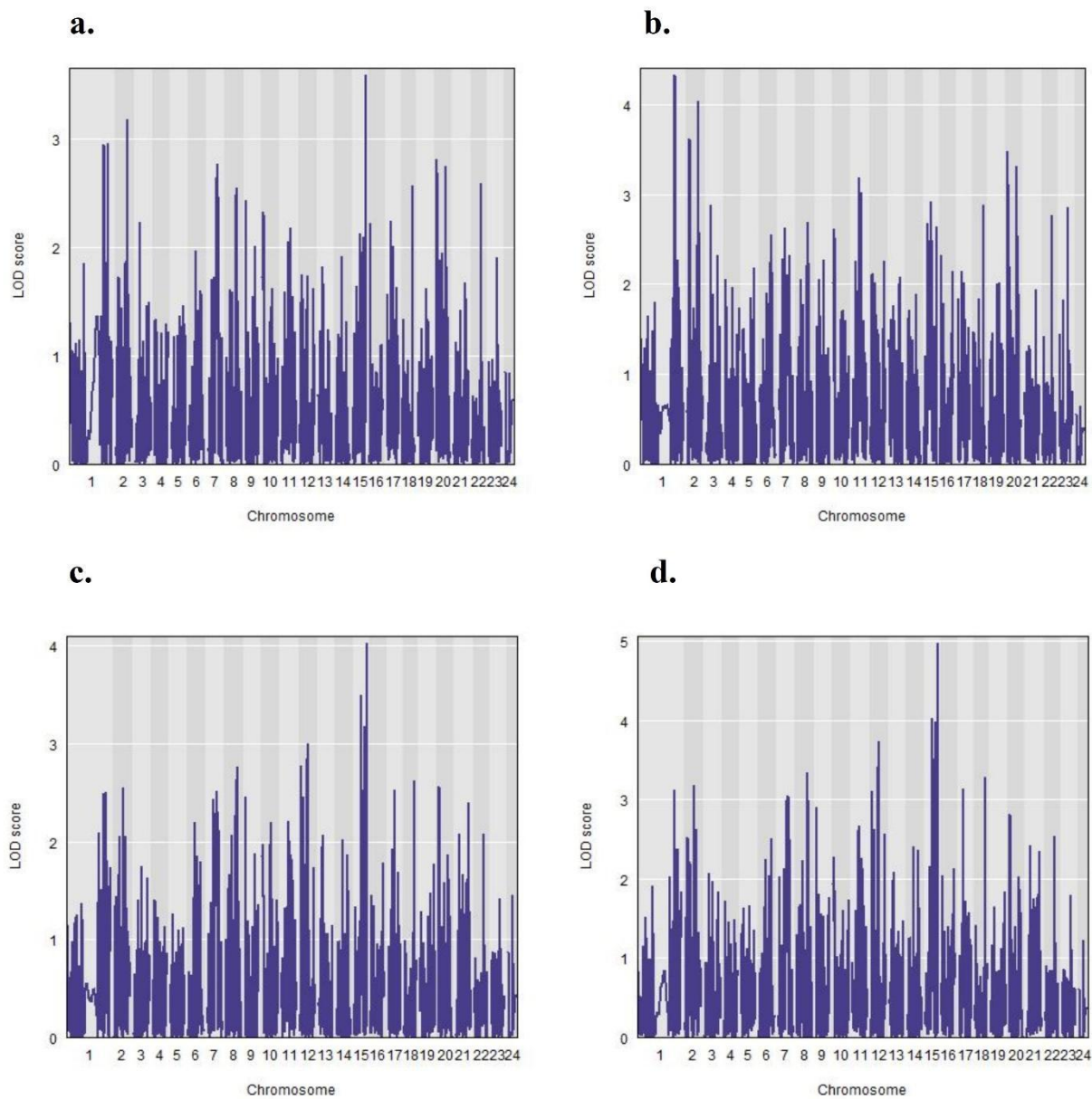
Έχει ιδιαίτερη σημασία να αναφέρουμε ορισμένα θεωρητικά και ερευνητικά στοιχεία που βοηθούν στην κατανόηση της δομής ενός γενετικού χάρτη, και ειδικότερα την δομή του καρυότυπου των τελεόστεων ιχθύων. Μία κεντρική σύντηξη (centric fusion) δυο ακροκεντρικών χρωμοσωμάτων των οποίων τα κεντρομερή βρίσκονται στην περιφέρεια, οδηγεί στην δημιουργία ενός μετακεντρικού χρωμοσώματος (Jones 1998; Nachman & Searle 1995). Αντίστροφα, μία κεντρική σχάση (centric fission) ενός μετακεντρικού χρωμοσώματος οδηγεί στην δημιουργία δύο ακροκεντρικών χρωμοσωμάτων. Ακόμη, το κεντρομερές σε ένα χρωμόσωμα, μπορεί να μετατοπιστεί μέσω περικεντρικής αναστροφής (pericentric inversion) (Imai & Crozier 1980) ή και μετατόπιση κεντρομερούς (Rocchi et al., 2012). Σε αντίθεση με την κεντρική σύντηξη και σχάση, η μετατόπιση του κεντρομερούς αλλάζει μόνο τον αριθμό των βραχίωνων σε ένα χρωμόσωμα και όχι τον αριθμό των χρωμοσωμάτων.

Επιπρόσθετα, έχειδειχθεί ότι στους τελεόστεους ιχθύες, από ένα δείγμα 2,587 γονιδιωμάτων, το ~30% των ειδών έχει καρυότυπους με ακροκεντρικά χρωμοσώματα, κατ' αποκλειστικότητα (Yoshida & Kitano 2021) και συγκεκριμένα δύο μεγάλες μονοφυλετικές ταξινομικές ομάδες τελεόστεων, οι Eurypterygii (1,368/16,000 είδη αναλύθηκαν) και οι Otophysi (1,026/7900 είδη αναλύθηκαν) έδειξαν διαφορές στην κατανομή των βραχίωνων στους καρυότυπους τους (τα Eurypterygii εμπεριέχουν τα είδη τα οποία αναφέρουμε στην παρούσα εργασία και απαρτίζουν την πλειοψηφία των σημερινών θαλάσσιων ιχθύων, ενώ η ομάδα Otophysi σε ιχθύς του γλυκού νερού). Το 45% των 1,368 καρυότυπων των Eurypterygii, που είναι η ομάδα ενδιαφέροντος με βάση την διατριβή, φάνηκε να αποτελούνται στο 100% από ακροκεντρικά χρωμοσώματα. Επιπλέον, οι Yoshida & Kitano (2021) έδειξαν ότι ο ρυθμός της κεντρικής σύντηξης / σχάσης εξαρτάται από τον αριθμό των χρωμοσωμάτων προέλευσης, τα ακροκεντρικά χρωμοσώματα για την σύντηξη και μετακεντρικά χρωμοσώματα για την σχάση, δείχνοντας ότι υπάρχει σύνδεση μεταξύ της εξέλιξης των χρωμοσωμάτων και του αριθμού βραχίωνων. Με βάση, τις παραπάνω παρατηρήσεις, είναι πιθανό ένας συνδυασμός παραγόντων, όπως ή ύπαρξη ενός μετακεντρικού χρωμοσώματος στον κρανίο (LG1) και 23 ακροκεντρικών να είναι μία δόκιμη υπόθεση. Μία υπόθεση θα μπορούσε επίσης να είναι, ο επιμερισμός τμημάτων του LG1 στα υπόλοιπα χρωμοσώματα ή και σε ένα φυλετικό χρωμόσωμα, επιμερισμός ο οποίος δεν ήταν εφικτός με βάση τα συγκεκριμένα δεδομένα προερχόμενα από ddRAD τα οποία έδωσαν την καλύτερη εφικτή διαμόρφωση με βάση τους λόγους των χρωμοσωμικών ανασυνδυασμών μεταξύ γονέων και απογόνων σε έναν καρυότυπο 24 χρωμοσωμάτων.

Η ανάλυση ανίχνευσης QTL έγινε για 2 φαινοτυπικούς χαρακτήρες. Το βάρος και το μήκος. Τα μοντέλα που υλοποιήθηκαν για τους 2 χαρακτήρες ήταν με και χωρίς την επίδραση του πολυγονδιακού παράγοντα. Η γενετική διακύμανση που συνδέεται με έναν γενετικό τόπο για την ανίχνευση QTL, μπορεί να οφείλεται στις αθροιστικές επιδράσεις πολλών γονιδίων ή και όχι. Η γενετική βάση των ποσοτικών χαρακτήρων, όπως το βάρος και το μήκος, δεν οφείλεται σε μονογονδιακές επιδράσεις, παρά στην αιτιώδη αλληλεπίδραση ομάδων γενετικών τόπων/γονιδίων. Τα αποτελέσματα που παρουσιάζονται σε αυτή τη μελέτη, ανίχνευσαν ένα στατιστικά σημαντικό QTL, για το βάρος χωρίς την επίδραση του πολυγονδιακού παράγοντα. Συγκεκριμένα στον γενετικό τόπο «36880» που στον γενετικό χάρτη σύνδεσης βρίσκεται στο LG15, βρέθηκε ένα QTL με LOD 4.97 (με όριο το 4.59). Ο συγκεκριμένος τόπος εντοπίζεται σε κωδικοποιούσες περιοχές, σε 3 είδη που συμπεριλαμβάνονται στην ανάλυση yellow croaker, medaka, τιλάπια. Στα 3 αυτά είδη, οι κωδικοποιούσες περιοχές με τις οποίες είναι ομόλογος ο QTL τόπος συνδέονται με το γονίδιο *eys*. Λειτουργικές περιοχές αυτού του γονιδίου που ονομάζονται EGF-1 like

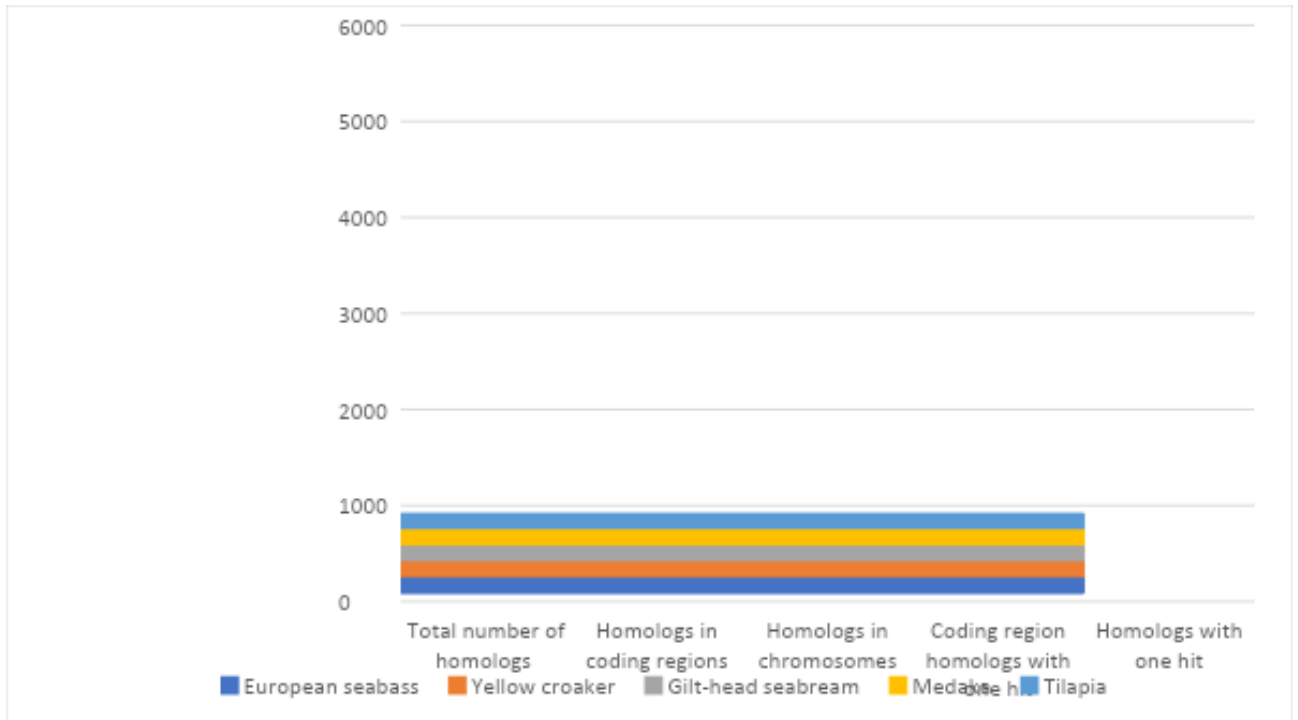
(epidermal growth factor 1 like domains), έχει δειχθεί ότι σχετίζονται με χαρακτήρες ανάπτυξης στην τιλάπια σε βοοειδή και αιγοπρόβατα (Yoshida, Yáñez 2021) **Σχήμα 5**.

Σχήμα 5. Γραφική αναπαράσταση των αποτελεσμάτων με και χωρίς την επίδραση του πολυγονιδιακού παράγοντα για το βάρος και το μήκος. α. μήκος με πολυγονιδιακό μοντέλο, β. μήκος χωρίς πολυγονιδιακό μοντέλο, γ. βάρος με πολυγονιδιακό μοντέλο και δ. βάρος χωρίς πολυγονιδιακό μοντέλο. Το QTL ανιχνεύθηκε στην εικόνα δ. στο LG 15



Η οικογένεια Sciaenidae, περιλαμβάνει περισσότερα από 270 είδη της τάξης των περκόμορφων (Perciformes). Ωστόσο, στο yellow croaker (Xiao et al., 2015) που ανήκει στην ίδια οικογένεια έχει περαιωθεί συγκριτική γονιδιωματική μελέτη χωρίς όμως την κατασκευή γενετικού χάρτη σύνδεσης υψηλής πυκνότητας. Για τον κρανίο, καθώς και για οποιοδήποτε μέλος της οικογένειας των Scianidae, είναι η πρώτη φορά που μελετώνται συγκριτικά οι ομολογίες του γονιδιώματος του, σε επίπεδο γενετικού χάρτη σύνδεσης με άλλα γονιδιώματα τελεόστεων ιχθύων. Προηγούμενη συγκριτική γονιδιωματική ανάλυση που βασίστηκε στη διατήρηση αλληλουχίων μέσω ορθόλογων γονιδίων και ομολογιών μεταξύ των χρωμοσωμάτων σε κωδικοποιούσες περιοχές έδειξε ότι το stickleback (*Gasterosteus aculeatus*, τάξη Gasterosteiformes) σχετίζεται στενότερα με το yellow croaker από το medaka (Ao et al., 2015). Ωστόσο, ο αριθμός χρωμοσωμάτων του yellow croaker (n = 24) είναι ίσος με του medaka (Kasahara et al., 2007) και διαφορετικός από του stickleback (n=21) (Ross et al., 2009), γεγονός το οποίο δείχνει ότι σημαντικές γονιδιωματικές αναδιατάξεις σημειώθηκαν μετά την απόκλιση αυτών των ειδών από τον κοινό τους πρόγονο. Το γονιδίωμα του κοινού προγόνου των τελεόστεων, υπεβλήθη σε τρεις γύρους διπλασιασμού (Glasauer & Neuhauss 2014), έναν παραπάνω από τα θηλαστικά, ενώ ο τρίτος γύρος πραγματοποιήθηκε πριν από περίπου 370 εκατομμύρια χρόνια (MYA) (Sato & Nishida 2010). Σε σχετικά σύντομο χρονικό διάστημα ~50 MYA μετά την εκδήλωση του τρίτου γονιδιωματικού διπλασιασμού, ο κοινός πρόγονος των περισσότερων τελεόστεων υπέστη οκτώ σημαντικές ενδοχρωμοσωμικές αναδιατάξεις, οδηγώντας σε έναν προγονικό καρυότυπο που αποτελείται από 24 χρωμοσώματα (Takeda 2008). Είναι ενδιαφέρον ότι το γονιδίωμα του medaka (*Oryzias latipes*) διατήρησε τον προγονικό καρυότυπο για περισσότερα από 300 MYA (Kasahara et al., 2007). Παρόλο που έχουν δημοσιευτεί πολλά γονιδιώματα, η συγκριτική ανάλυση αποκάλυψε την ύπαρξη ενδοχρωμοσωμικών αναδιατάξεων μετά από την απόκλιση από τον κοινό πρόγονο (Guyon et al., 2012, Zhang et al., 2013). Το κατά πόσο άλλα είδη διατηρούν τον προγονικό καρυότυπο, όπως το medaka, παραμένει άγνωστο. Ενώ, αξίζει να σημειωθεί ότι χρωμοσωμικές αναδιατάξεις έχει δειχθεί ότι έπονται ακόμη και μετά την ειδογένεση και την απόκλιση από το κοινό προγονικό είδος (Takeda 2008). Στην παρούσα μελέτη, πραγματοποιήθηκε μια συγκριτική γονιδιωματική ανάλυση σε ζεύγη, για την μελέτη των ομολογιών μεταξύ του γενετικού χάρτη υψηλής πυκνότητας του κρανιού, τα γονιδιώματα 5 τελεόστεων ιχθύων, καθώς και με τις κωδικοποιούσες περιοχές των γονιδιωμάτων αυτών. Καταρχήν, ο γενετικός χάρτης δείχνει σε μεγάλο βαθμό ομολογίες σε όλο το γονιδιωματικό εύρος των 5 ειδών, επιβεβαιώνοντας με αυτόν τον τρόπο την εγκυρότητα του. Η ανάλυση, δείχνει ότι ο κρανιός διατηρεί τον προγονικό καρυότυπο και τα 24 χρωμοσώματα, όπως αναφέρθηκε για το medaka.

Σχήμα 6. Σύνοψη των αποτελεσμάτων συγκριτικής γονιδιωματικής ανάλυσης σε όλο το γονιδιωματικό εύρος, καθώς και σε κωδικοποιούμενες περιοχές ανάμεσα στον γενετικό χάρτη σύνδεσης υψηλής πυκνότητας του κρανίου και των 5 τελεόστεων ιχθύων. Στον άξονα των y φαίνονται οι κατηγορίες : ομόλογες περιοχές με μία μόνο αντιστοιχία (Homologs with one hit), κωδικοποιούμενες περιοχές με μία μόνο αντιστοιχία (Coding region homologs with one hit), συνολικές χρωμοσωμικές ομολογίες (Homologs in chromosomes), ομολογίες σε κωδικοποιούμενες περιοχές (Homologs in coding regions), συνολικές ομολογίες (Total number of homologs). Ενώ στον άξονα των x βρίσκεται το αντίστοιχο μέγεθος, σε αντιστοιχίες, για κάθε ένα από τα 5 είδη

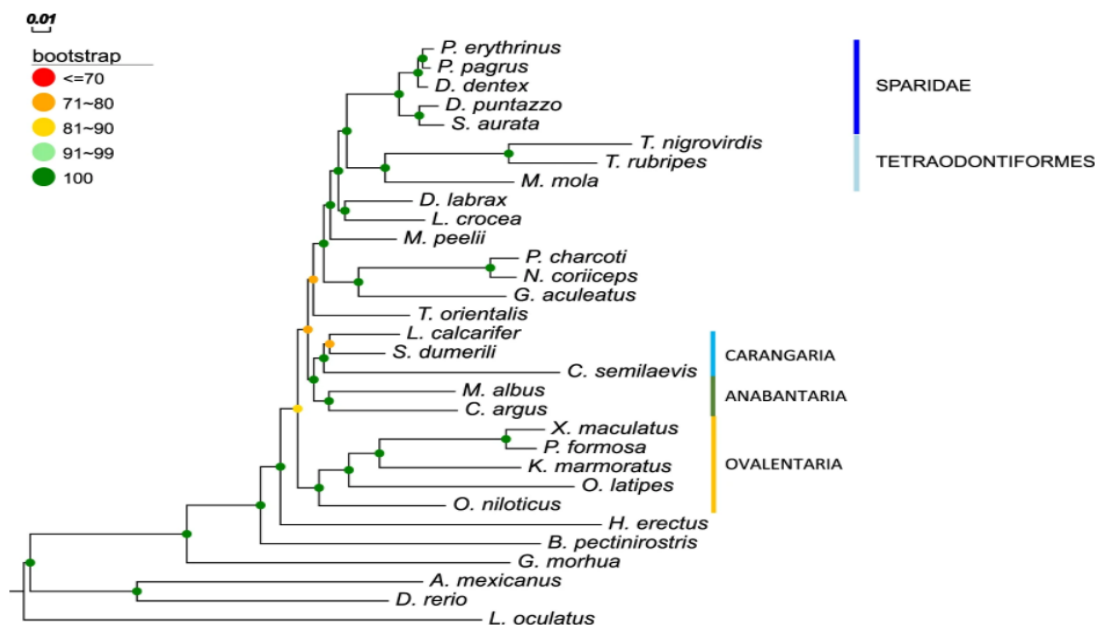


Από το **Σχήμα 6**, φαίνεται σε ένα πρώτο επίπεδο ότι τα αποτελέσματα των ομολογιών με μία αντιστοιχία είναι αρκετά υψηλά. Στον **Πίνακα 2**, παρουσιάζονται οι τιμές του Σχήματος 6. Το γεγονός ότι οι μοναδικές αντιστοιχίες των αλληλουχιών των γενετικών τόπων του χάρτη του κρανίου εμφανίζονται πολυάριθμες, δείχνει ότι η διαδικασία της ευθυγράμμισης με το γονιδίωμα αναφοράς του κρανίου (Papadogiannis et al., in preparation), μέσω του STACKS ήταν επιτυχής, ότι η ποιότητα της συναρμολόγησης του γονιδιώματος είναι σε πρώτο επίπεδο επαρκής, και ότι οι RAD τόποι που συνθέτουν τον χάρτη δεν εντοπίζονται, σε επαναλαμβανόμενες περιοχές των γονιδιωμάτων στόχων και πιθανότατα, στην πλειοψηφία τους ούτε και στο γονιδίωμα του κρανίου. Είναι επίσης εμφανές ότι, το yellow croaker εμφανίζει αυξημένη ομολογία, με τον κρανίο, γεγονός το οποίο είναι αναμενόμενο δεδομένου ότι ανήκουν στην ίδια οικογένεια (Scianidae) ενώ επίσης, το λαβράκι (seabass) και η τσιπούρα (gilthead seabream) εμφανίζουν επίσης αυξημένη ομολογία, όπως φαίνεται και στις μοναδικές αντιστοιχίες (Homologs with one hit) σε σχέση με το medaka και την τιλάπια. Αυτό αντανακλάται επίσης και στην φυλογονιδιωματική δουλειά από Natsidis et al., 2019 της οποίας μέρος, δείχνεται στο **Σχήμα 7**.

Πίνακας 2. Καταγραφή των αντιστοιχιών ανά κατηγορία ανάμεσα στον γενετικό χάρτη σύνδεσης του κρανίου και των 5 γονιδιωμάτων των τελεόστεων

Species	Total number of homologs	Homologs in coding regions	Homologs in chromosomes	Coding region homologs with one hit	Homologs with one hit
European seabass	4545	1774	2875	542	3338
Yellow croaker	4830	2046	2884	827	3657
Gilt-head seabream	4342	1743	2972	648	3301
Medaka	2417	1471	1686	607	1873
Tilapia	5692	1518	2309	423	2750

Από τα αποτελέσματα της συγκριτικής γονιδιωματικής ανάλυσης σε επίπεδο αριθμού μοναδικών ομολογιών καθώς και συνολικών ομολογιών, φαίνεται ότι ο κρανιός, εμφανίζει μεγαλύτερες ομολογίες με τον λαβράκι το οποίο φαίνεται να είναι φυλογενετικά κοντινότερα στο yellow croaker (*L. crocea*) από ότι η τσιπούρα (*S. aurata*). Κάτι το οποίο προκύπτει, αδρά και πρωτόλνα από τις ομολογίες, χωρίς να σημαίνει ότι αποτελεί επιβεβαίωση του γεγονότος, παρά μία ένδειξη.



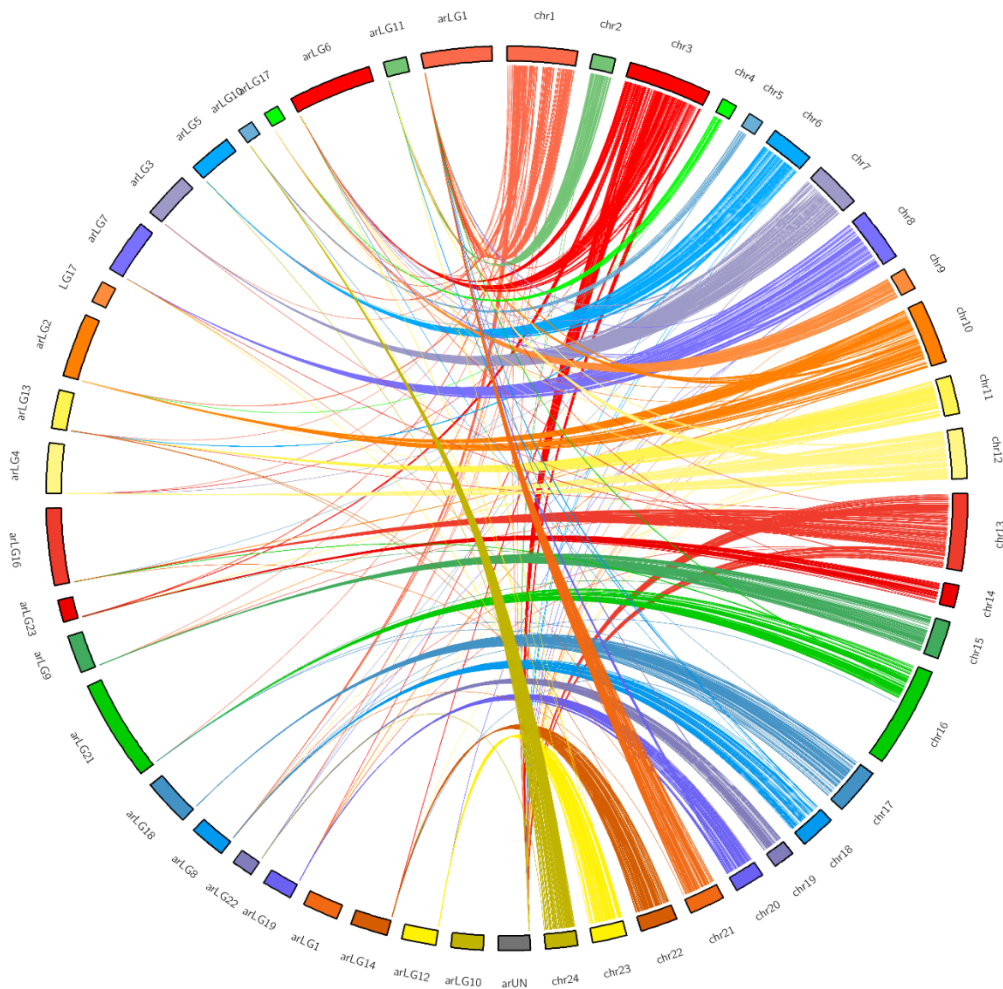
Σχήμα 7. Φυλογενετικές σχέσεις τελεόστεων ιχθύων, όπως παρουσιάστηκαν στην μελέτη Natsidis, Tsakogiannis, Pavlidis, Tsigenopoulos & Manousaki, 2019

Πίνακας 3. Ομολογίες σε επίπεδο χρωμοσωμάτων, όπως προέκυψαν από την συγκριτική γονιδιωματική ανάλυση. Στην πρώτη στήλη βρίσκονται τα Linkage group του κρανιού ενώ στις επόμενες 5 βρίσκονται οι αντιστοιχίες των ομόλογων χρωμοσωμάτων σε λαβράκι, yellow croaker, τσιπούρα (gilt-head seabream), medaka, tilapia. Οι ομολογίες έχουν σαν βάση τον κρανιό, για παράδειγμα, το LG1 του κρανιού είναι ομόλογο με τα LG2/LG6 του λαβρακιού, του CHR1/21 του yellow croaker, του CHR8/CHR18 της τσιπούρας, του CHR6/CHR10 για το medaka και του CHR2 στην τιλάπια. Το LG2 του κρανιού είναι ομόλογο στο LG7 του λαβρακιού κ.ο.κ. Με πράσινο χρώμα εμφανίζονται τα ομόλογα χρωμοσώματα που συμπίπτουν με τον καρύοτυπο του γενετικού χάρτη, ενώ με κίτρινο χρώμα, τα χρωμοσώματα τα οποία εμφανίζουν άλλη σειρά με τον γενετικό χάρτη αλλά ίδια μεταξύ τους. Για παράδειγμα το LG13 του λαβρακιού είναι ομόλογο με το CHR13 του medaka, ενώ και τα 2 είναι ομόλογα με το LG3 του κρανιού.

Meagre	European seabass	Yellow croaker	Gilt-head bream	Medaka	Tilapia
1	LG2/LG6	CHR1/CHR21	CHR08/CHR18	CHR6/CHR10	CHR2
2	LG7	CHR10	CHR1	CHR1	CHR6
3	LG13	CHR7	CHR2	CHR13	CHR14
4	LG8	CHR12	CHR23	CHR8	CHR4
5	LG1A	CHR6	CHR6	CHR5	CHR5
6	LG20	CHR3	CHR5	CHR9	CHR12
7	LG5	CHR8	CHR4	CHR3	CHR1
8	LG15	CHR18	CHR9	CHR21	CHR22
9	LG22/25	CHR15	CHR7	CHR7	CHR19
10	LG12	CHR5/CHR24	CHR16	CHR22	CHR18
11	LG10	CHR2	CHR21	CHR17	CHR17
12	LG18/21	CHR23	CHR19	CHR20	CHR9
13	LG17	CHR11	CHR22	CHR24	CHR15
14	LG14	CHR22	CHR13	CHR14	CHR10
15	LG11	N/A	CHR15	CHR15	CHR13
16	LG16	CHR13	CHR17	CHR16	CHR11
17	LG19	CHR4/CHR9	CHR12	CHR12	CHR7
18	LG4	CHR17	CHR11	CHR4	CHR16
19	LGx	CHR20	CHR14	CHR23	CHR16
20	LG9	CHR16	CHR3	CHR11	CHR20
21	LG1B	CHR19	CHR20	CHR19	CHR8
22	LG24	CHR14	CHR24	CHR2	CHR21
23	LG3	N/A	CHR10	CHR18	CHR3
24	N/A	N/A	N/A	N/A	-

Βλέπουμε ότι η σειρά των χρωμοσωμάτων στον γενετικό χάρτη σύνδεσης δεν ανταποκρίνεται συνολικά με την ταξινόμηση των χρωμοσωμάτων των 5 ειδών. Συγκεκριμένα, τα χρωμοσώματα ή LG τα οποία εμφανίζονται ομόλογα με τον κρανιό και ανταποκρίνονται στην ταξινόμηση του καρύοτυπου του γενετικού χάρτη είναι : το chr1 από yellow croaker, το LG16 και το LG14 από το λαβράκι, το LG15 από την τσιπούρα, τα chr5, chr14, chr15, chr16 από το medaka, και τα chr4, chr5, chr20 από την τιλάπια. Είναι ενδιαφέρον το γεγονός ότι η διατήρηση της σειράς των ομόλογων χρωμοσωμάτων είναι περισσότερο όμοια με αυτή του medaka το οποίο όπως αναφέρθηκε έχει διατηρήσει τον

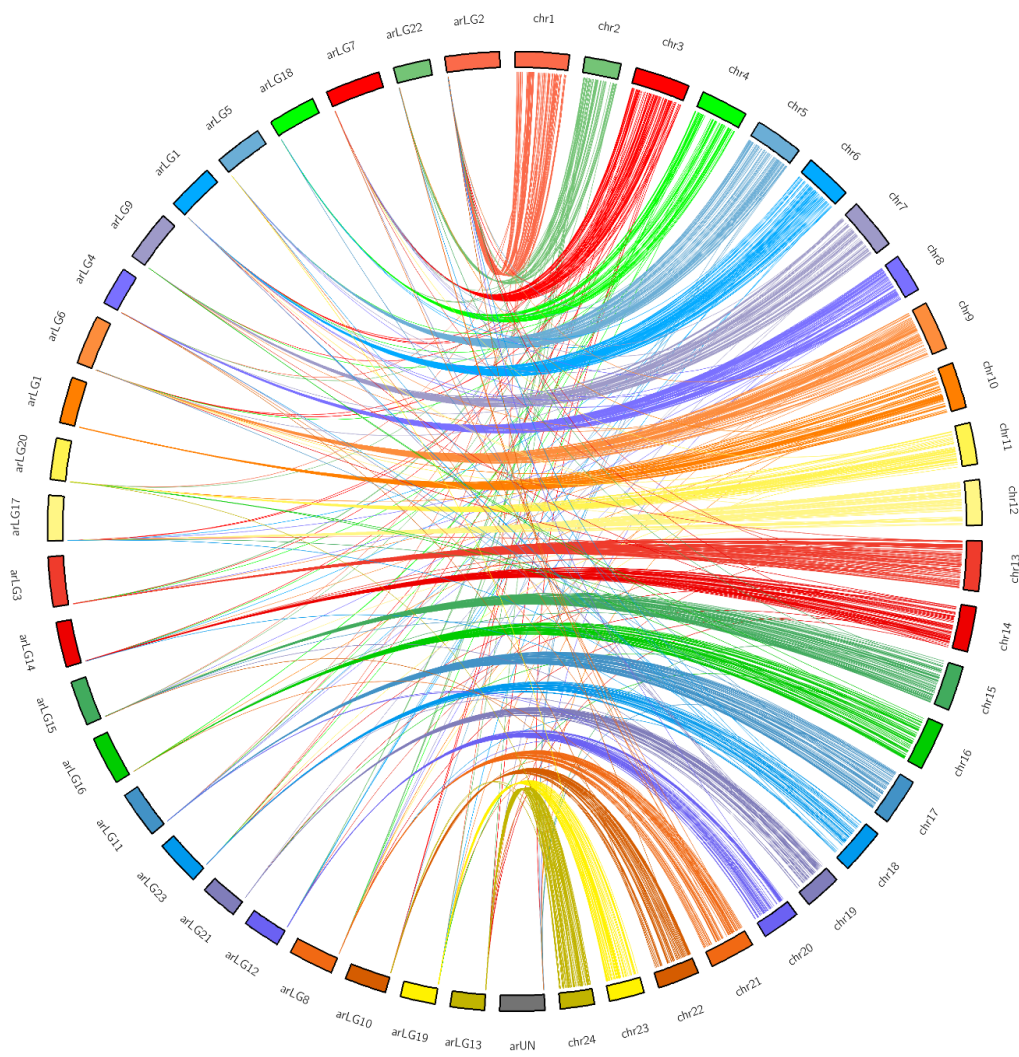
αρχικό αριθμό των χρωμοσωμάτων (n=24) του κοινού προγόνου των τελεόστεων για ~300 MYA. Για την οπτικοποίηση των ομολογιών του γενετικού χάρτη και των 5 γονιδιωμάτων χρησιμοποιήθηκε το λογισμικό Circos (Krzywinski et al. 2009). Παραθέτουμε εδώ, 3 ιδεογράμματα Circos (για όλα τα ιδεογραμματα βλ. Παράρτημα), συγκεκριμένα για το yellow croaker, το medaka και την τιλαπια.



Σχήμα 8. Στο αριστερό ημικύκλιο βρίσκονται τα LG του κρανιού και στο δεξί τα ομόλογα χρωμοσώματα του yellow croaker. Τα χρωμοσώματα είναι χρωματισμένα ανάλογα με την σειρά των χρωμοσωμάτων του yellow croaker, ούτως ώστε οι ομολογίες να ανταποκρίνονται και στο ίδιο χρώμα

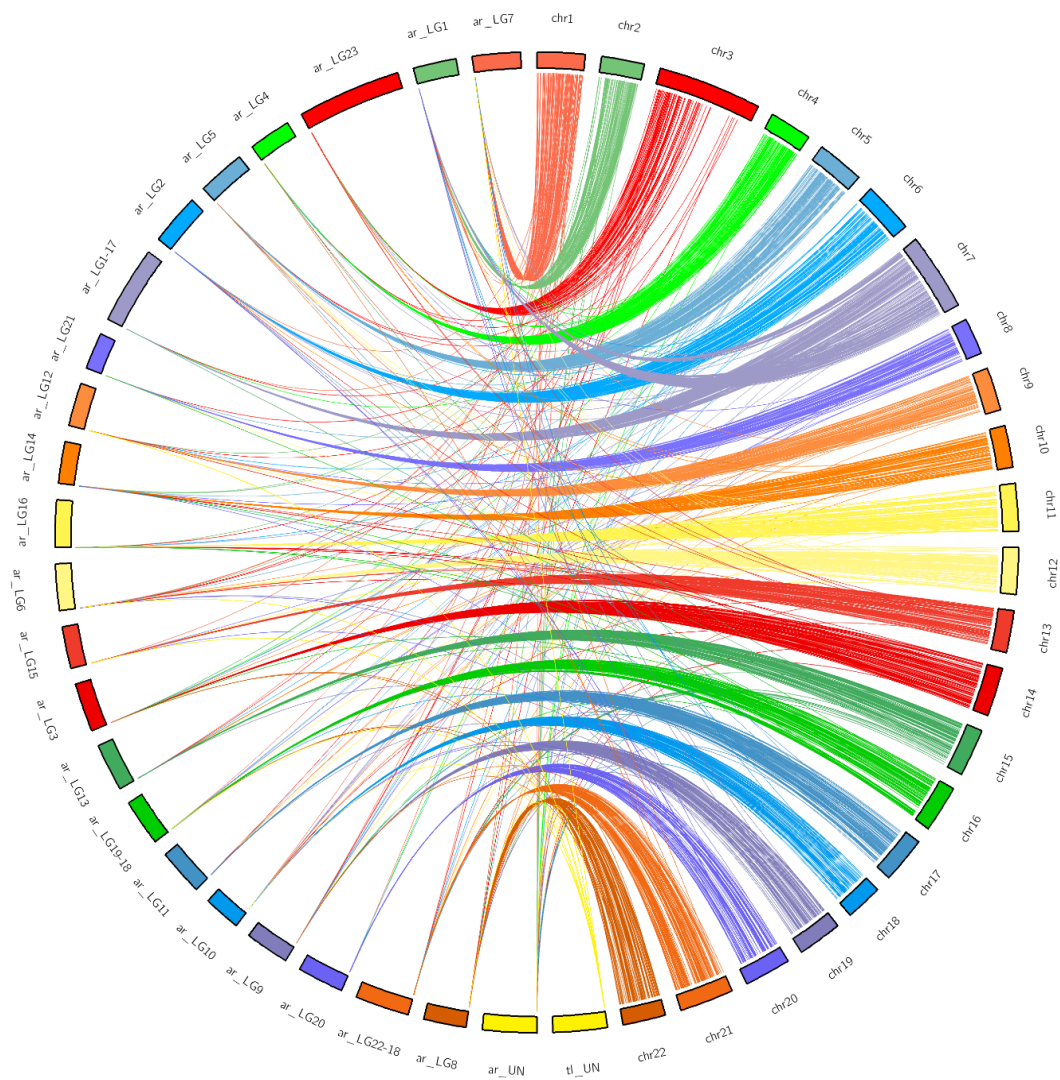
Στο **Σχήμα 8** βλέπουμε ότι οι ομολογίες μεταξύ του κρανιού και του yellow croaker, δεν είναι ιδανικές. Υπάρχει 1:1 ομολογία των χρωμοσωμάτων 4,5,6,7,14,15,16,17,18,19,20,22 του yellow croaker με τον κρανιό και φαίνεται ότι υπάρχει μία 2:1 σχέση μεταξύ συγκεκριμένων χρωμοσωμάτων του croaker : κρανιού. Συγκεκριμένα τα χρωμοσώματα 4 και 9 τα οποία αντιστοιχούν στο LG17 του κρανιού, όπως επίσης τα 1 και 21 αντιστοιχούν στο LG1 του κρανιού, και τα 5 και 24 αντιστοιχούν στο LG19 του κρανιού. Στην μελέτη των Xiao et al., 2015, κατασκευάζεται ο γενετικός χάρτης σύνδεσης προερχόμενος από μετάγραφα (transcripts) του yellow croaker, όπου παρατηρείται υψηλός βαθμός συντενίας με

το γονιδίωμα του medaka. Οι κωδικοποιούσες αλληλουχίες από τα μετάγραφα ευθυγραμμίζονται στο γονιδίωμα του yellow croaker και επιτυγχάνουν μία κάλυψη στο γονιδίωμα του yellow croaker 99.3%. Η κάλυψη των Xiao et al., (2015) είναι 3.5% μεγαλύτερη του 95.8% η οποία ήταν η κάλυψη που επιτεύχθη στην μελέτη του γονιδιώματος του yellow croaker στην αξιολόγηση της πληρότητας του assembly με 18,000 μετάγραφα. Συνεπώς, οι Xiao et al., χρησιμοποιώντας RNA-seq και κατασκευάζοντας έναν γενετικό χάρτη σύνδεσης, για την εύρεση συντενιών μεταξύ του yellow croaker και άλλων τελεόστεων, δεν λαμβάνουν υπόψη περιοχές στο γονιδίωμα του yellow croaker όπου το assembly είναι πιο πιθανό να περιέχει σφάλματα, όπως επαναλαμβανόμενες περιοχές, και παρέχουν έναν γενετικό χάρτη ο οποίος εμφανίζει υψηλό εύρος συντενίας με το medaka. Αυτά τα στοιχεία, θα μπορούσαν να εξηγήσουν την υψηλό βαθμό συντενίας που ανιχνεύσαμε μεταξύ του κρανιού και του medaka, αποτελέσματα που συγκλίνουν με τους Xiao et al., (2015) καθώς ο κρανιός είναι μέλος της οικογένειας των Sciaenidae, και από την άλλη μεριά, η μη αναμενόμενη μικρότερη συντενία του Κρανιού με το γονιδίωμα του yellow croaker είναι πιθανό να οφείλεται σε ατέλειες στο assembly του γονιδιώματος του yellow croaker. Αν και μία συγκριτική γονιδιωματική ανάλυση που ανιχνεύει ομολογίες σε κωδικοποιούσες και μή περιοχές είναι πιο πιθανό, να ανιχνεύσει χρωμοσομικές αναδιατάξεις καθώς εκπροσωπείται το εύρος του γονιδιώματος, δεν είναι φυσικά απόλυτο να συμβαίνει αυτό, καθώς μία σειρά παραμέτρων επηρεάζουν τα αποτελέσματα, όπως το βάθος αλληλούχισης των αναγνώσεων, η ποιότητα του base calling, ο αριθμός των αλληλουχιών σε σχέση με το μέγεθος του γονιδιώματος, η πολυπλοειδία.



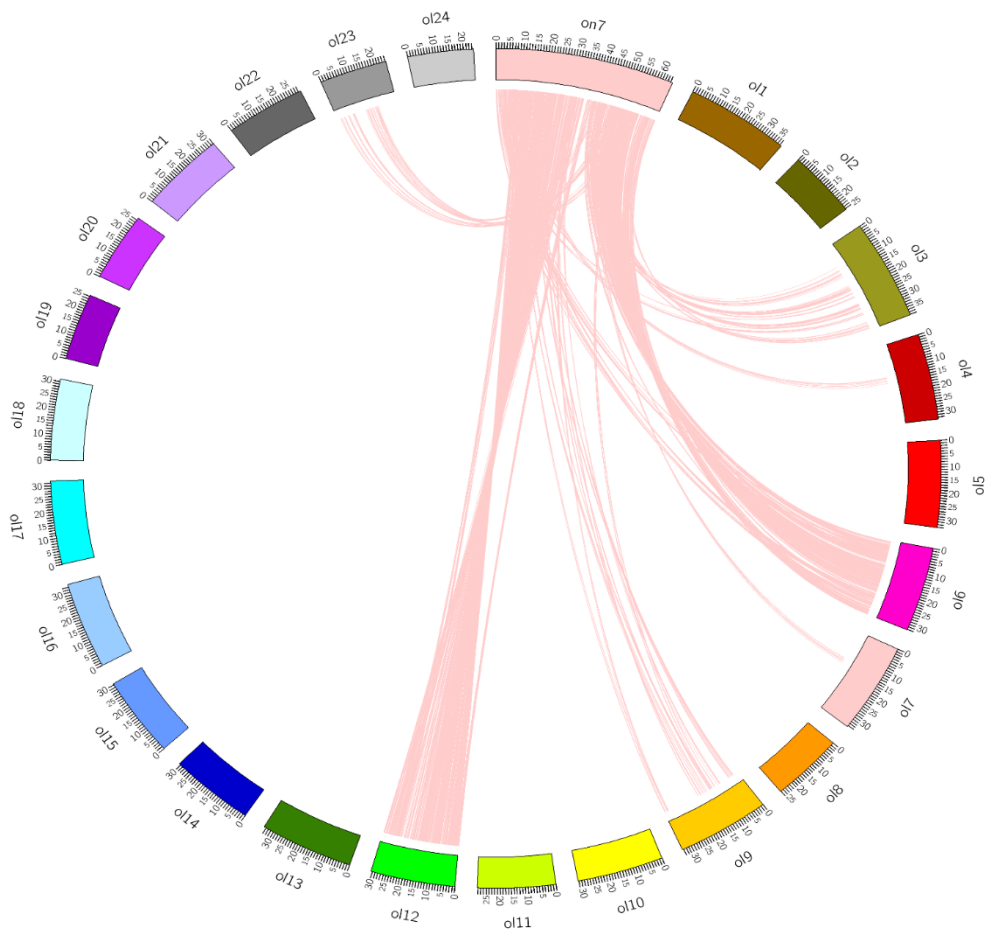
Σχήμα 9. Στο αριστερό ημικύκλιο βρίσκονται τα LG του κρανίου και στο δεξί τα ομόλογα χρωμοσώματα του medaka. Τα χρωμοσώματα είναι χρωματισμένα ανάλογα με την σειρά των χρωμοσωμάτων του medaka, ούτως ώστε οι ομολογίες να ανταποκρίνονται και στο ίδιο χρώμα.

Το ιδεόγραμμα των συνδέσεων του medaka με τον κρανίο στο **Σχήμα 9**, αντανακλά την συντηρημένη δομή του καρυότυπου των δύο ειδών με αρκετά αντιπροσωπευτικό τρόπο, όπου βλέπουμε μία 1:1 σχέση ανάμεσα σε όλα τα χρωμοσώματα των 2 ειδών.



Σχήμα 10. Στο αριστερό ημικύκλιο βρίσκονται τα LG του κρανίου και στο δεξί τα ομόλογα χρωμοσώματα της τιλάπιας. Τα χρωμοσώματα είναι χρωματισμένα ανάλογα με την σειρά των χρωμοσωμάτων του κρανίου, ούτως ώστε οι ομολογίες να ανταποκρίνονται και στο ίδιο χρώμα.

Ομοίως για την τιλάπια, μία παρόμοια εικόνα των 1:1 σχέσεων μεταξύ της τιλάπιας και του κρανίου με μία εξαίρεση. Το χρωμόσωμα 7 της τιλάπιας φαίνεται να εμφανίζει ομολογία σε 2 LG του κρανίου, στο LG1 και το LG17.



Σχήμα 11. Ιδεόγραμμα εύρεσης ομολογίας του χρωμοσώματος 7 της τιλάπια σε όλο το γονιδίωμα από το **medaka**. Η ανάλυση έχει γίνει στο [evol2circos](http://evol2circos.com) (203.190.147.148/evol2circos/) (Manmohan et al., 2020)

Το LG17 του κρανίου εμφάνισε ομολογία με το χρωμόσωμα 7 της τιλάπια, με το χρωμόσωμα 12 από το **medaka** καθώς επίσης και την σχέση 2:1 στο **yellow croaker** με τα χρωμοσώματα 4 και 9. Από το **Σχήμα 11**, βλέπουμε την ομολογία του χρωμοσώματος 7 στην τιλάπια με το χρωμόσωμα 12 του **medaka**. Λόγω του ότι τα χρωμοσώματα 4 και 9 του **yellow croaker** συνδέονται με το LG17 του κρανίου και πιθανότατα η διαφοροποίηση της οικογένειας των **Scianidae** από τον κοινό πρόγονο με το **medaka**, να έγινε αργότερα, συνεπώς εξελικτικά να αποτελεί νεότερη οικογένεια, το LG17 ίσως να αντανακλά ένα γεγονός χρωμοσωμικής αναδιάταξης στον κρανίο.

Με την κατασκευή ενός γενετικού χάρτη σχετικά υψηλής πυκνότητας για πρώτη φορά στον κρανίο, διερευνήσαμε πιθανά QTL για 2 χαρακτήρες ανάπτυξης, και ανιχνεύσαμε ένα QTL για το βάρος, καθώς επίσης αναδείξαμε τις ομολογίες μεταξύ των χρωμοσωμάτων ή LG του γενετικού χάρτη του κρανίου και άλλων 5 τελεόστεων ιχθύων. Τα αποτελέσματα αυτής της μελέτης, για την συγκριτική γονιδιωματική μελέτη τελεόστεων ιχθύων της οικογένειας **Scianidae**, καθώς επίσης και της ανίχνευσης γενετικών τόπων για ποσοτικούς χαρακτήρες ανάπτυξης, συμβάλλουν στην ευρύτερη κατανόηση της εξέλιξης των τελεόστεων καθώς επίσης παρέχουν χρήσιμα αποτελέσματα για την βελτίωση της παραγωγής του κρανίου ως είδος ανερχόμενο στην ιχθυοκαλλιέργεια.

5.1 Ευρήματα της διατριβής

Κεφάλαιο 2

- Αξιολόγηση της γενετικής παραλλακτικότητας σε αποθέματα με πολλαπλές γενιές, με χρήση παραμέτρων πληθυσμιακής γενετικής
- Εύρεση γενετικής ομαδοποίησης σε 11 αποθέματα Κρανιού σε εμπορικές συνθήκες με τον συνδυασμό δύο επιστημονικά έγκυρων μεθόδων
- Εκτίμηση σχέσεων συγγένειας, χρησιμοποιώντας τις βέλτιστες μεθόδους μέγιστης πιθανοφάνειας και εκτίμηση ατομικών συντελεστών ενδογαμίας
- Οικονομικά και χρονικά αποδοτική μεθοδολογία για την αξιολόγηση αποθεμάτων και σχεδιασμό νέων προγραμμάτων αναπαραγωγής

Κεφάλαιο 3

- Ανάπτυξη multiplex 12 γενετικών τόπων για τον Κρανιό
- Πρώτη προσπάθεια στον Κρανιό για την διερεύνηση της αθροιστικής γενετικής διακύμανσης σε σχέση με την φαινοτυπική, στους αναπτυξιακούς χαρακτήρες βάρος και μήκος
- Διερεύνηση της διακύμανσης της συνεισφοράς των γεννητόρων στις διαμορφούμενες οικογένειες
- Επιβεβαίωση της αποτελεσματικότητας της διαλογής. Οι απόγονοι της παρτίδας 1 έδειξαν 12% υψηλότερη κληρονομησιμότητα στο βάρος και 3% υψηλότερες εκτιμήσεις για το μήκος
- Εκτίμηση των ορίων της κληρονομησιμότητας για το βάρος και το μήκος στον Κρανιό

Κεφάλαιο 4

- Σύγκλιση αποτελεσμάτων με Natsidis et al., (2019) σε σχέση με το εύρος των ομολογιών ανάμεσα σε Κρανιό και άλλους τελεόστεους
- Συμφωνία με Tine et al., (2014) για 1:2 ομολογίες ανάμεσα σε Κρανιό – Λαβράκι
- Xiao et al., (2015) για την διατήρηση του καρυότυπου των 24 χρωμοσωμάτων Sciaenidae σε σχέση με μεντάκα
- Συμφωνία με Guyon et al., (2012), για ομολογίες Μεντάκα – Τιλάπια

5.2 Συνεισφορά της διατριβής και χρηματοδότηση

- **Multiplex γενετικών τόπων στον Κρανιό και Μαγιάτικο, τα οποία χρησιμοποιήθηκαν για πρώτη φορά**
 - α. Γενετική περιγραφή-δομή αποθεμάτων, εκτίμηση γενετικής παραλλακτικότητας
 - β. Ανάθεση γονέων στους απογόνους, εκτίμηση σχέσεων συγγένειας
 - γ. Εκτίμηση κληρονομησιμότητας, γενετικών και φαινοτυπικών συσχετίσεων
 - δ. Ατομικών συντελεστών ενδογαμίας
 - Πρώτος γενετικός χάρτης στον κρανιό, με υψηλής πυκνότητας NGS γενετικούς δείκτες
 - Ανάλυση ανίχνευσης QTL για το βάρος στον Κρανιό με στατιστικά σημαντικά αποτελέσματα
 - Συγκριτική γονιδιωματική ανάλυση για την βελτίωση του γονιδιώματος αναφοράς, πληροφορία για την γονιδιωματική σύνθεση (genomic makeup) του Κρανιού

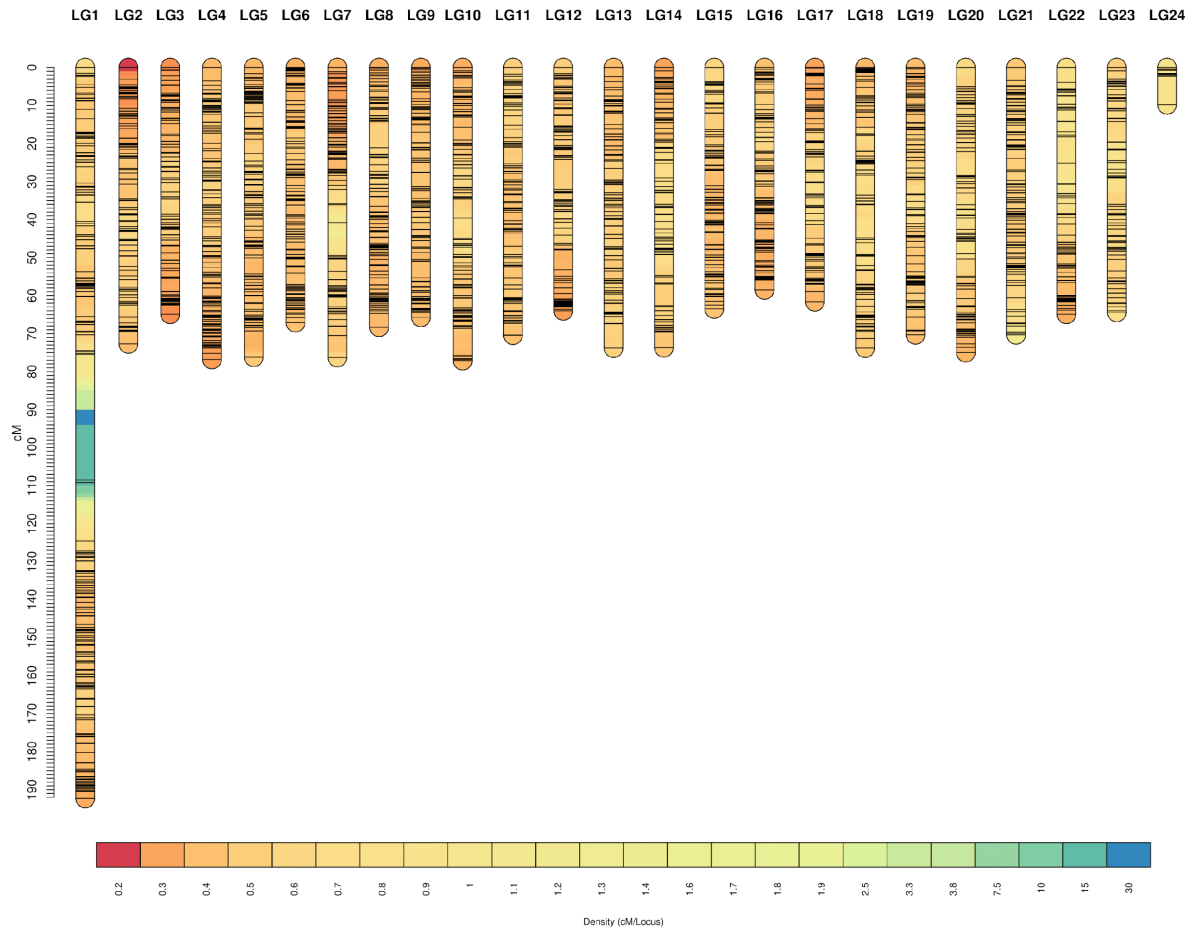
Η διατριβή χρηματοδοτήθηκε από

1. Το Ευρωπαϊκό πρόγραμμα PerformFISH

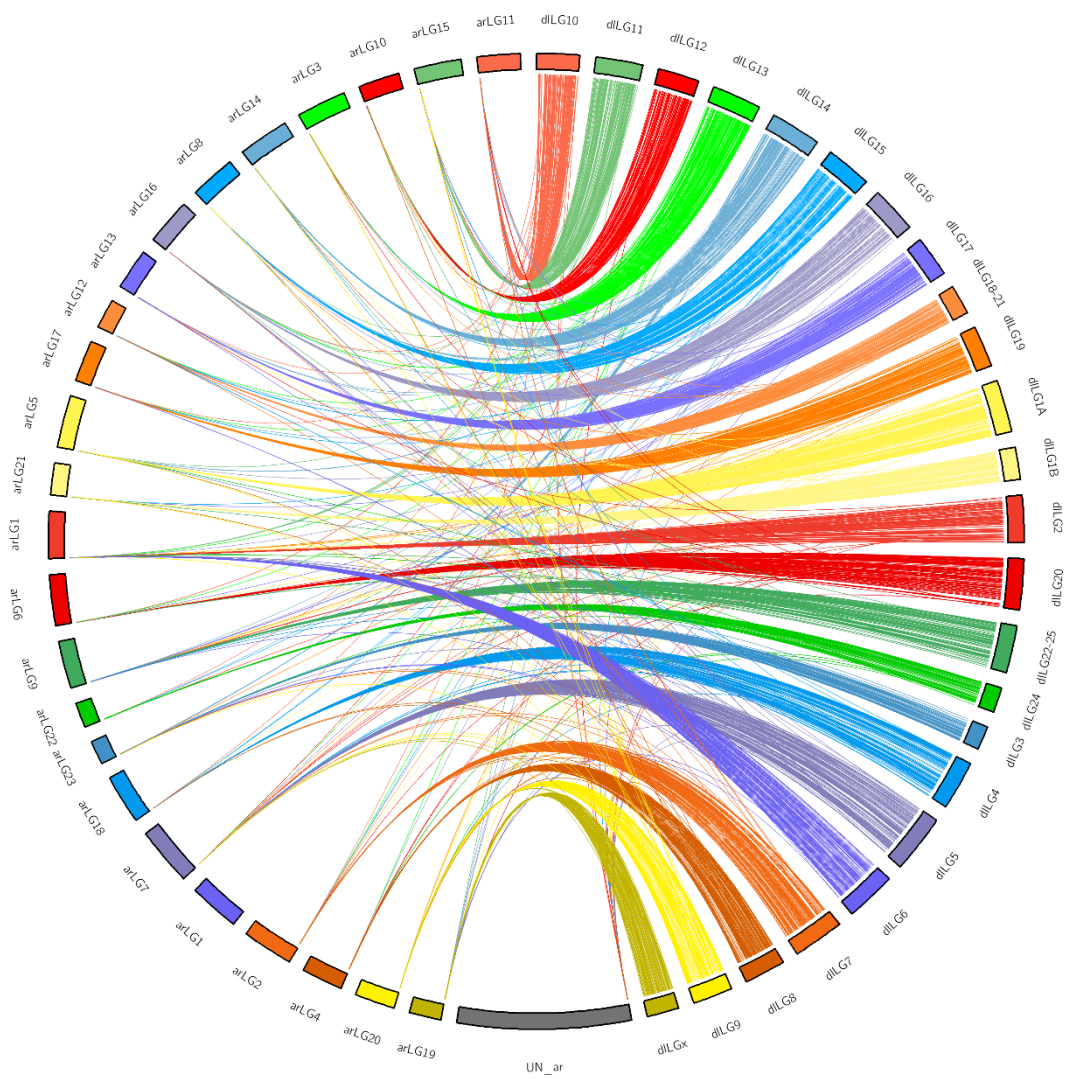
2. Το Ελληνικό πρόγραμμα ΕΠΑΛ "Διερεύνηση της παραλλακτικότητας μεγεθών σε καλλιεργούμενα ιχθύδια μαγιάτικου με στόχο την βελτίωση των πρακτικών παραγωγής και διαχείρισης" (Ε.Υ., Ν. Παπανδρουλακης)

3. Το Ελληνικό πρόγραμμα ΕΠΑΝΕΚ MeagreGEN "Γονιδιωματική Μελέτη του κρανιού (*Argyrosomus regius*) με σκοπό την χαρτογράφηση γονιδιακών τόπων ποσοτικών ιδιοτήτων και την έναρξη του πρώτου Ευρωπαϊκού προγράμματος γενετικής επιλογής" (Ε.Υ., Κ. Τσιγγενοπουλος)

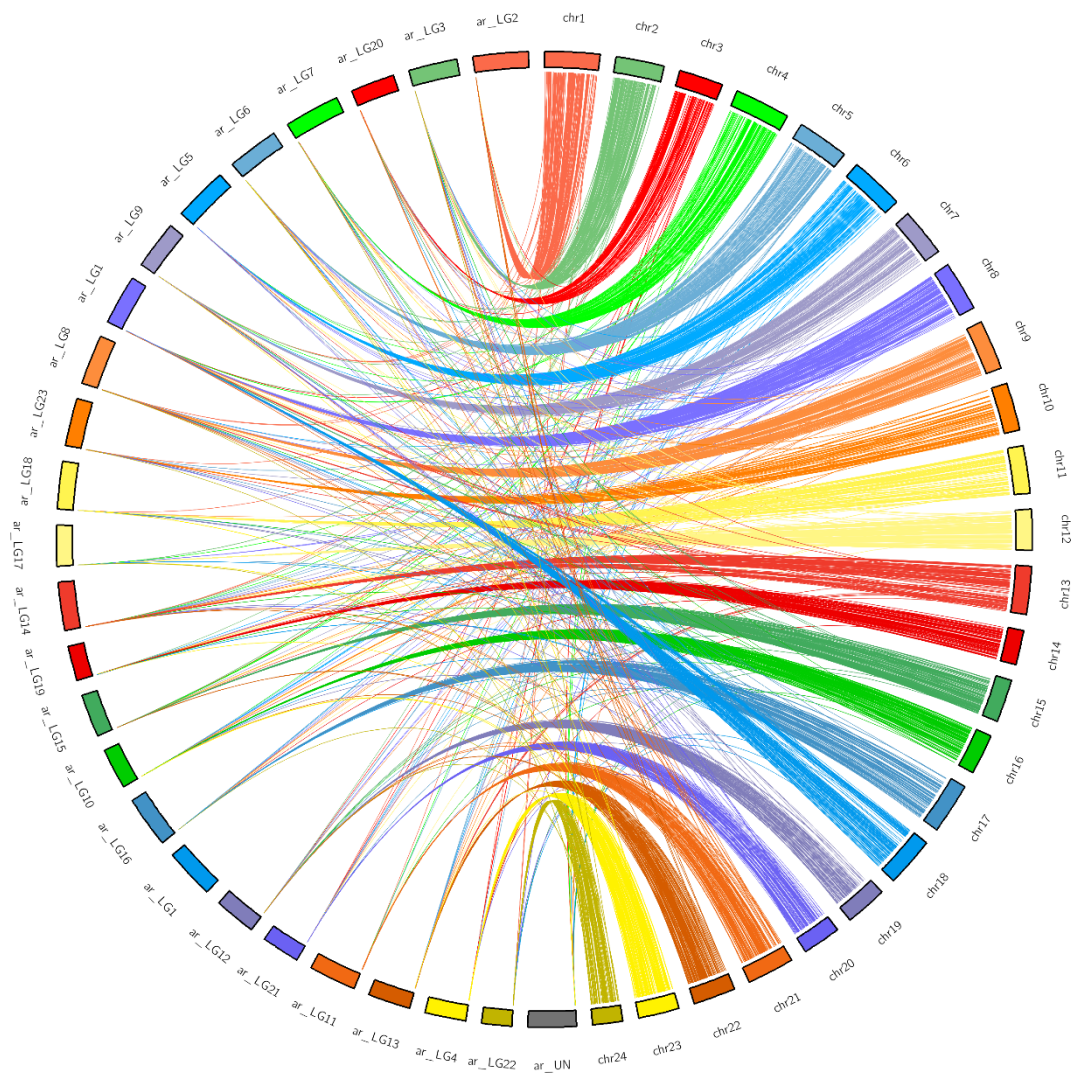
Παράρτημα



Σχήμα 1. Ο γενετικός χάρτης σύνδεσης προερχόμενος από τους τόπους με κάλυψη 90% και minor allele frequency 0.01, από τα δεδομένα προερχόμενα από την ddRAD βιβλιοθήκη στον κρανιό. Ο συγκεκριμένος χάρτης δεν έδωσε στατιστικά σημαντικά αποτελέσματα στην ανίχνευση QTL αλλά επέδειξε μικρότερα κενά στο LG1, γεγονός το οποίο αποδίδεται στην μεγαλύτερη κάλυψη σε κοινούς τόπους μεταξύ των δειγμάτων και συνεπώς στην υψηλότερη ανάλυση των χρωμοσωμικών ανασυνδυασμών όπως εκτιμήθηκαν με την διαδικασία που περιγράφηκε στην κατασκευή του γενετικού χάρτη σύνδεσης του κεφαλαίου 4.



Σχήμα 2. Στο αριστερό ημικύκλιο βρίσκονται τα LG του κρανίου και στο δεξί τα ομόλογα χρωμοσώματα του λαβρακιού. Τα χρωμοσώματα είναι χρωματισμένα ανάλογα με την σειρά των χρωμοσωμάτων του κρανίου, ούτως ώστε οι ομολογίες να ανταποκρίνονται και στο ίδιο χρώμα.



Σχήμα 3. Στο αριστερό ημικύκλιο βρίσκονται τα LG του κρανίου και στο δεξί τα ομόλογα χρωμοσώματα της τσιπούρας. Τα χρωμοσώματα είναι χρωματισμένα ανάλογα με την σειρά των χρωμοσωμάτων του κρανίου, ούτως ώστε οι ομολογίες να ανταποκρίνονται και στο ίδιο χρώμα.

Βιβλιογραφία

Abou Shabana, N. M., Abd El Rahman, S. H., Al Absawy, M. A., & Assem, S. S. (2012). Reproductive biology of *Argyrosomus regius* (Asso, 1801) inhabiting the south eastern Mediterranean Sea, Egypt. *The Egyptian Journal of Aquatic Research*, 38(2), 147-156.

Almudevar, A. (2003). A simulated annealing algorithm for maximum likelihood pedigree reconstruction. *Theoretical Population Biology*, 63, 63–75.

Alderson, L. Genetic conservation of domestic livestock. Wallingford, U.K.: CAB International, 1990. 242p.

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, 17, 81–92.

Andrews, S. "FastQC. A quality control tool for high throughput sequence data. Bioinformatics Group, Babraham Institute, Cambridge, United Kingdom." (2010).

Ao, Jingqun, et al. "Genome sequencing of the perciform fish *Larimichthys crocea* provides insights into molecular and genetic mechanisms of stress adaptation." *PLoS Genet* 11.4 (2015): e1005118.

Archangi, B., Chand, V., Mather, P.B., 2009. Isolation and characterization of 15 polymorphic microsatellite DNA loci from *Argyrosomus japonicus* (mulloway), a new aquaculture species in Australia. *Mol. Ecol. Resour.* 9, 412–414.

Ashton, David T., Peter A. Ritchie, and Maren Wellenreuther. "Fifteen years of quantitative trait loci studies in fish: challenges and future directions." *Molecular Ecology* 26.6 (2017): 1465-1476.

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One*, 3, e3376.

Bentsen, Hans B., et al. "Genetic improvement of farmed tilapias: Response to five generations of selection for increased body weight at harvest in *Oreochromis niloticus* and the further impact of the project." *Aquaculture* 468 (2017): 206-217.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, et al. (2008) Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PLoS ONE* 3(10): e3376.

Bansal, Vikas, Ali Bashir, and Vineet Bafna. "Evidence for large inversion polymorphisms in the human genome from HapMap data." *Genome research* 17.2 (2007): 219-230.

Bayona-Vásquez, N. J., Glenn, T. C., Kieran, T. J., Pierson, T. W., Hoffberg, S. L., Scott, P. A., Faircloth, B. C. (2019). Adapterama III: Quadrupleindexed, double/triple-enzyme RADseq libraries (2RAD/3RAD).

Belkhir, K_. "GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations." <http://www.genetix.univ-montp2.fr/genetix/genetix.htm> (2004).

Benfey, Philip N., and Thomas Mitchell-Olds. "From genotype to phenotype: systems biology meets natural variation." *Science* 320.5875 (2008): 495-497.

Beugin, Marie-Pauline, et al. "A fast likelihood solution to the genetic clustering problem." *Methods in Ecology and Evolution* 9.4 (2018): 1006-1016.

Birky, C. William, and J. Bruce Walsh. "Effects of linkage on rates of molecular evolution." *Proceedings of the National Academy of Sciences* 85.17 (1988): 6414-6418.

Blonk J.W., R., Komen, H., Kamstra, A., van Arendonk A.M., J., 2010. Effects of grading on heritability estimates under commercial conditions: A case study with common sole, *Solea solea*. *Aquaculture* 300 (1), 43–49.

Boudry, P., Collet, B., Cornette, F., Hervouet, V., Bonhomme, F., 2002. High variance in reproductive success of the Pacific oyster (*Crassostrea gigas*, Thunberg) revealed by microsatellite-based parentage analysis of multifactorial crosses. *Aquaculture* 204 (3–4), 283–296.

Broman, Karl W., et al. "Comprehensive human genetic maps: individual and sex-specific variation in recombination." *The American Journal of Human Genetics* 63.3 (1998): 861-869.

Broman, Karl W., and Saunak Sen. *A Guide to QTL Mapping with R/qtl*. Vol. 46. New York: Springer, 2009.

Broman, Karl W., et al. "R/qtl2: software for mapping quantitative trait loci with high-dimensional data and multiparent populations." *Genetics* 211.2 (2019): 495-502.

Broman, Karl W. "Fourteen years of R/qtl: Just barely sustainable." *Journal of Open Research Software* 2.1 (2014).

Brown, C.R., Woolliams, J.A., McAndrew, B.J., 2005. Factors influencing effective population size in commercial populations of gilthead seabream, *Sparus aurata*. *Aquaculture* 247, 219–225.

Catchen, Julian, et al. "Stacks: an analysis tool set for population genomics." *Molecular ecology* 22.11 (2013): 3124-3140.

Chakraborty, R., Shaw, M., & Schull, M. J. (1974). Exclusion of paternity: The current state of the art. *American Journal of Human Genetics*, 26, 477–488.

Chakraborty, R., & Schull, W. (1976). A note on the distribution of the number of exclusions to be expected in paternity testing. *American Journal of Human Genetics*.

Chatziplis, D., Batargias, C., Tsigenopoulos, C.S., Magoulas, A., Kollias, S., Kotoulas, G., Volckaert, F.A.M., Haley, C.S., 2007. Mapping quantitative trait loci in European sea bass (*Dicentrarchus labrax*): the BASSMAP pilot study. *Aquaculture* 272, S172–S182.

Cheng, Riyan, and Abraham A. Palmer. "A simulation study of permutation, bootstrap, and gene dropping for assessing statistical significance in the case of unequal relatedness." *Genetics* 193.3 (2013): 1015-1018.

Cockerham, C. Clark. "An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present." *Genetics* 39.6 (1954): 859.

Churchill, Gary A., and Rebecca W. Doerge. "Empirical threshold values for quantitative trait mapping." *Genetics* 138.3 (1994): 963-971.

D'Mello, Stacey AN, et al. "Signaling pathways in melanogenesis." *International journal of molecular sciences* 17.7 (2016): 1144.

De Jimenez, Monika K. Michalak, et al. "A radiation hybrid map of chromosome 1D reveals synteny conservation at a wheat speciation locus." *Functional & integrative genomics* 13.1 (2013): 19-32.

Dekkers JCM (2012) Application of genomics tools to animal breeding. *Current Genomics* 13, 207-212.

Devlin, B., Roeder, K., & Ellstrand, N. C. (1988). Fractional paternity assignment: Theoretical development and comparison to other methods. *Theoretical and Applied Genetics*, 76, 369–380.

DeWoody J. A. and Avise J. C. (2005) Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *Fish Biology*, 56, 3, 461-473.

Dong, S., Kong, J., Meng, X., Zhang, Q., Zhang, T., Wang, R., 2008. Microsatellite DNA markers associated with resistance to WSSV in *Penaeus (Fenneropenaeus) chinensis*. *Aquaculture* 282 (1–4), 138–141.

Dow, B. D., & Ashley, M. V. (1996). Microsatellite analysis of seed dispersal and parentage of saplings in bur oak, *Quercus macrocarpa*. *Molecular Ecology*, 5, 615–627.

Duncan, N.J., Estevez, A., Porta, J., Carazo, I., Norambuena, F., Aguilera, C., Gairin, I., Bucci, F., Valles, R., Mylonas, C.C., 2012. Reproductive development, GnRHa- induced spawning and egg quality of wild meagre (*Argyrosomus regius*) acclimatised to captivity. *Fish Physiol. Biochem.* 38, 1273–1286.

Duncan, N.J., Estevez, A., Fernandez-Palacios, H., Gairin, I., Hernandez-Cruz, C.M., Roo, F.J., Schuchardt, D., Valles, R., 2013. Aquaculture production of meagre (*Argyrosomus regius*): hatchery techniques, ongrowing and market. In: Allan, G., Burnell, G. (Eds.), *Advances in Aquaculture Hatchery Technology*. Woodhead Publishing Limited, Cambridge, UK, pp. 519–541.

Duncan, N.J., Mylonas, C.C., Milton Sullon, E., Karamanlidis, D., Franca Nogueira, M.C., Ibarra-Zatarain, Z., Chiumento, M., Aviles Carrillo, R.O., 2018. Paired spawning with male rotation of meagre *Argyrosomus regius* using GnRHa injections, as a method for producing multiple families for breeding selection programs. *Aquaculture* 495, 506–512.

Duncan, N., 2014. Dose-dependent effect of a single GnRH α injection on the spawning of meagre (*Argyrosomus regius*) broodstock reared in captivity. *Span. J. Agric. Res.* 12 (4), 1038–1048.

Dupont-Nivet, M., Vandeputte, M., Vergnet, A., Merdy, O., Haffray, P., Chavanne, H., Chatain, B., 2008. Heritabilities and G \times E interactions for growth in the European sea bass (*Dicentrarchus labrax* L.) using a marker-based pedigree. *Aquaculture* 275, 81–87.

Earl, D. A., and B. M. vonHoldt. "Structure Harvester: a website and program for visualizing STRUCTURE output and implementing the Evanno method Conservation Genetics Resources." *Resources* 3 (2011): 429-431.

Ellstrand, N. C. (1984). Multiple paternity within the fruits of the wild radish, *Raphanus sativus*. *The American Naturalist*, 123, 819–828.

Ellegren, H. (2000). Microsatellite mutations in the germline: Implications for evolutionary inference. *Trends in Genetics*, 12, 551–558.

Ellegren H (2014) Genome sequencing and population genomics in non-model organisms. *Trends in Ecology & Evolution* 29, 51-63.

Elshire RJ, Glaubitz JC, Sun Q, et al. (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6, e19379.

Elshire, R. J., Glaubitz, J. C., Sun, Q. I., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A Robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*, 6, e19379.

Evanno, Guillaume, Sebastien Regnaut, and Jérôme Goudet. "Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study." *Molecular ecology* 14.8 (2005): 2611-2620.

Fakriadis, I., & Mylonas, C. C. (2021). Sperm quality of greater amberjack *Seriola dumerili* throughout the reproductive season and in response to GnRH α treatment with controlled release implants. *Fish Physiology and Biochemistry*, 47(2), 281-292.

Falconer, D.S., Mackay, T.F.C., 1996. *Introduction to Quantitative Genetics*, fourth ed. Longman Group Limited, Harlow, Essex, U.K, pp. 464.

Falush, Daniel, Matthew Stephens, and Jonathan K. Pritchard. "Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies." *Genetics* 164.4 (2003): 1567-1587.

Farias, I.P., Muniz, L.B., Astolfi-Filho, S., Sampaio, I., 2006. Isolation and characterization of DNA microsatellite primers for *Cynoscion acoupa*, the most exploited sciaenid fish along the coast of Brazil. *Mol. Ecol. Notes* 6 (3), 660–663.

Ferguson, Moira M., and Roy G. Danzmann. "Role of genetic markers in fisheries and aquaculture: useful tools or stamp collecting?." *Canadian Journal of Fisheries and Aquatic Sciences* 55.7 (1998): 1553-1563.

Fernandes, T., Herlin, M., Belluga, M.D.L., Ballon, G., Martinez, P., Toro, M.A., Fernandez, J., 2017. Estimation of genetic parameters for growth traits in a hatchery population of gilthead sea bream (*Sparus aurata* L.). *Aquacult. Int.* 25, 499–514.

Fessehaye, Y., El-bialy, Z., Rezk, M.A., Crooijmans, R., Bovenhuis, H., Komen, H., 2006. Mating systems and male reproductive success in Nile tilapia (*Oreochromis niloticus*) in breeding hapas: a microsatellite analysis. *Aquaculture* 256, 148–158.

Fierst, Janna L. "Using linkage maps to correct and scaffold de novo genome assemblies: methods, challenges, and computational tools." *Frontiers in genetics* 6 (2015): 220.

Flint, Jonathan, and Trudy FC Mackay. "Genetic architecture of quantitative traits in mice, flies, and humans." *Genome research* 19.5 (2009): 723-733.

Fountoulaki, E., Grigorakis, K., Kounna, C., Rigos, G., Papandroulakis, N., Diakogeorgakis, J., Kokou, F., 2017. Growth performance and product quality of meagre (*Argyrosomus regius*) fed diets of different protein/lipid levels at industrial scale. *Ital. J. Anim. Sci.* 16, 685–694.

Franchini, P., Monné Parera, D., Kautt, A. F., & Meyer, A. (2017). quaddRAD: A new high-multiplexing and PCR duplicate removal ddRAD protocol produces novel evolutionary insights in a nonradiating cichlid lineage. *Molecular Ecology*, 26, 2783–2795.

Fu, Beide, et al. "A high-density genetic map and growth related QTL mapping in bighead carp (*Hypophthalmichthys nobilis*)." *Scientific reports* 6.1 (2016): 1-10.

Ganal, Martin W., Pierre Broun, and Steven D. Tanksley. "Genetic mapping of tandemly repeated telomeric DNA sequences in tomato (*Lycopersicon esculentum*)." *Genomics* 14.2 (1992): 444-448.

Garant, D., Kruuk, L.E.B., 2005. How to use molecular marker data to measure evolutionary parameters in wild populations: quantitative Genetics and selection estimates. *Mol. Ecol.* 14, 1843–1859.

Garbe, J. R. and Y. Da. (2008) Pedigraph: A Software Tool for the Graphing and Analysis of Large Complex Pedigree. User manual Version 2.4. Department of Animal Science, University of Minnesota.

Gessler, Damian DG, and Shizhong Xu. "Meiosis and the evolution of recombination at low mutation rates." *Genetics* 156.1 (2000): 449-456.

Ghozlan, A., Gaber, M.M., Zaki, M.A., Nour, A., 2018. Effect of stocking density on growth performance, production trait, food utilization and body composition, of meagre (*Argyrosomus regius*). *World J. Eng. Technol.* 6, 37–47.

Gibson, Greg. "Rare and common variants: twenty arguments." *Nature Reviews Genetics* 13.2 (2012): 135-145. Gjedrem, T., 2000. Genetic improvement of cold-water fish species. *Aquacult. Res.* 31, 25–33.

Grau, A., Crespo, S., Sarasquete, M. C., & De Canales, M. G. (1992). The digestive tract of the amberjack *Seriola dumerili*, Risso: a light and scanning electron microscope study. *Journal of Fish Biology*, 41(2), 287-303.

Gjedrem, Trygve. "The benefit of using selective breeding for aquatic species." *Annals of Aquaculture and Research* (2016).

T Gjedrem Selection and breeding programs in aquaculture, Å AKVAFORSK - 2005 – Springer. Gjedrem T., Gjøen H.M. & Gjerde B. (1991) Genetic origin of Norwegian farmed Atlantic salmon. *Aquaculture* 98, 41–50.

Gjedrem, T. and Robinson, N. (2014) Advances by Selective Breeding for Aquatic Species: A Review. *Agricultural Sciences*, 5, 1152-1158.

Glasauer, S.M.K., Neuhauss, S.C.F. Whole-genome duplication in teleost fishes and its evolutionary consequences. *Mol Genet Genomics* 289, 1045–1060 (2014).

Glasauer, Stella MK, and Stephan CF Neuhauss. "Whole-genome duplication in teleost fishes and its evolutionary consequences." *Molecular genetics and genomics* 289.6 (2014): 1045-1060.

Glaubitz, J. C., Rhodes, O. E. Jr, & DeWoody, J. A. (2003). Prospects for inferring pairwise relationships with single nucleotide polymorphisms. *Molecular Ecology*, 12, 1039–1047.

Goddard, Michael E., and Ben J. Hayes. "Mapping genes for complex traits in domestic animals and their use in breeding programmes." *Nature Reviews Genetics* 10.6 (2009): 381-391.

Gould SJ (2002). "Chapter 7, section "Synthesis as Hardening"". *The Structure of Evolutionary Theory*.

Guyon, Richard, et al. "A high-resolution map of the Nile tilapia genome: a resource for studying cichlids and other percomorphs." *BMC genomics* 13.1 (2012): 1-17.

Gowaty, P. A., & Karlin, A. A. (1984). Multiple maternity and paternity in single broods of apparently monogamous eastern bluebirds (*Sialia sialis*). *Behavioral Ecology and Sociobiology*, 15, 91–95.

Hadfield, J. D., Richardson, D. S., & Burke, T. (2006). Towards unbiased parentage assignment: Combining genetic, behavioural and spatial data in a Bayesian framework. *Molecular Ecology*, 15, 3715–3730.

Hanken, J., & Sherman, P. W. (1981). Multiple paternity in Belding's ground squirrel litters. *Science*, 212, 351–353.

Henderson, C. R. (1985). Best linear unbiased prediction of nonadditive genetic merits in noninbred populations. *Journal of animal science*, 60(1), 111-117.

Hemmer-Hansen J, Therkildsen NO, Pujolar JM (2014) Population genomics of marine fishes: next-generation prospects and challenges. *The Biological Bulletin* 227, 117-132. Tong J, Sun X (2015) Genetic and genomic analyses for economically important traits and their applications in molecular breeding of cultured fish. *Science China Life Sciences* 58, 178-186.

Herbinger, C.M., Doyle, R.W., Pitman, E.R., Paquet, D., Mesa, K.A., Morris, D.B., Wright, J.M., Cook, D., 1995. DNA fingerprint based analysis of paternal and maternal effects on offspring growth and survival in communally reared rainbow trout. *Aquaculture* 137, 245–256.

Herlin, M., Taggart, J.B., McAndrew, B.J., Penman, D.J., 2007. Parentage allocation in a complex situation: a large commercial Atlantic cod (*Gadus morhua*) mass spawning tank. *Aquaculture* 272, S195–S203.

Hill, W.G., 1979. A note on effective population size with overlapping generations. *Genetics* 92, 317–322 PMID:17248921 PMCID:PMC1213952. IBM Corp Released, 2016. IBM SPSS Statistics for Windows, Version 24.0. IBM Corp, Armonk, NY.

Hohenlohe, Paul A., et al. "Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout." *Molecular ecology resources* 11 (2011): 117-122.

Houston, R. D., Bean, T. P., Macqueen, D. J., Gundappa, M. K., Jin, Y. H., Jenkins, T. L., Robledo, D. (2020). Harnessing genomics to fast-track genetic improvement in aquaculture. *Nature Reviews Genetics*, 21(7), 389–409.

Hubby, J. L., & Lewontin, R. C. (1966). A molecular approach to the study of genic heterozygosity in natural populations. I. The number of alleles at different loci in *Drosophila pseudoobscura*. *Genetics*, 54, 577–594.

Imai HT, Crozier RH. Quantitative analysis of directionality in mammalian karyotype evolution. *The American Naturalist*. 1980; 116(4):537–69.

Janssen, Kasper, et al. "Impact of selective breeding on European aquaculture." *Aquaculture* 472 (2017): 8-16.

Jarne, P., & Lagoda, P. J. L. (1996). Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution*, 11, 424–429.

Jerez, S., Samper, M., Santamaría, F. J., Villamandos, J. E., Cejas, J. R., & Felipe, B. C. (2006). Natural spawning of greater amberjack (*Seriola dumerili*) kept in captivity in the Canary Islands. *Aquaculture*, 252(2-4), 199-207.

Jiang, D. L., Gu, X. H., Li, B. J., Zhu, Z. X., Qin, H., Meng, Z. N., ... Xia, J. H. (2019). Identifying a long QTL cluster across chrLG18 associated with salt tolerance in tilapia using GWAS and QTL-seq. *Marine Biotechnology*, 21, 250–261.

Jombart, Thibaut. "adegenet: a R package for the multivariate analysis of genetic markers." *Bioinformatics* 24.11 (2008): 1403-1405.

Jones, A. G., Small, C. M., Paczolt, K. A., & Ratterman, N. L. (2010). A practical guide to methods of parentage analysis. *Molecular Ecology Resources*, 10, 6–30.

Jones, A. G. (2001). gerud 1.0: A computer program for the reconstruction of parental genotypes from progeny arrays using multilocus DNA data. *Molecular Ecology Notes*, 1, 215–218.

Jones, A. G., & Avise, J. C. (1997a). Microsatellite analysis of maternity and the mating system in the Gulf pipefish *Syngnathus scovelli*, a species with male pregnancy and sex-role reversal. *Molecular Ecology*, 6, 203–213.

Jones, A. G., & Avise, J. C. (1997b). Polygynandry in the dusky pipefish *Syngnathus floridae* revealed by microsatellite DNA markers. *Evolution*, 51, 1611–1622.

Jones, A. G. (2005). gerud 2.0: A computer program for the reconstruction of parental genotypes from half-sib progeny arrays with known or unknown parents. *Molecular Ecology Notes*, 5, 708–711.

Jones, O., & Wang, J. (2010). COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources*, 10(3).

Jones, A. G., & Ardren, W. R. (2003). Methods of parentage analysis in natural populations. *Molecular Ecology*, 12, 2511–2523.

Jones, A.G., Ardren, W.R., 2003. Methods of parentage analysis in natural populations. *Mol. Ecol.* 12, 2511–2523.

Jones K. Robertsonian fusion and centric fission in karyotype evolution of higher plants. *The Botanical Review*. 1998; 64:273–89.

Kahraman A, Avramov A, Nashev LG, et al. (2005) PhenomicDB: a multi-species genotype/phenotype database for comparative phenomics. *Bioinformatics* 21, 418-420.

Kajungiro, R. A., Palaiokostas, C., Pinto, F. A. L., Mmochi, A. J., Mtolera, M., Houston, R. D., & de Koning, D. J. (2019). Population structure and genetic diversity of Nile tilapia (*Oreochromis niloticus*) strains cultured in Tanzania. *Frontiers in Genetics*, 10(1269), 1–12.

Kalinowski, S. T., Taper, M. L., & Marshall, T. C. (2007). Revising how the computer program cervus accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, 16, 1099–1106.

Kasahara, Masahiro, et al. "The medaka draft genome and insights into vertebrate genome evolution." *Nature* 447.7145 (2007): 714-719.

Kause, A., Ritola, O., Paananen, T., Mantysaari, E., Eskelinen, U., 2002. Coupling body weight and its composition: a quantitative genetic analysis in rainbow trout. *Aquaculture* 211, 65–79.

Kellogg, K. A., Markert, J. A., Stauffer, J. R., & Kocher, T. D. (1995). Microsatellite variation demonstrates multiple paternity in lekking cichlid fishes from Lake Malawi, Africa. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 260, 79–84.

Khaw, H.L., Bovenhuis, H., Ponzoni, R.W., Rezk, M.A., Charo-Karisa, H., Komen, H., 2009. Genetic analysis of Nile tilapia (*Oreochromis niloticus*) selection line reared in two input environments. *Aquaculture* 294, 37–42.

Kimura, M. (1968). Genetic variability maintained in a finite population due to mutational production of neutral and nearly neutral isoalleles. *Genetical Research*, 11(3), 247-270.

Kincaid Harold L. Rotational Line Crossing: An Approach to the Reduction of Inbreeding Accumulation in Trout Brood Stocks, *The Progressive Fish-Culturist*, 39, (1977).

Ko, A., & Nielsen, R. (2017). Composite likelihood method for inferring local pedigrees. *PloS Genetics*, 13, e1006963.

Kovač, M., Groeneveld, E., Garcia-Cortes, L.A., 2002. VCE-5: a Package for the Optimization of Dispersion Parameters. *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*. 20–23 August 2002, Montpellier, France.

Kozul, V., Skaramuca, B., Glamuzina, B., Glavic, N., Tutman, P., 2001. Comparative gonadogenesis and hormonal induction of spawning of cultured and wild Mediterranean amberjack (*Seriola dumerili*, Risso 1810). *Sci. Mar.* 65: 215-220.

Kreitman, Martin, and Richard R. Hudson. "Inferring the evolutionary histories of the Adh and Adh-dup loci in *Drosophila melanogaster* from patterns of polymorphism and divergence." *Genetics* 127.3 (1991): 565-582.

Kruglyak, Leonid. "The use of a genetic map of biallelic markers in linkage studies." *Nature genetics* 17.1 (1997): 21-24.

Krzywinski, Martin, et al. "Circos: an information aesthetic for comparative genomics." *Genome research* 19.9 (2009): 1639-1645.

Kucinski, M., Fopp-Bayat, D., Liszewski, T. *et al.* Genetic analysis of four European huchen (*Hucho hucho* Linnaeus, 1758) broodstocks from Poland, Germany, Slovakia, and Ukraine: implication for conservation. *J Appl Genetics* 56, 469–480 (2015).

Kulathinal, Rob J., Laurie S. Stevison, and Mohamed AF Noor. "The genomics of speciation in *Drosophila*: diversity, divergence, and introgression estimated using low-coverage genome sequencing." *PLoS Genet* 5.7 (2009): e1000550.

Lazzari, A., & Barbera, G. (1989). Farming the Mediterranean yellowtail, *Seriola dumerilii* (Risso 1810) in concrete ponds: results and perspectives.

Lee, Chaeyoung, and Jungdae Park. "Estimation of epistasis among finite polygenic loci for complex traits with a mixed model using Gibbs sampling." *Journal of biomedical informatics* 40.5 (2007): 500-506.

Lehner B (2013) Genotype to phenotype: lessons from model organisms for human genetics. *Nature Reviews Genetics* 14, 168-178.

Lenormand, Thomas. "The evolution of sex dimorphism in recombination." *Genetics* 163.2 (2003): 811-822. Lewin, Harris A., et al. "Every genome sequence needs a good map." *Genome research* 19.11 (2009): 1925-1928.

Li, Hong Lian, et al. "Genome-wide QTL analysis identified significant associations between hypoxia tolerance and mutations in the GPR132 and ABCG4 genes in Nile tilapia." *Marine biotechnology* 19.5 (2017): 441-453.

Li, B. J., Zhu, Z. X., Gu, X. H., Lin, H. R., & Xia, J. H. (2019). QTL mapping for red blotches in Malaysia Red Tilapia (*Oreochromis* spp.). *Marine Biotechnology*, 21, 384–395.

Li, H. L., Gu, X. H., Li, B. J., Chen, C. H., Lin, H. R., & Xia, J. H. (2017). Genome-wide QTL analysis identified significant associations between hypoxia tolerance and mutations in the GPR132 and ABCG4 genes in Nile Tilapia. *Marine Biotechnology*, 19, 441–453.

Li, Heng, and Richard Durbin. "Fast and accurate short read alignment with Burrows–Wheeler transform." *bioinformatics* 25.14 (2009): 1754-1760.

Liu, Haiyang, et al. "A high-density genetic linkage map and QTL fine mapping for body weight in crucian carp (*Carassius auratus*) using 2b-RAD sequencing." *G3: Genes, genomes, genetics* 7.8 (2017): 2473-2487.

Liu D, Zhou Y, Yang K, Zhang X, Chen Y, Li C, Li H, Song Z (2018) Low genetic diversity in broodstocks of endangered Chinese sucker, *Myxocyprinus asiaticus*: implications for artificial propagation and conservation. *ZooKeys* 792: 117–132.

Liu, X., Zhao, G., Wang, Z., Cai, M., Ye, H., Wang, Q., 2012. Parentage assignment and parental contribution analysis in large yellow croaker *Larimichthys crocea* using microsatellite markers. *Curr. Zool.* 58, 244–249.

Loughnan, Shannon R., et al. "Genetic diversity and relatedness estimates for captive barramundi (*Lates calcarifer*, Bloch) broodstock informs efforts to form a base population for selective breeding." *Aquaculture Research* 47.11 (2016): 3570-3584.

Loughnan, S.R., Domingos, J.A., Smith-Keune, C., Forrester, J.P., Jerry, D.R., Beheregaray, L.B., Robinson, N.A., 2013. Broodstock contribution after mass spawning and size grading in barramundi (*Lates calcarifer*, Bloch). *Aquaculture* 404–405, 139–149.

Lynch, M., Walsh, B., 1998. *Genetic and Analysis of Quantitative Traits*. Sinauer Associates Inc., Sunderland, USA.

Lynch, Michael, and Bruce Walsh. *Genetics and analysis of quantitative traits*. Vol. 1. Sunderland, MA: Sinauer, 1998.

Mackay TFC (2001) The genetic architecture of quantitative traits. *Annual Review of Genetics* 35,303-339.

Manolio, Teri A., et al. "Finding the missing heritability of complex diseases." *Nature* 461.7265 (2009): 747-753.

Manousaki, Tereza, et al. "Exploring a nonmodel teleost genome through rad sequencing—linkage mapping in Common Pandora, *Pagellus erythrinus* and comparative genomic analysis." *G3: Genes, genomes, genetics* 6.3 (2016): 509-519.

Marino, G., M. Porrello and S. Andaloro. – 1995. Aspects of reproductive biology of the Mediterranean amberjack (*Serioladumerili*Risso, 1810): Gonadal development, *Cah. OptionsMediterr.*16: 115-124.

Marshall, T. C., Slate, J., Kruuk, L. E. B., & Pemberton, J. M. (1998). Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology*, 7, 639–655.

Masel J, King OD, Maughan H (January 2007). "The loss of adaptive plasticity during long periods of environmental stasis". *The American Naturalist*. University of Chicago Press on behalf of the American Society of Naturalists. 169 (1): 38–46.

Meffe Gary K. *Conservation Genetics and the Management of Endangered Fishes, Fisheries*, Pages 14-23 , (2011).

Meagher, T. R., & Thompson, E. A. (1986). The relationship between single and parent pair genetic likelihoods in genealogy reconstruction. *Theoretical Population Biology*. Meagher, T. R. (1986). Analysis of paternity within a natural population of *Chamaelirium luteum*. 1. Identification of most-likely male parents. *The American Naturalist*.

Moses, M., Mtolera, M. S. P., Chauka, L. J., Lopes, F. A., de Koning, D. J., Houston, R. D., & Palaiokostas, C. (2019). Characterizing the genetic structure of introduced Nile tilapia (*Oreochromis niloticus*) strains in Tanzania using double digest RAD sequencing. *Aquaculture International*, 28, 477–492.

Micale, V., G. Maricchiolo and V. Genoveze. – 1999. The reproductive biology of the amberjack *Seriola dumerili*(Risso 1810). In:Oocyte development in captivity. *Aquacult. Res.*30(5): 349-355

Miller, Michael R., et al. "Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers." *Genome research* 17.2 (2007): 240-248.

Miller, Robert Hoover, J. E. Legates, and C. Clark Cockerham. "Estimation of nonadditive hereditary variance in traits of mice." *Genetics* 48.2 (1963): 177.

Miller, S.A., Dykes, D.D., Polesky, H.F., 1988. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 16 (3). I R L Press Limited, Oxford, England.

Miller, S.A., Dykes, D.D., Polesky, H.F., 1988. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 16 (3).

Milner, Pemberton, Brotherstone, Albon, 2000. Estimating variance components and heritabilities in the wild: a case study using the "animal model" approach. *J. Evol. Biol.* 13, 804–813.

Mirimin, L., Roodt-Wilding, R., 2015. Testing and validating a modified CTAB DNA extraction method to enable molecular parentage analysis of fertilized eggs and larvae of an emerging South African aquaculture species, the dusky kob *Argyrosomus japonicus*: dna extraction from *Argyrosomus japonicus* eggs and larvae. *J. Fish Biol.* 86, 1218–1223.

Moghadam, Hooman K., et al. "Quantitative trait loci for body weight, condition factor and age at sexual maturation in Arctic charr (*Salvelinus alpinus*): comparative analysis with rainbow trout (*Oncorhynchus mykiss*) and Atlantic salmon (*Salmo salar*)." *Molecular Genetics and Genomics* 277.6 (2007): 647-661.

Mohindra, Vindhya, et al. "Cross-species amplification of *Catla catla* microsatellite locus in *Labeo rohita*." *Indian Journal of Fisheries* 48.1 (2001): 103-108.

Montoya-López, Andrés F., et al. "Genetic diversity of four broodstocks of tilapia (*Oreochromis* sp.) from Antioquia, Colombia." *Revista Colombiana de Ciencias Pecuarias* 32.3 (2019): 201-213.

Moore, S.S., Whan, V., Davis, G.P., Byrne, K., Hetzel, D.J., Preston, N., 1999. The development and application of genetic markers for the Kuruma prawn *Penaeus japonicus*. *Aquaculture* 173, 19–32.

Mylonas, C.C., Mitrizakis, N., Papadaki, M., Sigelaki, I., 2013. Reproduction of hatcheryproduced meagre *Argyrosomus regius* in captivity I. Description of the annual reproductive cycle. *Aquaculture* 414–415, 309–317.

Morin, P. A., Wallis, J., Moore, J. J., & Woodruff, D. S. (1994). Paternity exclusion in a community of wild chimpanzees using hypervariable simple sequence repeats. *Molecular Ecology*, 3, 469–478.

Mylonas, C.C., Fatira, E., Karkut, P., Sigelaki, I., Papadaki, M., Duncan, N., 2015. Reproduction of hatchery-produced meagre *Argyrosomus regius* in captivity III. Comparison between GnRH_a implants and injections on spawning kinetics and egg/ larval performance parameters. *Aquaculture* 448, 44–53.

Mylonas, C.C., Salone, S., Biglino, T., de Mello, P.H., Fakriadis, I., Sigelaki, I., Duncan, N., 2016. Enhancement of oogenesis/spermatogenesis in meagre *Argyrosomus regius* using a combination of temperature control and GnRH_a treatments. *Aquaculture* 464, 323–330.

Nachman MW, Searle JB. Why is the house mouse karyotype so variable? *Trends in Ecology & Evolution.* 1995; 10:397–402.

Natsidis, Paschalis, et al. "Phylogenomics investigation of sparids (Teleostei: Spariformes) using high-quality proteomes highlights the importance of taxon sampling." *Communications biology* 2.1 (2019): 1-10.

Neff, B. D., Repka, J., & Gross, M. R. (2001). A Bayesian framework for parentage analysis: The value of genetic and other biological data. *Theoretical Population Biology*, 59, 315–331.

Nelder, J. A. "Regression, model-building and invariance." *Journal of the Royal Statistical Society: Series A (General)* 131.3 (1968): 303-315.

Nelder, J. A. "Weighted regression, quantal response data, and inverse polynomials." *Biometrics* (1968): 979-985.

Nielsen, R., Mattila, D. K., Clapham, P. J., & Palsbøll, P. J. (2001). Statistical approaches to paternity analysis in natural populations and applications to the North Atlantic humpback whale. *Genetics*, 157, 1673–1682.

Nielsen, Einar Eg, Michael Møller Hansen, and V. Loeschcke. "Analysis of microsatellite DNA from old scale samples of Atlantic salmon *Salmo salar*: a comparison of genetic composition over 60 years." *Molecular Ecology* 6.5 (1997): 487-492.

O'Connell, Michael, et al. "Differentiation of rainbow trout (*Oncorhynchus mykiss*) populations in Lake Ontario and the evaluation of the stepwise mutation and infinite allele mutation models using microsatellite variability." *Canadian Journal of Fisheries and Aquatic Sciences* 54.6 (1997): 1391-1399.

O'Malley, Kathleen G., et al. "Microsatellite DNA markers for kinship analysis and genetic mapping in red drum, *Sciaenops ocellatus* (Sciaenidae, Teleostei)." *Molecular Ecology Notes* 3.1 (2003): 155-158.

O'Malley, K.G., Abbey, C.A., Ross, K., Gold, J.R., 2003. Microsatellite DNA markers for kinship analysis and genetic mapping in red drum, *Sciaenops ocellatus* (Sciaenidae, Teleostei). *Mol. Ecol. Notes* 3 (1), 155–158.

O'Reilly, P., Wright, J.M., 1995. The evolving technology of DNA fingerprinting and its application to fisheries and aquaculture. *J. Fish Biol.* 47, 29–55.

Ott J, Kamatani Y, Lathrop M (2011) Family-based designs for genome-wide association studies. *Nature Reviews* 12, 465-474.

Otto, Sarah P., and Thomas Lenormand. "Resolving the paradox of sex and recombination." *Nature Reviews Genetics* 3.4 (2002): 252-261.

Ouellette, Lisa A., et al. "LinkageMapView—rendering high-resolution linkage and QTL maps." *Bioinformatics* 34.2 (2018): 306-307.

Ozaki, Akiyuki, et al. "Quantitative trait loci (QTL) associated with resistance to a monogenean parasite (*Benedenia seriola*) in yellowtail (*Seriola quinqueradiata*) through genome wide analysis." *PloS one* 8.6 (2013): e64987.

Palaiokostas, C., Bekaert, M., Khan, M., Taggart, J. B., Gharbi, K., McAndrew, B. J., & Penman, D. J. (2015). A novel sex-determining QTL in Nile tilapia (*Oreochromis niloticus*). *BMC Genomics*, 16, 171.

Pandey, Manmohan, et al. "Evol2Circos: A web-based tool for genome synteny and collinearity analysis and its visualization in fishes." *Journal of Heredity* 111.5 (2020): 486-490.

Paterson, Andrew H., et al. "Fine mapping of quantitative trait loci using selected overlapping recombinant chromosomes, in an interspecies cross of tomato." *Genetics* 124.3 (1990): 735-742.

Peakall R, Smouse PE. GenAEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. *Bioinformatics*. 2012 Oct 1;28(19):2537-9.

Peng, Wenzhu, et al. "An ultra-high density linkage map and QTL mapping for sex and growth-related traits of common carp (*Cyprinus carpio*)." *Scientific reports* 6.1 (2016): 1-16.

Perez-Enriquez, R., Takagi, M., Taniguchi, N., 1999. Genetic variability and pedigree tracing of a hatchery-reared stock of red sea bream *Pagrus major* used for stock enhancement, based on microsatellite DNA markers. *Aquaculture* 173 (1–4), 413–423.

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for De Novo SNP discovery and genotyping in model and non-model species. *PLoS One*, 7, e37135.

Poli, B.M., Parisi, G., Zampacavallo, G., Iurzan, F., Mecatti, M., Lupi, P., Bonelli, A., 2003. Preliminary results on quality and quality changes in reared meagre (*Argyrosomus regius*): body and fillet traits and freshness changes in refrigerated commercial-size fish. *Aquacult. Int.* 11 (3), 301–311.

Posada, David, and Keith A. Crandall. "The effect of recombination on the accuracy of phylogeny estimation." *Journal of molecular evolution* 54.3 (2002): 396-402.

Powell, W., Machray, G. C., & Provan, J. (1996). Polymorphism revealed by simple sequence repeats. *Trends in Plant Science*, 1, 215–222.

Rastas, Pasi, et al. "Construction of ultradense linkage maps with Lep-MAP2: stickleback F 2 recombinant crosses as an example." *Genome biology and evolution* 8.1 (2016): 78-93.

Rastas, Pasi, et al. "Lep-MAP: fast and accurate linkage map construction for large SNP datasets." *Bioinformatics* 29.24 (2013): 3128-3134.

Renshaw Mark A., John C. Patton, Caird E. Rexroad, John R. Gold, Isolation and characterization of dinucleotide microsatellites in greater amberjack, *Seriola dumerili* *Conserv Genet* (2007) 8:1009–1011.

Renshaw MA, Patton JC, Rexroad CE, Gold JR (2006) PCR primers for trinucleotide and tetranucleotide microsatellites in greater amberjack, *Seriola dumerili*. *Molecular Ecology Notes*.

Rice, William R. "Analyzing tables of statistical tests." *Evolution* 43.1 (1989): 223-225.
Rockman, Matthew V. "The QTN program and the alleles that matter for evolution: all that's gold does not glitter." *Evolution: International Journal of Organic Evolution* 66.1 (2012): 1-17.

Rocchi M, Archidiacono N, Schempp W, Capozzi O, Stanyon R. Centromere repositioning in mammals. *Heredity*. 2012; 108(1):59–67.

Rockman MV (2012) The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution* 66, 1-17.

Rutten, M.J.M., Bovenhuis, H., Komen, H., 2005. Genetic parameters for fillet traits and body measurements in Nile tilapia (*Oreochromis niloticus* L.). *Aquaculture* 246, 125–132.

Ruane, John, and Andrea Sonnino. "Marker-assisted selection as a tool for genetic improvement of crops, livestock, forestry and fish in developing countries: an overview of the issues." *Marker-assisted Selection-Current Status and Future Perspectives in Crops, Livestock, Forestry and Fish.-Rome* (Food and Agriculture Organization of the United Nations (FAO)) (2007): 3-13.

Saillant, E., Cizdziel, K., O'Malley, K.G., Turner, T.F., Pruett, c.L., Gold, R., 2004. Microsatellite markers for red drum, *Sciaenops ocellatus*. *Gulf Mexico Sci.* 22 (1).

Salem, Mohamed, et al. "RNA-Seq identifies SNP markers for growth traits in rainbow trout." *PloS one* 7.5 (2012): e36264.

Sato, Yukuto, and Mutsumi Nishida. "Teleost fish with specific genome duplication as unique models of vertebrate evolution." *Environmental Biology of Fishes* 88.2 (2010): 169-188.

Savolainen O, Lascoux M, Merila J (2013) Ecological genomics of local adaptation. *Nature Reviews Genetics* 14, 807-820.

Shannon, Claude E., and Warren Weaver. "The mathematical theory of information." Urbana: University of Illinois Press 97 (1949).

Slate, Jon, et al. "Gene mapping in the wild with SNPs: guidelines and future directions." *Genetica* 136.1 (2009): 97-107.

Song, Wentao, et al. "Construction of a high-density microsatellite genetic linkage map and mapping of sexual and growth-related traits in half-smooth tongue sole (*Cynoglossus semilaevis*)." *PLoS One* 7.12 (2012): e52097.

Song, Wentao, et al. "Construction of high-density genetic linkage maps and mapping of growth-related quantitative trait loci in the Japanese flounder (*Paralichthys olivaceus*)." *PLoS One* 7.11 (2012): e50404.

Soula, M., Zamorano, M.J., Navarro, A., Sanchez, J.J., Neil, D., Alejandro, G., Afonso, J.M., 2011. Diseño de dos nuevas PCRs multiplex para corvina (*Argyrosomus regius*). Congreso Nacional de Acuicultura, Barcelona España 2011.

Staples, J., Witherspoon, D. J., Jorde, L. B., Nickerson, D. A., The University of Washington Center for Mendelian Genomics, Below, J. E., & Huff, C. D. (2016). PADRE: Pedigree-aware distant-relationship estimation. *The American Journal of Human Genetics*, 99, 154–162.

Staples, J., Qiao, D., Cho, M. H., Silverman, E. K., University of Washington Center for Mendelian Genomics, Nickerson, D. A., & Below, J. E. (2014). PRIMUS: Rapid reconstruction of pedigrees from genome-wide estimates of identity by descent. *The American Journal of Human Genetics*, 95, 553–564.

Stranger BE, Stahl EA, Raj T (2011) Progress and promise of genome-wide association studies for human complex trait genetics. *Genetics* 187, 367-383.

Syaifudin, M., Bekaert, M., Taggart, J. B., Bartie, K. L., Wehner, S., Palaiokostas, C., Penman, D. J. (2019). Species-specific marker discovery in *Tilapia*. *Scientific Reports*, 9, 13001.

Symonds Jane E., Clarke Shannon M., King Nick, Walker Seumas P., Blanchard Brian, Sutherland David, Roberts Rodney, Preece Mark A., Tate Mike, Buxton Peter, Dodds Ken G. Developing Successful Breeding Programs for New Zealand Aquaculture: A Perspective on Progress and Future Genomic Opportunities *Frontiers in Genetics* 10, 27, (2019).

Takeda, Hiroyuki. "Draft genome of the medaka fish: a comprehensive resource for medaka developmental genetics and vertebrate evolutionary biology." *Development, growth & differentiation* 50 (2008): S157-S166.

Taslina, K., Wehner, S., Taggart, J. B., de Verdal, H., Benzie, J. A. H., Bekaert, M., Penman, D. J. (2020). Sex determination in the GIFT strain of tilapia is controlled by a locus in linkage group 23. *BMC Genetics*, 21, 49.

Tautz, Diethard, and Manfred Renz. "Simple sequences are ubiquitous repetitive components of eukaryotic genomes." *Nucleic acids research* 12.10 (1984): 4127-4138.

Tautz, Diethard. "Hypervariability of simple sequences as a general source for polymorphic DNA markers." *Nucleic acids research* 17.16 (1989): 6463-6471.

Tautz, D. (1989). Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research*, 17, 6463–6471.

Thodesen, Jørn, et al. "Genetic improvement of tilapias in China: genetic parameters and selection responses in growth of Nile tilapia (*Oreochromis niloticus*) after six generations of multi-trait selection for growth and fillet yield." *Aquaculture* 322 (2011): 51-64.

Thompson, E. A., & Meagher, T. R. (1987). Parental and sib likelihoods in genealogy reconstruction. *Biometrics*, 43, 585–600.

Toonen, R. J., Puritz, J. B., Forsman, Z. H., Whitney, J. L., Fernandez-Silva, I., Andrews, K. R., & Bird, C. E. (2013). ezRAD: A simplified method for genomic genotyping in non-model organisms. *PeerJ*, 1, e203.

Tsai, Hsin-Yuan, et al. "Genomic prediction of host resistance to sea lice in farmed Atlantic salmon populations." *Genetics Selection Evolution* 48.1 (2016): 1-11.

Tsai, Hsin Yuan, et al. "The genetic architecture of growth and fillet traits in farmed Atlantic salmon (*Salmo salar*)." *Bmc Genetics* 16.1 (2015): 1-11.

Turner, T. F., L. R. Richardson, and J. R. Gold. "Polymorphic microsatellite DNA markers in red drum (*Sciaenops ocellatus*)." *Molecular Ecology* 7.12 (1998): 1771-1773.

Valles, R., Estevez, A., 2013. Light conditions for larval rearing of meagre (*Argyrosomus regius*). *Aquaculture* 376–379, 15–19.

Vandeputte, M., Mauger, S., Dupont-Nivet, M., 2006. An evaluation of allowing mismatches as a way to manage genotyping errors in parentage assignment by exclusion. *Mol. Ecol. Notes* 6, 265–267.

Vandeputte, M., Mauger, S., Dupont-Nivet, M., 2006. An evaluation of allowing for mismatches as a way to manage genotyping errors in parentage assignment by exclusion. *Mol. Ecol. Notes* 6, 265–267.

Van Oosterhout, Cock, et al. "MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data." *Molecular Ecology Notes* 4.3 (2004): 535-538.

Visscher, Peter M., Jian Yang, and Michael E. Goddard. "A commentary on ‘common SNPs explain a large proportion of the heritability for human height’ by Yang et al.(2010)." *Twin Research and Human Genetics* 13.6 (2010): 517-524.

Visscher Peter M, Brown Matthew A, McCarthy Mark I, Yang J (2012) Five years of GWAS discovery. *The American Journal of Human Genetics* 90, 7-24.

Waldbieser, G.C., Wolters, W.R., 1999. Application of polymorphic microsatellite loci in a channel catfish *Ictalurus punctatus* breeding program. *J. World Aquac. Soc.* 30, 256–262.

Wang, S., Meyer, E., McKay, J. K., & Matz, M. V. (2012). 2b-RAD: A simple and flexible method for genome-wide genotyping. *Nature Methods*, 9, 808–810.

Wang, Le, et al. "Construction of a high-density linkage map and fine mapping of QTL for growth in Asian seabass." *Scientific reports* 5.1 (2015): 1-10.

Wang, Zhibin, et al. "Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes." *Cell* 138.5 (2009): 1019-1031.

Wang, Weiji, et al. "High-density genetic linkage mapping in turbot (*Scophthalmus maximus* L.) based on SNP markers and major sex-and growth-related regions detection." *Plos one* 10.3 (2015): e0120410.

Wang, C.M., Lo, L.C., Zhu, Z.Y., Lin, G., Feng, F., Li, J., Yang, W.T., Tan, J., Chou, R., Lim, H.S., Orban, L., Yue, G.H., 2008. Estimating reproductive success of brooders and

heritability of growth traits in Asian sea bass (*Lates calcarifer*) using microsatellites. *Aquacult. Res.* 39 (15), 1612–1619.

Wang, J. (2004). Sibship reconstruction from genetic data with typing errors. *Genetics*, 166, 1963–1979.

Wang, Jinliang. "COANCESTRY: a program for simulating, estimating and analysing relatedness and inbreeding coefficients." *Molecular ecology resources* 11.1 (2011): 141-145.

Wellenreuther, Maren, and Bengt Hansson. "Detecting polygenic evolution: problems, pitfalls, and promises." *Trends in Genetics* 32.3 (2016): 155-164.

Weir, Bruce S., and C. Clark Cockerham. "Estimating F-statistics for the analysis of population structure." *evolution* (1984): 1358-1370.

BS Weir - Inc., Sunderland, MA, 1996 *Genetic Data Analysis II* Sinauer Associates. Wray NR, Thompson R. Prediction of rates of inbreeding in selected populations. *Genet Res.* 1990 Feb;55(1):41-54.

Wringe, Brendan F., et al. "Growth-related quantitative trait loci in domestic and wild rainbow trout (*Oncorhynchus mykiss*)." *BMC genetics* 11.1 (2010): 1-14.

Wright, Sewall. *The theory of gene frequencies* Evolution and the genetics of populations: A Treatise in Three Volumes. No. 576.58 W9301t Ej. 1 025186. The University of Chicago Press; 1969.

Xiao, Shijun, et al. "Gene map of large yellow croaker (*Larimichthys crocea*) provides insights into teleost genome evolution and conserved regions associated with growth." *Scientific reports* 5.1 (2015): 1-12.

Yang, Jian, et al. "Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index." *Nature genetics* 47.10 (2015): 1114.

Yates, Frank. "The recovery of inter-block information in balanced incomplete block designs." *Annals of Eugenics* 10.1 (1940): 317-325.

Yoshida, Grazyella M., and José M. Yáñez. "Multi-trait GWAS using imputed high-density genotypes from whole-genome sequencing identifies genes associated with body traits in Nile tilapia." *BMC Genomics* 22.1 (2021): 1-13.

Yoshida, K., & Kitano, J. (2021). Tempo and mode in karyotype evolution revealed by a probabilistic model incorporating both chromosome number and morphology. *PLoS genetics*, 17(4), e1009502.

You, X., Shan, X., & Shi, Q. (2020). Research advances in the genomics and applications for molecular breeding of aquaculture animals. *Aquaculture*, 526, 735357.

Zardoya, Rafael, et al. "Evolutionary conservation of microsatellite flanking regions and their use in resolving the phylogeny of cichlid fishes (Pisces: Perciformes)." *Proceedings of the Royal Society of London. Series B: Biological Sciences* 263.1376 (1996): 1589-1598.

Zhang, Yu, et al. "Comparative genomic analysis of catfish linkage group 8 reveals two homologous chromosomes in zebrafish and other teleosts with extensive inter-chromosomal rearrangements." *BMC genomics* 14.1 (2013): 1-15.