



ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΤΡΟΦΙΜΩΝ

ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΤΡΟΦΙΜΩΝ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ ΜΕ ΘΕΜΑ:

«Εφαρμογές του next generation sequencing στα τρόφιμα»

DEGREE THESIS ON THE SUBJECT:

"Applications of next generation sequencing in foodstuff"



ΟΝΟΜΑ ΦΟΙΤΗΤΗ:

Παπαδοπούλου Σοφία-Φαίδρα/ Papadopoulou Sofia-Faidra (17080)

ΟΝΟΜΑ ΕΙΣΗΓΗΤΗ:

Μπατρίνου Ανθιμία/ Batrinou Anthimia

ΟΝΟΜΑΤΑ ΕΠΙΤΡΟΠΗΣ:

Α.Μπατρίνου/A.Batrinou,Επίκουρη Καθηγήτρια

Δ.Χούχουλα/D.Chouchoula, Αναπληρώτρια Καθηγήτρια

Σ.Κοντελές/S.Kodeles, Επίκουρος Καθηγητής

AIGALEO/AIGALEO 2022

Έγινε δεκτή

Οι υπογράφοντες δηλώνουμε ότι έχουμε εξετάσει τη πτυχιακή εργασία με τίτλο **«Εφαρμογές του next generation sequencing στα τρόφιμα»** που παρουσιάστηκε από τη φοιτήτρια **Παπαδοπούλου Σοφία-Φαίδρα με αριθμό μητρώου 17080** και βεβαιώνουμε ότι γίνεται δεκτή.

Ημερομηνία & Όνομα επιβλέποντος:

Ημερομηνία & Όνομα μέλους επιτροπής:

Ημερομηνία & Όνομα μέλους επιτροπής:

Δήλωση συγγραφέων πτυχιακής εργασίας περί λογοκλοπής/ Copyright

Η κάτωθεν υπογεγραμμένη Παπαδοπούλου Σοφία-Φαίδρα, με αριθμό μητρώου 17080, φοιτήτρια του Πανεπιστημίου Δυτικής Αττικής της Σχολής Επιστημών Τροφίμων του Τμήματος Επιστήμης και Τεχνολογίας Τροφίμων, δηλώνει υπεύθυνα ότι:

«Έχοντας πλήρη επίγνωση των συνεπειών του νόμου περί πνευματικής ιδιοκτησίας, δηλώνω ότι είμαι αποκλειστική συγγραφέας της παρούσας πτυχιακής εργασίας. Δηλώνω, επίσης, ότι αναλαμβάνω όλες τις συνέπειες, όπως αυτές νομίμως ορίζονται».

Υπογραφή φοιτήτριας:



Ευχαριστίες

Η παρούσα πτυχιακή εργασία πραγματοποιήθηκε στο Τμήμα Επιστήμης και Τεχνολογίας Τροφίμων, της σχολής Επιστημών Τροφίμων του Πανεπιστημίου Δυτικής Αττικής.

Θα ήθελα να ευχαριστήσω ιδιαίτερα την επιβλέπουσα καθηγήτρια, κύρια Μπατρίνου Ανθιμία, για την εμπιστοσύνη που μου έδειξε με την ανάθεση της εν λόγω εργασίας, και για τον χρόνο που διέθεσε για να μας καθοδηγήσει αλλά και για να προβεί στις διορθώσεις της.

Τέλος, ευχαριστώ θερμά την οικογένεια μου και τους φίλους μου, για την συμπαράσταση και την ψυχολογική στήριξη τους σε όλη την διάρκεια της φοίτησης μου στην σχολή αλλά και ιδιαίτερα κατά την συγγραφή της συγκεκριμένης πτυχιακής εργασίας.

Περίληψη

Από την ολοκλήρωση του έργου του ανθρώπινου γονιδιώματος το 2003, έχει σημειωθεί εξαιρετική πρόοδος στις τεχνολογίες προσδιορισμού αλληλουχίας γονιδιώματος, γεγονός που οδήγησε σε μειωμένο κόστος ανά βάση δεδομένων και σε αύξηση του αριθμού και της ποικιλομορφίας των αλληλουχούμενων γονιδιωμάτων. Ένα σημαντικό μέρος της αρχιτεκτονικής του γονιδιώματος έχει αποκαλυφθεί, καθιστώντας αυτές τις τεχνολογίες περισσότερο πολύτιμες. Ορισμένες προσεγγίσεις μεγιστοποιούν τον αριθμό των βάσεων που αλληλουχούνται στο ελάχιστο δυνατότερο χρονικό διάστημα, δημιουργώντας έναν πλούτο δεδομένων που μπορούν να χρησιμοποιηθούν για την κατανόηση ολόενα και πιο περίπλοκων φαινοτύπων. Εναλλακτικά, άλλες προσεγγίσεις στοχεύουν τώρα στην αλληλουχία μακρύτερων συνεχόμενων τμημάτων DNA, τα οποία είναι απαραίτητα για την επίλυση δομικά πολύπλοκων περιοχών. Αυτές και άλλες στρατηγικές παρέχουν στους ερευνητές, τους κλινικούς γιατρούς, αλλά και στη βιομηχανία τροφίμων, μια ποικιλία εργαλείων για την ανίχνευση γονιδιωμάτων σε μεγαλύτερο βάθος, οδηγώντας σε μια βελτιωμένη κατανόηση του τρόπου με τον οποίο οι παραλλαγές της αλληλουχίας του γονιδιώματος αποτελούν τη βάση του φαινοτύπου. Η ασφάλεια και η ποιότητα των τροφίμων αποτελούν στις μέρες μας μείζον μέλημα. Οποιαδήποτε περίπτωση αλλοίωσης τροφίμων, ειδικά όταν αναφέρεται από τα μέσα ενημέρωσης, έχει μεγάλο αντίκτυπο στην κοινή γνώμη. Υπάρχει μια αυξανόμενη ζήτηση για τη βελτίωση των ποιοτικών ελέγχων, ως εκ τούτου την επιστημονική έρευνα για την ανάπτυξη αξιόπιστων μοριακών εργαλείων για την ανάλυση τροφίμων. Το DNA Barcoding είναι ένα ευρέως χρησιμοποιούμενο σύστημα μοριακής βάσης, το οποίο μπορεί να αναγνωρίσει βιολογικά δείγματα και χρησιμοποιείται για την αναγνώριση τόσο των πρώτων υλών όσο και των επεξεργασμένων τροφίμων. Σε αυτή την πτυχιακή αναλύονται κριτικά τα αποτελέσματα πολλών ερευνών, προκειμένου να αξιοποιηθεί η αποτελεσματικότητα του DNA Barcoding στην ιχνηλασιμότητα των τροφίμων και να σκιαγραφηθούν ορισμένες βέλτιστες πρακτικές στην εφαρμογή του στη βιομηχανία τροφίμων. Συζητείται επίσης η χρήση του DNA Barcoding για την ασφάλεια των τροφίμων και για τον εντοπισμό εμπορικής απάτης.

Abstract

Since the completion of the Human Genome Project in 2003, tremendous progress has been made in genome sequencing technologies, which has led to reduced costs per database and an increase in the number and diversity of sequenced genomes. An important part of the genome architecture has been revealed, making these technologies more valuable. Some approaches maximize the number of bases that are interconnected in the least amount of time possible, creating a wealth of data that can be used to understand increasingly complex phenomena. Alternatively, other approaches now aim to sequence longer contiguous stretches of DNA, which are necessary to resolve structurally complex regions. These and other strategies provide researchers, clinicians, and the food industry with a variety of tools to probe genomes in greater depth, leading to an improved understanding of how genome sequence variations underlie phenotype. Food safety and quality are a major concern these days. Any case of food adulteration, especially when reported by the media, has a large impact in public opinion. There is a growing demand to improve quality controls, hence why scientific research is needed to develop reliable molecular tools for food analysis. DNA Barcoding is a widely used molecular-based system that can identify biological samples and is used to identify both raw materials and processed foods. In this thesis, the results of several researches are critically analyzed in order to exploit the effectiveness of DNA Barcoding in food traceability and to outline some best practices in the application in the food industry. The use of DNA Barcoding for food safety and commercial fraud detection is also discussed.

Περιεχόμενα

Δήλωση συγγραφέων πτυχιακής εργασίας περί λογοκλοπής/ Copyright.....	3
Ευχαριστίες.....	3
Περίληψη.....	4
Abstract	5
Περιεχόμενα.....	6
Κεφάλαιο 1: Εισαγωγή	8
Κεφάλαιο 2: Τεχνολογία NGS.....	11
2.1 Αλληλούχιση Σύντομης Ανάγνωσης.....	11
2.1.1 Αλληλούχιση Απολίνωσης.....	13
2.1.2 Αλληλούχιση Σύνθεσης	14
2.2 Αλληλούχιση Μακράς Ανάγνωσης.....	17
2.2.1 Αλληλούχιση Μακράς Ανάγνωσης Μονού Μορίου.....	18
2.2.2 Συνθετική Μακρά Ανάγνωση	19
Κεφάλαιο 3: Γονιδιακές Αλληλουχίες 16S rRNA και ITS	21
3.1 Ο βακτηριακός προσδιορισμός με τη χρήση της αλληλούχισης 16S rRNA	21
3.2 Βακτηριακή ονοματολογία σε σχέση με την γονιδιακή αλληλούχιση 16S rRNA	23
3.3 Επίλυση των προβλημάτων.....	24
3.4 Ιδιωτικές και δημόσιες βάσεις δεδομένων νουκλεοτιδίων	25
3.5 Επιλογή γονιδίων και εκκινήτων	26
3.6 Ακρίβεια αποτελεσμάτων	29
3.6.1 Μήκος αντιστοίχισης.....	29
3.6.2 Συνέπεια αντιστοίχισης.....	30
3.6.3 Διαφοροποίηση αντιστοίχισης.....	30
3.6.4 Αντιστοίχιση και βάσεις δεδομένων	32
3.7 Αναγνώριση παθογόνων βακτηρίων με γονιδιακή αλληλούχιση 16S rRNA	33

3.8	Τρέχοντες περιορισμοί της αλληλούχησης 16S rRNA για την αναγνώριση παθογόνων βακτηρίων	35
3.9	Τα <i>Staphylococcus spp.</i> και άλλοι αερόβιο gram-θετικοί μικροοργανισμοί	35
3.10	Τα <i>Streptococcus spp.</i> και άλλα όμοια βακτήρια	37
3.11	Αερόβιοι gram-θετικοί βάκιλοι	38
3.12	<i>Listeria spp.</i>	38
3.13	Εντεροβακτήρια (<i>Enterobacteriaceae</i>)	39
3.14	Καμπυλοβακτήρια (<i>Campylobacteriales</i>)	40
3.15	DNA barcoding	41
3.16	ITS «Internal Transcribed Spacer»	42
3.17	Σύμπλεγμα rDNA	44
3.18	Βιολογικός ρόλος των περιοχών ITS	45
3.19	Σύγκριση ITS1 και ITS2	46
Κεφάλαιο 4:	Εφαρμογή στη Βιομηχανία Τροφίμων	50
4.1	DNA Barcoding σε ιχθυηρά	54
4.2	Το DNA barcoding και η ιχνηλασιμότητα στα κρεατοσκευάσματα: το πρόβλημα της έλλειψης δεδομένων	56
4.3	Το DNA Barcoding στα γαλακτοκομικά προϊόντα: μια εν δυνάμει εφαρμογή	57
4.4	Το DNA Barcoding σε φυτικά προϊόντα	59
4.5	Το DNA Barcoding ως εργαλείο ιχνηλασιμότητας κατά την μεταποίηση τροφίμων	61
4.6	Χαρακτηρισμός «μεικτών» τροφίμων	63
4.7	Ασφάλεια και Απάτη (Fraud) στα τρόφιμα	64
Κεφάλαιο 5:	Συμπεράσματα	67
Κεφάλαιο 6:	Βιβλιογραφία	69

Κεφάλαιο 1: Εισαγωγή

Οι τεχνολογίες next-generation sequencing (NGS) είναι νέες μέθοδοι προσδιορισμού και τροποποίησης της αλληλουχίας των DNA και RNA, μία διεργασία γνωστή ως αλληλούχιση (sequencing) (Goodwin, McPherson, & McCombie, 2016). Οι ερευνητές χρησιμοποιούν, κυρίως, τις τεχνολογίες NGS στη βασική, εφαρμοσμένη και κλινική έρευνα, έχοντας μόνο πρόσφατα επεκταθεί στο τομέα της τεχνολογίας τροφίμων. Τη δεκαετία του 1970, η πρώτη αλληλούχιση DNA, που ονομάζεται αλληλούχιση Sanger ή πρωτότυπη αλληλούχιση DNA, αναπτύχθηκε από τους Frederick Sanger et al. (1977) (Sanger, Nicklen, & Coulson, 1977). Η αλληλούχιση DNA Sanger επέτρεψε την ανάλυση μεμονωμένων γονιδίων ή τμημάτων τους. Για τρεις δεκαετίες, η αλληλουχία Sanger χρησιμοποιήθηκε εκτενώς από τους βιολόγους. Αργότερα, εμφανίστηκε η τεχνολογία NGS που άνοιξε νέες δυνατότητες και προκλήσεις σε βιολόγους και κλινικούς ιατρούς στον τομέα της εξερεύνησης γονιδιωμάτων και της κλινικής διάγνωσης (Goodwin, McPherson, & McCombie, 2016). Η τεχνολογία NGS έχει πολλά πλεονεκτήματα σε σχέση με την αλληλούχιση Sanger. Η NGS μπορεί να παρέχει πολυπληθή δεδομένα υψηλής απόδοσης από πολλαπλά δείγματα, ανά ανάλυση, με σημαντικά μειωμένο κόστος σε σύγκριση με την αλληλούχιση Sanger (Goodwin et al., 2016). Η NGS, επιπλέον, παρέχει πιο ακριβή και αξιόπιστα δεδομένα αλληλούχισης από την Sanger. Το ανθρώπινο γονιδίωμα αναλύθηκε σε σχεδόν 13 χρόνια χρησιμοποιώντας την αλληλουχία Sanger με κόστος μεγαλύτερο από ένα εκατομμύριο δολάρια (Naidoo, Pawitan, Soong, Cooper, & Ku, 2011), ενώ τώρα, είναι δυνατός ο προσδιορισμός της αλληλουχίας του ανθρώπινου γονιδιώματος μέσα σε μια εβδομάδα με κόστος 100 δολαρίων. Λόγω αυτών των πλεονεκτημάτων, η NGS έχει γίνει ένα ισχυρό και πολύτιμο εργαλείο για έρευνα και κλινικές εφαρμογές. Οι βελτιώσεις στις τεχνολογίες NGS και στη βιοπληροφορική έχουν δημιουργήσει ευκαιρίες για τους ερευνητές να πραγματοποιούν αλληλούχιση και μεταγραφή ολόκληρου του γονιδιώματος (ιστών, κυττάρων και οργανισμών) (Jabeen, Ahmad, & Raza, 2018), προσδιορισμό αλληλουχίας μεθυλίωσης, αλληλούχιση ολόκληρου του εξονιώματος σε τρόφιμα (Gazara, et al., 2019) και ανοσοκατακρήμνιση χρωματίνης ακολουθούμενη από αλληλούχιση (ChIP-seq) (Wani & Raza, 2019). Υπάρχουν, ωστόσο, ορισμένες προκλήσεις των εφαρμογών των τεχνολογιών NGS που πρέπει να υπερνικηθούν ώστε να μπορούν να χρησιμοποιηθούν αποτελεσματικά σε κλινικές.

Η NGS σε συνδυασμό με ισχυρές βιοπληροφορικές προσεγγίσεις φέρνουν επανάσταση στη μικροβιολογία των τροφίμων. Η αλληλούχιση ολόκληρου του γονιδιώματος δειγμάτων επιτρέπει την υψηλά λεπτομερή σύγκριση των απομονωμένων στελεχών. Οι δύο βασικές προσεγγίσεις για τη διάκριση των στελεχών, η ανάλυση πολυμορφισμού απλού νουκλεοτιδίου (single nucleotide polymorphism, SNP) και ο γονιδιωματικός τύπος αλληλουχίας πολλαπλών θέσεων (multi-locus sequence typing, MLST) δείχνουν αναλυτικά αποτελέσματα για τη φυλογενετική ομαδοποίηση τους και είναι συμπληρωματικές μεταξύ τους. Οι μεταγονιδιωματικές (metagenomics) και οι αναλύσεις μεταγραμμικού κώδικα (metabarcoding), που εφαρμόζονται στο ολικό DNA που απομονώνεται είτε από υλικά τροφίμων είτε από το περιβάλλον παραγωγικής διαδικασίας, επιτρέπει την ταυτοποίηση των μικροβιακών πληθυσμών σε υψηλό βαθμό λεπτομέρειας. Οι μεταγονιδιωματικές αναλύσεις προσδιορίζουν ολόκληρο το γονιδιωματικό περιεχόμενο και, όταν συνδέεται με μεταγραφομικές (transcriptomics) ή πρωτεομικές (proteomics) αναλύσεις, επιτρέπουν την αναγνώριση της λειτουργικής ικανότητας και της βιοχημικής δραστηριότητας των μικροβιακών πληθυσμών. Η εφαρμογή της μεταγονιδιωματικής για την ασφάλεια των τροφίμων και τη βελτίωση της ποιότητας είναι ακόμη σε πρώιμα στάδια και προσφέρει μεγάλες ευκαιρίες για την πρόβλεψη της παρουσίας ή μελλοντικής εμφάνισης παθογόνων και αλλοιογόνων μικροοργανισμών με βάση τις αλλαγές που παρατηρούνται σε ολόκληρες μικροβιακές κοινότητες, καθώς και τη δυνατότητα χαρακτηρισμού άγνωστης μικροχλωρίδας (Goodwin, McPherson, & McCombie, 2016; Jagadeesan, et al., 2019).

Η αλληλούχιση σύντομης ανάγνωσης (Short-read sequencing) είναι εξαιρετικά ακριβής και παράγει μήκη ανάγνωσης 100–300 ζευγών βάσεων (base pairs, bp), τα οποία στη συνέχεια ανασυνδυάζονται σε ημιτελή ή αλλιώς «πρόχειρα» γονιδιώματα. Ολόκληρα γονιδιώματα δεν μπορούν να δημιουργηθούν από τις αλληλουχίσεις που λαμβάνονται σε μία μόνο ακολουθία, λόγω δυσκολιών στον ανασυνδυασμό επαναλαμβανόμενων περιοχών και μεγάλων γονιδιωματικών αναδιατάξεων, όπως εισαγωγές (insertions), διαγραφές (deletions) και αντιστροφές (inversions). Για πολλές εφαρμογές, συμπεριλαμβανομένης της συγκριτικής γονιδιωματικής (comparative genomics) και της φυλογένεσης (phylogeny), αυτό δεν είναι ένα ζήτημα, αλλά στις περιπτώσεις όπου απαιτούνται ολόκληρα γονιδιώματα και προσδιορισμός πολύπλοκων γονιδιωματικών περιοχών, απαιτούνται αλληλουχίσεις μακράς ανάγνωσης (long-

read sequencing). Η αλληλούχιση μακράς ανάγνωσης παράγει αναγνώσεις μήκους από 10 έως 50 Kb, αλλά αυτό είναι εις βάρος των υψηλότερων ποσοστών σφάλματος (Loman & Pallen, 2015).

Ο σκοπός της παρούσας ανασκόπησης είναι η ανάδειξη της πρόσφατης χρήσης και των μελλοντικών δυνατοτήτων της NGS, στη μικροβιολογία των τροφίμων, αλλά και των περιορισμών και προκλήσεών της. Θα ερευνηθούν οι τομείς που τη χρησιμοποιούν, όπως τα τμήματα δημόσιας υγείας και η βιομηχανία τροφίμων και θα ερευνηθεί ο τρόπος της ερμηνείας των αποτελεσμάτων και ο τρόπος που μπορούν να χρησιμοποιηθούν για να τοποθετηθούν, με τη σειρά τους, σε ένα ευρύτερο πλαίσιο. Η ανασκόπηση στοχεύει στην προώθηση της ευρύτερης εφαρμογής των τεχνολογιών NGS στη βιομηχανία τροφίμων, καθώς και στην ανάδειξη των κενών γνώσης και των σύγχρονων εφαρμογών της, με στόχο την προώθηση της μελλοντικής έρευνας και την αύξηση των αποτελεσμάτων, για την ασφάλεια των τροφίμων, από την ευρύτερη χρήση της. Θα αναλυθούν οι τεχνολογίες που χρησιμοποιούνται, το εύρος εφαρμογών, οι κύριες αρχές, τα διαφορετικά είδη, οι προσεγγίσεις και οι μέθοδοι της NGS.

Κεφάλαιο 2: Τεχνολογία NGS

Από την ολοκλήρωση του έργου του ανθρώπινου γονιδιώματος το 2003, έχει σημειωθεί εξαιρετική πρόοδος στις τεχνολογίες προσδιορισμού αλληλουχίας γονιδιώματος, γεγονός που οδήγησε σε μειωμένο κόστος, ανά βάση δεδομένων, και στην αύξηση του αριθμού της ποικιλομορφίας των αλληλουχούμενων γονιδιωμάτων. Έχει αποκαλυφθεί μια εκπληκτική πολυπλοκότητα της «αρχιτεκτονικής» του γονιδιώματος, προωθώντας αυτές τις τεχνολογίες αλληλουχίας σε ακόμη μεγαλύτερες προόδους. Ορισμένες προσεγγίσεις μεγιστοποιούν τον αριθμό των βάσεων που αλληλουχούνται στο ελάχιστο χρονικό διάστημα, δημιουργώντας μια πληθώρα δεδομένων που μπορούν να χρησιμοποιηθούν για την κατανόηση ολόενα και πιο περίπλοκων φαινοτύπων.

Εναλλακτικά, άλλες προσεγγίσεις στοχεύουν τώρα στην αλληλούχιση μακρύτερων συνεχόμενων μερών DNA, τα οποία είναι απαραίτητα για την επίλυση δομικά πολύπλοκων περιοχών. Αυτές και άλλες στρατηγικές παρέχουν στους ερευνητές και τους κλινικούς ιατρούς μια ποικιλία εργαλείων για την ανίχνευση γονιδιωμάτων σε μεγαλύτερο βαθμό, οδηγώντας σε μια βελτιωμένη κατανόηση του τρόπου με τον οποίο οι παραλλαγές της αλληλούχισης του γονιδιώματος αποτελούν τη βάση του φαινοτύπου και μίας νόσου.

2.1 Αλληλούχιση Σύντομης Ανάγνωσης

Οι προσεγγίσεις αλληλούχισης σύντομης ανάγνωσης εμπίπτουν σε δύο ευρείες κατηγορίες: αλληλούχιση απολίνωσης (Sequencing By Ligation, SBL) και αλληλούχιση σύνθεσης (Sequencing By Synthesis, SBS).

Στις προσεγγίσεις SBL, μια αλληλουχία βάσεων, που ονομάζεται ανιχνευτής, και που είναι συνδεδεμένη με μία φθοροφόρο ουσία, υβριδοποιείται σε ένα μόριο DNA και συνδέεται με ένα γειτονικό ολιγονουκλεοτίδιο, με σκοπό την απεικόνισή του. Το φάσμα εκπομπής της φθοροφόρου ουσίας χρησιμοποιείται για τον προσδιορισμό των βάσεων, ή των συμπληρωματικών τους, σε συγκεκριμένες θέσεις εντός του ανιχνευτή.

Σε προσεγγίσεις SBS, χρησιμοποιείται μια πολυμεράση ένα μέσο-σηματοδότης, όπως ένα φθοροφόρο μόριο ή μια αλλαγή στην ιοντική συγκέντρωση, το οποίο προσδιορίζει την ενσωμάτωση ενός νουκλεοτιδίου σε έναν επιμηκυνόμενο κλώνο. Στις περισσότερες προσεγγίσεις SBL και SBS, το DNA ενισχύεται κλωνικά σε μια στερεή επιφάνεια. Η ύπαρξη

πολλών χιλιάδων πανομοιότυπων αντιγράφων ενός μέρους, ενός μορίου DNA, σε μια καθορισμένη περιοχή διασφαλίζει ότι ο σηματοδότης μπορεί να διακριθεί από το θόρυβο του περιβάλλοντός του. Διευκολύνονται, επίσης, οι παράλληλες αλληλουχίσεις, με τη δημιουργία πολλών εκατομμυρίων μεμονωμένων «κέντρων αντίδρασης» SBL ή SBS, όπου το καθένα διαθέτει το δικό του μοναδικό πρότυπο DNA. Μια πλατφόρμα προσδιορισμού αλληλουχίας μπορεί, επιπλέον, να συλλέξει πληροφορίες από πολλά κέντρα αντίδρασης ταυτόχρονα, προσδιορίζοντας έτσι την αλληλουχία πολλών εκατομμυρίων μορίων DNA παράλληλα (Goodwin, McPherson, & McCombie, 2016).

Υπάρχουν πολλές διαφορετικές στρατηγικές που χρησιμοποιούνται για τη δημιουργία πληθυσμών από κλωνικά πρότυπα: με τη μέθοδο παραγωγής σφαιριδίων (bead-based generation), με τη μέθοδο παραγωγής στερεής φάσης (solid-state generation) και με τη μέθοδο παραγωγής νανοσφαιρών DNA (DNA nanoball generation). Το πρώτο βήμα της δημιουργίας προτύπου DNA είναι ο «κατακερματισμός» του δείγματος DNA (DNA fragmentation), που ακολουθείται από τη σύνδεση του σε έναν προσαρμογέα (adaptor) για κλωνική επιμήκυνση και αλληλούχιση. Για τη μέθοδο της παραγωγής σφαιριδίων, ένας προσαρμογέας είναι συμπληρωματικός με ένα ολιγονουκλεοτιδικό θραύσμα DNA, το οποίο έχει ακινητοποιηθεί σε ένα σφαιρίδιο. Χρησιμοποιώντας γαλακτωματική PCR (emulsion PCR, emPCR), το πρότυπο DNA ενισχύεται έτσι ώστε να ακινητοποιηθούν έως και ένα εκατομμύριο θραύσματα DNA σε ένα μόνο σφαιρίδιο (Shendure, et al., 2005). Τα σφαιρίδια μπορούν να διανεμηθούν σε μια γυάλινη επιφάνεια ή να παραταχθούν σε μια πικο-πλάκα (PicoTiterPlate) (Leamon, et al., 2003). Η παραγωγή στερεής φάσης αποφεύγει τη χρήση της emPCR, εφαρμόζοντας, αντί για αυτή, απευθείας ενίσχυση σε επιφάνεια (Harris, et al., 2008). Σε αυτή την προσέγγιση, πρόσθιοι (forward) και αντίστροφοι (reverse) εκκινητές (primers) συνδέονται ομοιοπολικά στην επιφάνεια της πλάκας, είτε τυχαία είτε με συγκεκριμένη σειρά. Αυτοί οι εκκινητές παρέχουν συμπληρωματικά άκρα στα οποία μπορούν να συνδεθούν πρότυπα μονόκλωνου DNA (Single-Stranded DNA, ssDNA). Ο ακριβής έλεγχος της συγκέντρωσης του προτύπου επιτρέπει την επιμήκυνση των προτύπων σε τοπικά, μη επικαλυπτόμενα κλωνικά συμπλέγματα, διατηρώντας έτσι την ακεραιότητα του χώρου τους. Πρόσφατα, πολλές πλατφόρμες NGS έχουν χρησιμοποιήσει ειδικές κυψέλες ροής. Με τον ακριβή προσδιορισμό του σημείου που οι εκκινητές συνδέονται στα πρότυπα, μπορούν να αναλυθούν περισσότερα πρότυπα DNA,

επιτρέποντας υψηλότερους αριθμούς κέντρων αντίδρασης και αυξάνοντας, έτσι, την απόδοση της αλληλούχησης.

2.1.1 Αλληλούχηση Απολίνωσης

Πρωταρχικά, οι προσεγγίσεις SBL περιλαμβάνουν τον υβριδισμό και τη σύνδεση (Tomkinson, Vijayakumar, Pascal, & Ellenberger, 2006) ανιχνευτή-σήματος και αλληλουχίες αναφοράς σε έναν κλώνο DNA. Οι ανιχνευτές κωδικοποιούν μία ή δύο γνωστές βάσεις και μια σειρά εκφυλισμένων ή καθολικών βάσεων, οδηγώντας σε συμπληρωματική σύνδεση μεταξύ του ανιχνευτή και του προτύπου, ενώ η αλληλουχία αναφοράς κωδικοποιεί μια γνωστή αλληλουχία που παρέχει μια θέση για την έναρξη της απολίνωσης. Μετά την απολίνωση, το πρότυπο απεικονίζεται (DNA imaging) και η γνωστή βάση ή βάσεις στον ανιχνευτή προσδιορίζονται (Landegren, Kaiser, Sanders, & Hood, 1988). Ένας νέος κύκλος ξεκινά μετά την πλήρη αφαίρεση του συμπλέγματος πρότυπο-ανιχνευτή, ή μέσω διάσπασης τους, για την αφαίρεση της φθοροφόρας ομάδας και για την αναγέννηση της θέσης απολίνωσης.

Πλατφόρμες αλληλούχησης όπως η «SOLID» χρησιμοποιούν ανιχνευτές με κωδικοποίηση δύο βάσεων, στους οποίους κάθε φθορομετρικό σήμα αντιπροσωπεύει ένα δινουκλεοτίδιο (Alouene, 2008). Κατά συνέπεια, η παραγωγή κλώνων δεν σχετίζεται απαραίτητα με την ενσωμάτωση ενός γνωστού νουκλεοτιδίου. Επειδή οι 16 πιθανοί συνδυασμοί δινουκλεοτιδίων δεν μπορούν να συσχετιστούν μεμονωμένα με φασματικά διαλυτά φθοροφόρα μόρια, χρησιμοποιούνται τέσσερα φθορίζοντα σήματα, το καθένα από τα οποία αντιπροσωπεύει ένα υποσύνολο τεσσάρων συνδυασμών δινουκλεοτιδίων. Έτσι, κάθε σήμα αντιπροσωπεύει ένα από τα πολλά πιθανά δινουκλεοτίδια, οδηγώντας στον όρο «χρωματικός χώρος» (colour-space), ο οποίος πρέπει να αποσυνελικωθεί κατά την ανάλυση δεδομένων.

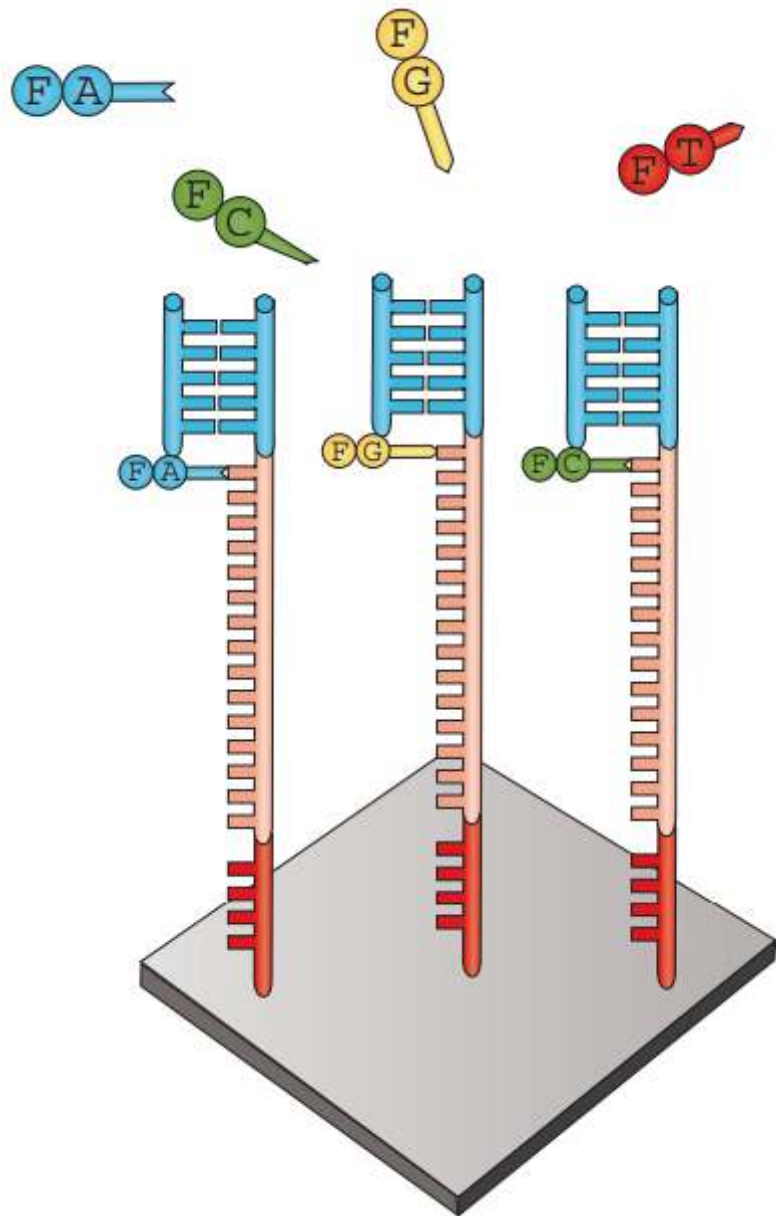
Η διαδικασία προσδιορισμού αλληλουχίας «SOLID» αποτελείται από μια σειρά κύκλων δέσμευσης ανιχνευτή-προτύπου, απολίνωσης, απεικόνισης και διάσπασης για την επιμήκυνση του συμπληρωματικού κλώνου.

2.1.2 Αλληλούχιση Σύνθεσης

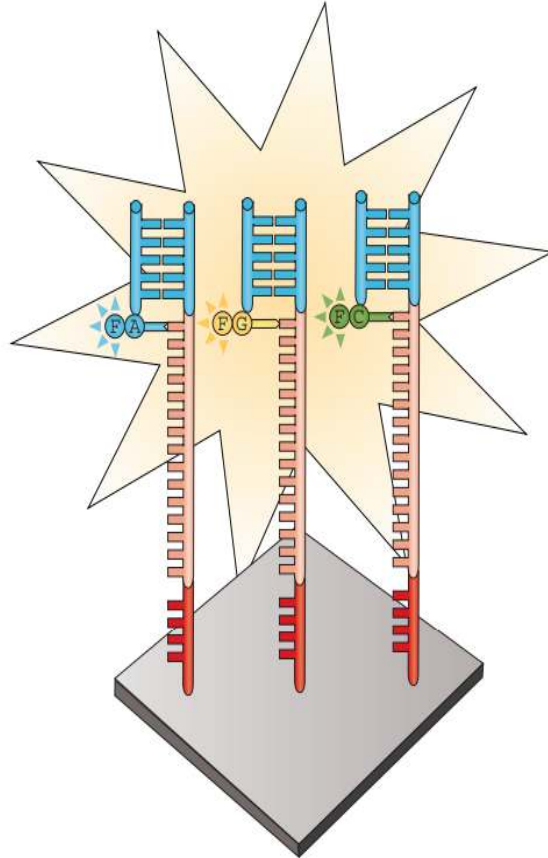
2.1.2.1 Προσεγγίσεις CRT

Το SBS είναι ένας όρος που χρησιμοποιείται για να περιγράψει πολλές μεθόδους που εξαρτώνται από την DNA πολυμεράση στη βιβλιογραφία, αλλά δεν περιγράφει τους διαφορετικούς μηχανισμούς που εμπλέκονται σε αυτές τις προσεγγίσεις. Οι προσεγγίσεις SBS ταξινομούνται στον κυκλικό αναστρέψιμο τερματισμό (Cyclic, Reversible Termination, CRT) και στη προσθήκη μονονουκλεοτιδίου (Single-Nucleotide Addition, SNA) (Metzker, 2010)..

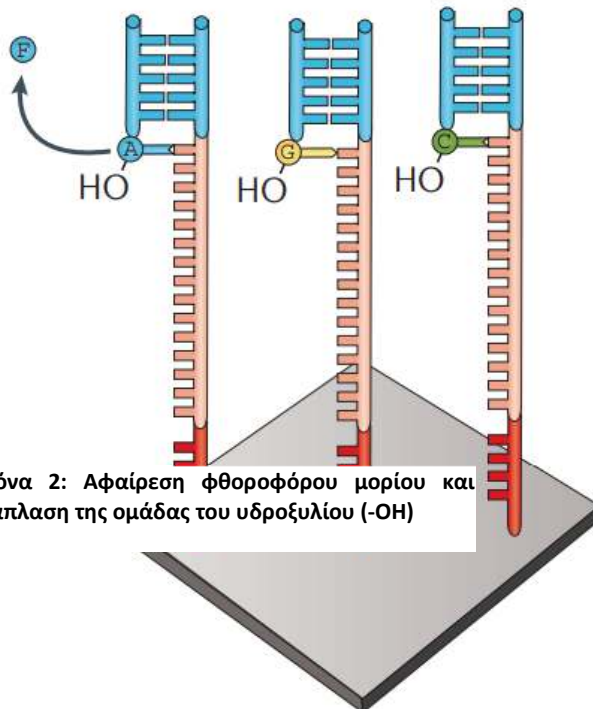
Οι προσεγγίσεις CRT ορίζονται από τη χρήση μορίων τερματισμού που είναι παρόμοια με εκείνα που χρησιμοποιούνται στην αλληλουχία Sanger, στην οποία η ομάδα ριβόζης 3'-OH είναι αποκλεισμένη, αποτρέποντας έτσι την επιμήκυνση του μορίου (Guo, 2008). Για να ξεκινήσει η διαδικασία, ένα πρότυπο DNA εκκινείται από μια αλληλουχία που είναι συμπληρωματική σε μια περιοχή προσαρμογέα, το οποίο θα ξεκινήσει τη σύνδεση της πολυμεράσης σε αυτή την περιοχή δίκλωνου DNA (Double-Stranded DNA, dsDNA). Κατά τη διάρκεια κάθε κύκλου, προστίθεται ένα μείγμα και των τεσσάρων μεμονωμένα επισημασμένων και 3'-δεσμευμένων δεοξυνουκλεοτιδίων (Dye-Labelled Normal Deoxynucleotides, dNTPs). Μετά την ενσωμάτωση ενός μόνο dNTP σε κάθε επιμηκυνόμενο συμπληρωματικό κλώνο, τα μη δεσμευμένα dNTP αφαιρούνται και η επιφάνεια απεικονίζεται για να προσδιοριστεί ποιο dNTP ενσωματώθηκε σε κάθε σύμπλεγμα. Το φθοροφόρο μόριο και η ομάδα αποκλεισμού μπορούν στη συνέχεια να αφαιρεθούν και να ξεκινήσει ένας νέος κύκλος.



Εικόνα 1: Προσθήκη νουκλεοτιδίων (A,T,C,G), ενωμένων με φθοροφόρο μόριο (F), σε κλώνο DNA



Εικόνα 3: Απεικόνιση DNA (DNA Imaging) με σκανάρισμα 2-4 καναλιών λέιζερ. Διαφορετικό χρώμα εκπομπής για κάθε αζωτούχο βάση.



Εικόνα 2: Αφαίρεση φθοροφόρου μορίου και ανάπλαση της ομάδας του υδροξυλίου (-OH)

2.1.2.2 Προσεγγίσεις SNA

Σε αντίθεση με το CRT, οι προσεγγίσεις SNA βασίζονται σε ένα μόνο σήμα για να επισημάνουν την ενσωμάτωση ενός dNTP σε μια επιμήκη αλυσίδα. Κατά συνέπεια, καθένα από τα τέσσερα νουκλεοτίδια πρέπει να προστεθεί επαναληπτικά σε μια αντίδραση προσδιορισμού αλληλουχίας για να διασφαλιστεί ότι μόνο ένα dNTP είναι υπεύθυνο για το σήμα. Επιπλέον, αυτό δεν απαιτεί αποκλεισμό των dNTP, καθώς η απουσία του επόμενου νουκλεοτιδίου στην αντίδραση προσδιορισμού αλληλουχίας αποτρέπει την επιμήκυνση. Η εξαίρεση σε αυτό είναι οι περιοχές ομοπολυμερούς όπου προστίθενται πανομοιότυπα dNTP, με την ταυτοποίηση της αλληλουχίας να βασίζεται σε μια αναλογική αύξηση του σήματος καθώς ενσωματώνονται πολλαπλά dNTP.

2.2 Αλληλούχιση Μακράς Ανάγνωσης

Έχει καταστεί προφανές ότι τα γονιδιώματα είναι εξαιρετικά πολύπλοκα με πολλά επαναλαμβανόμενα στοιχεία, αλλοιώσεις αριθμού αντιγράφων και δομικές παραλλαγές που σχετίζονται με την εξέλιξη, την προσαρμογή και τις ασθένειες (McCarroll & Altshuler, 2007). Ωστόσο, πολλά από αυτά τα πολύπλοκα στοιχεία είναι τόσο μεγάλα σε έκταση, όπου οι τεχνολογίες αλληλούχισης σύντομης ανάγνωσης δεν επαρκούν για την επίλυσή τους.

Η αλληλούχιση μακράς ανάγνωσης παρέχει αναγνώσεις της τάξης πολλών κιλοβάσεων, επιτρέποντας την ανάλυση μιας μεγάλης πληθώρας δομικών χαρακτηριστικών. Τέτοιες μεγάλες αναγνώσεις μπορούν να εκτείνονται σε σύνθετες ή επαναλαμβανόμενες περιοχές με μία μόνο συνεχή ανάγνωση, εξαλείφοντας έτσι την ασάφεια στις θέσεις ή το μέγεθος των γονιδιωματικών στοιχείων. Οι μακρές αναγνώσεις μπορούν επίσης να είναι χρήσιμες για μεταγραφικές έρευνες, καθώς είναι ικανές να καλύπτουν ολόκληρα μεταγραφήματα mRNA, επιτρέποντας στους ερευνητές να προσδιορίσουν την ακριβή συνδεσιμότητα των εξονίων και να διακρίνουν τους πολυμορφισμούς των γονιδίων.

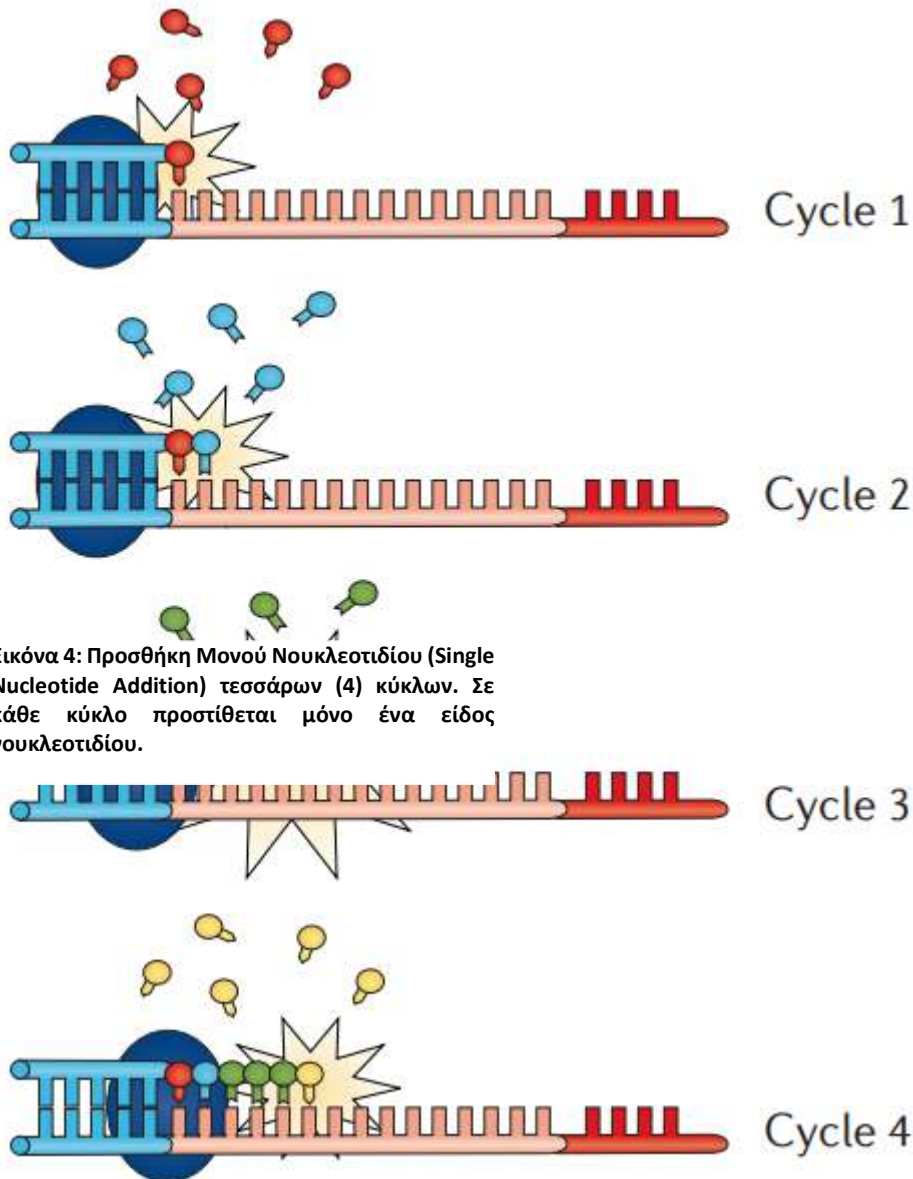
Επί του παρόντος, υπάρχουν δύο κύριοι τύποι τεχνολογιών μακράς ανάγνωσης: προσεγγίσεις αλληλουχίας ενός μορίου σε πραγματικό χρόνο και συνθετικές προσεγγίσεις που βασίζονται σε υπάρχουσες τεχνολογίες σύντομης ανάγνωσης για την κατασκευή μακρών αναγνώσεων *in silico*. Οι προσεγγίσεις «μονού μορίου» (single-molecule approaches) διαφέρουν από τις προσεγγίσεις σύντομης ανάγνωσης στο ότι δεν βασίζονται σε έναν κλωνικό πληθυσμό

ενισχυμένων μορίων DNA για τη δημιουργία ανιχνεύσιμου σήματος, ούτε απαιτούν κύκλο αλληλούχισης για κάθε dNTP που προστίθεται.

2.2.1 Αλληλούχιση Μακράς Ανάγνωσης Μονού Μορίου

Επί του παρόντος, η πιο ευρέως χρησιμοποιούμενη πλατφόρμα μακράς ανάγνωσης είναι η προσέγγιση αλληλουχίας μονού μορίου σε πραγματικό χρόνο (Single-Molecule Real-Time, SMRT) που χρησιμοποιείται από την Pacific Biosciences (PacBio) (Eid, 2009). Η συσκευή χρησιμοποιεί μια εξειδικευμένη κυψέλη ροής με πολλές χιλιάδες μεμονωμένα πηγάδια πικολίτρων με διαφανείς κυματοδηγούς μηδενικής λειτουργίας (Zero-Mode Waveguides, ZMW). Οι τεχνολογίες SBS σύντομης ανάγνωσης δεσμεύουν το DNA και επιτρέπουν στην πολυμεράση να κινείται κατά μήκος του προτύπου DNA, ενώ το σύστημα PacBio στερεώνει την πολυμεράση στο κάτω μέρος του και επιτρέπει στον κλώνο του DNA να προχωρήσει μέσω του ZMW. Έχοντας μια σταθερή θέση ενσωμάτωσης λόγω του στατικού ενζύμου, το σύστημα μπορεί να επικεντρωθεί σε ένα μόνο μόριο. Η ενσωμάτωση dNTP σε κάθε πρότυπο μεμονωμένου μορίου ανά πηγάδι οπτικοποιείται συνεχώς με ένα σύστημα λέιζερ και κάμερας που καταγράφει το χρώμα και τη διάρκεια του εκπεμπόμενου φωτός καθώς το επισημασμένο νουκλεοτίδιο σταματά στιγμιαία κατά την ενσωμάτωση στο κάτω μέρος του ZMW.

Η πολυμεράση διασπά το φθοροφόρο μόριο που είναι συνδεδεμένο με dNTP κατά τη διάρκεια της ενσωμάτωσης, επιτρέποντάς του να διαχέεται μακριά από την περιοχή του αισθητήρα προτού ενσωματωθεί το επόμενο σημασμένο dNTP. Η πλατφόρμα SMRT χρησιμοποιεί επίσης ένα μοναδικό κυκλικό πρότυπο που επιτρέπει σε κάθε πρότυπο να αναλύει την επιθυμητή αλληλουχία πολλές φορές καθώς η πολυμεράση διασχίζει επανειλημμένα το κυκλικό μόριο. Αν και είναι δύσκολο για πρότυπα DNA μεγαλύτερα από ~3 kb να αναλυθούν πολλές φορές, τα μικρότερα πρότυπα DNA μπορούν να αναλυθούν πολλές φορές ως συνάρτηση του μήκους του προτύπου (Eid, 2009). Αυτά τα πολλαπλά περάσματα χρησιμοποιούνται για τη δημιουργία μιας συναινετικής ανάγνωσης του δείγματος, γνωστή ως αλληλουχία κυκλικής συναίνεσης (Circular Consensus Sequence, CCS).



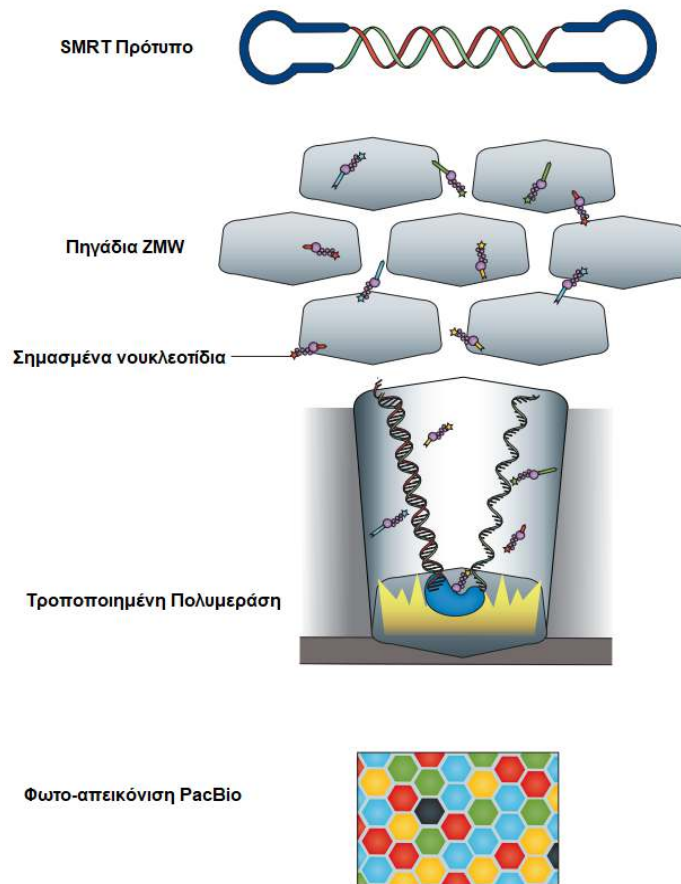
Εικόνα 4: Προσθήκη Μονού Νουκλεοτιδίου (Single Nucleotide Addition) τεσσάρων (4) κύκλων. Σε κάθε κύκλο προστίθεται μόνο ένα είδος νουκλεοτιδίου.

2.2.2 Συνθετική Μακρά Ανάγνωση

Σε αντίθεση με τις πραγματικές πλατφόρμες αλληλούχισης, η συνθετική τεχνολογία μακράς ανάγνωσης βασίζεται σε ένα σύστημα γραμμικής κωδικοποίησης για τη συσχέτιση μορίων που αλληλουχούνται σε υφιστάμενες ακολουθίες σύντομης ανάγνωσης (Voskobougnik, 2013). Αυτές οι προσεγγίσεις κατανέμουν μεγάλα μόρια DNA είτε σε πηγάδια μικροτιτλοδότησης (microtitre wells) είτε σε μορφή γαλακτώματος, έτσι ώστε να υπάρχουν πολύ λίγα μόρια σε κάθε δείγμα.

Μέσα σε κάθε δείγμα, τα επιθυμητά μέρη του προτύπου DNA κόβονται και κωδικοποιούνται με γραμμικό κώδικα. Αυτή η προσέγγιση επιτρέπει την αλληλούχιση σε υπάρχοντα όργανα

σύντομης ανάγνωσης, μετά την οποία τα δεδομένα χωρίζονται κώδικα και συναρμολογούνται εκ νέου με τη γνώση ότι τα δείγματα με ίδιους κωδικούς προέρχονται από το ίδιο αρχικό μόριο (Schatz, Delcher, & Salzberg, 2010). Παρόμοια με μια παλαιότερη τεχνολογία, την αλληλούχιση BAC-by-BAC, οι συνθετικές αναγνώσεις με γραμμικό κώδικα παρέχουν μια συσχέτιση μεταξύ μικρών μορίων που προέρχονται από ένα μεγαλύτερο. Διαχωρίζοντας τα μόρια, μπορούν να απομονωθούν επαναλαμβανόμενες ή περίπλοκες περιοχές, επιτρέποντας σε καθεμία να συναρμολογηθεί τοπικά. Αυτό αποτρέπει στα μη επιλύσιμα σημεία διακλάδωσης να συναρμολογηθούν, γεγονός που οδηγεί σε σπασίματα (κενά) και μικρότερες συναρμολογημένες συνεχόμενες ακολουθίες.



Εικόνα 5: Αλληλούχιση μακράς ανάγνωσης πραγματικού χρόνου.

Κεφάλαιο 3: Γονιδιακές Αλληλουχίες 16S rRNA και ITS

Η χρήση της γονιδιακής αλληλούχησης 16S rRNA για τη μελέτη της βακτηριακής φυλογένεσης και ταξινόμησης ήταν μακράν ο πιο κοινός in-house γενετικός δείκτης που χρησιμοποιείται για διάφορους λόγους. Αυτοί οι λόγοι περιλαμβάνουν:

- i. την παρουσία του σε όλα σχεδόν τα βακτήρια, που συχνά αποτελούν μια πολυγονιδιακή οικογένεια, ή σε οπερόνια.
- ii. το γεγονός ότι η λειτουργία του γονιδίου 16S rRNA με την πάροδο του χρόνου δεν έχει αλλάξει, υποδηλώνοντας ότι οι τυχαίες αλλαγές της ακολουθίας είναι το πιο ακριβές μέτρο της εξέλιξης των βακτηρίων, και ότι
- iii. το γονίδιο 16S rRNA (1.500 bp) είναι αρκετά μεγάλο σε εύρος για σκοπούς πληροφορικής μετεπεξεργασίας

3.1 Ο βακτηριακός προσδιορισμός με τη χρήση της αλληλούχησης 16S rRNA

Μία από τις πιο ενδιαφέρουσες πιθανές χρήσεις της πληροφορικής για την αλληλούχηση 16S rRNA είναι η παροχή αναγνώρισης γένους και είδους για απομονωμένα δείγματα που δεν ταιριάζουν με κανένα αναγνωρισμένο βιοχημικό προφίλ, για στελέχη που δημιουργούν μόνο «χαμηλή πιθανότητα» ή «αποδεκτή» αναγνώριση σύμφωνα με συστήματα του εμπορίου, ή για τάξεις που σπάνια συνδέονται με ανθρώπινες μολυσματικές ασθένειες. Τα αποτελέσματα από έναν περιορισμένο αριθμό μελετών μέχρι σήμερα υποδηλώνουν ότι η αλληλουχία του γονιδίου 16S rRNA παρέχει ταυτοποίηση γένους στις περισσότερες περιπτώσεις (>90%), αλλά παρέχει ταυτοποίηση του είδους σε μικρότερο βαθμό (65 έως 83%), με το 1 έως 14% των απομονωμένων δειγμάτων να παραμένουν άγνωστα μετά τη δοκιμή (Mignard & Fladhnrois, 2006). Οι δυσκολίες που αντιμετωπίζονται στην απόκτηση αναγνώρισης γένους και είδους περιλαμβάνουν την αναγνώριση νέων ταξινομικών κατηγοριών, τις εξαιρετικά λίγες αλληλουχίες που έχουν κατατεθεί σε βάσεις δεδομένων νουκλεοτιδίων, τα είδη που μοιράζονται παρόμοια ή/και πανομοιότυπα 16S rRNAs και τα προβλήματα αλληλουχίας και ονοματολογίας που προκύπτουν από πολλαπλές γονιδιωματικές δοκιμές που έχουν εκχωρηθεί σε μεμονωμένα είδη ή σύμπλοκα ειδών.

Οι έρευνες εξέτασαν τη σκοπιμότητα ταυτοποίησης κλινικών απομονώσεων ρουτίνας ή συγκεκριμένων ομάδων κλινικά σημαντικών βακτηρίων χρησιμοποιώντας δεδομένα γονιδιακής αλληλουχίας μικρής υπομονάδας (Small Subunit, SSU). Σε κάθε μία από αυτές τις μελέτες, τα

δεδομένα αλληλουχίας SSU συγκρίθηκαν με τα αποτελέσματα αναγνώρισης που ελήφθησαν είτε σε συμβατικές είτε σε εμπορικές μορφές δοκιμής (Πίνακας 1).

Πίνακας 1: Προσδιορισμός ειδών μέσω 16S αλληλούχησης (Janda & Abbott, 2007)

Αριθμός στελέχους	Ομάδα υπο μελέτη	16S		Μέγεθος (bp)	Βάση Δεδομένων	Κριτήρια (%)	Εμπορικά Συστήματα	Προσδιορισμός είδους (%)	
								16S	Conv
72	GNB	1,189, 527, 418	MicroSeq	CM	Conv, MIDI, Biolog	90	67.7-84.6	89.2	16
328	Mycobacteria	500	MicroSeq	≥99	Conv.	42		62.5	8
83	GNB, GPB	527	MicroSeq	CM	Vitek 2, Phoenix		77.1	100	6
231	<i>Bacteroides</i>	899, 711	GenBank	≥99	Conv	74.5		83.1	14
47	CNS	1,500	GenBank	>97	API StaphID, Phoenix		63.8-85.1	87.2	9
20	GPA	1,500	GenBank	≥98	Vitek ANA, RapID ANA II,		20-45	65	10
			MicroSeq		API 20A				
107	GNNFB	796	GenBank, EMBL, DDBJ	≥99	API 20NE, Vitek 2		53.2-54.2	91.6	2

Μπορούν να γίνουν μερικές γενικές παρατηρήσεις από αυτές τις έρευνες, συγκεκριμένα,

- (i) ένα υψηλότερο ποσοστό ταυτοποιήσεων ειδών λήφθηκε χρησιμοποιώντας αποτελέσματα αλληλουχίας SSU από ό,τι με συμβατικές ή εμπορικές μεθόδους και

- (ii) οι περισσότερες μελέτες, με εξαίρεση μία μελέτη από τους Fontana et al., έχουν βρει ότι η τεχνολογία 16S απέδωσε ποσοστά αναγνώρισης ειδών από 62 έως 91%. Στη μελέτη των Fontana et al., η πλησιέστερη αντιστοίχιση στη βάση δεδομένων MicroSeq 500 θεωρήθηκε ως «επιτυχής αναγνώριση» ανεξάρτητα από τη βαθμολογία της δοκιμής (Fontana, Favaro, Pelliccioni, Pistoia, & Favalli, 2005).

Για τα βακτήρια που είναι δύσκολο να αναπτυχθούν ή να αναγνωριστούν, τα ποσοστά ταυτοποίησης ήταν χαμηλότερα με την αλληλουχία 16S rRNA (62 έως 83%) από τις τιμές που ήταν αποδεκτές σε κλινικό εργαστήριο (δηλαδή, $\geq 90\%$) (Patel, 2001). Τα προβλήματα και πάλι αφορούσαν τις ακριβείς βάσεις δεδομένων και τις ομάδες που δεν διακρίνονται εύκολα από την αλληλούχηση του γονιδίου 16S rRNA (Hall, Doerr, Wohlfiel, & Roberts, 2003).

Είναι σαφές από τις πληροφορίες που παρατίθενται στον Πίνακα 1 ότι οι πληροφορίες αλληλούχησης γονιδίου 16S rRNA έχουν έναν διευρυνόμενο ρόλο στην ταυτοποίηση βακτηρίων σε κλινικά ή δημόσια περιβάλλοντα υγείας. Ωστόσο, τα δεδομένα δείχνουν επίσης ξεκάθαρα ότι η τεχνολογία αυτή δεν είναι αλάνθαστη και ότι δεν πρέπει να εφαρμόζεται σε κάθε περίπτωση.

3.2 Βακτηριακή ονοματολογία σε σχέση με την γονιδιακή αλληλούχηση 16S rRNA

Υπήρχαν περισσότερα από 1.700 είδη στους Εγκεκριμένους Καταλόγους του 1980, αλλά αυτό δεν σημαίνει ότι όλες οι καταγεγραμμένες τάξεις είναι έγκυρες. Πολλά ονόματα περιελάμβαναν προγενέστερες μελέτες υβριδισμού DNA-DNA και φυλογενετικές έρευνες. Επομένως, τα στελέχη για πολλά είδη μπορεί να μην αντικατοπτρίζουν με ακρίβεια ολόκληρη τη γονιδιωματική σύνθεση του κάθε είδους. Σε τέτοιες περιπτώσεις ο προσδιορισμός του είδους έχει άμεση σχέση με τις μελέτες SSU με αναφορά τη μικροβιακή ταυτοποίηση. Ορισμένα βακτηριακά είδη υπάρχουν ως «φαινοείδη» ή «συμπλέγματα», δηλαδή υπάρχουν περισσότερα από ένα γονιδίωμα σε αυτό το είδος και δεν μπορούν να διαχωριστούν φαινοτυπικά. Παραδείγματα τέτοιων καταστάσεων αποτελούν τα *Enterobacter cloacae* (τουλάχιστον 7 γονιδιωματικά αρχικά), το *Pseudomonas stutzeri* (αρχικά 18 γονιδιωματικά) και το γένος *Acinetobacter* (22 γονιδιωματικά αρχικά) (Janda & Abbott, 2007).

3.3 Επίλυση των προβλημάτων

Αν και η γονιδιακή αλληλούχηση 16S rRNA είναι πολύ χρήσιμη όσον αφορά την ταξινόμηση των βακτηρίων, έχει χαμηλή φυλογενετική ισχύ σε επίπεδο είδους και φτωχή διακριτική δύναμη για ορισμένα γένη, ενώ είναι απαραίτητες οι μελέτες σχετικά με το DNA για την παροχή απόλυτης επίλυσης σε αυτά τα ταξινομικά προβλήματα. Το γένος *Bacillus* είναι ένα καλό παράδειγμα αυτού.

Τα στελέχη τύπου του *B. globisporus* και του *B. psychrophilus* μοιράζονται περί το 99,5% της αλληλουχίας τους όσον αφορά τα γονίδια 16S rRNA τους, και, σε επίπεδο DNA, εμφανίζουν μόνο 23 έως 50% συσχέτιση στις αντιδράσεις αμοιβαίας υβριδοποίησης (Fox, Wisotzkey, & Jurtshuk, 2005). Τέτοια παραδείγματα υποδεικνύουν ότι η ομοιότητα της αλληλουχίας SSU ακόμη και σε πολύ υψηλό επίπεδο δεν υποδηλώνει σε κάθε περίπτωση ταυτότητα ή ακρίβεια στις μικροβιακές ταυτοποιήσεις.

Πολλοί ερευνητές έχουν βρει προβλήματα σε επίπεδο γένους και/ή είδους με δεδομένα αλληλουχίας γονιδίου 16S rRNA (Πίνακας 2). Αυτές οι ομάδες περιλαμβάνουν (όχι αποκλειστικά), την οικογένεια Enterobacteriaceae (ιδίως, *Enterobacter* και *Pantoea*), τα ταχέως αναπτυσσόμενα μυκοβακτήρια, το *Acinetobacter baumannii*-A. σύμπλοκο *calcoaceticus*, *Achromobacter*, *Stenotrophomonas* και *Actinomyces*. Μερικά από αυτά τα προβλήματα σχετίζονται με την ονοματολογία και την ταξινόμηση των βακτηρίων, ενώ άλλα σχετίζονται με διαφορετικά ζητήματα που αναφέρονται παρακάτω.

Πίνακας 2: Γένη και είδη βακτηρίων με προβλήματα ταξινόμησης (Janda & Abbott, 2007)

Γένος	Είδος
<i>Aeromonas</i>	<i>A. veronii</i>
<i>Bacillus</i>	<i>B. anthracis</i> , <i>B. cereus</i> , <i>B. globisporus</i> , <i>B. psychrophilus</i>
<i>Bordetella</i>	<i>B. bronchiseptica</i> , <i>B. parapertussis</i> , <i>B. pertussis</i>
<i>Burkholderia</i>	<i>B. cocovenenans</i> , <i>B. gladioli</i> , <i>B. pseudomallei</i> , <i>B. thailandensis</i>
<i>Campylobacter</i>	Non-jejuni-coli group
<i>Edwardsiella</i>	<i>E. tarda</i> , <i>E. hoshinae</i> , <i>E. ictaluri</i>
<i>Enterobacter</i>	<i>E. cloacae</i>
<i>Neisseria</i>	<i>N. cinerea</i> , <i>N. meningitidis</i>
<i>Pseudomonas</i>	<i>P. fluorescens</i> , <i>P. jessenii</i>
<i>Streptococcus</i>	<i>S. mitis</i> , <i>S. oralis</i> , <i>S. pneumoniae</i>

Ένα επιπλέον πρόβλημα σχετικά με την ανάλυση της αλληλουχίας του γονιδίου 16S rRNA αφορά την ταυτότητα αλληλουχίας ή τις πολύ υψηλές βαθμολογίες ομοιότητας. Οι βιβλιογραφικές αναφορές έχουν τεκμηριώσει ομοιότητες ή ταυτότητα αλληλουχίας γονιδίου 16S rRNA για την ομάδα *Streptococcus mitis* και άλλους μη ζυμωτικούς μικροοργανισμούς (Πίνακας 2). Σε τέτοιες περιπτώσεις, τα δεδομένα αλληλούχησης γονιδίου 16S rRNA δεν μπορούν να δώσουν μια οριστική απάντηση αφού δεν μπορούν να διακρίνουν μεταξύ των ειδών. Σε άλλες περιπτώσεις, η διαφορά μεταξύ της πλησιέστερης και της επόμενης πλησιέστερης αντιστοίχισης με το άγνωστο στέλεχος διαθέτει απόκλιση κοντά στο 0,5% (ή κατά 99,5% ομοιότητα). Υπό αυτές τις συνθήκες, τέτοιες μικρές διαφορές δεν μπορούν να δικαιολογήσουν την επιλογή της πλησιέστερης αντιστοιχίας ως οριστικής ταυτοποίησης, αν και σε ορισμένες μελέτες πραγματοποιήθηκε αυτό ακριβώς (Fontana, Favaro, Pelliccioni, Pistoia, & Favalli, 2005).

3.4 Ιδιωτικές και δημόσιες βάσεις δεδομένων νουκλεοτιδίων

Η χρησιμότητα του προσδιορισμού αλληλουχίας του γονιδίου 16S rRNA ως εργαλείο στη μικροβιακή ταυτοποίηση εξαρτάται από δύο βασικά στοιχεία:

- i. την εναπόθεση πλήρων αδιαμφισβήτητων αλληλουχιών νουκλεοτιδίων σε δημόσιες ή ιδιωτικές βάσεις δεδομένων και
- ii. την εφαρμογή της σωστής σήμανσης σε κάθε αλληλουχία.

Πριν από χρόνια η συνολική ποιότητα των αλληλουχιών νουκλεοτιδίων που κατατέθηκαν σε δημόσιες βάσεις δεδομένων ήταν αμφίβολη, καθώς πολλές καταθέσεις ήταν κακής ποιότητας (Petti, 2007). Μεγάλο μέρος αυτής της παραπληροφόρησης που υπήρχε αρχικά σε τέτοιες βάσεις δεδομένων θεωρήθηκε ότι είχε διορθωθεί. Ωστόσο, μια μελέτη από το Ηνωμένο Βασίλειο εκτίμησε, επιφυλακτικά, ότι τουλάχιστον το 5% από τις 1.399 αλληλουχίες που αναζητήθηκαν είχαν σημαντικά σφάλματα που σχετίζονται με σφάλματα χιμαιρισμού (64%) έως και σφάλματα ή ανωμαλίες αλληλούχησης (35%) (Ashelford, Chuzhanova, Fry, Jones, & Weightman, 2005). Μια μελέτη του 1995 από τους Clayton et al. αποκάλυψε επίσης ότι τουλάχιστον το 26% των ζευγών αλληλουχιών γονιδίου 16S rRNA (δύο αλληλουχίες που κατατέθηκαν για το ίδιο είδος) στο GenBank είχαν >1% τυχαία σφάλματα αλληλουχίας και, από αυτά, σχεδόν τα μισά είχαν >2% τυχαία σφάλματα αλληλουχίας (Clayton, Sutton, Hinkle, Bult, & Fields, 1995).

3.5 Επιλογή γονιδίων και εκκινητών

Το γονίδιο 16S rRNA έχει μήκος περίπου 1.500 νουκλεοτίδια (~1,5 kb, αν και αυτός είναι ένας μέσος όρος και ορισμένοι οργανισμοί μπορεί να έχουν μικρότερες ή μεγαλύτερες αλληλουχίες 16S) και είναι μέρος της μικρής υπομονάδας 30S των προκαρυωτικών ριβοσωμάτων που συνδέεται με την αλληλουχία Shine-Dalgarno στο 3' τέλος (Baker, Smith, & Cowan, 2003). Το 16S rRNA έχει διάφορες λειτουργίες, συμπεριλαμβανομένου ενός δομικού ρόλου καθώς και ρόλους καθοριστικής σημασίας για τη σύνθεση πρωτεϊνών. Μαζί με το 23S, παρέχει ένα «ικρίωμα» (scaffold) για να βοηθήσει στη σύνδεση των ριβοσωμικών υπομονάδων 50S και 30S, καθώς και στον καθορισμό των θέσεων της ριβοσωματικής πρωτεΐνης (Woese, Gutell, Gupta, & Noller, 1983).

Το 3' άκρο του 16S RNA συνδέεται επίσης με τις πρωτεΐνες S1 και S21, που είναι γνωστό ότι εμπλέκονται στην έναρξη της πρωτεϊνοσύνθεσης με διασύνδεση RNA-πρωτεΐνης (Odom, Deng, Dabbs, & Hardesty, 1984). Όλοι οι μικροοργανισμοί έχουν τουλάχιστον ένα αντίγραφο του 16S, καθιστώντας το πανταχού παρόν, και καθώς είναι εξαιρετικά συντηρημένο και εξελίσσεται αργά, είναι ο πιο ευρέως χρησιμοποιούμενος μεμονωμένος στόχος για φυλογενετικές μελέτες βακτηρίων και αρχαίων, εδώ και πολλά χρόνια (Doolittle, 1999). Πλέον, έχει βρεθεί ότι πολλαπλές αλληλουχίες του 16S μπορεί να υπάρχουν σε ένα μόνο βακτήριο και ορισμένα αντίγραφα μπορεί να διαφέρουν (Louca, Doebeli, & Parfrey, 2018). Οι μελέτες γονιδιωματικής αλληλουχίας δείχνουν επίσης ότι πολλά βακτηριακά είδη έχουν ενδογονική ετερογένεια (δηλ. φιλοξενούν πολλαπλά αντίγραφα γονιδίου 16S και πολυμορφισμούς μεταξύ αυτών των αντιγράφων) που επιτρέπουν την υποτυποποίηση μεταξύ των ειδών μέσω μερικής ή πλήρους αλληλουχίας του 16S (Sun, Jiang, Wu, & Zhou, 2013). Η διάκριση μεταξύ των ειδών με βάση την ενδογονική ετερογένεια έχει αποδειχθεί για μια ποικιλία ανθρώπινων παθογόνων, συμπεριλαμβανομένων των *Neisseria*, *Haemophilus*, *Salmonella*, και *Listeria* (Deidre, et al., 2020). Οριζόντια μεταφορά του 16S συμβαίνει επίσης, αν και σπάνια και μόνο σε επίπεδο ενδογόνου ή ενδοειδούς. Αν και αυτό είναι πολύ πιο περιορισμένο από τη βακτηριακή οριζόντια μεταφορά λειτουργικών γονιδίων (δηλ. γονίδια που κωδικοποιούν τα ένζυμα), ορισμένοι ερευνητές αμφισβητούν εάν το 16S θα πρέπει να είναι ο μόνος στόχος που χρησιμοποιείται για ταυτοποίηση ή φυλογενετικούς σκοπούς. Ωστόσο, το 16S έχει μελετηθεί και εφαρμοστεί εκτενώς

για την καθιέρωση μιας περιγραφής ειδών, ταξινόμησης ειδών και φυλογενετικών σχέσεων. Επομένως, το 16S είναι ο μοριακός στόχος επιλογής για την αναγνώριση του γένους ή του είδους αναγνώριση, σε κλινικό επίπεδο, λόγω της οικουμενικής του φύσης μεταξύ βακτηρίων και αρχαίων (~ 10% έως 15%) και της πληθώρας δεδομένων αλληλουχίας σε σύγκριση με αυτών για άλλους στόχους (Deidre, et al., 2020).

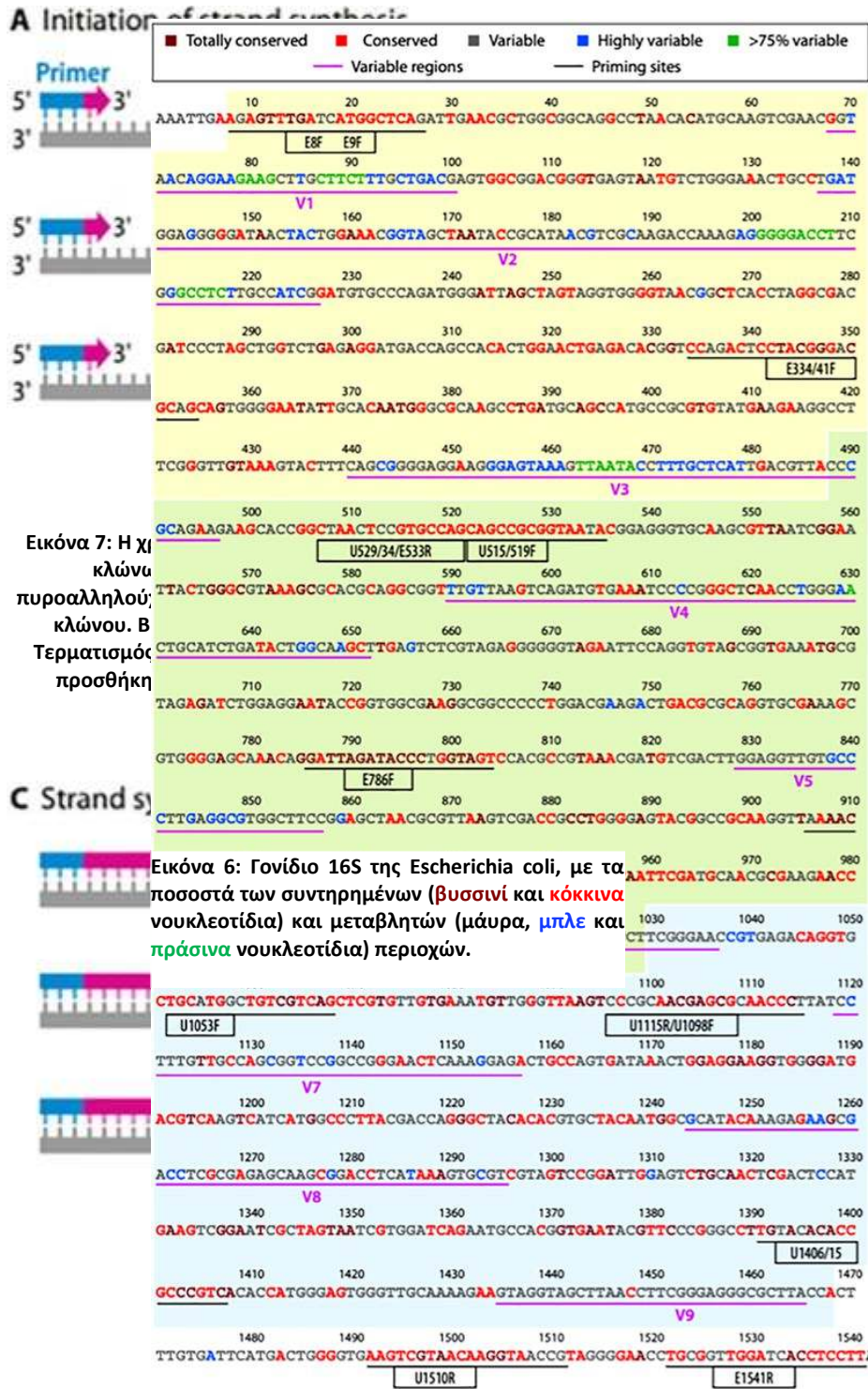
Το γονίδιο 16S περιέχει «μωσαϊκά» (mosaics) αλληλουχίας που κυμαίνονται από εξαιρετικά συντηρημένες, μεταβλητές και υπερμεταβλητές περιοχές, όπως επεξηγήθηκε προηγουμένως από τους Baker et al. στο σχηματικό τους σχήμα του *Escherichia coli* (βλ. Εικόνα 6) (Baker, Smith, & Cowan, 2003).

Εντός ορισμένων περιοχών, το 16S παρέχει περιοχές ειδικές για το γένος ή/και το είδος που καθιστούν δυνατή την ακριβή ταυτοποίηση ανάλογα με τις στοχευόμενες γονιδιακές περιοχές για ένα συγκεκριμένο βακτήριο/ομάδες μικροοργανισμών.

Οι εκκινητές (primers) του γονιδίου 16S μπορούν να σχεδιαστούν για να στοχεύουν τις συντηρημένες περιοχές του, από τις οποίες εξάγονται ορισμένα μοτίβα που είναι κοινά σε ολόκληρο το βασίλειο των βακτηρίων (Οι εκκινητές αυτοί είναι γνωστοί ως «eubacterial primers»). Στην εργαστηριακή πρακτική, αυτοί οι εκκινητές στοχεύουν τα πρώτα 500 περίπου ζεύγη βάσεων του γονιδίου της μικρής ριβοσωματικής υπομονάδας, επειδή η ανάλυση των περιοχών V1-V3 (βλ. Εικόνα 6) θεωρούνται επαρκείς για να επιτρέπουν την ακριβή αναγνώριση των περισσότερων γενών/ειδών. Ως εκ τούτου, οι περισσότερες αλληλουχίες 16S που κατατίθενται επί του παρόντος σε δημόσιες βάσεις δεδομένων αντιστοιχούν σε αυτό το τμήμα του γονιδίου.

Είναι σημαντικό να ληφθεί υπόψιν ότι ο εκκινητής «προς τα εμπρός» (Forward, ορίζεται ως η ολιγονουκλεοτιδική αλληλουχία που είναι συμπληρωματική προς τον αντιπληροφοριακό κλώνο του δίκλωνου DNA) και ο αντίστροφος (Reverse, ορίζεται ως αλληλουχία ολιγονουκλεοτιδίου που είναι συμπληρωματική προς τον πληροφοριακό κλώνο του δίκλωνου DNA) σύνολα που στοχεύουν στις V1-V3 περιοχές του 16S, σχεδιάστηκαν ιστορικά για ανάλυση περιβαλλοντικής κοινότητας μικροβιώματος και όχι για κλινικά απομονωμένα στελέχη (Εικ. 1) (Deidre, et al., 2020). Η πιο αυστηρή αναγνώριση των ανθρώπινων παθογόνων απαιτεί τη χρήση άλλων ζευγών εκκινητών, Forward/Reverse, ειδικών για το 16S, όπως αναφέρθηκε προηγουμένως. Αν και προηγούμενες αναλύσεις δείχνουν ότι είναι δύσκολο να σχεδιαστούν

εκκινητές για την καθολική ανίχνευση όλων των προκαρυωτικών αλληλουχιών γονιδίου 16S rRNA, αυτό έχει επιτευχθεί ως έναν βαθμό.



Εικόνα 7: Η χημική δομή των κλώνων πυροαλληλούκλωνου. Β. Τερματισμός προσθήκη

Εικόνα 6: Γονίδιο 16S της Escherichia coli, με τα ποσοστά των συντηρημένων (βυσσινί και κόκκινα νουκλεοτίδια) και μεταβλητών (μάρια, μπλε και πράσινα νουκλεοτίδια) περιοχών.

V1-V3

V4-V6

V7-V9

where the dNTP is d by -H

E

3.6 Ακρίβεια αποτελεσμάτων

Η ακρίβεια αντιστοίχισης είναι ο βαθμός ομοιότητας μεταξύ της αλληλουχίας, του προς ανάλυση δείγματος και της ακολουθίας αναφοράς. Όσο μεγαλύτερη είναι η ομοιότητα, τόσο καλύτερη η αντιστοίχιση. Ωστόσο, πρέπει να ληφθούν υπόψιν τα ακόλουθα ζητήματα:

- οι κωδικοί ασάφειας (IUB, IUPAC) ερμηνεύονται από τον αλγόριθμο αναζήτησης BLAST ως πλήρεις αναντιστοιχίες, επομένως μια ακολουθία που περιέχει έναν αριθμό διφορούμενων βάσεων θα κατατάσσεται χαμηλότερα στη λίστα παρά την παρουσία μερικών αντιστοιχιζόμενων βάσεων.
- Οι βαθμολογίες BLAST, επιπλέον, κατατάσσουν τις αντίστοιχες αλληλουχίες αναφοράς σύμφωνα με το μήκος της ακολουθίας πάνω από τον αριθμό αναντιστοιχίας. Έτσι, μια ανακτημένη ακολουθία αναφοράς μπορεί να εμφανιστεί σε υψηλότερη θέση κατάταξης, σύμφωνα με τη βαθμολογία BLAST, με βάση το συνολικό μήκος της αντιστοιχίας, παρόλο που έχει περισσότερες αναντιστοιχίες από μια άλλη ακολουθία χαμηλότερης κατάταξης με βάση τη μικρότερη διάρκεια αγώνα.

Στο παρελθόν, ένα όριο ομοιότητας >98,5% είχε προταθεί ως εμπειρικός κανόνας για την ανάθεση μιας ακολουθίας δειγμάτων 16S σε ένα συγκεκριμένο είδος. Αυτή η απλή αποκοπή καθιστά ευκολότερη την ερμηνεία των αποτελεσμάτων στα διαγνωστικά εργαστήρια, αλλά μπορεί ταυτόχρονα να αποτελέσει παραπλανητική (Clarridge, 2004).

3.6.1 Μήκος αντιστοίχισης

Οι αλληλουχίες αναφοράς πρέπει να αντιστοιχούνται κατά το περισσότερο δυνατό, ή ιδανικά εξ' ολοκλήρου, με την ακολουθία, του προς ανάλυση δείγματος. Σε ίσο αριθμό αναντιστοιχιών, θα πρέπει να προτιμώνται οι μεγαλύτερες αντιστοιχίσεις για αναγνώριση, διασφαλίζοντας έτσι καλύτερη αξιοπιστία. Πρέπει να ληφθεί υπόψη ότι ο αλγόριθμος BLAST τείνει να περικόπτει τις αντιστοιχίσεις ακολουθίας όταν η στοίχιση γίνεται αβέβαιη στο άκρο 5' και 3', λόγω αναντιστοιχιών ή εισαγωγών/διαγραφών. Αυτό μπορεί να οδηγήσει σε αντιστοιχίσεις προτύπων μεταξύ τους, παρά την παρουσία αναντιστοιχιών στα άκρα. Θα πρέπει, επομένως, πάντα να επαληθεύεται το μήκος αντιστοίχισης μιας ακολουθίας ενός δείγματος με μια αλληλουχία αναφοράς πριν εξαχθούν συμπεράσματα για το δείγμα. Σε περιπτώσεις αμφιβολίας, ότι

υπάρχουν αναντιστοιχίες στις άκρες των προτύπων, θα πρέπει να πραγματοποιηθεί μια κατά-ζεύγη, ή μια πολλαπλή στοίχιση που περιλαμβάνει την ακολουθία του δείγματος.

3.6.2 Συνέπεια αντιστοίχισης

Ο κατάλογος των αλληλουχιών αναφοράς που αντιστοιχούνται (λ.χ. τα ονόματα των ειδών και των γενών) θα πρέπει να αναθεωρηθεί ως προς τη συνοχή της ονομασίας εντός του είδους και εντός του γένους (σημειώνεται, επίσης, ότι οι συνεχιζόμενες ταξινομικές αλλαγές ονομάτων έχουν δημιουργήσει σύγχυση για τα κλινικά εργαστήρια σε κλινικά σημαντικούς μικροοργανισμούς). Έτσι, το πρότυπο με το μεγαλύτερο ποσοστό αντιστοίχισης θα πρέπει να υποσημειώνεται με το ίδιο όνομα είδους (υπό την προϋπόθεση ότι η βάση δεδομένων αναφοράς περιέχει πολλαπλές εγγραφές αυτού του είδους) ή με το ίδιο γένος (υπό την προϋπόθεση ότι η βάση δεδομένων αναφοράς περιέχει μόνο μία ή λίγες εγγραφές ανά είδος).

Οι αναντιστοιχίες μπορεί να αντικατοπτρίζουν τη φυσική μεταβλητότητα μεταξύ των ειδών. Εάν τα πρότυπα με την καλύτερη αντιστοίχιση περιέχουν άλλα είδη με ίσο αριθμό αναντιστοιχιών, ενδέχεται να μην είναι δυνατός ο προσδιορισμός του είδους σε αυτό το πλαίσιο. Αντίθετα, σε τέτοιες περιπτώσεις, θα μπορούσε να πραγματοποιηθεί ένας σχετικά ακριβής προσδιορισμός γένους.

Αν τα πρότυπα με τις περισσότερες αντιστοιχίσεις προέρχονται από διαφορετικά γένη με ίσες αναντιστοιχίες και βαθμολογίες, θα πρέπει να πραγματοποιείται ταυτοποίηση μόνο σε επίπεδο οικογένειας (π.χ., βλ. *E. coli* και *Shigella spp.* που ερμηνεύονται ως «*Enterobacteriaceae*» ή «*Enterobacteriales*»). Η συνέπεια αντιστοίχισης μπορεί να είναι ένδειξη προβλημάτων ανεπαρκούς κάλυψης ενός είδους ή ενός γένους σε μια βάση δεδομένων, ή ένδειξη κακής επιλογής της περιοχής αλληλουχίας, η οποία πρέπει να είναι αρκετά μεταβλητή για να υπάρχει διαφοροποίηση μεταξύ ορισμένων ειδών και γενών.

3.6.3 Διαφοροποίηση αντιστοίχισης

Η διαφοροποίηση αντιστοίχισης είναι η ικανότητα να ονομάζουμε ταυτοποίηση ενός είδους, ενώ διασφαλίζουμε ότι κανένα άλλο είδος δεν ταιριάζει επίσης. Λαμβάνεται υπόψη ο αριθμός των αντιστοιχιών και των αναντιστοιχιών μεταξύ της αλληλουχίας απομόνωσης και των αναφορών με την καλύτερη αντιστοίχιση των πλησιέστερων ειδών.

Η λίστα αντιστοίχισης πρέπει να εμφανίζει αντιπροσωπευτικές αλληλουχίες περισσότερων του ενός ειδών για να είναι δυνατή η διαφοροποίηση από το επόμενο πλησιέστερο. Η διαφοροποίηση αντιστοίχισης αναφέρεται στη μεταβλητότητα μεταξύ των ειδών (άρα, ανεκτές αναντιστοιχίες) και στις αναντιστοιχίες μεταξύ των ειδών (επιτρέποντας έτσι τη διαφοροποίηση των ειδών).

Είναι σημαντικό ο κατάλογος αντιστοίχισης να περιλαμβάνει και τα επόμενα πιο κοντινά είδη. Μια πολλαπλή ευθυγράμμιση της αλληλουχίας απομόνωσης με τις καλύτερα αντιστοιχισμένες αλληλουχίες αναφοράς των πλησιέστερων ειδών (δύο ή τρία) είναι συχνά χρήσιμη για την αξιολόγηση της διαφοροποίησης σχετικά με τη θέση των αναντιστοιχιών.

Οι αναντιστοιχίες μιας αλληλουχίας απομόνωσης που εμφανίζονται σε μεταβλητές περιοχές, όπου τα είδη αυτού του γένους διαφέρουν συνήθως, είναι ενδεικτικά μιας μη αντιστοιχίας με ένα είδος και πρέπει να τεκμηριώνονται. Σε αυτές τις περιπτώσεις, μπορεί κανείς να αναφέρει «κοντά» στο πλησιέστερο αντίστοιχο είδος.

Πίνακας 3: Σημαντικοί παράμετροι του αλγορίθμου BLAST για την αλληλούχιση του γονιδίου 16S

Παράμετρος	Ορισμός
Ακρίβεια αντιστοίχισης	Οι αλληλουχίες αναφοράς που ταιριάζουν καλύτερα θα πρέπει να δείχνουν το χαμηλότερο αριθ. αναντιστοιχιών
Μήκος αντιστοίχισης	Οι αλληλουχίες αναφοράς που ταιριάζουν πρέπει να καλύπτουν εξ ολοκλήρου μια απομονωμένη ακολουθία. Μικρότερες αλληλουχίες αντιστοίχισης από τον αλγόριθμο BLAST θα πρέπει να επαληθεύονται με αντιστοίχιση νουκλεοτιδίου-νουκλεοτιδίου, μεταξύ της αλληλουχίας του δείγματος και του κλώνου αναφοράς, για εύρεση των αναντιστοιχιών στις άκρες των κλώνων.
Συνέπεια αντιστοίχισης	Μια ακολουθία που ταιριάζει με ένα αριθμό των αλληλουχιών αναφοράς, μέσα στο ίδιο είδος αυξάνει την εμπιστοσύνη της ακριβούς αναγνώρισής της.
Διαφοροποίηση αντιστοίχισης	Για να είναι δυνατή η εκτίμηση του βαθμού διαφοροποίησης ενός απομονωμένου στελέχους στο επόμενο πλησιέστερο είδος, ακολουθίες που προέρχονται από στενά συγγενικά αλλά διαφορετικά είδη θα πρέπει να εμφανίζονται στη λίστα των αντίστοιχων αλληλουχιών αναφοράς, που θα αξιολογηθούν εντός αυτού του πλαισίου.

3.6.4 Αντιστοίχιση και βάσεις δεδομένων

Η ποσότητα κάλυψης της ακολουθίας 16S, όπως ορίζεται από την επαρκή αναπαράσταση μεταβλητών και διατηρημένων περιοχών που σχετίζονται με το γένος, επηρεάζει την ομοιότητα ενός δείγματος με μια ακολουθία αναφοράς. Εάν οι διατηρημένες τάσεις είναι κυρίαρχες στην ακολουθία δειγμάτων, η ομοιότητα αντιστοίχισης με την καλύτερη.

Η αλληλουχία ταιριάσματος θα υπερβεί την αποκοπή χωρίς να δώσει μια σαφή αναγνώριση. Επομένως, μια βαθμολογία ομοιότητας θα απαιτούσε μια ελάχιστη κάλυψη μεταβλητών περιοχών για το εμπλεκόμενο γένος ή γένη. Η ποικιλομορφία του γένους παίζει επίσης ρόλο, επειδή ορισμένα γένη έχουν μεγάλη ποικιλία ενώ άλλα όχι (π.χ. *Mycobacterium*). Αργώς- αναπτυσσόμενα μυκοβακτήρια, όπως το *M. genavense*, δεν θα διαχωριστούν από άλλα άτυπα μυκοβακτήρια χρησιμοποιώντας αλληλουχία 16S με αποκοπή 98,5% (123). Στην περίπτωση μιας τυπικής αλληλουχίας των πρώτων 500 bp, το *M. genavense* και το *M. triplex* διαφέρουν γενετικά μόνο κατά 4 αναντιστοιχίες ή <1% της αλληλουχίας V1-V3, αλλά μπορούν να διαφοροποιηθούν σαφώς σε αυτή τη βάση όπως περιγράφεται παρακάτω. Η ποικιλομορφία των ειδών παίζει επίσης ρόλο επειδή είδη με μεγάλη ποικιλία, όπως το *Fusobacterium nucleatum*, το οποίο περιλαμβάνει έναν αριθμό τουλάχιστον 5 υποειδών παρουσιάζουν ενδοειδική ποικιλομορφία 10 έως 12 αναντιστοιχιών (>2%) εντός των πρώτων 500 bp μεταξύ των παραλλαγών. Εάν η βάση δεδομένων αναφοράς που χρησιμοποιείται δεν καλύπτει ρητά τα σχετικά υποείδη και παραλλαγές όπως περιγράφονται παρακάτω, η ακολουθία του δείγματος δεν θα αντιστοιχιστεί με αρκετά υψηλή βαθμολογία για την αναγνώριση των ειδών. Η εφικτή ομοιότητα αντιστοίχισης εξαρτάται επίσης από την επαρκή κάλυψη ειδών, υποειδών και παραλλαγών στη βάση δεδομένων αναφοράς που χρησιμοποιείται. Εάν η αντίστοιχη παραλλαγή δεν υπάρχει στη βάση δεδομένων, μια ακολουθία δειγμάτων μπορεί να ταιριάζει μόνο με αναφορές κάτω από την αποκοπή, δίνοντας ένα ασαφές αποτέλεσμα (Church et al., 2020).

Οι αλληλουχίες αναφοράς που λείπουν μπορεί επίσης να οδηγήσουν σε μια σαφή αντιστοίχιση πάνω από το σημείο αποκοπής με μια ακολουθία αναφοράς παρούσα, ενώ το σωστό αποτέλεσμα θα έπρεπε να ήταν διφορούμενο και όχι οριστικό. (Λάβετε υπόψη ότι οι λίστες αντιστοίχισης BLAST θα πρέπει πάντα να ελέγχονται για άλλες πιθανές αντιστοιχίσεις και

σε πολλαπλή ευθυγράμμιση με την ακολουθία δειγμάτων, η συνάφεια των αναντιστοιχιών θα γίνει διαφανής όσον αφορά τις μεταβλητές περιοχές που σχετίζονται με το εμπλεκόμενο γένος).

3.7 Αναγνώριση παθογόνων βακτηρίων με γονιδιακή αλληλούχηση 16S rRNA

Τα κλινικά μικροβιολογικά εργαστήρια πρέπει να είναι σε θέση να εντοπίζουν γρήγορα και με ακρίβεια ένα ευρύ φάσμα βακτηριακών απομονώσεων, προκειμένου να διαγνώσουν την αιτιολογία της λοίμωξης και να παρέχουν καθοδήγηση σχετικά με την κατάλληλη αντιβιοτική θεραπεία. Η μερική ή πλήρης αλληλουχία του 16S έχει αποδειχθεί ότι είναι ένα ανεκτίμητο εργαλείο για την παροχή αξιόπιστης αναγνώρισης λοιμώξεων που προκαλούνται από ασυνήθιστα ή σπάνια συναντώμενα βακτήρια, ιδιαίτερα στην εποχή πριν από το MALDI-TOF MS (Church et al., 2020).

Η ταυτοποίηση επιπέδου των βακτηριακών απομονώσεων λήφθηκε με αλληλούχηση του γονιδίου 16S rRNA σε >90% και 65% έως 83%, αντίστοιχα, ανάλογα με την ομάδα βακτηρίων και τα κριτήρια που χρησιμοποιούνται για τον ορισμό του είδους σε περιπτώσεις όπου οι συμβατικές φαινοτυπικές μέθοδοι είχαν αποτύχει (Church et al., 2020).

Με την εμφάνιση της αλληλουχίας του γονιδίου 16S rRNA ως εργαλείο ταυτοποίησης τα τελευταία 20 χρόνια, η χρησιμότητα των εμπορικών βάσεων δεδομένων έχει επίσης υποστεί περιορισμένη κλινική αξιολόγηση. Το σύστημα αναγνώρισης που βασίζεται σε rDNA MicroSeq 500 16S μπορεί να αναγνωρίσει αξιόπιστα >80% κλινικά σχετικών βακτηριακών απομονώσεων με άτυπα φαινοτυπικά προφίλ και 89.2% ασυνήθιστων αερόβιων Gram-αρνητικών βακίλλων. Έχει επίσης αποδειχθεί χρήσιμο για την ταυτοποίηση ορισμένων βραδέως αναπτυσσόμενων βακτηρίων, όπως τα είδη *Mycobacterium*, παρά τους περιορισμούς της χρήσης αυτού του στόχου (δηλαδή, οι αλληλουχίες 16S δεν μπορούν να διαφοροποιήσουν είδη εντός του συμπλέγματος *M. tuberculosis avium intracellulare*, ή το *M. chelonae*/M. σύμπλεγμα αποστήματος) (Church et al., 2020). Οι Simmons et al. συνέκρινε την ταυτοποίηση με συμβατικές μεθόδους μιας διαφορετικής ομάδας βακτηριακών κλινικών απομονώσεων με αλληλουχίες γονιδίων που διερευνήθηκαν από τις βάσεις δεδομένων SmartGene και MicroSeq (Simmon et al., 2006). Από 300 απομονώσεις, το SmartGene εντόπισε 295 (98%) σε επίπεδο γένους και 262 (87%) σε επίπεδο είδους, με 5 (2%) να είναι ασαφείς. Το MicroSeq προσδιόρισε 271 (90%) στο επίπεδο του γένους και 223 (74%) στο επίπεδο του είδους, με 29 (10%) να είναι ασαφείς. Οι SmartGene

και MicroSeq συμφώνησαν για το γένος για 233 (78%) απομονώσεις και το είδος για 212 (71%) απομονώσεις.

Οι συμβατικές μέθοδοι εντόπισαν 291 (97%) απομονωμένα στελέχη σε επίπεδο γένους και 208 (69%) σε επίπεδο είδους, με 9 (3%) να είναι ασαφή. Τα SmartGene, MicroSeq και συμβατικές ταυτοποιήσεις συμφωνήθηκαν για 193 (64%) των αποτελεσμάτων. Η χρήση της 16S PCR/αλληλούχισης για τον εντοπισμό κλινικά σχετικών βακτηρίων που προηγουμένως θα ήταν λανθασμένα ή μη ταυτοποιημένα από κλινικά δείγματα, παρείχε επίσης πληροφορίες για το επιδημιολογικό και παθογόνο δυναμικό σπάνιων ή ασυνήθιστων βακτηρίων σε ανθρώπινες λοιμώξεις.

Ο Woo και οι συνεργάτες του συνόψισαν τα νέα βακτηριακά είδη που ανακαλύφθηκαν από ανθρώπινα δείγματα σε μόλις 7 χρόνια, από το 2001 έως το 2007 (Woo et al., 2008). Αναφέρθηκαν συνολικά 215 νέα είδη, 29 που ανήκουν σε νέα γένη. Επιπλέον, 100 (15 νέα γένη) νέα είδη βρέθηκαν σε 4 ή περισσότερους ασθενείς, και οι μεγαλύτεροι αριθμοί ήταν των γενών *Mycobacterium* και *Nocardia*. Τότε και τώρα, η στοματική κοιλότητα/σχετικά με τα οδοντικά δείγματα και η γαστρεντερική οδός ήταν οι πιο σημαντικές δεξαμενές για την ανακάλυψη νέων ειδών (Woo et al., 2008).

Αυτό συμφωνεί με την τεράστια ποικιλία μικροβίων που εντοπίστηκε σε αυτές τις σημαντικές τοποθεσίες του σώματος από το έργο ανθρώπινου μικροβιώματος (Church et al., 2020). Από την ανακάλυψή τους, οι *Streptococcus sinensis*, *Laribacter hongkongensis*, *Clostridium hathewayi* και *Borrelia spielmanii* έχουν χαρακτηριστεί πληρέστερα, συμπεριλαμβανομένης της επιδημιολογίας και των οδών μετάδοσης τους (Woo et al., 2008). Η μελλοντική τοπική εμπειρία με τον προσδιορισμό αλληλουχίας 16S μπορεί επίσης να βοηθήσει στον προσδιορισμό της περιφερειακής επιδημιολογίας των νέων ευκαιριακών παθογόνων.

Η απόδοση της αλληλουχίας 16S σε μεγάλο αριθμό κλινικά σχετικών παθογόνων την τελευταία δεκαετία στο εργαστήριό, μας αποκάλυψε την επιδημιολογία επεμβατικών λοιμώξεων, όπως η βακτηριαμία, λόγω αρκετών ασυνήθιστων βακτηρίων, συμπεριλαμβανομένων των ειδών *Eggerthella lenta* και *Peptoniphilus* και *Actinomyces*.

3.8 Τρέχοντες περιορισμοί της αλληλούχησης 16S rRNA για την αναγνώριση παθογόνων βακτηρίων

Μία σημαντική κλινική εργαστηριακή κατευθυντήρια γραμμή παρέχει ερμηνευτικά κριτήρια για την ταυτοποίηση ενός ευρέος φάσματος κλινικά σχετικών βακτηρίων και μυκήτων στα επίπεδα του γένους και του είδους χρησιμοποιώντας μερική ή πλήρη αλληλουχία 16S. Συγκεκριμένα, έχουν πραγματοποιηθεί ολοκληρωμένες πολυευθυγραμμίσεις για να αναλύθούν οι σχετικές αλληλουχίες 16S για τα περισσότερα κλινικά σχετικά παθογόνα και στενά συνδεδεμένα περιβαλλοντικά είδη. Αν και περισσότερα δεδομένα αλληλουχίας 16S είναι διαθέσιμα για ανθρώπινα παθογόνα σε δημόσιες/ιδιωτικές βάσεις δεδομένων από άλλους στόχους γονιδίων, πρέπει να αναγνωριστεί ότι λίγες έως καθόλου αλληλουχίες (δηλ. που ορίζονται ως ≤ 5 μεμονωμένες αλληλουχίες/είδη 16S που έχουν κατατεθεί επί του παρόντος στη βάση δεδομένων GenBank [NCBI]) έχουν δημοσιευθεί για μια μεγάλη ποικιλία παθογόνων οργανισμών/ομάδων μικροοργανισμών που περιγράφονται εδώ. Τα στατιστικά στοιχεία σχετικά με την ομολογία του γένους δημιουργήθηκαν χρησιμοποιώντας τις καλύτερες αντιπροσωπευτικές αλληλουχίες για κάθε είδος βέλτιστου μήκους (όπου είναι διαθέσιμο) από την GenBank/NCBI (όπου είναι διαθέσιμο), ομαδοποιώντας τις ανά γένος και στη συνέχεια ευθυγραμμίζοντάς τις ώστε να καλυφθούν τουλάχιστον ~50 έως 1.200 bp του γονιδίου 16S χρησιμοποιώντας MAFFT V7, εξαιρουμένων των αλληλουχιών ειδών που δεν καλύπτουν αυτές τις θέσεις (Petti et al., 2008).

Κάθε στοίχιση αναλύθηκε ανά στήλη/θέση: μια στήλη όπου όλα τα είδη-αλληλουχίες έχουν το ίδιο νουκλεοτίδιο μετρήθηκε ως πανομοιότυπη θέση και μια στήλη όπου τουλάχιστον ένα είδος-αλληλουχία είχε κενό ή ένα διαφορετικό νουκλεοτίδιο μετρήθηκε ως αποκλίνουσα θέση. Η καταμέτρηση ξεκίνησε στην πρώτη κοινή θέση όλων των ακολουθιών και σταμάτησε στην τελευταία κοινή θέση για να αποφευχθεί η καταγραφή της ποικιλομορφίας όπου οι ακολουθίες ήταν μικρότερες. Τα ποσοστά δίνουν μια ιδέα για την ομολογία ενός γένους. Γενικά, τα γένη με λίγα είδη τείνουν να εμφανίζουν υψηλότερο βαθμό ομολογίας.

3.9 Τα *Staphylococcus spp.* και άλλοι αερόβιο gram-θετικοί μικροοργανισμοί

Ο σταφυλόκοκκος είναι ένα πολύ ομοιογενές γένος και η αλληλουχία μεγαλύτερης έκτασης (έως 1.060 bp) του 16S συνιστάται για τη διαφοροποίηση των ειδών με αρκετές αναντιστοιχίες ζευγών

βάσεων για αύξηση της βεβαιότητας (Petti et al., 2008). Ο *S. aureus* και ο *S. lugdunensis* μπορούν, ωστόσο, να διαφοροποιηθούν από τη μεταβλητότητα της αλληλουχίας 16S στα πρώτα 500 bp του 16S. Πολλές κοαγουλάσες-οι αρνητικοί σταφυλόκοκκοι (CoNS) σχετίζονται στενά, και λόγω της γενετικής ομοιότητας μεταξύ του γονιδίου 16S, δεν μπορούν να διαφοροποιηθούν με βεβαιότητα, ακόμη και με μεγαλύτερη αλληλουχία 16S (Petti et al., 2008). Ορισμένα είδη, όπως το *S. capitis/S. caprae* ή *S. agnetis/S. hyicus*, έχουν πανομοιότυπες ακολουθίες 16S, εκτός από ορισμένες προαιρετικές αναντιστοιχίες ζευγών βάσεων στις περιοχές V6 και V7. Άλλοι, όπως ο *S. pasteurii/S. warneri* ή *S. carnosus/S. piscifermentans*, έχουν είτε πανομοιότυπες είτε σχεδόν πανομοιότυπες αλληλουχίες 16S, επομένως αυτό το γονίδιο στόχος δεν μπορεί να χρησιμοποιηθεί για διαφοροποίηση.

Περιορισμένα δεδομένα αλληλουχίας 16S είναι επί του παρόντος διαθέσιμα για αρκετά είδη *Staphylococcus* που απομονώθηκαν από ανθρώπινες (*S. massiliensis*), ζωικές ή/και περιβαλλοντικές πηγές (*S. felis*, *S. fleuretti*, *S. lutrae*, *S. microti*, *S. muscae*, *S. rostri*, *S. simiae* και *S. stepanovicii*). Θα πρέπει επίσης να σημειωθεί ότι το *S. massiliensis* σχετίζεται στενά με τα *S. piscifermentans*, *S. condimenti*, *S. carnosus subsp. carnosus*, *S. carnosus subsp. utilis* και *S. simulans*.

Ο *Micrococcus* και ο *Citricoccus* είναι στενά συνδεδεμένα γένη, αλλά μπορούν να διακριθούν με βάση τη μεταβλητότητα της αλληλουχίας 16S εντός των πρώτων ~500 bp του 16S (Petti et al., 2008).

Οι υπερμεταβλητές περιοχές εντός των V2 και V6 επιτρέπουν τη διαφοροποίηση του *Micrococcus* spp. και τα περισσότερα *Citricoccus* spp. Το *C. muralis* και το *C. nitrophenolicus* δεν μπορούν να διακριθούν με 16S, καθώς οι αλληλουχίες τους διαφέρουν μόνο από ένα μόνο ζεύγος βάσεων αναντιστοιχία στο V2. Περιορισμένα δεδομένα αλληλουχίας 16S είναι επί του παρόντος διαθέσιμα για τον περιβαλλοντικό οργανισμό *Micrococcus lactis* (Prakash et al., 2015), που πρόσφατα μεταφέρθηκε σε ένα νέο γένος, το *Neomicrococcus*, το οποίο περιλαμβάνει επίσης το *N. aestuarii* (παλαιότερα γνωστό ως *Zhihengliuella aestuarii*) (Prakash et al., 2015). Περιορισμένα δεδομένα αλληλουχίας 16S είναι επί του παρόντος διαθέσιμα για τα περισσότερα περιβαλλοντικά *Citricoccus* spp. (δηλαδή, *C. muralis*, *C. nitrophenolicus*, *C. parietis* και *C. zhacaiensis*) εκτός από το *C. alkalitolerans*. «*C. massiliensis*» είναι ένα νέο βακτηριακό είδος που

απομονώθηκε πρόσφατα από το ανθρώπινο δέρμα από την *culturomics* του οποίου η αλληλουχία 16S έχει υψηλό βαθμό ταυτότητας (98,61%) με το *C. nitrophenolicus* (Prakash et al., 2015).

3.10 Τα *Streptococcus spp.* και άλλα όμοια βακτήρια

Η σύγκριση του αριθμού των πανομοιότυπων με τις αποκλίνουσες θέσεις 16S για διάφορα γένη δείχνει ένα ευρύ φάσμα ποσοστιαίας ταυτότητας, με τους *Streptococcus* και *Lactococcus* να είναι τα πιο αποκλίνοντα γένη. Ωστόσο, σε καθένα από αυτά τα κλινικά σημαντικά γένη υπάρχουν αρκετά είδη που δεν μπορούν να αναγνωριστούν αξιόπιστα με βάση την ανάλυση 16S.

Περίπου το 20% των ειδών *Streptococcus* δεν μπορεί να διακριθεί χρησιμοποιώντας το 16S. Οι υπερμεταβλητές περιοχές στις αλληλουχίες του είδους *Streptococcus* 16S εμφανίζονται εντός των V1-3 και V6, έτσι ώστε πολλά είδη μπορούν να διαφοροποιηθούν μόνο με διακριτές αναντιστοιχίες ζευγών βάσεων σε αυτές τις μεταβλητές περιοχές. Επομένως, μια μακρά ακολουθία 16S (δηλαδή, ~1.060 bp) συνιστάται για αξιόπιστη ταυτοποίηση σε επίπεδο είδους (5). Τα κλινικά σημαντικά παθογόνα *S. pneumoniae* και *S. pseudopneumoniae* σχετίζονται πολύ στενά, αλλά μπορούν να διαφοροποιηθούν μεταξύ τους και από το *S. mitis* μέσω μιας ή δύο αναντιστοιχιών στην περιοχή μεταξύ ~600 και 900 bp (Petti et al., 2008)

Η διαφοροποίηση με *groB* διακυβεύεται επίσης από την υψηλή γενετική ομοιότητα μεταξύ αυτών των στενά συνδεδεμένων *Streptococcus spp.* (Church et al., 2020). Οι περισσότερες ομάδες *S. viridans* (*S. mitis*, *S. salivarius*, *S. bovis* και *S. mutans*), με εξαίρεση την ομάδα *S. anginosus*, σχετίζονται στενά και, λόγω ομοιότητας μεταξύ του 16S, δεν μπορούν να διαφοροποιηθούν, ακόμη και με μεγαλύτερη ακολουθία 16S (Petti et al., 2008). Περιορισμένα δεδομένα αλληλουχίας 16S είναι επί του παρόντος διαθέσιμα για το *S. acidominimus*, το *S. devriesei* και το *S. massiliensis*. Οι β-αιμολυτικοί στρεπτόκοκκοι εντός του σχήματος τυποποίησης Lancefield μπορούν να αναγνωριστούν από τη μεταβλητότητα της αλληλουχίας εντός των πρώτων ~500 bp του 16S (Petti et al., 2008). *Aerococcus spp.* μπορεί να διαφοροποιηθεί με τη μεταβλητότητα 16S στις περιοχές V2 και V7, αλλά τα κλινικά παθογόνα *A. viridans* και *A. urinae* έχουν πανομοιότυπες αλληλουχίες 16S. Περιορισμένα δεδομένα αλληλουχίας 16S είναι επίσης διαθέσιμα επί του παρόντος για *A. sanguinicola*, *A. suis* και *A. urinaehominis* (Petti et al., 2008).

3.11 Αερόβιοι gram-θετικοί βάκιλοι

Ο βάκιλος μπορεί γενικά να διαφοροποιηθεί από τη μεταβλητότητα στο 16S εντός των περιοχών V1-3 και V6 (Petti et al., 2008). Πολλά *Bacillus spp.* σχετίζονται σε μεγάλο βαθμό και μεγάλες εκτάσεις 16S (~1.060 bp) θα πρέπει να αναλυθούν για οριστική ταυτοποίηση. *B. anthracis*, *B. cereus*, *B. wiedmannii* και *B. thuringensis* έχουν σχεδόν πανομοιότυπες αλληλουχίες 16S και δεν μπορούν να διαφοροποιηθούν χρησιμοποιώντας το 16S. Ο *Geobacillus* σπάνια απομονώνεται από κλινικά δείγματα και είναι ένα πολύ γενετικά πανομοιότυπο γένος, με πολλά στενά συγγενικά είδη. Η διαφοροποίηση σε επίπεδο είδους απαιτεί αναλύσεις μιας μεγαλύτερης ακολουθίας 16S που εκτείνεται σε μεταβλητές περιοχές V1-3, V6, και V7 αλλά περιορίζεται σε πολύ λίγες αναντιστοιχίες σε πολλές μεταβλητές περιοχές (Petti et al., 2008). Περιορισμένα δεδομένα αλληλουχίας 16S είναι επί του παρόντος διαθέσιμα για πολλά περιβαλλοντικά είδη, συμπεριλαμβανομένων των *G. galactosidasium*, *G. jurassicus*, *G. thermantarcticus* και *G. vulcani*.

Αρκετά *Raenibacillus spp.* έχουν αναφερθεί ότι προκαλούν μόλυνση στον άνθρωπο, αν και άλλοι αναγνωρίζονται ως συνήθεις μολυντές των κλινικών δειγμάτων (Church et al., 2020). *Raenibacillus spp.* μπορεί να διαφοροποιηθεί από τη μεταβλητότητα 16S εντός των περιοχών V1, V2, V3 και V6, αλλά τα στενά συγγενικά είδη μπορεί να απαιτούν ανάλυση πλήρους μήκους αλληλουχιών 16S (Petti et al., 2008). Περιορισμένα δεδομένα αλληλουχίας 16S είναι επί του παρόντος διαθέσιμα για πολλά περιβαλλοντικά είδη, συμπεριλαμβανομένου του *P. Brasilensis* (Petti et al., 2008).

3.12 *Listeria spp.*

Η διαφοροποίηση θα πρέπει να πραγματοποιείται με αναλύσεις της μεταβλητότητας 16S εντός των περιοχών V1, V2, V6 και V8. Τα πολύ συγγενικά είδη, όπως το *L. monocytogenes* και το *L. innocua*, διαφέρουν μόνο κατά μία ξεχωριστή αναντιστοιχία ζεύγους βάσεων εντός των περιοχών V2 και V8 (Petti et al., 2008).

Περιορισμένα δεδομένα αλληλουχίας 16S είναι επί του παρόντος διαθέσιμα για πολλά περιβαλλοντικά (γεωργικά και φυσικά περιβάλλοντα) είδη, συμπεριλαμβανομένων των *L.*

aquatic, *L. booriae*, *L. cornelliensis*, *L. fleischmanii*, *L. floridensis*, *L. grandensis*, *L. marthii*, *L. newyorkensis*, *L. riparia*, *L. rocourtiae* και *L. weihenstephanensis* (Church et al., 2020).

3.13 Εντεροβακτήρια (*Enterobacteriaceae*)

Τα *Enterobacteriaceae* είναι μια μεγάλη, πολύπλοκη οικογένεια που περιέχει σήμερα περισσότερα από 30 γένη και πάνω από 100 είδη. Αν και η συστηματική ταξινόμηση των *Enterobacteriaceae* εξακολουθεί να συζητείται, μια νέα ταξινομική ταξινόμηση έχει προταθεί πρόσφατα για αυτή τη μεγάλη, πολύπλοκη ομάδα οργανισμών, η οποία περιέχει πολλά σημαντικά Gram-αρνητικά εντερικά παθογόνα. Οι Alnajjar et al. έχουν πρόσφατα προτείνει την τοποθέτηση των *Enterobacteriaceae* στην τάξη *Enterobacteriales*, στην κατηγορία *Gamma*proteobacteria, με βάση τη φυλογενετική ανάλυση των πολλών διαφορετικών ειδών (Adeolu et al., 2016). Το έργο τους υποστηρίζει την ύπαρξη επτά διακριτών μονοφυλετικών κλάδων γενών εντός της τάξης, καθιστώντας την ταξινομική σημασία να διαιρεθεί η πρώτη οικογένεια *Enterobacteriaceae* σε επτά οικογένειες, συμπεριλαμβανομένων των οικογενειών *Enterobacteriaceae*, *Erwiniaceae*, *Pectobacteriaceae*, *Yersiniaceae*, *Hafniaceae*, *Morganellaceae* και *Budviciaceae*.

Επιπλέον, αυτό το σύστημα ταξινόμησης θα διαχωρίσει και θα κατανείμει πολλά κλινικά σχετικά παθογόνα μεταξύ αυτών των οικογενειών. Για παράδειγμα, το *Enterobacter*-*Escherichia* clade είναι η μεγαλύτερη ομάδα της τάξης των *Enterobacterale* και αποτελείται από τα γένη «*Atlantibacter*», *Buttiauxella*, *Cedecea*, *Citrobacter*, *Cronobacter*, *Enterobacter*, *Escherichia*, *Franconibacter*, *Klebsiella*, *Kluuyvera*, *Kosakonia*, *Leclercia*, *Lelliottia*, *Mangrovibacter*, *Pluralibacter*, *Raoultella*, *Salmonella*, *Shigella*, *Shimwellia*, *Siccibacter*, *Trabulsiella* και *Yokenella*.

Ο κλάδος *Erwinia*-*Pantoea*, ο οποίος υπάρχει σε μονοφυλετική ομάδα με το κλάδο *Enterobacter*-*Escherichia*, αποτελείται από τα γένη *Erwinia*, *Pantoea*, *Phaseolibacter*, και *Tatumella*. Ο κλάδος *Pectobacterium*-*Dickeya* αποτελείται από τα γένη *Brenneria*, *Dickeya*, *Lonsdalea*, *Pectobacterium* και *Sodalis*. Ο κλάδος *Yersinia*-*Serratia* αποτελείται από τα γένη *Chania*, *Ewingella*, *Rahnella*, *Rouxiella*, *Serratia* και *Yersinia*, ενώ ο *Hafnia*-*Edwardsiella* αποτελείται από τα γένη *Edwardsiella*, *Hafnia* και *Obesumbacterium*.

Ο *Proteus-Zenorhabdus* αποτελείται από τα γένη *Arseorhonus*, *Moellerella*, *Morganella*, *Photorhabdus*, *Proteus*, *Providencia* και *Xenorhabdus* και ο κλάδος *Budvicia* αποτελείται από τα γένη *Budvicia*, *Leminorella* και *Pragia* (Petti et al., 2008).

3.14 Καμπυλοβακτήρια (*Campylobacteriales*)

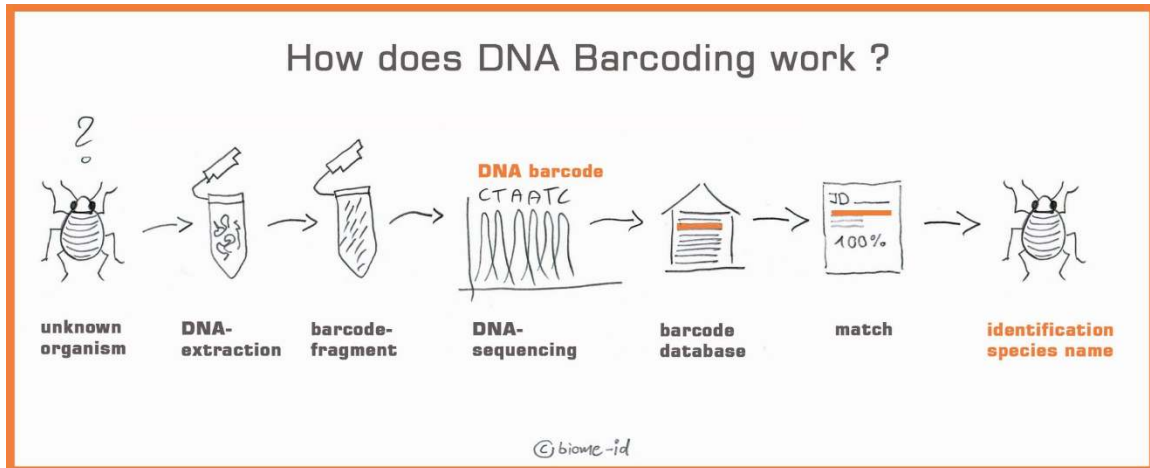
Η σύγκριση του αριθμού των πανομοιότυπων με τις αποκλίνουσες θέσεις 16S για διάφορα γένη δείχνει ένα ευρύ φάσμα ποσοστιαίας ταυτότητας. Τα περισσότερα δεδομένα αλληλουχίας 16S είναι διαθέσιμα για το *Helicobacter*, και αυτό, όπως και το *Campylobacter*, είναι πολύ αποκλίνοντα γένη. Ωστόσο, σε καθένα από αυτά τα κλινικά σημαντικά γένη υπάρχουν πολλά είδη που δεν μπορούν να αναγνωριστούν αξιόπιστα με βάση την ανάλυση 16S. Ορισμένα είδη *Campylobacter* είναι σημαντικά παθογόνα των ζώων και των ανθρώπων. Το *C. coli* και το *C. jejuni* είναι σχεδόν πανομοιότυπα στο 16S, με πολύ λίγες, ίσως ασυνεπείς, αναντιστοιχίες ζευγών βάσεων εντός της περιοχής V5, ενώ το *C. jejuni* εμφανίζει κάποια διακριτή ενδοειδική μεταβλητότητα στις περιοχές V1 και V4. Μεγαλύτερες εισαγωγές (περίπου 30 έως 40 bp) παρατηρήθηκαν για το *C. curvus*, το *C. sputorum* και το *C. rectus* πέρα από την περιοχή V2 (Petti et al., 2008). Περιορισμένα δεδομένα αλληλουχίας 16S είναι επί του παρόντος διαθέσιμα για αρκετά γεωργικά και περιβαλλοντικά είδη, συμπεριλαμβανομένων των *C. anium*, *C. insulaenigrae*, *C. peloridis*, και *C. volucris*. Το ελικοβακτηρίδιο του πυλωρού προκαλεί πεπτικό έλκος. Το *H. acinonychis* και το ελικοβακτηρίδιο του πυλωρού είναι στενά συνδεδεμένα, αλλά μπορούν να διαφοροποιηθούν με την αλληλούχιση μιας μεγαλύτερης έκτασης 16S συμπεριλαμβανομένων των περιοχών V2 και V4-V6.

Ορισμένα είδη, όπως τα *H. bilis*, *H. canis*, *H. fennelliae*, *H. macacae*, *H. marmotae*, *H. mastomyrinus* και *H. tyhlonius*, δείχνουν μια σημαντική εισαγωγή μερικές φορές >150 bp μετά την περιοχή V2 (5). Περιορισμένα δεδομένα αλληλουχίας 16S είναι επί του παρόντος διαθέσιμα για πολλά ζωικά είδη, συμπεριλαμβανομένων των *H. acinonychis*, *H. anseris*, *H. aurati*, *H. baculiformis*, *H. branta*, *H. cholecystus*, *H. cynogastricus*, *H. equorum*, *H. marmotae*, *H. mastomyrinus*, *H. mesocricetorum*, *H. muridarum*, *H. pametensis*, *H. rodentium*, *H. salmononis* και *H. tyhlonius* (Church et al., 2020).

3.15 DNA barcoding

Η απουσία ενός παγκοσμίως αποδεκτού γραμμικού κώδικα DNA για τους Μύκητες, το δεύτερο πιο συγκεκριμένο ευκαρυωτικό βασίλειο (Blackwell, 2011; Mora et al., 2011), αποτελεί σοβαρό περιορισμό για πολυταξονικές μελέτες οικολογίας και βιοποικιλότητας. Η γραμμική κωδικοποίηση DNA χρησιμοποιεί τυποποιημένες αλληλουχίες 500 έως 800 bp για την αναγνώριση ειδών όλων των ευκαρυωτικών βασιλείων χρησιμοποιώντας εκκινητές που ισχύουν για την ευρύτερη δυνατή ταξινόμική ομάδα. Οι γραμμωτοί κώδικες αναφοράς πρέπει να προέρχονται από εξειδικευμένες πηγές που έχουν κατατεθεί σε βιολογικές συλλογές με διαδικτυακά μεταδεδομένα και επικυρωμένα από διαθέσιμα ηλεκτρονικά χρωματογραφήματα αλληλουχίας. Η διαειδική παραλλαγή θα πρέπει να υπερβαίνει την ενδοειδική παραλλαγή (το χάσμα του γραμμικού κώδικα) και η γραμμική κωδικοποίηση να είναι βέλτιστη όταν μια ακολουθία είναι σταθερή και μοναδική για ένα είδος (Sulakvelidze, Z Alavidze, 2001),(Schoch et al., 2012). Στην ιδανική περίπτωση, η θέση του γραμμικού κώδικα θα ήταν η ίδια για όλους τους οργανισμούς. Μια περιοχή του μιτοχονδριακού γονιδίου που κωδικοποιεί την υπομονάδα 1 της οξειδάσης του κυτοχρώματος c (CO1) είναι ο γραμμωτός κώδικας για τα ζώα (Sulakvelidze, Z Alavidze, 2001),(Schoch et al., 2012) και ο προεπιλεγμένος δείκτης που υιοθετήθηκε από την Κοινοπραξία για το Barcode of Life για όλες τις ομάδες οργανισμών, συμπεριλαμβανομένων των μυκήτων (Schindel & Miller, 2005). Το CO1 λειτουργεί αρκετά καλά ως γραμμικός κώδικας σε ορισμένα γένη μυκήτων, όπως το *Penicillium*, με αξιόπιστους εκκινητές και επαρκή ανάλυση ειδών (67%) (Seifert, 2007). Ωστόσο, τα αποτελέσματα στις λίγες άλλες ομάδες που εξετάστηκαν πειραματικά είναι ασυνεπή και συχνά απαιτείται κλωνοποίηση. Το πυρηνικό rRNA σιστρόνιο έχει χρησιμοποιηθεί για μυκητιακά διαγνωστικά και φυλογενετικά για περισσότερα από 20 χρόνια (Begerow et al., 2010) και τα συστατικά του συζητούνται συχνότερα ως εναλλακτικές του CO1 (Rossman, 2007), Eberhardt 2010). Το ευκαρυωτικό σιστρόνιο rRNA αποτελείται από τα γονίδια rRNA 18S, 5.8S και 28S που μεταγράφονται ως μονάδα από την RNA πολυμεράση I. Οι μεταμεταγραφικές διεργασίες διασπούν το σιστρόνιο, αφαιρώντας δύο εσωτερικά μεταγραφόμενα διαχωριστικά. Αυτοί οι δύο διαχωριστές, συμπεριλαμβανομένου του γονιδίου 5.8S, αναφέρονται συνήθως ως περιοχή ITS. Το γονίδιο rRNA πυρηνικής ριβοσωματικής μικρής

υπομονάδας 18S (SSU) χρησιμοποιείται συνήθως στη φυλογενετική, και παρόλο που το ομόλογό του (16S) χρησιμοποιείται συχνά ως διαγνωστικό είδος για βακτήρια (Loessner, 2005), έχει λιγότερες υπερμεταβλητές περιοχές στους μύκητες. Το 28S πυρηνικό ριβοσωμικό γονίδιο rRNA μεγάλης υπομονάδας (LSU) μερικές φορές διακρίνει είδη από μόνο του ή σε συνδυασμό με το ITS.

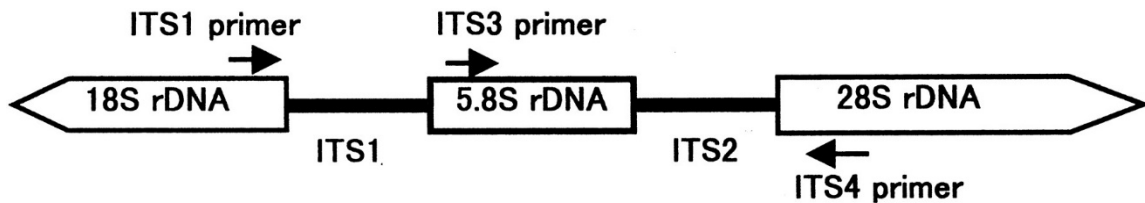


Εικόνα 8: DNA Barcoding

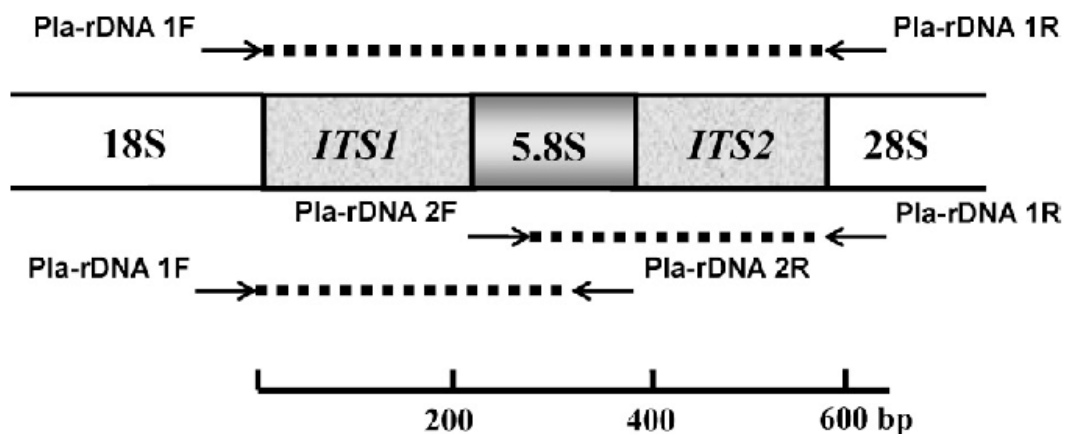
3.16 ITS «Internal Transcribed Spacer»

Οι πρόοδοι της μοριακής τεχνολογίας δείχνουν μεγάλες δυνατότητες για την ταχεία ανίχνευση και ταυτοποίηση μυκήτων για ιατρικούς, επιστημονικούς και εμπορικούς σκοπούς. Έχουν αξιολογηθεί πολυάριθμοι στόχοι εντός του μυκητιακού γονιδιώματος, Αν και δεν μεταφράζονται σε πρωτεΐνες, οι περιοχές κωδικοποίησης ITS διαδραματίζουν κρίσιμο ρόλο στην ανάπτυξη λειτουργικού rRNA, με παραλλαγές αλληλουχίας μεταξύ των ειδών που δείχνουν υποσχόμενες περιοχές για μοριακούς προσδιορισμούς. Το τμήμα του γονιδιώματος περιλαμβάνει τα γονίδια 18S, 5.8S και 28S που κωδικοποιούν το ριβοσωμικό RNA (rRNA) και τα οποία έχουν μια σχετικά διατηρημένη νουκλεοτιδική αλληλουχία μεταξύ των μυκήτων (Iwen et al., 2002). Ο εσωτερικός μεταγραφόμενος διαχωριστής (ITS) του πυρηνικού ριβοσωμικού DNA (nrDNA) έχει χρησιμοποιηθεί ευρέως τόσο σε μοριακές, συστηματικές όσο και σε οικολογικές μελέτες μυκήτων με αποτέλεσμα η περιοχή του εσωτερικού μεταγραφόμενου διαχωριστή (ITS) rDNA να επιλέγεται ως ο πιο χρησιμοποιούμενος γραμμωτός κώδικας για τη μελέτη της ποικιλότητας των μυκήτων (Yang et al., 2018), (Mbareche et al., 2020). Η περιοχή ITS περιέχει τρία διαμερίσματα: ITS1, 5.8S και ITS2. Η περιοχή 5.8S δεν περιέχει επαρκή αριθμό πληροφοριών που μπορούν να

χρησιμοποιηθούν για φυλογενετικές μελέτες και barcoding DNA (Khot & Fredrick, 2009). Το μήκος της αλληλουχίας ITS είναι πολύ μεταβλητό από το ένα είδος μύκητα στο άλλο και εξαρτάται σε μεγάλο βαθμό από τους εκκινητές που χρησιμοποιούνται για τη στόχευση της αλληλουχίας (Mbareche et al., 2020). Λόγω της πολυαντιγραφικής φύσης του, το ITS επιτρέπει την εύκολη ενίσχυση από δείγματα που περιέχουν χαμηλές συγκεντρώσεις DNA. Επιπλέον, χιλιάδες αλληλουχίες ITS διαφορετικών ειδών είναι άμεσα διαθέσιμες από διάφορες διαδικτυακές βάσεις δεδομένων, συμπεριλαμβανομένης της UNITE και της Διεθνούς Βάσης δεδομένων αλληλουχιών νουκλεοτιδίων (GenBank, EMBL και DDBJ), που παρέχει μια μεγάλη συλλογή αναφοράς για ταξινομική ταξινόμηση. Οι πιο συχνά χρησιμοποιούμενες αλληλουχίες υψηλής απόδοσης έχουν μέγιστο περιορισμό μήκους ανάγνωσης. Αυτός ο περιορισμός αναγκάζει τη χρήση μόνο μιας από τις δύο υπομονάδες του ITS (ITS1 ή ITS2) κατά την εφαρμογή της προσέγγισης HTS με βάση το αμπλικόνιο για τον προσδιορισμό της μυκητιακής ποικιλομορφίας (Yang et al., 2018).



Εικόνα 9: PCR χρησιμοποιώντας Internal Transcribed Spacer 1 και 2 Regions για ταχεία ανίχνευση



Εικόνα 10: Γενική δομή ενός πυρηνικού ριβοσωμικού γονιδίου RNA ITS1 και ITS2

3.17 Σύμπλεγμα rDNA

Οι περιοχές ITS βρίσκονται στο σύμπλεγμα γονιδίων rDNA όλων των ευκαρυωτικών κυττάρων μεταξύ των γονιδίων rDNA 18S και 5.8S (ITS1) και των γονιδίων rDNA 5.8S και 28S (ITS2). Το συνολικό σύμπλεγμα γονιδίων rDNA περιλαμβάνει τη διαγονιδιακή περιοχή διαχωρισμού (IGS), η οποία περιέχει την περιοχή εξωτερικού μεταγραφόμενου διαχωριστή ETS1 (ονομάζεται επίσης μη μεταγραφόμενος διαχωριστής) στο 50 άκρο και την περιοχή ETS2 στο άκρο 30, μαζί με μια μεταβλητή τοποθέτηση του γονιδίου 5S rDNA, του γονιδίου 18S rDNA, της περιοχής ITS1, του γονιδίου 5.8S rDNA, της περιοχής ITS2 και του γονιδίου 28S rDNA. Κατά τη διάρκεια της πρώιμης μεταγραφής στον πυρήνα του κυττάρου, ο αρχικός πρόδρομος στο σχηματισμό του RNA περιέχει νουκλεοτίδια για την υπομονάδα 5'-ETS1/18S/ITS1/5.8S/ITS2/28S/ETS2-3', που αναφέρεται ως η μονάδα μεταγραφής rDNA 35S σε 45S. Αυτή η γονιδιωματική μονάδα, η οποία περιέχει τα γονίδια που είναι απαραίτητα για το σχηματισμό του ριβοσωμικού RNA (rRNA), έχει χαρακτηριστεί ως σύμπλεγμα γονιδίων rDNA του μυκητιακού γονιδιώματος. Καθώς το γονίδιο 5S rDNA υφίσταται επεξεργασία κατά τη διάρκεια της μεταγραφής σε ένα βήμα διαφορετικό από τα άλλα προϊόντα γονιδίου rDNA, δεν περιλαμβάνεται ως μέρος της μεταγραφής του πρόδρομου RNA. Επομένως, το γονίδιο 5S θα αποκλειστεί όσον αφορά το σύμπλεγμα rDNA, αν και είναι ζωτικής σημασίας για το σχηματισμό ριβοσωμάτων. Παραδείγματος χάριν στον ζυμομύκητα *Saccharomyces cerevisiae*, το πρώτο ευκαρυωτικό κύτταρο στο οποίο έγινε η αλληλουχία του πλήρους γονιδιώματός του, το σύμπλεγμα γονιδίων rDNA έχει μέγεθος περίπου 9 kb και βρίσκεται στο χρωμόσωμα XII. Αυτό το σύμπλεγμα γονιδίων, το οποίο επαναλαμβάνεται περίπου 140 φορές σε αυτό το χρωμόσωμα, έχει συνολικό αριθμό επαναλήψεων βάσεων περίπου 1260 kb. Καθώς το μέγεθος του γονιδιώματος του *S. cerevisiae* είναι περίπου 13,4 μεγαβάσεις, το σύμπλεγμα γονιδίου rDNA επομένως περιλαμβάνει περίπου το 10% του συνολικού γονιδιώματος. Αυτό το σύμπλεγμα πολλαπλών αντιγράφων γονιδίων, που περιέχει τόσο διατηρημένες όσο και μεταβλητές περιοχές αλληλουχίας νουκλεοτιδίων, χρησιμεύει ως σημείο αναφοράς για μελέτες εξελικτικής απόκλισης λόγω της καθολικής παρουσίας του σε όλα τα ευκαρυωτικά κύτταρα (Iwen et al., 2002).

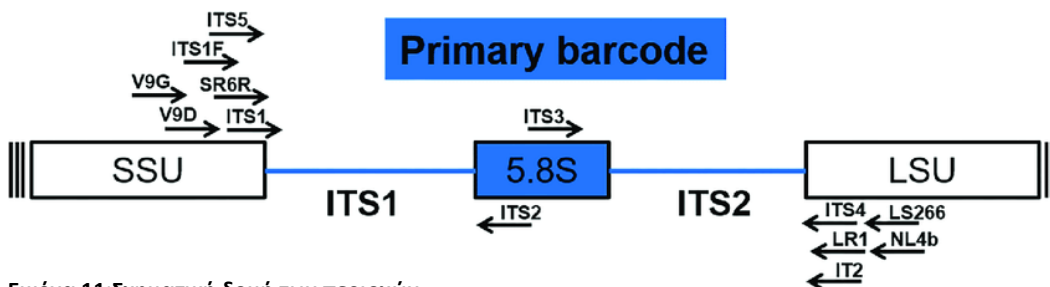
3.18 Βιολογικός ρόλος των περιοχών ITS

Ο βιολογικός ρόλος των περιοχών ITS κατά την επεξεργασία των rRNA παραμένει ένα ενδιαφέρον παζλ, παρόλο που έχει προταθεί ένας αριθμός λειτουργιών. Κατά την έναρξη της μεταγραφής, το πρωτεύον 35S σχηματίζει μονάδα μεταγραφής 45S rRNA. Αυτή η πρόδρομη μεταγραφή υφίσταται ενζυμική απομάκρυνση από τις τέσσερις αλληλουχίες διαχωρισμού για να γίνει ένα πρωτεύον μεταγράφημα rRNA που τροποποιείται μέσω μεθυλίωσης RNA και άλλων μετατροπών βάσεων και συναρμολογείται σε ένα σωματίδιο πυρηνικής ριβονουκλεοπρωτεΐνης 80S (RNP). Το σωματίδιο 80S RNP υφίσταται διάσπαση για να σχηματίσει τα ώριμα rRNA 18S, 5.8S και 28S που τελικά συνδυάζονται με πρωτεΐνες για να γίνουν λειτουργικά ριβοσώματα. Αυτά τα λειτουργικά ριβοσώματα συνδέονται με mRNA και tRNA οδηγώντας σε παραγωγή πρωτεΐνης κατά τη μετάφραση. Μελέτες έχουν δείξει ότι η αφαίρεση μιας ή και των δύο περιοχών ITS πριν από την έναρξη της μεταγραφής έχει επιβλαβή επίδραση στον σχηματισμό ώριμων rRNAs. Κατά τον Musters ανέφεροταν ότι η αφαίρεση του ITS1 εμπόδιζε τον σχηματισμό ώριμου 18S rRNA και συνεπώς την παραγωγή πρωτεΐνης κατά τη μετάφραση. Επιπλέον, οι Van der Sande έδειξαν ότι η αφαίρεση της αλληλουχίας ITS2 εμπόδιζε την ωρίμανση των 5.8S και 28S rRNAs. Περαιτέρω εργασία των Good et al. επιβεβαίωσε τη σημασία της ακολουθίας ITS2 στον σχηματισμό λειτουργικών ριβοσωμάτων. Πρότειναν ότι κατά την ωρίμανση του rRNA, φάνηκε να υπάρχει μια αλληλεξάρτηση στον ρόλο της περιοχής ITS ως μηχανισμού ποιοτικού ελέγχου για να διασφαλιστεί ότι μόνο λειτουργικό rRNA ενσωματώθηκε στο ριβόσωμα. Καθώς το DNA των περιοχών ITS αφαιρείται και δεν αποτελεί μέρος του μορίου του ώριμου RNA, θεωρούνται μη κωδικοποιητικές περιοχές του γονιδιώματος. Ωστόσο, δεν αναγνωρίζονται συνήθως ως εσώνια,

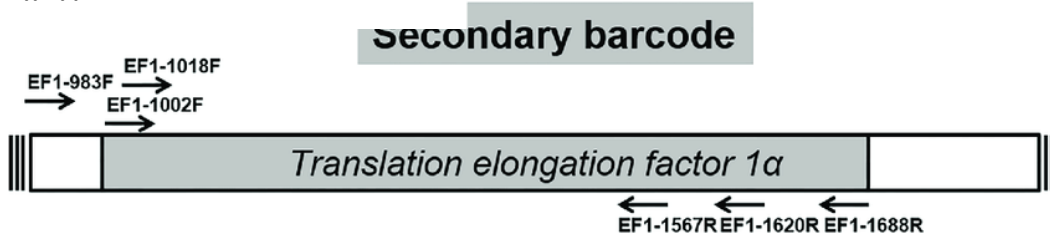
αλλά ως παρεμβαλλόμενες μη κωδικοποιητικές ακολουθίες ή ψευδοιντρόνια λόγω τους σημασία ως πρόδρομες ουσίες στην παραγωγή πρωτεϊνών (Iwen et al., 2002).

3.19 Σύγκριση ITS1 και ITS2

Υπάρχουν κάποιες διαφωνίες στην επιλογή των δεικτών για την αλληλουχία. Οι συγκρίσεις μεταξύ των ευκαρυωτικών γονιδιωματικών ριβοσωμικών περιοχών, ITS1 και ITS2 έχουν αξιολογηθεί σε διάφορες μελέτες, όπως για παράδειγμα για την διαφοροποίηση των προφίλ των μυκήτων. Το αποτέλεσμα ότι το ITS1 ήταν πιο μεταβλητό από το ITS2 ήταν σχεδόν συνεπές. Μέσα στην περιοχή του, το ITS1 εξελίσσεται πιο γρήγορα και έχει πιο μεταβλητό μήκος από το ITS2. Σε ορισμένες αναφορές βρέθηκε ότι τα ITS1 και ITS2 αποδίδουν παρόμοια αποτελέσματα ομαδοποίησης και ταξινόμησης. Ωστόσο, η ταξινομική ανάλυση δεν ήταν ίση σε διαφορετικά ταξινομικά επίπεδα όσον αφορά την ταυτοποίηση ταξινομικών κατηγοριών με τα ITS1 και ITS2. Τα αποτελέσματα που λαμβάνονται είναι ξεκάθαρα ως προς την υπεραπόδοση του ITS1 έναντι του ITS2 όσον αφορά τον πλούτο και την ταξινομική κάλυψη. Η ανάλυση διαφορικής αφθονίας έδειξε ότι ορισμένες συστηματικές ταξινομήσεις διαφορετικών οργανισμών ανιχνεύθηκαν αποκλειστικά μόνο από το ITS2 και αντίστροφα για το ITS1. Ωστόσο, η μεταγονιδιωματική προσέγγιση έδειξε ένα ταξινομικό προφίλ που μοιάζει περισσότερο με το ITS1 παρά το ITS2 (Mbareche et al., 2020). Έχει διαπιστωθεί επίσης (Bazzicalupo et al.) ότι το ITS2 φαίνεται καταλληλότερο για την ανάδειξη του πλούτου των λειτουργικών ταξινομικών μονάδων (OTU:



Εικόνα 11: Σχηματική δομή των περιοχών γραμμικού κώδικα DNA, ITS και TEF1α

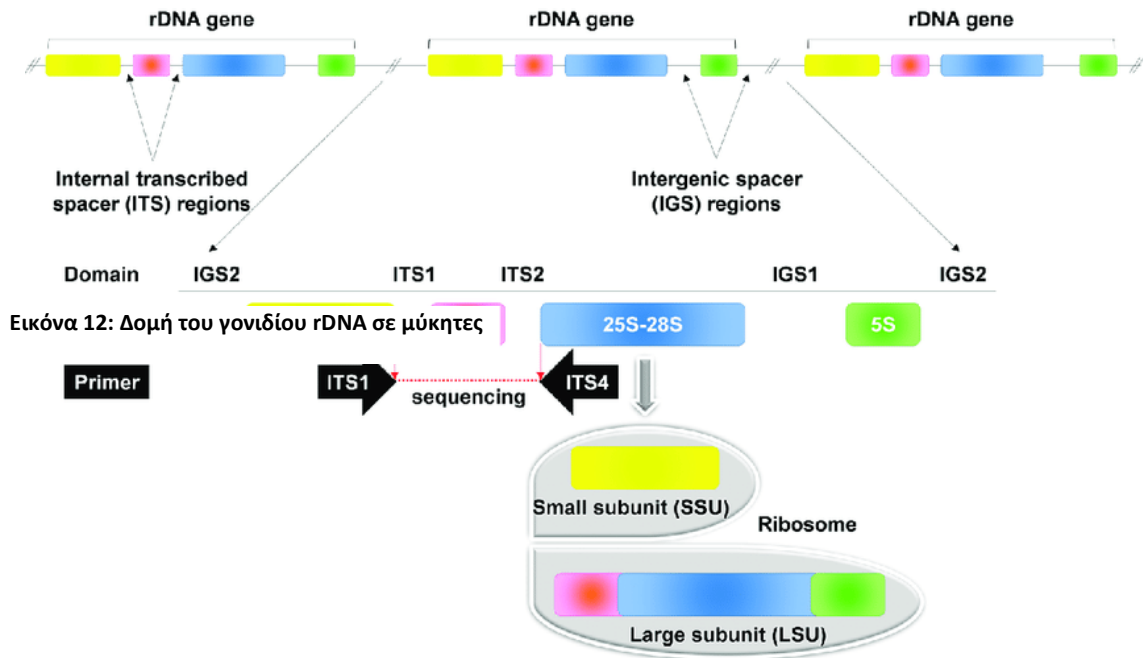


operational taxonomic units) σε κοινότητες μυκήτων. Ωστόσο, έχει γίνει εξίσου αναφορά (Mello

et al. και Wang et al.) ότι το ITS1 ήταν ίσως η καλύτερη επιλογή για τη μελέτη μυκήτων ή ευκαρυωτικών ειδών. Με βάση τις μελέτες, κανένας από τους γραμμωτούς κώδικες (Barcodes) που αξιολογήθηκαν δεν είναι τέλειος όσον αφορά τη διάκριση όλων των ειδών. Η χρήση και των δύο γραμμωτών κωδικών προσφέρει μια ευρύτερη εικόνα του πληθυσμού των μυκήτων. Ωστόσο συνιστάται ανεπιφύλακτα η χρήση του ITS1 ως καθολικού γραμμικού κώδικα μυκήτων για γρήγορες γενικές αναλύσεις της διαφορετικότητας των οικογενειών των μυκήτων (Mbareche et al., 2020). Σε προηγούμενες μελέτες, οι συγκρίσεις των ITS1 και ITS2 διεξήχθησαν κυρίως με βάση την αλληλουχία ITS1 και ITS2 ξεχωριστά ή την εξαγωγή των ITS1 και ITS2 από τη βάση δεδομένων ITS. Δεν προέρχονταν όλα τα ITS1 και ITS2 από το ίδιο ITS. Ωστόσο στη μελέτη που διεξήχθη από τον (Yang et al., 2018) τα ITS1 και ITS2, ανεξάρτητα από τα σύνολα δεδομένων *insilico* ή *pyrosequencing*, εξήχθησαν όλα από το ίδιο ITS, το οποίο μπορεί να είναι πιο κατάλληλο για την αξιολόγηση διαφορετικών τμημάτων του ITS και τη διερεύνηση μυκητιακών κοινοτήτων για την αποσαφήνιση των αποκλίσεων μεταξύ τους. Επιπρόσθετα, οι περισσότερες από τις μελέτες έχουν ως επίκεντρο την επίδραση της ποικιλομορφίας άλφα και βήτα. Η μελέτη τους προσπάθησε να εξετάσει την ομοιότητα στη ομαδοποίηση χρησιμοποιώντας διαφορετικά τμήματα των ακολουθιών ITS από τις υπάρχουσες βάσεις δεδομένων *insilico* ή *pyrosequencing*. Τα αποτελέσματα της έρευνας έδειξαν ότι κατά την *insilico* ανάλυση βρέθηκαν συνολικά 83.120 πλήρους μήκους αλληλουχίες ITS από τους μύκητες σύμφωνα με την βάση δεδομένων UNITE. Οι αλληλουχίες διαχωρίστηκαν σε πέντε φυλές. Οι βάσεις δεδομένων για αυτές τις αλληλουχίες αναφέρονται ως *insilicoITS*, *insilicoITS1* και *insilicoITS2*. Όσον αφορά την διακύμανση μήκους των ITS1 και ITS2 σε διαφορετικές ομάδες, το ITS2 ήταν μεγαλύτερο από το ITS1 και στις 5 φυλές. Το μήκος των εξαγόμενων τμημάτων ITS1 κυμαινόταν από 9 bp έως 1181 bp και το μήκος των εξαγόμενων τμημάτων ITS2 κυμαινόταν από 14 bp έως 730 bp μεταξύ των μυκήτων. Το μήκος του ITS1 είχε ευρύτερο εύρος διακύμανσης. Κατά την ταξινομική ανάλυση των ITS, ITS1 και ITS2, πραγματοποιήθηκαν επίσης συγκρίσεις μεταξύ των διαφορετικών τμημάτων του ITS με βάση την ανάλυση κατά την οποία τα δεδομένα που συλλέχθηκαν τοποθετήθηκαν σε συστηματικές ταξινομήσεις διαφορετικών οργανισμών. Όταν συγκρίθηκε η βάση δεδομένων BLAST (Basic Local Alignment Search Tool) με την UNITE, το 53,27% των ακολουθιών ITS επέστρεψαν οι ίδιοι, το οποίο ήταν πολύ υψηλότερο ποσοστό από ό,τι για το ITS1 (35,55%) και το ITS2 (35,73%). Τα

αποτελέσματα αποκάλυψαν ότι η ομαδοποίηση και η ταξινόμηση για το ITS2 ήταν πιο παρόμοια με εκείνα για το ITS παρά με εκείνα για το ITS1. Το μικρότερο μήκος, η χαμηλότερη διακύμανση περιεχομένου GC και το μεγαλύτερο ταξινομικό περιεχόμενο πληροφοριών του ITS2 μπορεί να το καταστήσουν πιο κατάλληλο από το ITS1 για μελέτες σε βάθος αλληλουχίας σε κοινότητες μυκήτων. Ορισμένες αναφορές έχουν δείξει ότι το ITS1 και το ITS2 δημιούργησαν παρόμοια μοτίβα δομής κοινότητας όταν χρησιμοποιούνται ως μεταγραμμικοί κώδικες DNA. Ωστόσο, τα κοινά στοιχεία μεταξύ ITS1 και το ITS2 είναι πολύ λίγα. Τα προηγούμενα χρόνια, δεν υπήρξε ομοφωνία σχετικά με την επιλογή του ITS1 ή του ITS2 σε μελέτες για την ποικιλότητα των μυκήτων. Σύμφωνα με μελέτες των Nilsson, Ihrmark και Alanagre το ITS2 ήταν η καλύτερη επιλογή για pyrosequencing ή αλληλουχία με πλατφόρμες Illumina. Ωστόσο, αρκετές είναι και οι μελέτες που αποδεικνύουν το αντίθετο όπως του Wang δείχνοντας ότι το ITS1 μπορεί να είναι καλύτερος ταξινομικός γραμμωτός κώδικας DNA από το ITS2 στους ευκαρυώτες, με βάση τις αναλύσεις *silico*. Αυτή η ασάφεια και η διαφορετικότητα των αποτελεσμάτων μπορεί να εξηγηθεί με δύο πιθανούς τρόπους. Αρχικά, οι αλληλουχίες που χρησιμοποιήθηκαν στη μελέτη των Wang et al. ανήκαν σε όλους τους ευκαρίους, συμπεριλαμβανομένων των μυκήτων, των φυτών και των ζώων, σε αντίθεση με την μελέτη των (Yang et al., 2018), όπου μελετήθηκαν λεπτομερώς μόνο οι αλληλουχίες μυκήτων. Εν συνεχεία, οι Wang επικεντρώθηκαν κυρίως στην ταυτοποίηση ειδών σχετικά με τους σχολιασμούς των ακολουθιών που είναι αποθηκευμένες στη βάση δεδομένων του NCBI, ενώ στην μελέτη των (Yang et al., 2018) επικυρώθηκαν οι ταυτοποιήσεις ειδών, αλλά ταυτόχρονα ελήφθη υπόψη και η ομαδοποίηση OTU με βάση τις αλληλουχίες ITS, ITS1 και ITS2.

Αν και το περιεχόμενο GC του ITS2 ήταν ελαφρώς υψηλότερο από αυτό του ITS1, λιγότερες αλληλουχίες φιλτραρίστηκαν με αυτό το κριτήριο (<20% και >80%). Το ITS2 μπορεί να έχει θετική επίδραση στην αποτελεσματικότητα της PCR και της αλληλουχίας (Yang et al., 2018). Το ITS2 μπορεί να είναι ο καταλληλότερος δείκτης για την ποικιλομορφία των μυκήτων λόγω του μικρότερου μήκους, της χαμηλότερης παραλλαγής GC, της μεγαλύτερης αφθονίας αναφορών σε δημόσιες βάσεις δεδομένων, της ευρύτερης επιλογής εκκινητών ειδικών για τη σειρά και του μεγαλύτερου μέρους του μήκους του μπορεί να παρέχει ταξινομικές πληροφορίες. Ορισμένοι παράγοντες μπορεί να έχουν ως αποτέλεσμα την εμφάνιση συγκρούσεων, όπως μήκος, περιεχόμενο GC, διαειδικές παραλλαγές και ομαδοποιήσεις και ταξινομικές προτιμήσεις των γονιδίων-στόχων. Συγκεκριμένα, χαρακτηριστικά των ITS1 και ITS2 μπορεί να θεωρηθούν ως περιορισμοί κατά τη χρήση τους σε κοινότητες μυκήτων (Yang et al., 2018). Όπως επισημάνθηκε από ορισμένους ερευνητές, π.χ., οι Nilsson, η τυποποίηση συγκεκριμένων υποπεριοχών ITS για την επίτευξη της αλληλούχισης απαιτεί περαιτέρω μελέτη.



Εικόνα 12: Δομή του γονιδίου rDNA σε μύκητες

Κεφάλαιο 4: Εφαρμογή στη Βιομηχανία Τροφίμων

Οι υψηλής ποιότητας πρώτες ύλες είναι θεμελιώδεις για την παραγωγή τροφίμων με επαρκή θρεπτική αξία και επιθυμητή γεύση (Konczak & Rouille, 2011; Pereira, Barros, Carvalho, & Ferreira, 2011). Η βιομηχανία τροφίμων έχει αναπτύξει διάφορες τεχνολογικές (π.χ. μικροδιήθηση, υπερθερμική επεξεργασία) και βιοτεχνολογικές (π.χ. ζύμωση) διεργασίες για τη διατήρηση και ενίσχυση των οργανοληπτικών ιδιοτήτων των προϊόντων της. Οι ποιοτικοί έλεγχοι γίνονται με διάφορες εργαστηριακές δοκιμές, οι οποίες αντιπροσωπεύουν το υποχρεωτικό σημείο εκκίνησης για ένα σωστό σύστημα ιχνηλασιμότητας τροφίμων. Οι κυβερνήσεις έχουν διαφορετικές εθνικές κατευθυντήριες γραμμές για την παραγωγή και τη συντήρηση τροφίμων (βλ., για παράδειγμα, τις συστάσεις του Παγκόσμιου Οργανισμού Υγείας — www.who.int/foodsafety/fs_management/infosan/en/ ή κανονισμούς όπως η Ευρωπαϊκή EC/178/2002), ενώ ο ορισμός των δοκιμών που πρέπει να χρησιμοποιούνται για την αξιολόγηση της ποιότητας και της ασφάλειας των τροφίμων είναι ευθύνη αρκετών ανεξάρτητων φορέων, όπως ο Αμερικανικός Οργανισμός Τροφίμων και Φαρμάκων και της Ευρωπαϊκής Αρχής για την Ασφάλεια των Τροφίμων. Η ζήτηση για αξιόπιστα συστήματα ιχνηλασιμότητας τροφίμων έχει αντιμετωπίσει την επιστημονική έρευνα, δημιουργώντας έτσι διαφορετικές αναλυτικές προσεγγίσεις στο πρόβλημα (Bottero & Dalmaso, 2011; Fajardo, Gonzalez, Rojas, Garcia, & Martin, 2010; Hellberg & Morrisey, 2011; Mafra, & Oliveira, 2008). Η επικύρωση της γνησιότητας των τροφίμων βασίζεται κυρίως στην ανάλυση πρωτεϊνών ή/και αλληλουχιών DNA. Οι μέθοδοι που βασίζονται σε πρωτεΐνες περιλαμβάνουν ανοσολογικές δοκιμασίες, ηλεκτροφορητικές και χρωματογραφικές τεχνικές όπως HPLC και TLC (Fügel, Carle, & Schieber, 2005; Kurtz, Leitenberger, Carle, & Schieber, 2010). Αν και είναι αποτελεσματικές στη δοκιμή φρέσκων προϊόντων, οι προσεγγίσεις που βασίζονται σε πρωτεΐνες έχουν χαμηλή αποτελεσματικότητα όταν εφαρμόζονται στην ανάλυση βαριά επεξεργασμένων τροφίμων. Σε αυτές τις περιπτώσεις, οι μέθοδοι που βασίζονται στο DNA είναι πιο αποτελεσματικές και μπορούν επίσης να εφαρμοστούν σε διαφορετικές μήτρες τροφίμων (Lockley & Bardsley, 2000; Mafra et al., 2008). Επιπλέον, το DNA είναι πιο ενημερωτικό από τις πρωτεΐνες και μπορεί εύκολα να εξαχθεί και με την παρουσία μικρών ιχνών οργανικού υλικού (Hellberg & Morrisey, 2011). Χάρη στις πρόσφατες εξελίξεις στη μοριακή βιολογία, οι δείκτες DNA έχουν γίνει το πιο αποτελεσματικό όργανο στην ανάλυση του DNA φυτικών ποικιλιών και φυλών ζώων και χρησιμοποιούνται επίσης για την

παρακολούθηση των πρώτων υλών στις διαδικασίες της βιομηχανίας τροφίμων (Kumar, Gupta, Misra, Modi, & Pandey, 2009· Mafra et al., 2008· Woolfe & Primrose, 2004). Ο στόχος της παρούσας ανασκόπησης είναι να συνοψίσει την τελευταία λέξη της τεχνολογίας σχετικά με τη χρήση της γραμμικής κωδικοποίησης DNA ως καθολικού εργαλείου για την ιχνηλασιμότητα των τροφίμων. Γενικά, οι μέθοδοι που βασίζονται σε DNA χρησιμοποιούν συγκεκριμένες αλληλουχίες DNA ως δείκτες και μπορούν να χωριστούν σε i) δείκτες που βασίζονται σε υβριδισμό και ii) δείκτες που βασίζονται στην αλυσιδωτή αντίδραση πολυμεράσης (PCR). Σε μεθόδους που βασίζονται στον υβριδισμό, τα ειδικά για είδη προφίλ DNA ανακαλύπτονται με υβριδισμό DNA που έχει υποστεί πέψη από ένζυμα περιορισμού και συγκρίνοντάς το με επισημασμένους ανιχνευτές (τεμάχια DNA γνωστής προέλευσης ή αλληλουχίας). Οι μέθοδοι που βασίζονται σε PCR περιλαμβάνουν την ενίσχυση των τόπων-στόχων με τη χρήση ειδικών ή αυθαίρετων εκκινητών και ενός ενζύμου DNA πολυμεράσης. Θραύσματα, στη συνέχεια διαχωρίζονται ηλεκτροφορητικά και τα μοτίβα λωρίδων ανιχνεύονται με διαφορετικές μεθόδους χρώσης, όπως η αυτοραδιογραφία. Οι μέθοδοι που βασίζονται στην PCR είναι εξαιρετικά ευαίσθητες, συχνά ταχύτερες από άλλες τεχνολογίες και χρησιμοποιούνται ευρέως στη γεωργία και τη ζωοτεχνία (Doulaty Baneh et al., 2007; Grassi, Labra, & Minuto, 2006; Labra et al., 2004; Mane, Tanwar, Girish, & Dixit, 2006; Teletchea, Maudet, & Hänni, 2005). Ασυνεχείς μοριακοί δείκτες όπως RAPD, AFLPs, καθώς και οι παραλλαγές τους (δηλαδή ISSR, SSAP, SAMPL) έχουν χρησιμοποιηθεί με επιτυχία στον χαρακτηρισμό διαφορετικών ειδών πρώτης ύλης (Chuang, Lur, Hwu, & Chang, 2011; Fajardo et al., 2010· Mafra et al., 2008· Nijman et al., 2003). Τα τελευταία χρόνια, η PCR-μετουσιωτική ηλεκτροφόρηση γέλης βαθμίδωσης (PCR-DGGE) έχει χρησιμοποιηθεί σε μεγάλο βαθμό στον τομέα της ιχνηλασιμότητας και της ασφάλειας των τροφίμων για τον χαρακτηρισμό βακτηρίων και ζυμομυκήτων σε προϊόντα που έχουν υποστεί ζύμωση (Dalmacio, Angeles, Larcia, Balolong, & Estacio, 2011· Muyzer, De Waal, & Uitterlinden, 1993· Peres, Barlet, Loiseau, & Montet, 2007· Zheng et al., 2012). Με τη χρήση αυτής της τεχνικής, η σύνθεση μικροοργανισμών ορίζεται με βάση το πρότυπο μετανάστευσης των PCR-τμημάτων που ανήκουν σε συγκεκριμένες γονιδιωματικές περιοχές όπως το 16S και το 26S rDNA (El Sheikh et al., 2009). Η PCR-DGGE χρησιμοποιήθηκε επίσης για την παρακολούθηση της βακτηριακής μόλυνσης σε προϊόντα διατροφής όπως τα ποτά που έχουν υποστεί ζύμωση (Hosseini, Hippe, Denner, Kollegger, & Haslberger, 2012) και για τον καθορισμό της προέλευσης της πρώτης ύλης

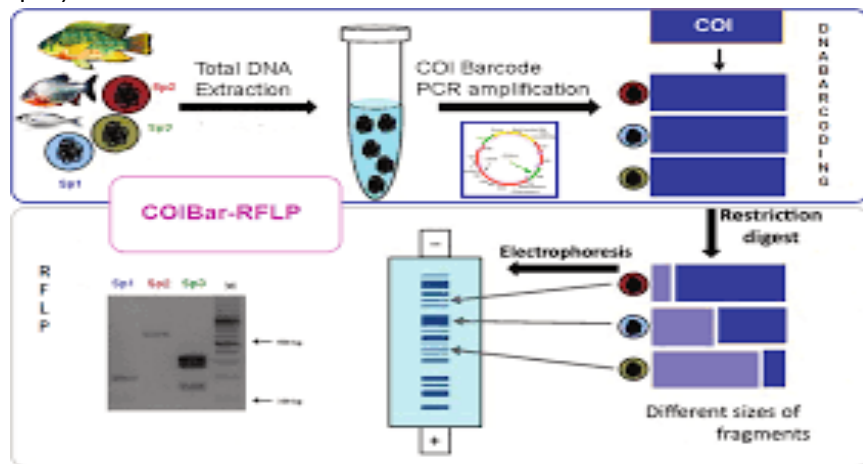
ξεκινώντας από τα χαρακτηριστικά της ζύμης ή των βακτηριακών κοινοτήτων της, όπως στην περίπτωση των φρούτων (El Sheikha, Bouvet, & Montet, 2011; El Sheikha, Durand, Sarter, Okullo, & Montet, 2012; El Sheikha, Métayer, & Montet, 2011) και των ψαριών (Le Nguyen, Ha, Dijoux, Loiseauet, & Montet, 2008; Montet, Le Nguyen, & El Sheikha, 2008). Η επιλογή της καταλληλότερης μοριακής προσέγγισης εξαρτάται από διάφορες πτυχές, συμπεριλαμβανομένης της ποσότητας γενετικής διαφοροποίησης του αναλυόμενου είδους, του χρόνου που απαιτείται για την ανάλυση, της αναλογίας κόστους/αποτελεσματικότητας και της τεχνογνωσίας των εργαστηρίων. Επιπλέον, οι γονιδιωματικές τεχνικές απαιτούν DNA υψηλής ποιότητας για να λειτουργήσει επιτυχώς, επειδή η αποτελεσματικότητα μπορεί να επηρεαστεί αρνητικά από αλλοιωμένο ή κατακερματισμένο DNA (Hellberg & Morrissey, 2011; Meusnier et al., 2008; Pafundo, Agrimonti, Maestri, & Marmioli, 2007). Όσον αφορά τα συστήματα που βασίζονται στην αλληλουχία, τα Single Nucleotide Polymorphisms (SNPs) και Simple Sequence Repeats (SSRs), χρησιμοποιούνται ευρέως στις μέρες μας λόγω του υψηλού επιπέδου πολυμορφισμού και της υψηλής αναπαραγωγιότητάς τους (Kumar et al., 2009). Αυτές οι προσεγγίσεις χρησιμοποιούνται τόσο για τον εντοπισμό ποικιλιών φυτών (Labra et al., 2003; Pasqualone, Lotti, & Blanco, 1999) και φυλών ζώων (Nijman et al., 2003), όσο και για την πρόληψη δόλιων εμπορικών δραστηριοτήτων (Chuang et al., 2011). Ωστόσο, επειδή είναι ιδιαίτερα ειδικές για τα διάφορα είδη, αυτές οι προσεγγίσεις απαιτούν πρόσβαση στη σωστή αλληλουχία DNA των οργανισμών (π.χ. στελέχη/ποικιλίες ή οικότυποι) και η εφαρμογή τους συχνά περιορίζεται σε ένα μεμονωμένο ταξινομικό ή σε στενά συγγενή ταξινομικές ομάδες. Η έλλειψη τυποποίησης και καθολικότητας είναι το πιο σημαντικό πρόβλημα των προσεγγίσεων που βασίζονται στο DNA. Το 2003, ένα νέο σύστημα αναγνώρισης, το DNA barcoding, αναπτύχθηκε από ερευνητές στο Πανεπιστήμιο του Guelph (Καναδάς). Αυτή η προσέγγιση βασίζεται στην ανάλυση της μεταβλητότητας σε μια τυπική περιοχή του γονιδιώματος που ονομάζεται «γραμμωτός κώδικας DNA» (Hebert, Ratnasingham, & deWaard, 2003). Αυτή η προσέγγιση αποδείχθηκε χρήσιμη στην επίλυση ταξινομικών προβλημάτων σε πολλές θεωρητικές και πρακτικές εφαρμογές (Hollingsworth, Graham, & Little, 2011; Rasmussen, Morrissey, & Hebert, 2009; Valentini, Pompanon, & Taberlet, 2009). Με μια αυστηρή έννοια, το barcoding DNA δεν είναι εντελώς καινοτόμο, επειδή οι προσεγγίσεις μοριακής ταυτοποίησης ήταν ήδη σε χρήση. Ωστόσο, έχει το

πλεονέκτημα ότι συνδυάζει τρεις σημαντικές καινοτομίες: τη μοριακή διαμόρφωση των διαδικασιών ταυτοποίησης (δηλαδή τη διερεύνηση της μεταβλητότητας του DNA για τη διάκριση μεταξύ ταξινομικών ταξινομήσεων), την τυποποίηση της διαδικασίας (από τη συλλογή δειγμάτων έως την ανάλυση των μοριακών εκρμών) και τη μηχανογράφηση (δηλ. όχι περιττή μεταφορά των δεδομένων με χρήση πληροφορικής) (Casiraghi, Labra, Ferri, Galimberti, & De Mattia, 2010). Το όνομα γραμμωτός κώδικας DNA αναφέρεται μεταφορικά στον τρόπο με τον οποίο ένας σαρωτής υπερύθρων αναγνωρίζει μονοσήμαντα ένα προϊόν χρησιμοποιώντας τις μαύρες λωρίδες του Universal Product Code (UPC). Ένας ιδανικός γραμμωτός κώδικας DNA απαιτεί δύο θεμελιώδη χαρακτηριστικά: υψηλή ταξινομική κάλυψη και υψηλή ανάλυση (Hebert et al., 2003). Η υψηλή ταξινομική κάλυψη (ονομάζεται επίσης «καθολικότητα») αναφέρεται στη σωστή ενίσχυση της γονιδιωματικής περιοχής που επιλέχθηκε ως γραμμωτός κώδικας DNA στο ευρύτερο πλαίσιο των ταξινομικών κατηγοριών. Από την άλλη πλευρά, η υψηλή ανάλυση διασφαλίζει την αναγνώριση διαφορετικών ταξινομικών κατηγοριών, με βάση τις διαειδικές διαφορές στις αλληλουχίες γραμμωτού κώδικα DNA. Ως γενική αρχή, οι περιοχές γραμμικού κώδικα DNA πρέπει να έχουν υψηλή ενδοειδική και χαμηλή ενδοειδική μεταβλητότητα. Η αναγνώριση των οργανισμών είναι θεμελιώδης για τη διασφάλιση υψηλών προτύπων ποιότητας για τη βιομηχανία τροφίμων και την αγορά (Myers, 2011· Novak, Gruber-Gréger, & Lukas, 2007). Η γραμμική κωδικοποίηση DNA είναι αποτελεσματική στην πιστοποίηση τόσο της προέλευσης όσο και της ποιότητας των πρώτων υλών και στον εντοπισμό νοθευμάτων. Ωστόσο, η απόδοσή του επηρεάζεται έντονα από τη μοριακή μεταβλητότητα των οργανισμών και επιτυγχάνεται υψηλό επίπεδο ανάλυσης όταν ένας οργανισμός έχει χαμηλό ενδοειδικό πολυμορφισμό, γεγονός που τον κάνει να διακρίνεται καλά από τα στενά συγγενικά ταξινομικά (Casiraghi et al., 2010; Hebert et al., 2003). Ένα άλλο κρίσιμο στοιχείο μπορεί να είναι η διαθεσιμότητα υψηλής ποιότητας αποθετηρίων ακολουθιών αναφοράς. Για το λόγο αυτό, ένας μεγάλος αριθμός αλληλουχιών γραμμωτού κώδικα DNA από ζώα και φυτά (συμπεριλαμβανομένων των εκτρεφόμενων ειδών) έχει υποβληθεί τα τελευταία 10 χρόνια και στις βάσεις δεδομένων NCBI και BOLD (www.barcodeoflife.org), ακολουθώντας τις οδηγίες που παρέχονται από το Database Working Group.

4.1 DNA Barcoding σε ιχθυηρά

Το barcoding DNA αποδείχθηκε ότι είναι ιδιαίτερα αποτελεσματικό στην ιχνηλασιμότητα των θαλασσινών (Becker, Hanner, & Steinke, 2011). Ο όρος «θαλασσινά» χρησιμοποιείται συνήθως για να δηλώσει βρώσιμες μορφές υδρόβιας ζωής, συμπεριλαμβανομένων των ψαριών, των μαλακίων, των καρκινοειδών και των εχινόδερμων τα οποία διατίθενται στην αγορά ως ολόκληροι οργανισμοί ή ως μεταποιημένα προϊόντα. Τα είδη θαλασσινών προσδιορίζονται γενικά σύμφωνα με την περιοχή προέλευσής τους και με διάφορους μορφολογικούς περιγραφείς. Ωστόσο, η αυξημένη ζήτηση θαλασσινών και η παγκοσμιοποίηση της αγοράς, έχουν κάνει πιο δύσκολο τον έλεγχο τόσο των εμπορικών οδών όσο και των συστημάτων βιομηχανικής επεξεργασίας (δηλαδή συστήματα αποθήκευσης, κατάψυξη, ξήρανση). Επιπλέον, αρκετά νέα είδη έχουν εισαχθεί στην αγορά. Μερικές φορές, αυτά τα «νέα ψάρια» έχουν την ίδια εμπορική ονομασία με οργανισμούς που κυκλοφορούσαν προηγουμένως στην αγορά, αλλά δεν αντιστοιχούν στο ίδιο είδος. Θα μπορούσαν επίσης να έχουν διαφορετικές διατροφικές αξίες ή/και να είναι δυνητικά αντιγονικά (Barbuto et al., 2010). Η γραμμική κωδικοποίηση DNA είναι επιτυχής όταν εφαρμόζεται στα θαλασσινά επειδή: (1) σε σύγκριση με άλλες ζωικές πηγές (π.χ. βοοειδή, πρόβατα, κασίκες, άλογο) ο αριθμός των ειδών είναι υψηλότερος, έτσι ώστε η αποτελεσματικότητα αυτής της τεχνικής ενισχύεται. (2) οι κλασικές προσεγγίσεις αναγνώρισης δεν είναι χρήσιμες σε πολλές περιπτώσεις, ιδίως με τα επεξεργασμένα τρόφιμα. (3) στα θαλασσινά περισσότερο από ό,τι σε άλλες ζωντανές ομάδες, η μοριακή ταυτοποίηση μπορεί να υπερβεί το επίπεδο των ειδών, επιτρέποντας σε αρκετές περιπτώσεις την αναγνώριση τοπικών ποικιλιών και ως εκ τούτου τον προσδιορισμό της προέλευσης ενός συγκεκριμένου προϊόντος. Αρκετοί ερευνητές έχουν συζητήσει τη δυνατότητα της γραμμικής κωδικοποίησης DNA ως εγκληματολογικού εργαλείου για την ιχνηλασιμότητα των βρώσιμων ψαριών (βλ. για παράδειγμα Barbuto et al., 2010· Smith, McVeagh, & Steinke, 2008· Yancy et al., 2008). Η περιοχή *cox1* έδειξε μια καλή διακριτική δύναμη στην αναγνώριση των ειδών ψαριών (98% των ερευνηθέντων θαλάσσιων ειδών και 93% των ειδών του γλυκού νερού εντοπίστηκαν με επιτυχία, Ward, Hanner, & Hebert, 2009). Επιτυχή αποτελέσματα λήφθηκαν επίσης ξεκινώντας από μικρή μερίδα φρέσκου ή επεξεργασμένου υλικού χρησιμοποιώντας λίγους συνδυασμούς γενικών ασταριών (βλ. Steinke & Hanner, 2011). Ο γραμμωτός κώδικας DNA προτάθηκε από την

Υπηρεσία Τροφίμων και Φαρμάκων των ΗΠΑ για τον έλεγχο ταυτότητας εμπορικών προϊόντων με βάση τα ψάρια (Yancy et al., 2008). Ειδικότερα, η FDA των ΗΠΑ σχεδίαζε να συμπεριλάβει δεδομένα γραμμικού κώδικα DNA στη Ρυθμιστική Εγκυκλοπαίδεια Ψαριών, προκειμένου να βοηθήσει στη διερεύνηση της λανθασμένης επισήμανσης και της υποκατάστασης ειδών ψαριών. Το barcoding DNA αποδείχθηκε επίσης αποτελεσματικό στην παρακολούθηση των θαλασσιών μετά από βιομηχανική επεξεργασία. Ορισμένα είδη απαιτούν μόνο μια πρωτογενή επεξεργασία, όπως η κατάψυξη του φρέσκου ψαριού για διανομή σε λιανοπωλητές φρέσκων ψαριών και καταστήματα εστίασης, διατηρώντας επομένως μορφολογικούς χαρακτήρες χρήσιμους για ακριβή αναγνώριση. Ωστόσο, όταν απαιτείται μια περίπλοκη διαδικασία παραγωγής (δηλαδή προϊόντα με απλή ψύξη, κατεψυγμένα και κονσερβοποιημένα προϊόντα για το εμπόριο λιανικής και τροφοδοσίας), ή στην περίπτωση ψαριών που πωλούνται σε μέρη (π.χ. μπριζόλες, τεμάχια, σουρίμι, μπασοούνια και πτερύγια), κλασικές διαδικασίες αναγνώρισης δεν είναι αποτελεσματικά και η γραμμική κωδικοποίηση DNA μπορεί να είναι χρήσιμη για την απόκτηση αναγνώρισης. Παρά την αποδεδειγμένη αποτελεσματικότητά του, ελάχιστες μελέτες έχουν γίνει σχετικά με την εφαρμογή γραμμικής κωδικοποίησης DNA σε άλλες κατηγορίες θαλασσιών (π.χ. καβούρια: Haye, Segovia, Vera, Gallardo, & Gallardo-Escàrate, 2012, holothurians: Uthicke et al., 2010, αστακοί : Naro-Maciel et al., 2011). Επιπλέον, η προσέγγιση γραμμικού κώδικα DNA που βασίζεται στο *cox1* δεν είναι πάντα κατάλληλη για τον εντοπισμό ορισμένων οργανισμών, όπως τα γαστερόποδα (Meyer & Paulay, 2005). Απαιτούνται εκτενέστερες μελέτες για να επιβεβαιωθεί η πιθανή χρήση αυτής της τεχνικής σε όλα τα είδη θαλασσιών ως αξιόπιστο «εργαλείο ιχνηλασιμότητας»



Εικόνα 13: Αναγνώριση Ειδών Ιχθυερών με DNA Barcoding

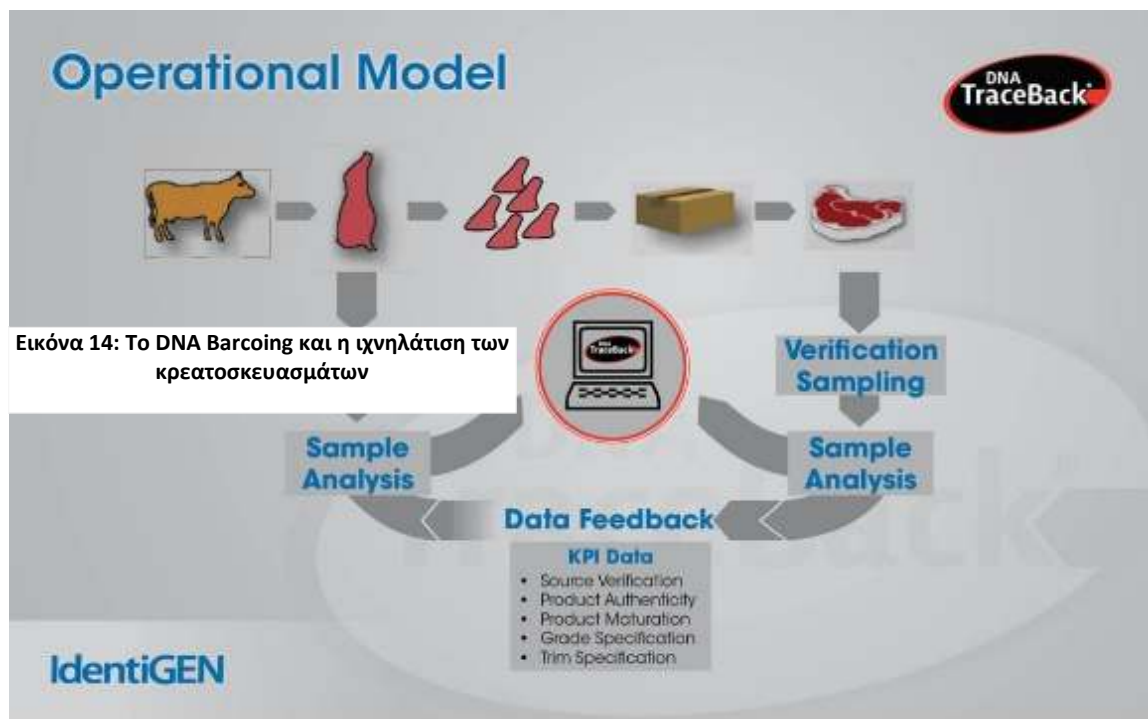
4.2 Το DNA barcoding και η ιχνηλασιμότητα στα κρεατοσκευάσματα: το πρόβλημα της έλλειψης δεδομένων

Το κρέας υπόκειται συνήθως σε μεγάλες αλυσίδες παραγωγής και διανομής, κάτι που απαιτεί κατάλληλα συστήματα ιχνηλασιμότητας. Παθολογίες που σχετίζονται με το κρέας ως τρόφιμο (π.χ. ΣΕΒ, γρίπη των πτηνών) και κακές πρακτικές ορισμένων παραγωγών, έχουν αυξήσει την ευαισθητοποίηση του κοινού σχετικά με την προέλευση και την ποιότητα του κρέατος. Ως εκ τούτου, είναι απαραίτητος ο ορισμός ακριβών και αξιόπιστων μεθόδων για τον προσδιορισμό της σύνθεσης του κρέατος τροφίμων, εκτός από τη χρήση ετικετών, οι οποίες δεν παρέχουν επαρκείς εγγυήσεις για το πραγματικό περιεχόμενο ενός προϊόντος. Αυτές οι νέες μέθοδοι θα πρέπει να προστατεύουν τόσο τους καταναλωτές όσο και τους παραγωγούς από απάτες και τα ζωικά είδη από την υπερεκμετάλλευση ή το παράνομο εμπόριο (Manel, Berthier, & Luikart, 2002). Μια ποικιλία προσεγγίσεων που βασίζονται στο DNA για την ιχνηλασιμότητα του κρέατος, όπως PCR-RFLP, PCR και PCR αλληλουχίας ειδικών για είδη, έχουν αναπτυχθεί (Mane et al., 2006; Teletchea et al., 2005). Αυτές οι προσεγγίσεις περιλαμβάνουν τη χρήση μιτοχονδριακών διαφορετικών από πυρηνικούς δείκτες. Πρόσφατα, οι Teletchea, Bernillon, Duffraisie, Laudet και ο Hänni (2008) πρότεινε μια μέθοδο βασισμένη σε μικροσυστοιχίες, η οποία χρησιμοποιεί ανιχνευτές προερχόμενους από το κυτόχρωμα β, ως εργαλείο για τον εντοπισμό εμπορικών και απειλούμενων ειδών σπονδυλωτών τόσο σε τρόφιμα όσο και σε ιατροδικαστικά δείγματα κρέατος. Η περιοχή του κυτοχρώματος b εμφανίζει μεγάλη ενδοειδική και χαμηλή ενδοειδική ποικιλομορφία, καθώς και διατηρημένες πλευρικές περιοχές, επομένως είναι τυπική υποψήφια ως περιοχή γραμμικού κώδικα DNA. Η επιλογή του cyt b αντί του cox1 οφείλεται κυρίως σε πρακτικούς λόγους. Αρκετές χιλιάδες αλληλουχίες cyt b κατατίθενται σε δημόσιες βάσεις δεδομένων για ένα μεγάλο εύρος βρώσιμων ειδών θηλαστικών, ενώ μόνο λίγες αλληλουχίες cox1 είναι διαθέσιμες στο BOLD και στο GenBank. Ωστόσο, παρά την έλλειψη δεδομένων, η τεχνική γραμμικής κωδικοποίησης DNA που βασίζεται στο cox1 μπορεί να θεωρηθεί μια αξιόπιστη μέθοδος για την ιχνηλασιμότητα του κρέατος θηλαστικών (βλ. Cai et al., 2011; Francis et al., 2010; Luo et al., 2011). Όπως εφαρμόζεται στην αγορά κρέατος, οι σχέσεις μεταξύ αλληλουχιών γραμμικής κωδικοποίησης DNA και ονομάτων ειδών θα πρέπει να αξιολογούνται κριτικά, επειδή η εμπορική ονομασία ενός προϊόντος κρέατος θα μπορούσε να αναφέρεται σε

διαφορετικές μοριακές μονάδες (οι λεγόμενες Molecular Operational Taxonomic Units, ή MOTUs, Casiraghi et al. , 2010). Για παράδειγμα, οι Ludt, Schroeder, Rottmann και Kuehn (2004) έδειξαν ξεκάθαρα σταθερές μοριακές διαφορές μέσα στο είδος *Cervus elaphus*. Κατά συνέπεια, το κρέας ελαφιού θα πρέπει να ταυτιστεί με δύο διαφορετικές αλληλουχίες DNA που αντιστοιχούν στο *Cervus canadensis* (που εμφανίζεται στην Ασία και τη Βόρεια Αμερική) και το *C. elaphus* (που κατοικεί στην Ευρώπη). Παρόμοια κατάσταση συνέβη σε είδη πτηνών όπως στην περίπτωση των αγγλικών και των ΗΠΑ φυλών γαλοπούλας (*Meleagris gallopavo*) που παρουσίασαν σταθερές γενετικές διαφορές (Hird, Goodier, & Hill, 2003). Υπάρχουν επίσης αρκετές περιπτώσεις ειδών ή φυλών με το ίδιο προφίλ DNA. Σε αυτήν την περίπτωση, η προσέγγιση γραμμικού κώδικα DNA δεν θα είναι σε θέση να επιστρέψει μια σωστή αναγνώριση, καθιστώντας επομένως αδύνατη την παρακολούθηση ορισμένων προϊόντων κρέατος. Αυτό το φαινόμενο, λόγω του υβριδισμού, που προκαλεί γενετική εισβολή, είναι σύνηθες στα ζώα. Τα βοοειδή, όπου πολλές φυλές προέρχονται από γεγονότα υβριδισμού (βλ. Kikkawa et al., 2003; Nijman et al., 2003; Verkaar et al., 2003), είναι ένα χαρακτηριστικό παράδειγμα.

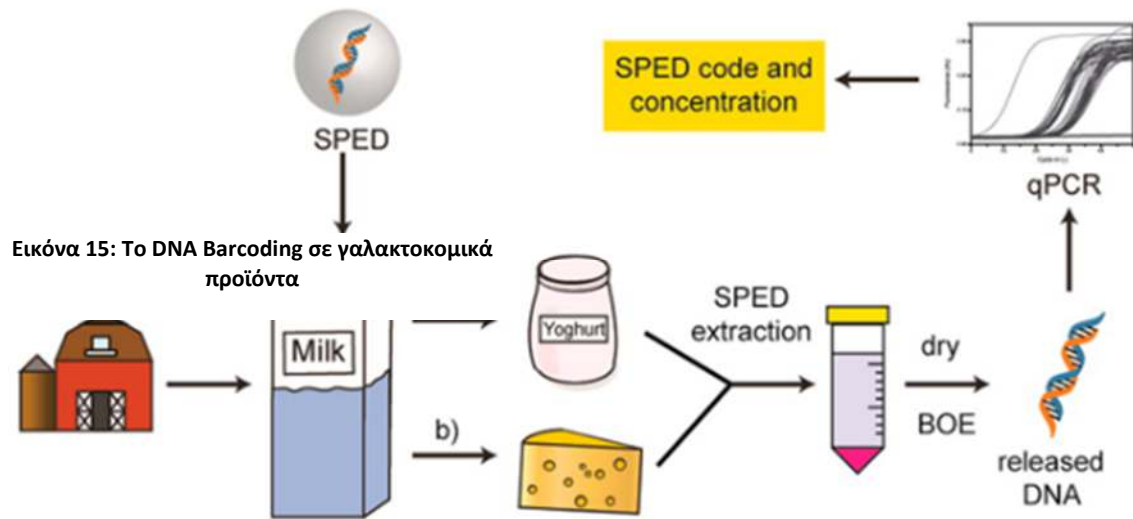
4.3 Το DNA Barcoding στα γαλακτοκομικά προϊόντα: μια εν δυνάμει εφαρμογή

Τα γαλακτοκομικά προϊόντα ορίζονται γενικά ως τρόφιμα που παρασκευάζονται από γάλα



θηλαστικών. Λόγω της οικονομικής συνάφειας, του κινδύνου αλλεργιών και των θρησκευτικών

πρακτικών που σχετίζονται με αυτή την κατηγορία προϊόντων, η ανάπτυξη τεχνικών για την αξιολόγηση της γνησιότητας και της νοθείας των τροφίμων που προέρχονται από το γάλα είναι ένα θέμα πρωταρχικής σημασίας (Mafra et al., 2008). Η χρήση μοριακών εργαλείων για τον χαρακτηρισμό και τον εντοπισμό των γαλακτοκομικών προϊόντων κερδίζει μεγάλη αποδοχή (Ponzoni, Mastromauro, Gianì, & Breviario, 2009) ακόμη και αν δεν υπάρχουν μελέτες που να βασίζονται σε μια αυστηρή προσέγγιση γραμμικού κώδικα DNA. Ωστόσο, η ειδική για είδος PCR έχει αποδειχθεί αξιόπιστη μέθοδος για τον έλεγχο της αυθεντικότητας αυτής της κατηγορίας τροφίμων, επειδή μια συγκεκριμένη αλληλουχία στόχος (π.χ. 12S rRNA, 16S rRNA, cytb) μπορεί να ανιχνευθεί σε μήτρες που περιέχουν μια δεξαμενή ετερογενούς γονιδιωματικού DNA, όπως γάλα (Mafra et al., 2008). Μεταξύ των εφαρμογών αυτών των μοριακών εργαλείων, υπάρχει η δυνατότητα ανίχνευσης νοθείας γάλακτος υψηλότερης αξίας από μη δηλωμένο αγελαδινό γάλα ή παράλειψης δηλωμένου είδους γάλακτος. Όσον αφορά τον χαρακτηρισμό της προέλευσης και της ποιότητας του γάλακτος, πρόσφατα περιγράφηκε μια ενδιαφέρουσα εφαρμογή του γραμμικού κώδικα DNA. Το plastidial rbcL - ο πιο καθολικός δείκτης για τη γραμμική κωδικοποίηση του DNA των φυτών - βρέθηκε ικανός να ανιχνεύσει ίχνη θραυσμάτων φυτικού DNA που προέρχονται από ζωοτροφές στο νωπό αγελαδινό γάλα και στα κλάσματά του (Nemeth et al., 2004; Ponzoni et al., 2009). Αυτό ανοίγει νέες προοπτικές για την ιχνηλασιμότητα του γάλακτος και των γαλακτοκομικών προϊόντων γενικότερα. Συνολικά, για να επιτευχθεί ένας ακριβής χαρακτηρισμός της ποιότητας των γαλακτοκομικών προϊόντων, είναι απαραίτητη μια πολυεπίπεδη μοριακή προσέγγιση. Συγκεκριμένα, τεχνικές που μοιάζουν με barcoding DNA είναι χρήσιμες για την παροχή αξιόπιστου χαρακτηρισμού της σύνθεσης του νωπού γάλακτος, ενώ άλλες προσεγγίσεις όπως η PCR-DGGE μπορούν να είναι χρήσιμες για την αξιολόγηση της μικροβιακής σύνθεσης και προέλευσης των επεξεργασμένων γαλακτοκομικών προϊόντων (Arcuri, El Sheikha, Rychlik, Piro-Métayer, & Montet, 2013; Borelli, Ferreira, Lacerda, Franco, & Rosa, 2006; Coppola, Blaiotta, Ercolini, & Moschetti, 2001; Dolci, Alessandria, Rantsiou, Bertolino, & Coolini, 2010; , 2004; Ercolini, Frisso, Mauriello, Salvatore, & Coppola, 2008).



4.4 Το DNA Barcoding σε φυτικά προϊόντα

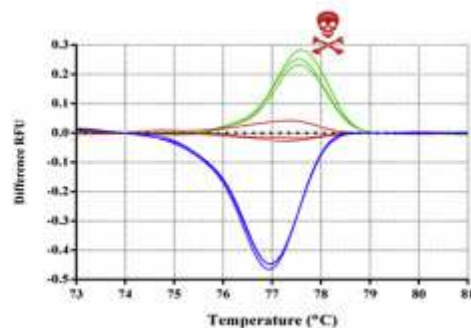
Τα φυτά αποτελούν ουσιαστικό στοιχείο της ανθρώπινης διατροφής, τόσο άμεσα (τα δημητριακά είναι η βάση της διατροφικής πυραμίδας και ακολουθούν τα φρούτα και τα λαχανικά) όσο και έμμεσα (τα φυτικά προϊόντα χρησιμοποιούνται για τη διατροφή των βοοειδών). Επιπλέον, πολλά φυτά χρησιμοποιούνται ως πρόσθετα τροφίμων (π.χ. σόγια). Η αξιόπιστη αναγνώριση των ειδών των καλλιεργειών, καθώς και η προέλευση και η ιχνηλασιμότητα τους, αποτελούν βασικά στοιχεία στον τομέα της ασφάλειας των τροφίμων. Τα τελευταία 20 χρόνια έχουν δοκιμαστεί αρκετές μέθοδοι PCR σε διάφορες ποικιλίες καλλιεργειών, όπως ρύζι, καλαμπόκι, σόργο, κριθάρι, σίκαλη (Pasqualone et al., 1999; Ren, Zhu, Warndorff, Bucheli, & Shu, 2006; Salem et al., 2007; Terzi et al., 2005). Αυτές οι μέθοδοι είναι χρήσιμες τόσο για τους παραγωγούς που ενδιαφέρονται για την προστασία και την πιστοποίηση των καλλιεργειών τους (De Mattia, Imazio, Grassi, & Labra, 2008; Labra et al., 2003; Ren et al., 2006), όσο και για τους καταναλωτές που ενδιαφέρονται για την ποιότητα και την προέλευση των τροφίμων τους. Η αυξανόμενη διάχυση των γενετικά τροποποιημένων (GM) καλλιεργειών έχει αυξήσει περαιτέρω τη ζήτηση για μοριακές τεχνικές παρακολούθησης διαγονιδίων (Auer, 2003). Τα τελευταία χρόνια, έχει αναπτυχθεί ένα τσιπ πολλαπλής μικροσυστοιχίας DNA για ταυτόχρονη ταυτοποίηση ΓΤΟ, με βάση κανονισμούς διαφορετικών χωρών (Leimanis et al., 2006; Nikolic, Taski-Ajdukovic, Jevtic, & Marinkovic, 2009), καθώς και παρόμοια συστήματα αφιερωμένα στην αναγνώριση ειδών και ποικιλιών φυτών (Agrimonti, Vietina, Pafundo, & Marmioli, 2011; Xie,

McNally, Li, Leung, & Zhu, 2006). Ωστόσο, όπως αναφέρθηκε προηγουμένως, αυτές οι μοριακές μέθοδοι έχουν έναν κοινό περιορισμό στην υψηλή εξειδίκευσή τους σε είδη. Λόγω της παγκοσμιοποίησης, ένας αυξανόμενος αριθμός φυτών που προέρχονται από διάφορες περιοχές του κόσμου προσφέρονται πλέον στους καταναλωτές, αλλά δεν υπάρχουν αξιόπιστα, καθολικά εργαλεία για την αναγνώρισή τους. Η γραμμική κωδικοποίηση DNA θα μπορούσε να είναι μια αξιόπιστη εναλλακτική λύση στις προσεγγίσεις δακτυλικών αποτυπωμάτων DNA στην ταυτοποίηση φυτών, με υψηλότερη αναλογία αποτελεσματικότητας/κόστους. Στην πραγματικότητα, το barcoding DNA δεν απαιτεί εκτενή γνώση του γονιδιώματος κάθε οργανισμού, καθώς βασίζεται στη χρήση ενός ή λίγων καθολικών δεικτών (Hollingsworth et al., 2011). Η έρευνα για το barcoding του DNA στον τομέα της βοτανικής μετατοπίζεται από την ανάλυση της απόδοσης διαφορετικών δεικτών σε πιο πρακτικές εφαρμογές. Μεταξύ των βρώσιμων φυτών, αυτή η προσέγγιση έχει χρησιμοποιηθεί για την παρακολούθηση των μπαχαρικών (De Mattia et al., 2011). Τα είδη του γένους μέντας, ώκιμον, ρίγανης, σάλβια, θύμους και δεντρολίβανου αναλύθηκαν χρησιμοποιώντας την περιοχή του γραμμικού κώδικα (matK+rbcl) και τον διαγονιδιακό διαχωριστή trnH-psbA. Με το barcoding DNA, μπορούν να εντοπιστούν τα πιο κοινά μπαχαρικά, με εξαίρεση τη μαντζουράνα και τη ρίγανη, που ανήκουν στο ίδιο γένος. Η ρίγανη, έχει ενδοειδική ποικιλομορφία η οποία είναι μεγαλύτερη από τη μεσοειδική, λόγω αρκετών περιπτώσεων υβριδισμού. Η γραμμική κωδικοποίηση DNA έχει δείξει υψηλές επιδόσεις στη διάκριση των ειδών βασιλικού: τα matK και trnH-psbA μπόρεσαν να διακρίνουν τον εμπορικό βασιλικό (*Ocimum basilicum* L) από άλλα είδη *Ocimum*, καθώς και από διαφορετικές ποικιλίες βασιλικού. Ο γραμμωτός κώδικας DNA χρησιμοποιήθηκε επίσης για τη διερεύνηση της γενετικής σχέσης μεταξύ άγριων και καλλιεργούμενων φυτών, καθώς και για την προέλευσή τους. Οι Nicolè et al. (2011) χρησιμοποίησε γραμμωτό κώδικα DNA σε βλάστημα φασολιών (*Phaseolus vulgaris* L.) παρατηρώντας διακριτούς απλότυπους για προσαρτήσεις φασολιών που αντιστοιχούν σε περιοχές της Μεσοαμερικανικής ή των Άνδεων. Ωστόσο, αυτή η μελέτη τόνισε επίσης τα όρια προσέγγισης στην επίλυση γενετικών σχέσεων μεταξύ φυλών και αυστηρά συγγενών ποικιλιών. Τα όρια της υιοθέτησης καθολικών δεικτών γραμμικού κώδικα είναι εμφανή σε επίπεδο ποικιλίας, όπου η γενετική μεταβλητότητα είναι περιορισμένη και υπάρχουν επιπλοκές λόγω γεγονότων αναπαραγωγής. Για να ξεπεραστούν αυτά τα όρια, οι Kane

και Cronk (2008) πρότειναν τη μεθοδολογία υπερ-γραμμικού κώδικα, η οποία βασίζεται στην αλληλουχία ολόκληρου του πλαστιδικού γονιδιώματος, μαζί με μεγάλα τμήματα του πυρηνικού γονιδιώματος. Αυτός ο συνδυασμός παρέχει αρκετές πληροφορίες για την απόδειξη της γενετικής ποικιλότητας κάτω από το επίπεδο των ειδών, διακρίνοντας τα υβρίδια από τις καθαρές γραμμές, επομένως είναι πολύ πιο ευαίσθητος από την παραδοσιακή γραμμοκώδικα DNA (Nock et al., 2011; Parks, Cronn, & Liston, 2009; Steele & Pires, 2011). Οι Kane και Cronk (2008) αξιολόγησαν την αποτελεσματικότητα της υπερ-γραμμικής κωδικοποίησης στο κακάο (*Theobroma cacao* L.) και βρήκαν αρκετά πλαστικά και πυρηνικά SNP, τα οποία ήταν χρήσιμα για την αναγνώριση διαφορετικών ποικιλιών. Αυτή η τεχνική είναι πολλά υποσχόμενο, αλλά είναι δύσκολο να εφαρμοστεί σε μεγάλη κλίμακα λόγω του υψηλού του κόστους και την υπερβολική ιδιαιτερότητά του ως προς το είδος.



Εικόνα 16: Η χρήση DNA Barcoding φυτών σε συνδυασμό με ανάλυση HRM για τη διαφοροποίηση των βρώσιμων λαχανικών από τα δηλητηριώδη φυτά για την ασφάλεια των τροφίμων



4.5 Το DNA Barcoding ως εργαλείο ιχνηλασιμότητας κατά την μεταποίηση τροφίμων

Ένα «ιδανικό» σύστημα ιχνηλασιμότητας θα ακολουθούσε το «ιστορικό» ενός προϊόντος από την προέλευσή του έως τη στιγμή που χρησιμοποιείται, λαμβάνοντας υπόψη όλα τα στάδια μετατροπής και εμπορευματοποίησης. Συστήματα μοριακής ταυτοποίησης και ιχνηλασιμότητας αναπτύχθηκαν για να λειτουργούν σε πρώτες ύλες. Ωστόσο, οι σπόροι, τα φρούτα και διάφορα μέρη φυτών και ζώων μετατρέπονται σε τρόφιμα με καθορισμένο σχήμα, γεύση και οσμή μέσω φυσικών (π.χ. θέρμανση, βρασμό, ακτινοβολία UV) ή χημικών (δηλαδή προσθήκη συντηρητικών τροφίμων, τεχνητών γλυκαντικών) επεξεργασιών, οι οποίες θα μπορούσαν να αλλάξουν τη δομή του DNA. Για το λόγο αυτό, η εφαρμογή τεχνικών αναγνώρισης με βάση το DNA (μεταξύ των οποίων η γραμμοκώδικα DNA) σε μετασχηματισμένα εμπορεύματα μπορεί να είναι

αναποτελεσματική λόγω του επιπέδου αποικοδόμησης του DNA και της ταυτόχρονης παρουσίας πολλών γονιδιωμάτων που ανήκουν σε διαφορετικούς οργανισμούς. Το DNA είναι συνήθως πιο ανθεκτικό σε βιομηχανικές διεργασίες από άλλα μόρια, όπως οι πρωτεΐνες (Martinez et al., 2003) και οι μέθοδοι δακτυλικών αποτυπωμάτων DNA μπορούν να χρησιμοποιηθούν με επιτυχία στην αναγνώριση ζωικών φυτικών υλικών, ακόμη και όταν υπάρχουν μικρά ίχνη. Παρόλα αυτά, η επεξεργασία τροφίμων προκαλεί χημικές και φυσικές αλλοιώσεις, ενώ η αποδόμηση και ο κατακερματισμός είναι τα πιο κοινά αποτελέσματα (Bauer, Weller, Hammes, & Hertel, 2003). Η ακεραιότητα του DNA επηρεάζει σε μεγάλο βαθμό την αποτελεσματικότητα των μοριακών μεθοδολογιών (Hellberg & Morrisey, 2011; Meusnier et al., 2008; Pafundo et al., 2007). Η γραμμική κωδικοποίηση DNA μπορεί να έχει δύο πλεονεκτήματα σε σύγκριση με τις προσεγγίσεις δακτυλικών αποτυπωμάτων DNA: i) απαιτεί την ενίσχυση ενός μικρού θραύσματος DNA (επομένως υπάρχει μικρότερος κίνδυνος κατακερματισμού) και ii) βασίζεται σε μιτοχονδριακό ή πλαστιδιακό γονιδίωμα (περισσότερο διατηρημένο κατά τη διάρκεια επεξεργασία). Οι Aslan, Hamill, Sweeney, Reardon και Mullen (2009) έδειξαν ότι το πυρηνικό DNA διατηρείται λιγότερο από το μιτοχονδριακό στο μαγειρεμένο κρέας. Η ανάλυση διαφορετικών περιοχών mtDNA μπορεί να είναι αποτελεσματική στον εντοπισμό του βοείου, του πρόβειου και του χοιρινού κρέατος ακόμη και μετά το βράσιμο, το μαγείρεμα υπό πίεση ή το τηγάνισμα (Arslan, Ilhak, & Caliciogiu, 2006; Aslan et al., 2009; Mane et al., 2009). Οι πλήρεις αλληλουχίες *cox1* ελήφθησαν από καπνιστά προϊόντα ψαριών όπως ο μπακαλιάρος, το *groper*, το σκουμπρί, ο σολομός και ο τόνος (Smith et al., 2008). Διαπιστώθηκαν ορισμένες δυσκολίες στη λήψη γραμμικών κωδικών DNA πλήρους μήκους από κονσέρβες ψαριών. Σε αυτές τις περιπτώσεις, η χρήση μικρότερων ακολουθιών γραμμικού κώδικα θεωρήθηκε κατάλληλη επιλογή, όταν περιοριζόταν σε σκοπούς ιχνηλασιμότητας (Rasmussen et al., 2009). Ομοίως με το mtDNA, το γονιδίωμα του πλαστιδίου διατηρείται στα περισσότερα επεξεργασμένα τρόφιμα που προέρχονται από φυτά. Ο γραμμικός κώδικας DNA χρησιμοποιήθηκε για την αναγνώριση διαφορετικών αρωματικών ειδών μετά από βιομηχανική ξήρανση και τεμαχισμό (De Mattia et al., 2011). Οι δείκτες γραμμικού κώδικα DNA χρησιμοποιήθηκαν επίσης αποτελεσματικά για την αναγνώριση του τσαγιού του εμπορίου (Stoeckle et al., 2011), των ειδών φρούτων στο γιαούρτι (Knight, Ortola-Vidal, Schnerr, Rojmyr, & Lysholm, 2007) και των υπολειμμάτων φρούτων (π.χ.

μπανάνα) στους χυμούς, πουρές, σοκολάτες, μπισκότα κ.λπ. (Sakai et al., 2010). Ως εκ τούτου, η προσέγγιση γραμμικού κώδικα DNA θα μπορούσε να χρησιμοποιηθεί για την ανάλυση διαφορετικών μητρών τροφίμων, με κύριους περιορισμούς: i) το επίπεδο αποικοδόμησης του DNA. ii) η ανάπτυξη αξιόπιστων μεθόδων για την εξαγωγή DNA, και iii) η αποτελεσματικότητα των διαφορετικών μεθοδολογιών barcode (Hellberg & Morrisey, 2011).

4.6 Χαρακτηρισμός «μεικτών» τροφίμων

Οι περιοχές γραμμικού κώδικα DNA και οι εκκινητές που χρησιμοποιούνται από την ενίσχυση DNA είναι καθολικές (Hebert et al., 2003). Λαμβάνοντας υπόψη αυτές τις υποθέσεις, οι ενισχύσεις PCR που πραγματοποιούνται σε δείγματα DNA που προέρχονται από μικτές μήτρες τροφίμων παράγουν πολλά θραύσματα γραμμικού κώδικα DNA, τα οποία αντιστοιχούν σε διαφορετικά είδη. Ως εκ τούτου, η αλληλουχία DNA με βάση το Sanger, αν και είναι αποτελεσματική όταν χρησιμοποιείται για γραμμικού κώδικα DNA, δεν είναι μια εφικτή προσέγγιση για μικτά τρόφιμα, εκτός εάν προηγηθεί μια προσέγγιση κλωνοποίησης, η οποία θα μπορούσε να προκαλέσει προκαταλήψεις, λόγω της συνένισης θραυσμάτων DNA από διαφορετικά άτομα ή τάξεις. Διάφορες τεχνικές, όπως η πέψη με συγκεκριμένα περιοριστικά ένζυμα (δηλαδή RFLP), ή η ηλεκτροφορητική ανάλυση (Colombo, Chess, Cattaneo, & Bernardi, 2011; Mane et al., 2009; Teletchea, 2009) χρησιμοποιήθηκαν για τον διαχωρισμό διαφορετικών θραυσμάτων DNA πριν από την διαδικασία αλληλουχίας. Ωστόσο, αυτές οι μέθοδοι είναι αποτελεσματικές μόνο όταν η μήτρα τροφίμων αποτελείται από λίγα είδη και όταν έχουν σχετικά σχετικές διαφορές στους γραμμικούς κώδικες DNA τους (δηλαδή διαφορετικές περιοχές στόχου για ένζυμα περιορισμού και αλληλουχίες διαφορετικού μήκους). Σε άλλες περιπτώσεις, τα αμπλικόνια θα πρέπει να κλωνοποιούνται σε φορείς πλασμιδίου και να εισάγονται σε βακτηριακά ικανά κύτταρα (Zeale, Butlin, Barker, Lees, & Jones, 2011), προκειμένου να ληφθούν μεμονωμένα θραύσματα. Μέχρι σήμερα, αυτή η διαδικασία έχει χρησιμοποιηθεί σε διατροφικές μελέτες αφιερωμένες σε ορισμένα ζώα, όπως θηλαστικά και πτηνά, ή για την αναγνώριση φυτών που υπάρχουν στα εντερικά δείγματα των μαμούθ (Van Geel et al., 2011). Μια πιθανή αποτελεσματική προσέγγιση στην εφαρμογή γραμμικού κωδικοποίησης DNA σε σύνθετες μήτρες τροφίμων θα μπορούσε να είναι η μεθοδολογία πυροαλληλουχίας 454, η οποία παράγει αρκετές εκατοντάδες χιλιάδες αλληλουχίες ανά εκτέλεση, που αντιστοιχούν σε ολόκληρο το

μείγμα των μορίων DNA που εξάγονται από τη μήτρα. Αυτή η προσέγγιση επιτρέπει τον εντοπισμό όλων των πρώτων υλών, συμπεριλαμβανομένων των ρύπων, ή των στοιχείων που εμφανίζονται μόνο σε ίχνη. Η Pyrosequencing χρησιμοποιήθηκε για πολλές αναλύσεις γραμμικού κώδικα DNA. Ο περιορισμός αυτής της προσέγγισης είναι το μειωμένο μήκος των αλληλουχιών barcode, οι οποίες κυμαίνονται από 250 έως 400 bp (Valentini et al., 2009). Αυτό το όριο έχει επιλυθεί εν μέρει χρησιμοποιώντας minibarcodes: μικρότερα τμήματα του *cox1* περίπου 150 bp (Hellberg & Morrisey, 2011; Meusnier et al., 2008; Shokralla et al., 2011), τα οποία μπορούν επίσης να ληφθούν μέσω πυροαλληλουχίας 454. Η προσέγγιση του minibarcode παρέχει αρκετές πληροφορίες για τον εντοπισμό ειδών από διαφορετικές μήτρες (Hajibabaei, Singer, Clare, & Hebert, 2007; Hajibabaei et al., 2006, 2011; Meusnier et al., 2008), καθώς και για τον προσδιορισμό του περιεχομένου των θαλασσινών προϊόντων (Becker et al., 2011; Botti & Giuffra, 2010; Rasmussen et al., 2009). Ωστόσο, το μειωμένο μήκος των ακολουθιών minibarcode μπορεί να μην είναι αρκετά ενημερωτικό για τον εντοπισμό στενά συγγενών ειδών.

4.7 Ασφάλεια και Απάτη (Fraud) στα τρόφιμα

Η ασφάλεια των τροφίμων σχετίζεται αυστηρά με τη χημική και μικροβιολογική ποιότητα της πρώτης ύλης. Άλλες σημαντικές πτυχές περιλαμβάνουν τις πρακτικές υγιεινής που υιοθετούνται κατά τις βιομηχανικές διεργασίες και τη διανομή των τελικών προϊόντων. Αρκετοί κανονισμοί, όπως ο Ευρωπαϊκός EC/178/2002 περιλαμβάνει τις κύριες παραδοχές και κανόνες που σχετίζονται με την ασφάλεια των τροφίμων γενικά και με τις αρχές που εμπλέκονται στις διαδικασίες ελέγχου των τροφίμων. Πρόσφατες μελέτες στην Ευρώπη και τη Βόρεια Αμερική ανέφεραν ότι οι εμπορικές απάτες κυμαίνονται από 15% έως 43% των συνολικών εμπορικών θαλασσινών, με το 75% των απατών στην περίπτωση λιθρινιού (*Lutjanus campechanus*, Hellberg & Morrisey, 2011; Rasmussen & Morrisey, 2008). Η μεθοδολογία barcoding DNA θα μπορούσε να χρησιμοποιηθεί για την ανακάλυψη αντικατάστασης ειδών, αποδεικνύοντας εμπορικές απάτες. Η μιτοχονδριακή περιοχή *cox1* υιοθετήθηκε σε μεγάλο βαθμό για την αναγνώριση θαλασσινών ειδών. Η διαθεσιμότητα μιας καλά οργανωμένης βάσης δεδομένων αναφοράς όπως το BOLD υποδηλώνει ότι στις μέρες μας, η γραμμική κωδικοποίηση DNA, μπορεί να χρησιμοποιηθεί ως τυπικό εργαλείο δοκιμής από ρυθμιστικούς φορείς. Σε πολλές περιπτώσεις,

οι εμπορικές απάτες σχετίζονται με ταξινομικά προβλήματα: τα είδη μπορούν να αναγνωριστούν από μια κοινή δημοτική ονομασία, η οποία μπορεί να αντιστοιχεί σε διαφορετικές ταξινομικές κατηγορίες. Σε άλλες περιπτώσεις, διαφορετικές ονομασίες συνδέονται με το ίδιο είδος σε διαφορετικές περιοχές (Burgess et al., 2005). Για να αποφευχθούν απάτες και εσφαλμένη επισήμανση, η ονομασία των βρώσιμων ειδών θα πρέπει να γράφεται μαζί με τη σωστή επιστημονική ονομασία και την αναφορά στην ακολουθία γραμμικού κώδικα DNA. Το barcoding DNA έδειξε υψηλή αποτελεσματικότητα στην αξιολόγηση της παρουσίας αλλεργιογόνων ειδών, τόσο σε νωπά όσο και σε επεξεργασμένα τρόφιμα. Οι ξηροί καρποί θεωρούνται μία από τις κύριες πηγές αλλεργιογόνων (Hubalkova & Rencova, 2011) και η παρουσία τους στα τρόφιμα (επίσης σε ίχνη) είναι ανιχνεύσιμη με μοριακή ανάλυση που βασίζεται σε διαφορετικούς δείκτες, συμπεριλαμβανομένων των περιοχών γραμμικού κώδικα DNA (π.χ. matK) (Yano et al. , 2007). Το αμύγδαλο (*Prunus dulcis*), που χρησιμοποιείται συνήθως σε πολλά προϊόντα διατροφής (αρτοποιεία, ζαχαροπλαστεία, σνακ), είναι επίσης ένα πιθανό αλλεργιογόνο. Σε αυτή την περίπτωση, η διάκριση των ιχνών DNA αμυγδάλου από άλλα ομογενή βρώσιμα είδη όπως το κεράσι (*P. cerasus*), τα δαμάσκηνα (*P. domestica* L.) και το ροδάκινο (*Prunus persica*) θα μπορούσε να είναι πρόβλημα. Ωστόσο, οι αναλύσεις που έγιναν στο πλαστιδικό γονιδίωμα αυτών των συγγενών ειδών κατέδειξαν ορισμένες διαφορές, οι οποίες μπορούν να χρησιμοποιηθούν στην ταυτοποίηση (Badenes & Parfitt, 1995). Η αναγνώριση αλλεργιογόνου υλικού είναι μια από τις πιο ενδιαφέρουσες εφαρμογές της γραμμικής κωδικοποίησης DNA. Μπορεί να χρησιμοποιηθεί για την ικανοποίηση των απαιτήσεων του FAO και της Ευρωπαϊκής Επιτροπής, οι οποίες απαριθμούν αλλεργιογόνα είδη που πρέπει να δηλώνονται στις ετικέτες των τροφίμων (Οδηγία 2003/89/ΕΚ.1). Παρόμοιες προσεγγίσεις θα μπορούσαν επίσης να εφαρμοστούν στην τροφική δυσανεξία ως συνέπεια ουσιών που υπάρχουν σε ορισμένα γένη ή είδη, όπως η γλουτένη για άτομα με κοιλιοκάκη. Πρόσφατα έχουν αναπτυχθεί μέθοδοι PCR για τον εντοπισμό της παρουσίας σίκαλης, σίτου και κριθαριού σε προϊόντα που φέρουν ετικέτα «χωρίς γλουτένη», με βάση την ανάλυση των πλαστιδικών δεικτών (π.χ. trnL) (Maskova, Paulickova, Rysova, & Gabrovska, 2012). Τα τρόφιμα που έρχονται σε αντίθεση με τον ατομικό τρόπο ζωής ή τους θρησκευτικούς κανόνες μπορούν επίσης να συμπεριληφθούν στην κατηγορία των διατροφικών απατών. Αυτή είναι η περίπτωση της προσθήκης κρέατος ή των υποπροϊόντων

του σε τρόφιμα που καταναλώνουν οι χορτοφάγοι ή η αδήλωτη χρήση χοιρινού κρέατος, η οποία απαγορεύεται από τις εβραϊκές και μουσουλμανικές θρησκείες (Ibrahim, 2008; Kesmen et al., 2007; Montiel-Sosa et al., 2000). Το barcoding DNA μπορεί να είναι ένα αποτελεσματικό εργαλείο για την ανακάλυψη αυτών των απατών. Σε όλα αυτά τα παραδείγματα, η αποτελεσματικότητα της μεθόδου γραμμικής κωδικοποίησης DNA σχετίζεται αυστηρά με την παρουσία αξιόπιστων και προσβάσιμων αλληλουχιών αναφοράς, οι οποίες μπορούν να βρεθούν μόνο σε μια αξιόπιστη βάση δεδομένων αναφοράς, που αναπτύχθηκε από κοινή προσπάθεια επιστημόνων από όλο τον κόσμο. Αυτό ισχύει ιδιαίτερα στην περίπτωση των φυτών, για τα οποία πρακτικά απουσιάζουν βάσεις δεδομένων αναφοράς ή υποπληθυσμένα.

Κεφάλαιο 5: Συμπεράσματα

Η μοριακή ταυτοποίηση έχει τεράστιες δυνατότητες να προωθήσει την κατανόησή μας για τη βιοποικιλότητα, τους οικολογικούς ρόλους και τη γεωγραφική κατανομή πολλών τροφίμων. Έχει επίσης μεγάλες δυνατότητες να επιταχύνει τη βιολογική αναζήτηση και άλλα πεδία εφαρμοσμένης έρευνας. Το barcoding DNA μπορεί να χρησιμοποιηθεί ως ένα καθολικό εργαλείο για την ιχνηλασιμότητα των τροφίμων και συγκεκριμένα η περιοχή ITS χρησιμοποιείται ως de facto τυπικός δείκτης «γραμμικής κωδικοποίησης» για πολλά χρόνια και υπάρχει κάθε λόγος να πιστεύουμε ότι θα παραμείνει τουλάχιστον ένας από τους κύριους δείκτες γραμμικής κωδικοποίησης για πολλά χρόνια ακόμη. Ακόμα κι αν, από απλή τεχνική άποψη, δεν είναι εντελώς καινοτόμο, μέσα σε λίγα μόλις χρόνια έχει γίνει ευρέως χρησιμοποιούμενο. Αυτό εξασφαλίστηκε από έναν συνδυασμό παραγόντων:

- i) το μειωμένο κόστος των μοριακών αναλύσεων.
- ii) την αυξανόμενη διαθεσιμότητα εξοπλισμένων εργαστηρίων και ειδικευμένου προσωπικού.
- iii) την παρουσία ελεύθερα διαθέσιμων πόρων που βασίζονται στο διαδίκτυο.
- iv) ο αυξανόμενος αριθμός ενημερωμένων καταναλωτών που απαιτούν υψηλά πρότυπα ποιότητας στα τρόφιμα.

Αυτό το σενάριο δημιούργησε το αίτημα για μια τεχνική που βασίζεται στη μοριακή ανάπτυξη, τυποποίηση και μηχανογράφηση. Υπό αυτή την έννοια, το barcoding DNA δεν είναι μόνο ενημερωμένο, αλλά είναι το φυσικό προϊόν της δεκαετίας του 2000. Αυτές οι περιπτωσιολογικές μελέτες και οι τεχνικές εξελίξεις δείχνουν ξεκάθαρα ότι η γραμμική κωδικοποίηση DNA είναι μια ευαίσθητη, γρήγορη, φθηνή και αξιόπιστη μέθοδος για τον εντοπισμό και την παρακολούθηση μιας ευρείας ομάδας πρώτων υλών και παραγώγων προϊόντων διατροφής και για την ανίχνευση αλλεργιογόνων ή δηλητηριωδών συστατικών που ενδέχεται να εμφανίζονται σε μήτρες τροφίμων. Όπως για παράδειγμα κατά την αναγνώριση ειδών κρέατος και πουλερικών σε μια ποικιλία επεξεργασμένων προϊόντων με ένα μόνο είδος στην ετικέτα, με εξαίρεση τα κονσερβοποιημένα τρόφιμα. Ωστόσο, οι εκκινητές μίνι-γραμμικού κώδικα δεν έχουν καλή απόδοση με αρκετά από τα είδη που έχουν δοκιμαστεί, ιδίως με κοτόπουλο, βοδινό και κρέας βίσωνα ή βουβαλιού. Ως εκ τούτου, συνιστάται μελλοντική έρευνα για την ανάπτυξη ενός μίνι σετ γραμμικού κώδικα σχετικά με τη χρήση mini-barcoding σε συνδυασμό με αλληλούχιση

επόμενης γενιάς για να καταστεί δυνατή η αναγνώριση με βάση την αλληλουχία των μιγμάτων ειδών στα τρόφιμα. Λόγω της καθολικότητάς του, το DNA barcoding μπορεί να χρησιμοποιηθεί σε διαφορετικά περιβάλλοντα και από διαφορετικούς χειριστές. Διεθνείς φορείς ή θεσμοί, που είναι υπεύθυνοι για τον ποιοτικό έλεγχο των πρώτων υλών ή των προϊόντων διατροφής, μπορούν να συνεργαστούν ανταλλάσσοντας τα δεδομένα τους, δημιουργώντας βάσεις δεδομένων αναφοράς πληθυσμού, η έλλειψη των οποίων είναι το μόνο πραγματικό όριο της μεθόδου. Στην πραγματικότητα, ενώ ορισμένες ομάδες οργανισμών (π.χ. ψάρια) εκπροσωπούνται καλά, απαιτείται πολλή δουλειά για να παρασχεθεί μια αξιόπιστη πηγή δεδομένων αναφοράς DNA barcoding για ομάδες που δεν έχουν διερευνηθεί ελάχιστα. Για το λόγο αυτό, στο εγγύς μέλλον το DNA barcoding είναι πιθανό να γίνει μια δοκιμή ρουτίνας σε πολλούς τομείς, και ειδικότερα στον ποιοτικό έλεγχο και την ιχνηλασιμότητα των τροφίμων.

Κεφάλαιο 6: Βιβλιογραφία

- Adeolu, M., Alnajjar, S., Naushad, S., & Gupta, R. S. (2016). Genome-based phylogeny and taxonomy of the 'Enterobacteriales': proposal for Enterobacterales ord. nov. divided into the families Enterobacteriaceae, Erwiniaceae fam. nov., Pectobacteriaceae fam. nov., Yersiniaceae fam. nov., Hafniaceae fam. nov., Morgane. *International Journal of Systematic and Evolutionary Microbiology*, 66(12), 5575–5599.
- Alouev, A. (2008). A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome Res.*, 1051-1063.
- Ashelford, K. E., Chuzhanova, N. A., Fry, J. C., Jones, A. J., & Weightman, A. J. (2005). At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain substantial anomalies. *Appl. Environ. Microbiol.*, 7724-7736.
- Baker, G. C., Smith, J. J., & Cowan, D. A. (2003). Review and re-analysis of domain-specific 16S primers. *J. Microbiol. Methods*, 541-555.
- Begerow, D., Nilsson, H., Unterseher, M., & Maier, W. (2010). Current state and perspectives of fungal DNA barcoding and rapid identification procedures. *Applied Microbiology and Biotechnology*, 87(1), 99–108. <https://doi.org/10.1007/s00253-010-2585-4>
- Blackwell, M. (2011). The fungi: 1, 2, 3 ... 5.1 million species? *American Journal of Botany*, 98(3), 426–438. <https://doi.org/10.3732/ajb.1000298>
- Church, D. L., Cerutti, L., Gürtler, A., Griener, T., Zelazny, A., & Emler, S. (2020). Performance and application of 16S rRNA gene cycle sequencing for routine identification of bacteria in the clinical microbiology laboratory. *Clinical Microbiology Reviews*, 33(4), e00053--19.
- Clarridge, J. E. (2004). Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. *Clin. Microbiol. Rev.*, 840-862.
- Clayton, R. A., Sutton, G., Hinkle, P. S., Bult, C., & Fields, C. (1995). Intraspecific variation in small-subunit rRNA sequences in GenBank: why single sequences may not adequately represent prokaryotic taxa. *Int. J. Syst. Bacteriol.*, 5, 595-599.

- Deidre, L. C., Cerutti, L., Gurtler, A., Griener, T., Zelazny, A., & Emler, S. (2020). Performance and Application of 16S rRNA Gene Cycle Sequencing for Routine Identification of Bacteria in the Clinical Microbiology Laboratory. *Clinical Microbiology Reviews*.
- Doolittle, W. F. (1999). Phylogenetic classification and the universal tree. *Science*, 2124-2129.
- Eid, J. (2009). Real-time DNA sequencing from single polymerase molecules. *Science*, 133-138.
- Fontana, C., Favaro, M., Pelliccioni, M., Pistoia, E. S., & Favalli, C. (2005). Use of the MicroSeq 16S rRNA gene-based sequencing for identification of bacterial isolates that commercial automated systems failed to identify correctly. *J. Clin. Microbiol.*, 615-619.
- Fox, G. E., Wisotzkey, J. D., & Jurtshuk, P. J. (2005). Use of the MicroSeq 16S rRNA gene-based sequencing for identification of bacterial isolates that commercial automated systems failed to identify correctly. *J. Clin. Microbiol.*, 615-619.
- Gazara, R. K., De Oliveira, E. A., Rodrigues, B. C., Da Fonseca, R. N., Elenir, A., Oliveira, A., & Venancio, T. M. (2019). Transcriptional landscape of soybean (*Glycine max*) embryonic axes during germination in the presence of paclobutrazol, a gibberellin biosynthesis inhibitor. *Scientific Reports*, 9.
- Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, 17, 333–351.
- Guo, J. (2008). Four-color DNA sequencing with 3'-O-modified nucleotide reversible terminators and chemically cleavable fluorescent dideoxynucleotides. *Proc. Natl Acad. Sci.*, 9145-9150.
- Hall, L., Doerr, K. A., Wohlfiel, S. L., & Roberts, G. D. (2003). Evaluation of the MicroSeq system for identification of mycobacteria by 16S ribosomal DNA sequencing and its integration into a routine clinical mycobacteriology laboratory. *J. Clin. Microbiol.*, 1447-1453.

- Harris, T. D., Buzby, P. R., Babcock, H., Beer, E., Bowers, J., Braslavsky, I., . . . Giladi, E. (2008). Single-molecule DNA sequencing of a viral genome. *Science*, 106-109.
- Iwen, P. C., Hinrichs, S. H., & Rupp, M. E. (2002). Utilization of the internal transcribed spacer regions as molecular targets to detect and identify human fungal pathogens. *Medical Mycology*, 40(1), 87–109. <https://doi.org/10.1080/mmy.40.1.87.109>
- Jabeen, A., Ahmad, N., & Raza, K. (2018). Machine Learning-Based State-of-the-Art Methods for the Classification of RNA-Seq Data. In *Classification in BioApps* (pp. 133-172).
- Jagadeesan, B., Gerner-Smidt, P., Allard, M. W., Leuillet, S., Winkler, A., Xiao, Y., . . . Grant, K. (2019). The use of next generation sequencing for improving food safety: Translation into practice. *Food Microbiology*, 79, 96-115.
- Janda, J. M., & Abbott, S. L. (2007). 16S rRNA Gene Sequencing for Bacterial Identification in the Diagnostic Laboratory: Pluses, Perils, and Pitfalls. *Journal Of Clinical Microbiology*.
- Khot, P. D., & Fredrick, D. N. (2009). PCR-based diagnosis of human fungal infections. *Expert Review of Anti-Infective Therapy*, 7(10), 1201–1221. <https://doi.org/10.1586/ERI.09.104>
- Landegren, U., Kaiser, R., Sanders, J., & Hood, L. (1988). A ligase-mediated gene detection technique. *Science*, 1077-1080.
- Leamon, J. H., Lee, W. L., Tartaro, K. R., Lanza, J. R., Sarkis, G. J., deWinter, A. D., . . . Lohman, K. L. (2003). A massively parallel PicoTiterPlate based platform for discrete picoliter-scale polymerase chain reactions. *Electrophoresis*, 3769-3777.
- Loessner, M. J. (2005). Bacteriophage endolysins - Current state of research and applications. *Current Opinion in Microbiology*, 8(4), 480–487. <https://doi.org/10.1016/j.mib.2005.06.002>
- Loman, N. J., & Pallen, M. J. (2015). Twenty years of bacterial genome sequencing. *Nat. Rev. Microbiol.*, 13(12), 787-794.
- Louca, S., Doebeli, M., & Parfrey, L. W. (2018). Correcting for 16S rRNA gene copy numbers in microbiome surveys remains an unsolved problem. *Microbiome*, 41.

- Mbareche, H., Veillette, M., Bilodeau, G., & Duchaine, C. (2020). Comparison of the performance of ITS1 and ITS2 as barcodes in amplicon-based sequencing of bioaerosols. *PeerJ*, 2020(2), 1–36. <https://doi.org/10.7717/peerj.8523>
- McCarroll, S. A., & Altshuler, D. M. (2007). Copy-number variation and association studies of human disease. *Nat. Gen.*, 37-42.
- Metzker, M. L. (2010). Sequencing technologies — the next generation. *Nat. Rev. Genet.*, 31-46.
- Mignard, S., & Fladnois, J. P. (2006). 16S rRNA sequencing in routine bacterial identification: a 30-month experiment. *J. Microbiol. Methods*, 574-581.
- Naidoo, N., Pawitan, Y., Soong, R., Cooper, D. N., & Ku, C. -S. (2011). Human genetics and genomics a decade after the release of the draft sequence of the human genome. *Hum Genomics*, 5(6), 577-622.
- Odom, O. W., Deng, H. Y., Dabbs, E. R., & Hardesty, B. (1984). Binding of S21 to the 50S subunit and the effect of the 50S subunit on nonradiative energy transfer between the 3' end of 16S RNA and S21. *Biochemistry*, 5069-5076.
- Patel, J. (2001). 6S rRNA gene sequencing for bacterial pathogen identification in the clinical laboratory. *Mol. Diagn.*, 313-321.
- Petti, C. A. (2007). Detection and identification of microorganisms by gene amplification and sequencing. *Clin. Infect. Dis.*, 44, 1108-1114.
- Petti, C. A., Bosshard, P. P., Brandt, M. E., Clarridge, J. E., Feldblyum, T. v, Foxall, P., Furtado, M. R., Pace, N., & Procop, G. (2008). Interpretive criteria for identification of bacteria and fungi by DNA target sequencing; approved guideline. *Clinical and Laboratory Standards Institute (CLSI) Documents*, 28, 11898–19087.
- Prakash, O., Sharma, A., Nimonkar, Y., & Shouche, Y. S. (2015). Proposal for creation of a new genus *Neomicrococcus* gen. nov. to accommodate *Zhihengliuella aestuarii* Baik et al. 2011 and *Micrococcus lactis* Chittipurna et al. 2011 as *Neomicrococcus aestuarii* comb.

- nov. and *Neomicrococcus lactis* comb. nov. *International Journal of Systematic and Evolutionary Microbiology*, 65(Pt\11), 3771–3776.
- Rossman, A. (2007). Report of the planning workshop for all fungi DNA Barcoding.
 - Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.*, 74(12), 5463-5467.
 - Schatz, M. C., Delcher, A. L., & Salzberg, S. L. (2010). Assembly of large genomes using second-generation sequencing. *Genome Res.*, 1165-1173.
 - Schindel, & Miller. (2005). correspondence DNA barcoding a useful tool for taxonomists. *Group*, 37012–37012.
 - Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Levesque, C. A., Chen, W., Bolchacova, E., Voigt, K., Crous, P. W., Miller, A. N., Wingfield, M. J., Aime, M. C., An, K. D., Bai, F. Y., Barreto, R. W., Begerow, D., Bergeron, M. J., Blackwell, M., ... Schindel, D. (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences of the United States of America*, 109(16), 6241–6246. <https://doi.org/10.1073/pnas.1117018109>
 - Seifert, K. (2007). Prospects for fungus identification using CO1 DNA barcodes, with *Penicillium* as a test case.
 - Shendure, J., Porreca, G. J., Reppas, N. B., Lin, X., McCutcheon, J. P., Abraham, M. R., . . . Church, G. M. (2005). Accurate multiplex polony sequencing of an evolved bacterial genome. *Science*, 1728-1732.
 - Simmon, K. E., Croft, A. C., & Petti, C. A. (2006). Application of SmartGene IDNS software to partial 16S rRNA gene sequences for a diverse group of bacteria in a clinical laboratory. *Journal of Clinical Microbiology*, 44(12), 4400–4406.
 - Sulakvelidze, Z Alavidze, J. M. J. (2001). Bacteriophage therapy, *Antimicrob.*
 - Sun, D. L., Jiang, X., Wu, Q. L., & Zhou, N. Y. (2013). Intragenomic heterogeneity of 16S rRNA genes causes overestimation of prokaryotic diversity. *Appl. Environ. Microbiol.*, 5962-5969.

- Tomkinson, A. E., Vijayakumar, S., Pascal, J. M., & Ellenberger, T. (2006). DNA ligases: structure, reaction mechanism, and function. *Chem. Rev.*, 106, 687-699.
- Voskoboynik, A. (2013). The genome sequence of the colonial chordate. *Botryllus schlosseri* eLife, 00569.
- Wani, N., & Raza, K. (2019). Integrative approaches to reconstruct regulatory networks from multi-omics data: A review of state-of-the-art methods. *Comput. Biol. Chem.*, 83.
- Woese, C. R., Gutell, R., Gupta, R., & Noller, H. F. (1983). Detailed analysis of the higher-order structure of 16S-like ribosomal ribonucleic acids. *Microbiol. Rev.*, 621-669.
- Woo, P. C. Y., Lau, S. K. P., Teng, J. L. L., Tse, H., & Yuen, K.-Y. (2008). Then and now: use of 16S rDNA gene sequencing for bacterial identification and discovery of novel bacteria in clinical microbiology laboratories. *Clinical Microbiology and Infection*, 14(10), 908–934.
- Yang, R. H., Su, J. H., Shang, J. J., Wu, Y. Y., Li, Y., Bao, D. P., & Yao, Y. J. (2018). Evaluation of the ribosomal DNA internal transcribed spacer (ITS), specifically ITS1 and ITS2, for the analysis of fungal diversity by deep sequencing. *PLoS ONE*, 13(10), 1–17. <https://doi.org/10.1371/journal.pone.0206428>