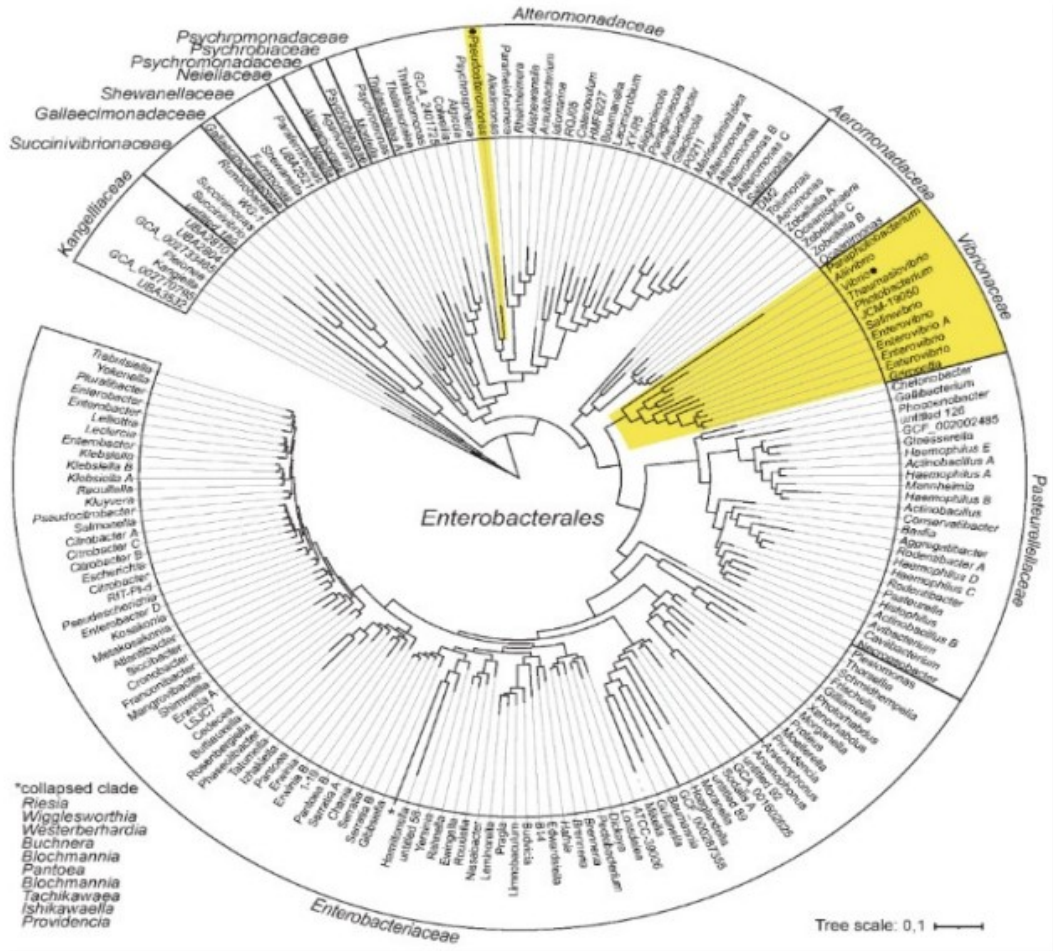


# Ανάλυση γονιδιώματος πυρήνα στα βακτήρια του γένους *Vibrio* με μεθόδους βιοπληροφορικής

## Core genome analysis of *Vibrio* genus using Bioinformatics methods



Παπαδόπουλος Οδυσσέας-Αχιλλέας του Χρήστου

Τμήμα Βιοχημείας και Βιοτεχνολογίας  
 Πανεπιστήμιο Θεσσαλίας

Η συγκεκριμένη διπλωματική εργασία πραγματοποιήθηκε και ολοκληρώθηκε στο τμήμα Βιοχημείας και Βιοτεχνολογίας του Πανεπιστημίου Θεσσαλίας στο εργαστήριο Βιοπληροφορικής.

### **Επιβλέπων καθηγητής:**

Αμούτζιας Γρηγόριος : Αναπληρωτής Καθηγητής Βιοπληροφορικής με έμφαση στην Μικροβιολογία στο τμήμα Βιοχημείας και Βιοτεχνολογίας του Σχολή επιστημών υγείας Πανεπιστημίου Θεσσαλίας

### **Τριμελής Συμβουλευτική Επιτροπή:**

Αμούτζιας Γρηγόριος : Αναπληρωτής Καθηγητής Βιοπληροφορικής με έμφαση στην Μικροβιολογία, Τμήμα Βιοχημείας Και Βιοτεχνολογίας, Σχολή Επιστημών Υγείας, Πανεπιστήμιο Θεσσαλίας

Μόσιαλος Δημήτριος : Αναπληρωτής καθηγητής Βιοτεχνολογίας Μικροβίων, Τμήμα Βιοχημείας Και Βιοτεχνολογίας, Σχολή Επιστημών Υγείας, Πανεπιστήμιο Θεσσαλίας

Ηλιόπουλος Ιωάννης : Αναπληρωτής καθηγητής Βιοπληροφορικής/Μοριακής Βιολογίας Ιατρική Σχολή Πανεπιστημίου Κρήτης

## **ΕΥΧΑΡΙΣΤΙΕΣ**

Η εκπόνηση και η ολοκλήρωση της παρούσας πτυχιακής εργασίας δεν θα μπορούσε να επιτευχθεί χωρίς την συμμετοχή καταλυτικών προσώπων και παραγόντων που με την αμέριστη συμπαράσταση τους με βοήθησαν να φτάσω στο παρόν αποτέλεσμα, με πρώτο και βασικότερο το Πανεπιστήμιο Θεσσαλίας και ειδικότερα το τμήμα Βιοχημείας και Βιοτεχνολογίας για την παροχή των υποδομών και των τεχνικών μέσων.

Πρωταρχικό ρόλο τόσο στην προπαρασκευαστική εκπαιδευτική περίοδο ενσωμάτωσης μου στο θέμα της εργασίας αλλά και στο δύσκολο πρακτικό σκέλος της κατείχε ο επιβλέπων καθηγητής κύριος Αμούτζιας Γρηγόριος, ο οποίος με την συνεχή του στήριξη και την προσφορά του μέσα στο διάστημα παραγωγής ενός ποιοτικού έργου όχι μόνο με βοήθησε να φέρω εις πέρας τους ερευνητικούς στόχους αλλά υπήρξε και πρόσωπο μεγάλης υπομονής, γενναιοδωρίας ανιδιοτέλειας και ανεκτικότητας απέναντι σε προσωπικές δύσκολες στιγμές και για αυτό τον ευχαριστώ από καρδιάς. Την αμέσως πιο καταλυτική ιδιότητα είχε ο πανταχού παρών βοηθός στο έργο αυτό διδακτορικός φοιτητής Νικολαΐδης Μάριος, που αμέτρητες φορές παρείχε την γνώση πάντα με φιλική και πρόσχαρη διάθεση δίνοντας καίριας σημασίας λύσεις και δημιουργώντας ένα γόνιμο κλίμα συνεργασίας για την διεξαγωγή και ολοκλήρωση της πτυχιακής μου και τον ευχαριστώ θερμά. Τέλος θα ήθελα να ευχαριστήσω τα μέλη της Τριμελούς επιτροπής που αφιερώνουν το χρόνο για να αξιολογήσουν και να τιμήσουν με την παρουσία τους την εργασία μου προσφέροντας την συγκεντρωμένη γνώση τους δίνοντας την ευκαιρία για κριτική και βελτίωση του έργου μου.

## ΠΕΡΙΛΗΨΗ

Το γένος *Vibrio* περιλαμβάνει βακτηριακά είδη τα οποία, κατέχουν σημαντικό ρόλο στην ισορροπία κυρίως υδροφόρων οικοσυστημάτων ενώ ταυτόχρονα χαρακτηρίζονται από υψηλή επικινδυνότητα στην δημόσια υγεία. Με τον δημοφιλέστερο αντιπρόσωπο του (*Vibrio Cholerae*) να ευθύνεται για μια από τις μεγαλύτερες πανδημίες, το ενδιαφέρον γύρω από την φυσιολογία, την διαβίωση, την μολυσματικότητα αλλά και την εξελικτική διαδρομή του συγκεκριμένου γένους είναι εκτεταμένο. Σε αυτήν την μελέτη συγκεντρώθηκαν 530 διαφορετικά βακτηριακά στελέχη του γένους *Vibrio* στα οποία πραγματοποιήθηκε φυλογενετική και συγκριτική γονιδιωματική ανάλυση ανακτώντας το γονιδίωμα πυρήνα, το παν-γονιδίωμα και τα χαρακτηριστικά γονίδια του κάθε είδους ονομαζόμενα ως fingerprints. Τα αποτελέσματα της ανάλυσης έδειξαν ότι 399 από τα εξεταζόμενα στελέχη κατατάσσονται σε 64 διαφορετικά είδη και ανακτήθηκε το πρωτέωμα-πυρήνας το οποίο αποτελείται από 174 ορθόλογες πρωτεΐνες. Στα 15 από αυτά τα είδη υπολογίσθηκε ο αριθμός των ορθόλογων γονιδίων που διατηρούνται εντός του κάθε είδους καθώς ο αριθμός των μοναδικών πρωτεϊνικών αποτυπωμάτων, fingerprint, που φέρει αποκλειστικά κάθε είδος.

## **ABSTRACT**

The genus *Vibrio* includes bacterial species which play an important role in the balance of mainly water-bearing ecosystems while at the same time being characterized by a high risk to public health. With its most popular representative (*Vibrio Cholerae*) being responsible for one of the largest pandemics, interest in the physiology, habitat, infectivity and evolutionary path of this genus is extensive. In this study, 530 different bacterial strains of the genus *Vibrio* were collected and phylogenetic and comparative genomic analysis was performed by recovering the core genome, the pan-genome and the unique characteristic genes of each species called fingerprints. The results of the analysis showed that 399 of the examined strains belong to 64 different species and the core proteome was recovered which consists of 174 orthologous proteins. In 15 of these species, the number of orthologous genes conserved within each species was calculated as well as the number of unique protein fingerprints, exclusively carried by each species.

## Table of Contents

1. ΕΙΣΑΓΩΓΗ.....	7
1.1. Παν-γονιδίωμα και γονιδίωμα-πυρήνας (Pan- genome and core-genome).....	7
1.1.1. Παν-γονιδίωμα.....	7
1.1.2. Γονιδίωμα-πυρήνας (Core-genome).....	9
1.2. Διαχωρισμός των βακτηριακών ειδών.....	10
1.2.1. Το γονίδιο της 16S rRNA ριβοσωμικής υπομονάδας.....	11
1.2.2. Η τεχνική MLST.....	12
1.2.3. Η τεχνική MLSA.....	13
1.2.4. Η τεχνική DDH.....	13
1.2.5. Η μέθοδος ANI.....	14
1.3. Το γένος <i>Vibrio</i> .....	17
1.3.1. Γενικά χαρακτηριστικά.....	17
1.3.2. Ταξινόμηση του γένους <i>Vibrio</i> .....	18
1.3.3. Το γονιδίωμα του γένους.....	19
1.3.4. Είδη του γένους <i>Vibrio</i> ως παθογόνα στον άνθρωπο.....	21
1.3.4.1. Η νόσος της χολέρας.....	21
1.3.4.2. Βιμπρίωση (vibriosis) και το <i>V. vulnificus</i> .....	26
1.3.5. Προηγούμενες αναλύσεις του γένους.....	28
2. ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΙ.....	32
2.1. ΛΟΓΙΣΜΙΚΟ.....	32
2.1.1. Linux-Ubuntu 20.04 LTS.....	32
2.1.2. PERL5.....	32
2.1.3. Python 3.6.....	32
2.1.4. Λογισμικό εύρεσης γονιδιώματος πυρήνα.....	33
2.1.5. IQ-Tree2.....	33
2.1.6. FastANI.....	34
2.1.7. Markov Cluster Algorithm (MCL).....	34
2.1.8. Interactive Tree of Life (iTOL).....	34
2.1.9. eggNOG-mapper v2.....	34
2.2. NCBI.....	35
3. Αποτελέσματα-Συζήτηση.....	36
4. Βιβλιογραφία.....	52

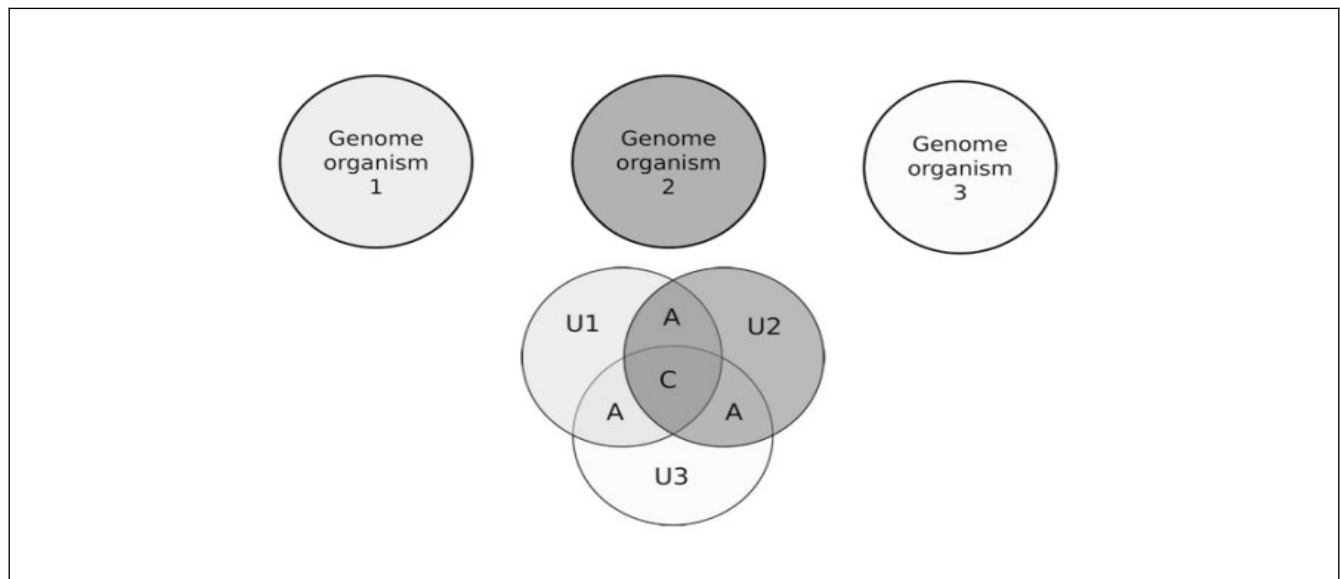
# 1. ΕΙΣΑΓΩΓΗ

## 1.1. Παν-γονιδίωμα και γονιδίωμα-πυρήνας (Pan- genome and core-genome).

### 1.1.1. Παν-γονιδίωμα

Σε έρευνα του 2005 μια ερευνητική ομάδα επιστημόνων μελετώντας παθογόνα στελέχη του *Streptococcus agalactiae* όρισε ως παν-γονιδίωμα το σύνολο όλων των γονιδίων που μπορεί να βρεθούν σε ένα συγκεκριμένο γένος μεταξύ των ειδών που το αποτελούν (Tettelin et al., 2005). Η έννοια αυτή προσδίδει σημαντικά χαρακτηριστικά που αφορούν την γενετική ποικιλομορφία ενός γένους και πως αυτή επηρεάζεται από την γονιδιακή δεξαμενή που προκύπτει όταν αθροιστούν όλες οι γενετικές πληροφορίες που φέρουν εξελικτικά παρόμοιες ομάδες οργανισμών. Μέσα στο γενετικό αυτό σύνολο περιλαμβάνονται : γονίδια που είναι συντηρημένα και εμφανίζονται σε όλους τους εκπροσώπους ενός γένους, γνωστά ως γονιδίωμα-πυρήνα (core-genome), γονίδια τα οποία είναι συντηρημένα και εμφανίζονται μεταξύ κάποιων ειδών αλλά όχι όλου του γένους, συμπληρωματικό γονιδίωμα(accessory genome) αλλά και γονίδια τα οποία εμφανίζονται αποκλειστικά σε ένα είδος και σε κανένα άλλο εντός του γένους, γόνιδιακά αποτυπώματα (fingerprints).

Κατανοώντας καλύτερα τους όρους αυτούς ας φανταστούμε το παν-γονιδίωμα ως το υπερσύνολο που περιέχει το γονιδίωμα πυρήνα το accessory genome και τα γονίδια fingerprints. Στην εικόνα 1 διακρίνεται μια απεικόνιση διαγράμματος Venn που εξηγεί αυτούς τους όρους.

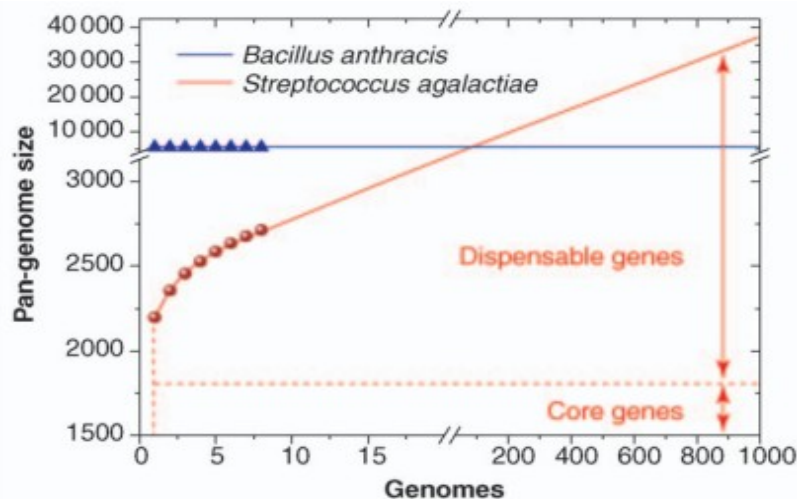


**Εικόνα 1 :** Έστω τρεις διαφορετικοί οργανισμοί ο καθένας με το δικό του γονιδίωμα το οποίο αναπαρίσταται με την μορφή ενός τέλειου κύκλου (Genome organism 1, 2, 3). Η τομή και των

τριών αυτών γονιδιωμάτων περιλαμβάνει το γονιδίωμα-πυρήνα, core-genome (C). Οι τομές που συμβαίνουν μεταξύ δύο εκ των τριών κύκλων συνεπώς και γονιδιωμάτων αποτελούν το συμπληρωματικό γονιδίωμα, accessory-genome (A). Ενώ τα τμήματα των κύκλων τα οποία δεν εμφανίζουν κανένα κοινό επίπεδο μεταξύ των υπολοίπων αποτελούν τα γονιδιακά αποτυπώματα, fingerprints (U1, U2, U3) (Bosi et al., 2015).

Το παν-γονιδίωμα διακρίνεται σε δυο βασικές μορφές που αφορούν την ιδιότητα του να αλλάζει σημαντικά το μέγεθος του ανάλογα τον αριθμό των ξεχωριστών γονιδιωμάτων που προσθέτονται κάθε φορά. Έτσι μπορούμε να διακρίνουμε το παν-γονιδίωμα ως κλειστό η ανοιχτό. Κλειστό θεωρείται ένα παν-γονιδίωμα όπου το μέγεθος του δεν αλλάζει σημαντικά και είναι σχετικά απίθανο να προκύψουν νέα γονίδια από την αλληλούχιση παραπάνω διαφορετικών στελεχών. Ενώ ανοιχτό θεωρείται ένα παν-γονιδίωμα όταν η προσθήκη καινούργιων γονιδιωμάτων αλλάζει δραματικά το μέγεθος του και έτσι καθιστά πιο πιθανή την εύρεση νέων γονιδίων (Bosi et al., 2015). Το παν-γονιδίωμα δεν συναντάται μόνο σε επίπεδο γένους μπορεί να βρεθεί και ως όρος που περιγράφει το σύνολο των γονιδίων μεταξύ και ενός είδους, το οποίο θα περιλαμβάνει διαφορετικά γονιδιώματα από διαφορετικά στελέχη του είδους (εικόνα.2).

Το αν ένα παν-γονιδίωμα είναι κλειστό η ανοιχτό έχει τεράστια σημασία για την εξέλιξη της υπό εξέταση φυλογενετικής ομάδας καθώς παν-γονιδιώματα ανοιχτά υποδηλώνουν ότι υπάρχει μεγαλύτερη γενετική ποικιλομορφία, η οποία πολύ πιθανόν συνήθως βασίζεται στους εξελικτικούς μηχανισμούς που συμβαίνουν εντός αυτών των ομάδων (γονιδιακοί διπλασιασμοί ομόλογοι ανασυνδυασμοί κτλ.). Επομένως το παν-γονιδίωμα είναι ένας δείκτης της προέλευσης και της εξέλιξης των υπό διερεύνηση ταξινομικών ομάδων.



**Εικόνα 2 :** Σύγκριση μεταξύ των παν-γονιδιωμάτων δύο διαφορετικών ειδών, των *Bacillus anthracis* (μπλε τρίγωνα) και του *Streptococcus agalactiae* (κόκκινες κουκίδες). Οι καμπύλες αποτελούν μια μαθηματική συσχέτιση ανάμεσα στο μέγεθος του παν γονιδιώματος και των



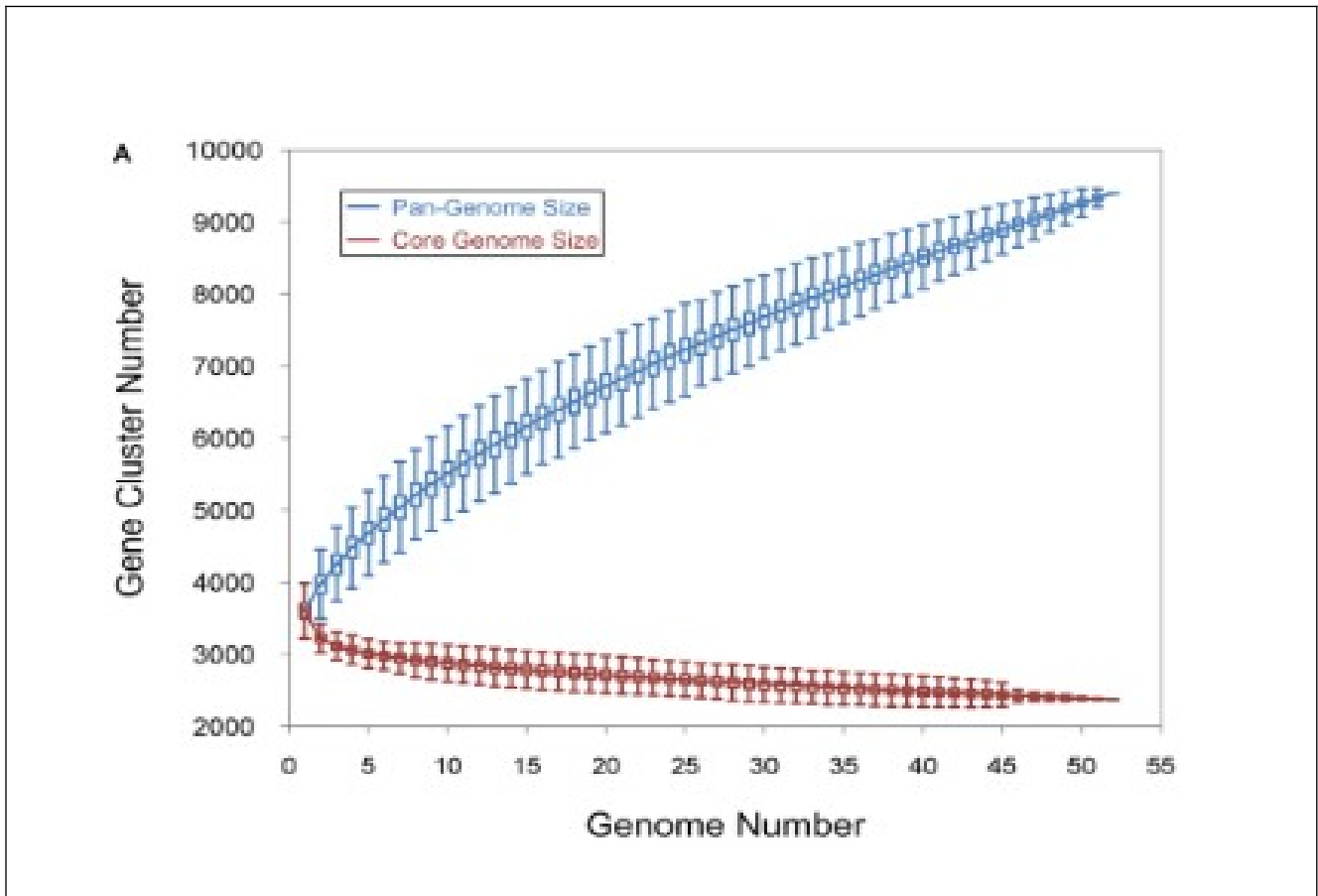
αριθμών των γονιδιωμάτων που έχουν αλληλουχισθεί. Το μέγεθος του παν-γονιδιώματος μπορεί να μεγαλώνει με την αλληλούχιση νέων γονιδιωμάτων δηλαδή να είναι ανοιχτό (*Streptococcus agalactiae*), ή να φτάνει ένα πλατό στο οποίο το μέγεθος του παραμένει αμετάβλητο με την πρόσθετη ενσωμάτωση νέων ακολουθιών δηλαδή να είναι κλειστό (*Bacillus Anthracis*). Ο αριθμός των νέων γονιδίων που βρίσκονται μεταξύ των διάφορων στελεχών (dispensable genes) μεγαλώνει σε ένα ανοιχτό παν-γονιδίωμα ενώ ο αριθμός των γονιδίων στο γονιδίωμα πυρήνα παραμένει σταθερός (Medini et al., 2005).

### 1.1.2. Γονιδίωμα-πυρήνας (Core-genome)

Όπως ήδη έχει περιγραφεί ο όρος γονιδίωμα πυρήνας αναφέρεται στο σύνολο των γονιδίων που είναι κοινά σε όλα τα μέλη μιας ταξινομικής ομάδας (πχ γένος). Μια αναβαθμισμένη διατύπωση του ορισμού θα αναφερόταν στο σύνολο των ορθόλογων γονιδίων. Εύλογα λοιπόν το γονιδίωμα-πυρήνας αποτελεί ένα κριτήριο το οποίο καθορίζει την γενετική και συνεπώς μοριακή ταυτότητα των μελών του και συγχρόνως το διαφοροποιεί από υπόλοιπες ομάδες μικροοργανισμών. Επιπλέον το σύνολο των συντηρημένων γονιδίων φαίνεται να αποτελεί σημαντικό δείκτη για την διαβίωση του είδους που τα φέρει καθώς πολύ συχνά είναι υπεύθυνα για τον μεταβολισμό και τις κυτταρικές λειτουργίες των μικροοργανισμών. Άρα η εμβάθυνση και η ανάλυση του γονιδιώματος-πυρήνα μπορεί να προσφέρει νέες ικανότητες στον διαχωρισμό των ειδών και ταυτόχρονα να συμβάλλει στον ποιοτικότερο χαρακτηρισμό των λειτουργιών κάθε ταξινομικής ομάδας.

Σημαντικό ρόλο στο μέγεθος του γονιδιώματος πυρήνα αποτελεί ο αριθμός των διαφορετικών αντιπροσώπων που κάθε φορά συνεισφέρουν στην δημιουργία της κοινής αυτή γονιδιακής δεξαμενής αλλά εξίσου σημαντικό ρόλο παίζει και ο αριθμός των γονιδίων που φέρει το κάθε μέλος ξεχωριστά. Έτσι ο αριθμός των κοινών γονιδίων πρώτα πρώτα θα εξαρτάται άμεσα από το πλήθος των πρωτεϊνών που υπάρχουν κατά μέσο όρο σε ένα είδος, με αυτό να σημαίνει ότι η προσθήκη ενός στελέχους που έχει σημαντικά μειωμένο αριθμό πρωτεϊνών από το μέσο όρο θα περιορίσει σημαντικά και το μέγεθος του γονιδιώματος πυρήνα.

Για αυτό και είναι σημαντικό σε τέτοιες αναλύσεις να περιλαμβάνονται οργανισμοί που έχουν αλληλουχισθεί σωστά και που έχουν μικρές διαφορές στον αριθμό των γονιδίων τους. Ταυτόχρονα διαπιστώνεται ότι καθώς προστίθενται παραπάνω μέλη στην κατασκευή του γονιδιώματος-πυρήνα το μέγεθος του θα φθίνει και εν τέλει μετά από κάποιον αριθμό συγκρινόμενων γονιδιωμάτων δεν θα αλλάζει σημαντικά έως καθόλου φτάνοντας ένα πλατό. Άρα όσο αυξάνεται ο αριθμός των γονιδιωμάτων που συγκρίνονται τόσο μειώνεται και το μέγεθος του γονιδιώματος-πυρήνα (Medini et al., 2005).



**Εικόνα 3 :** Ανάλυση του γκρουπ *Bacillus pumilus* μιας ομάδας βακτηριακών ειδών του γένους των Βακίλων που περιλαμβάνει τα είδη *B. pumilus*, *B. safensis*, *B. altitudinis*, *B. xiamenensis*, *B. Zhangzhouensis* και *B. australimaris*. Διακρίνεται η σχέση του παν-γονιδιώματος και του γονιδιώματος πυρήνα με την πρόσθεση νέων γονιδιωμάτων. Σε 52 διαφορετικά στελέχη υπολογίστηκε αριθμός των ορθογωνίων γονιδίων που συντηρούνται ανάμεσα τους (γονιδίωμα πυρήνας) και ο αριθμός των συνολικών γονιδίων του φυλογενετικού αυτού γκρουπ (παν-γονιδίωμα). Παρατηρείται ότι καθώς αυξάνεται ο αριθμός των γονιδιωμάτων αυξάνεται και το μέγεθος του παν-γονιδιώματος (το οποίο χαρακτηρίζεται ως ανοικτό) που απεικονίζεται με μπλε χρώμα, ενώ το μέγεθος του γονιδιώματος πυρήνα φθίνει μέχρι να φτάσει ένα σταθερό αριθμό και απεικονίζεται με κόκκινο χρώμα (Fu et al., 2021).

## 1.2. Διαχωρισμός των βακτηριακών ειδών

Τα βακτήρια παρόλο το μικροσκοπικό τους μέγεθος τους παρουσιάζουν μεγάλες διαφορές στην κυτταρική δομή, την φυσιολογία, την λειτουργία τους και γενικότερα διαφορές στον τρόπο ανάπτυξης και επιβίωσης. Έτσι κρίθηκε αναγκαία η εύρεση ενός τρόπου ταξινόμησης και κατηγοριοποίησης τους βάσει κοινών χαρακτηριστικών. Η μικροβιολογία παρόμοια με όλους τους βιολογικούς κλάδους αποσκοπεί μέσω της ταξινόμησης και ενός

συστημικού τρόπου ομαδοποίησης να καθιερώσει ένα σύστημα που να αντικατοπτρίζει την “τάξη στην φύση” όσο το δυνατόν καλύτερο, έχοντας ως απώτερο στόχο της την περιγραφή των μηχανισμών της εξελικτικής διαδικασίας και την εξερεύνηση του μυστηρίου της προέλευσης της ζωής (Kämpfer & Glaeser, 2011).

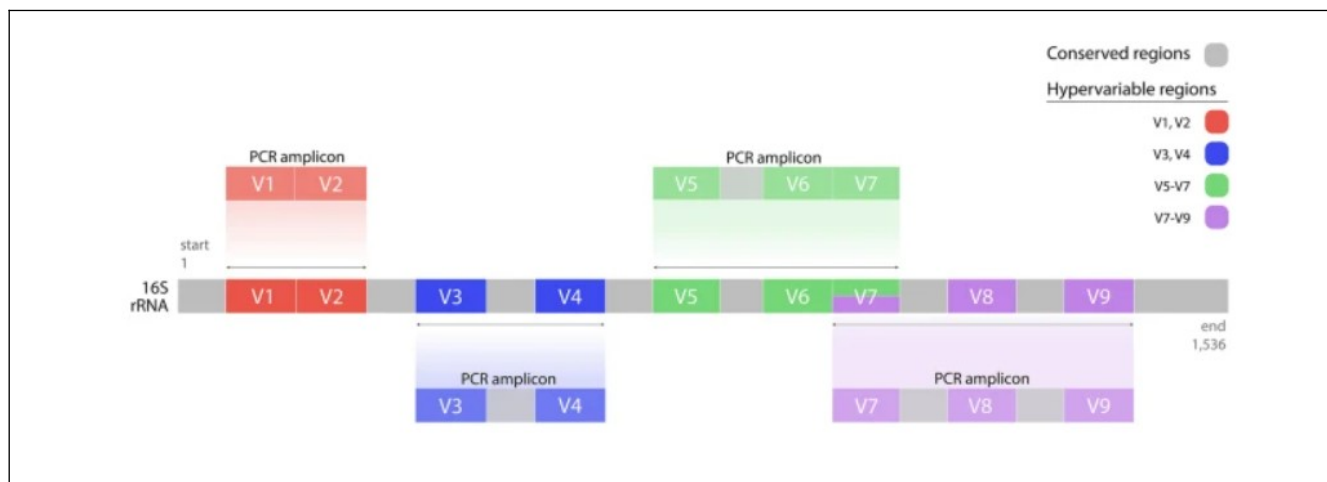
Ο αρχικός τρόπος ταξινόμησης των βακτηριακών ειδών προήλθε από τα πλεονεκτήματα που προσέφερε η ανακάλυψη του μικροσκοπίου. Το αποτέλεσμα λοιπόν της κατάταξης ήταν εξαιτίας φαινοτυπικών χαρακτηριστικών που εξαρτώνταν κυρίως από το σχήμα την κίνηση και το μέγεθος των παρατηρούμενων μικροβίων ενώ και σε άλλες περιπτώσεις εξαιτίας παθογονικού χαρακτήρα (Drews, 2000).

Ωστόσο με την πρόοδο της γενετικής και της γονιδωματικής καθώς και την παράλληλη ανάπτυξη της βιοπληροφορικής γεννήθηκαν νέα εργαλεία που χρησιμοποιούσαν πλέον και την ικανότητα κατηγοριοποίησης βάσει γονοτυπικής ομοιότητας. Η επανάσταση ήρθε με την αλληλούχιση του γονιδίου 16S rRNA .

### **1.2.1. Το γονίδιο της 16S rRNA ριβοσωμικής υπομονάδας**

Το γονίδιο 16S rRNA, κωδικοποιεί μία ριβοσωμική υπομονάδα η οποία συμμετέχει στην μεταφραστική διαδικασία των βακτηρίων και των αρχαίων. Επειδή τα γονίδια 16S rRNA είναι παρόντα σε όλα τα βακτήρια και τα αρχαία και παρουσιάζουν τόσο συντηρημένες όσο και μεταβλητές περιοχές, χρησιμεύουν ως εξαιρετικοί φυλογενετικοί δείκτες.

Με την αλληλουχία του γονιδίου 16S rRNA από έναν άγνωστο μικροοργανισμό και τη σύγκρισή του με ένα γονιδίωμα αναφοράς, οι επιστήμονες μπορούν να προσδιορίσουν την εξελικτική σχέση και την ταξινομική ομάδα στο οποίο ανήκει. Ο τρόπος με τον οποίο συνήθως γίνεται αυτή η ανάλυση είναι μέσω των σχεδιασμών εκκινητών στις συντηρημένες περιοχές και την χρήση PCR για την ταυτοποίηση της αλληλουχίας του γονιδίου ανάμεσα σε διαφορετικά είδη. Συγκρίνοντας έπειτα τις μεταβλητές περιοχές του γονιδίου αυτού μπορεί να γίνει βάσει της ομοιότητας των μεταβλητών περιοχών ανάλογη κατηγοριοποίηση. Ο όρος NI αναφέρεται στην αγγλική φράση nucleotide identity, περιγράφει την νουκλεοτιδική ταύτιση μιας αλληλουχίας, και χρησιμοποιείται για να περιγράψει πόσο όμοιες είναι δυο αλληλουχίες. Υπολογίζοντας έτσι την γενετική ταύτιση μεταξύ ακολουθιών 16S rRNA υπάρχει μια αριθμητική κρίση για την ομοιότητα των αλληλουχιών, και έτσι την οριοθέτηση ανάμεσα σε διαφορετικά είδη.



**Εικόνα 4 :** Απεικόνιση της δομής του γονιδίου 16s rRNA της *E. Coli*. Παρουσιάζονται οι συντηρημένες περιοχές οι οποίες εμφανίζονται με γκρι χρώμα, ενώ επίσης παρουσιάζονται και οι υπερ-μεταβλητές περιοχές V1, V2, V3, V4, V5, V6, V7, V8, V9.. Κάθε φορά γίνεται η ενίσχυση του τμήματος των υπερ-μεταβλητών περιοχών μέσω PCR και ακολουθεί η σύγκριση τους με ένα γονιδίωμα αναφοράς για την κατάταξη τους σε μια φυλογενετική ομάδα (López-Aladid et al., 2023).

Αν και η χρήση του γονιδίου αυτού είναι αρκετά αποτελεσματική στο διαχωρισμό ανώτερων ταξινομικών ομάδων (φύλο, γένος), όταν εφαρμόζεται σε επίπεδο ειδών πολλές φορές κρίνεται ακατάλληλη για την διαφοροποίηση στενών συγγενικών στελεχών εντός ενός γένους. Για αυτό το λόγο υπήρξε ανάγκη για την εύρεση νέων γενετικών τόπων και άλλων τεχνικών που θα βοηθούσαν στην αύξηση της διακριτικής ικανότητας στο διαχωρισμό των βακτηριακών ειδών.

### 1.2.2. Η τεχνική MLST

Μια ακόμα τεχνική η οποία εξυπηρετεί την ταξινόμηση των βακτηριακών ειδών είναι η MLST (multi locus sequence typing). Η μέθοδος αυτή βασίζεται στην νουκλεοτιδική αλληλούχιση εσωτερικών τμημάτων γονιδίων που ευθύνονται για βασικές κυτταρικές διαδικασίες (house keeping genes) και βρίσκονται σε κάθε στέλεχος από ένα συγκεκριμένο είδος. Μετά την αλληλούχιση το τμήμα του κάθε γονιδίου που έχει προκύψει συγκρίνεται με όλες τις προϋπάρχουσες ακολουθίες που έχουν κατοχυρωθεί για αυτό το τμήμα, και αποκτά έναν αριθμό αλληλομόρφου. Το σύνολο των διαφορετικών επτά αλληλόμορφων καθορίζει το αλληλικό προφίλ του κάθε στελέχους και κάθε ένα από αυτά τα προφίλ ορίζεται ως sequence type (Aanensen & Spratt, 2005). Η σχέση των διαφορετικών ειδών βασίζεται στην σύγκριση αυτών των αλληλικών προφίλ.

### 1.2.3. Η τεχνική MLSA

Μια παραλλαγή αυτής της μεθόδου η MLSA (multi-locus sequence analysis) χρησιμοποιεί ιδιοσύστατα γονίδια (house-keeping genes) όπως και η MLST με την διαφορά ότι δημιουργείται κάθε φορά και ένα φυλογενετικό δέντρο που απεικονίζει τις εξελικτικές αποστάσεις των υπό μελέτη στελεχών. Σε αυτήν την περίπτωση δημιουργείται για το κάθε στέλεχος μία υπερακολουθία η οποία περιέχει κάθε νουκλεοτιδική αλληλουχία, την μία δίπλα στην άλλη (concatenated), από τα ιδιοσύστατα γονίδια που χρησιμοποιούνται κάθε φορά (Gevers et al., 2005). Έπειτα οι υπερακολουθίες που προκύπτουν συγκρίνονται για την κατασκευή ενός δέντρου το οποίο μπορεί να διακρίνει στενά συγγενικά είδη εντός του γένους (Fang et al., 2019). Έτσι για την κατάταξη ενός άγνωστου στελέχους η διαδικασία αυτή περιλαμβάνει δύο βήματα. Αρχικά την αλληλούχιση 16s rRNA για την κατάταξη αυτού του στελέχους σε ένα γένος ή μια οικογένεια η οποία θα καθορίσει και το σετ γονιδίων που θα χρησιμοποιηθούν για την MLSA ανάλυση και την κατάταξη του ανώνυμου στελέχους σε κάποιο είδος (Glaeser & Kämpfer, 2015).

### 1.2.4. Η τεχνική DDH

Η τεχνική DDH (DNA-DNA hybridization) χρησιμοποιείται με σκοπό να διαπιστωθεί το ποσοστό νουκλεοτιδικής ταύτισης μεταξύ δύο διαφορετικών αλληλουχιών DNA (πρόκειται για μέθοδο εργαστηρίου wet lab). Το πλεονέκτημα με τις προηγούμενες μεθόδους είναι ότι προσφέρει την σύγκριση σε όλο εύρος του γονιδιώματος μεταξύ των στελεχών και όχι μόνο συγκεκριμένων γενετικών τόπων, γεγονός που την καθιστά πιο αξιόπιστη μέθοδο στο διαχωρισμό των ειδών. Βασίζεται στις αρχές της υβριδοποίησης του DNA και εξαρτάται κυρίως από το ποσοστό των αζωτούχων βάσεων γουανίνης (G) και κυτοσίνης (C) (GC content), και από το σημείο τήξεως του DNA ( $T_m$ ). Όσο πιο γενετικά όμοιες είναι δυο ακολουθίες τόσο μεγαλύτερη θερμοκρασία θα χρειάζεται για τον διαχωρισμό των δύο γενετικών αλληλουχιών DNA. Θεωρείται ύστερα από σχετικές μελέτες ότι ποσοστό ομοιότητας από 70% και πάνω ομαδοποιεί τα στελέχη στο ίδιο είδος (Wayne et al., 1987). Εκτιμάται ότι πάνω από 5000 είδη έχουν κατοχυρωθεί με την χρήση αυτής της τεχνικής (Rosselló-Móra, 2006). Παρόλο αυτά αυτή η μέθοδος δεν είναι πανάκεια καθώς όταν εφαρμόστηκε σε είδη *Rickettsias* το ποσοστό 70 % και πάνω εμφανιζόταν ανάμεσα σε στελέχη που ανήκουν σε διαφορετικά είδη *Rickettsia rickettsii*, *Rickettsia conorii*, *Rickettsia sibirica*, και *Rickettsia montanensis* (Drancourt & Raoult, 1994). Επιπλέον μειονεκτήματα της μεθόδου είναι ότι είναι ιδιαίτερα χρονοβόρα διαδικασία, εκτελείται από συγκεκριμένα μόνο εργαστήρια, είναι ακατάλληλη για την ταχεία ομαδοποίηση των προκαρυωτών, απαιτεί οι οργανισμοί να μπορούν να αναπτυχθούν σε καλλιέργειες ενώ το μεγαλύτερο μειονέκτημα το συναντάμε στην απαίτηση της μεθόδου να χρησιμοποιεί μόνο δυο γονιδιώματα κάθε φορά για σύγκριση χωρίς λοιπόν να γίνεται πολλαπλή σύγκριση σε βάσεις δεδομένων με άλλα γονιδιώματα (Mahato et al., 2017).

Η πρόοδος του κλάδου της γονιδιωματικής μέσω τεχνικών που αλληλουχούν όλο το γονιδίωμα (whole genome sequencing) έφερε επανάσταση στην εξελικτική βιολογία και ειδικότερα στην ταξινομική διάκριση των ειδών. Αυτές οι τεχνικές αλλά και οι βελτιωμένες μορφές τους τεχνολογίες NGS (next generation sequencing), απαιτούσαν την αποθήκευση μεγάλων όγκων πληροφοριών σε βάσεις δεδομένων. Το γενωμικό αυτό οπλοστάσιο αξιοποιήθηκε με την επινόηση μεθόδων *in silico* που συγκρίνουν πολλαπλά γονιδιώματα ταυτόχρονα για τον εντοπισμό της ομοιότητας τους και την επακόλουθη φυλογενετική ταξινόμηση των ειδών που τα περιέχουν (Tsai et al., 2019). Ένας όρος που χρησιμοποιείται για να περιγράψει αυτήν την ομοιότητα είναι ο όρος overall genome relatedness index (OGRI) που προτάθηκε για πρώτη φορά το 2014 (Chun & Rainey, 2014). Πλέον ήρθε η απελευθέρωση από την χρήση συγκεκριμένων γενετικών τόπων για τον διαχωρισμό των ειδών και δόθηκε η ικανότητα αξιοποίησης όλου του εύρους του γονιδιώματος για την ποιοτικότερη φυλογενετική ταξινόμηση των στελεχών. Έχουν αναπτυχθεί πολλά μετρικά συστήματα που υπολογίζουν το OGRI με την πιο διαδεδομένη να είναι η μέθοδος ANI.

### 1.2.5. Η μέθοδος ANI

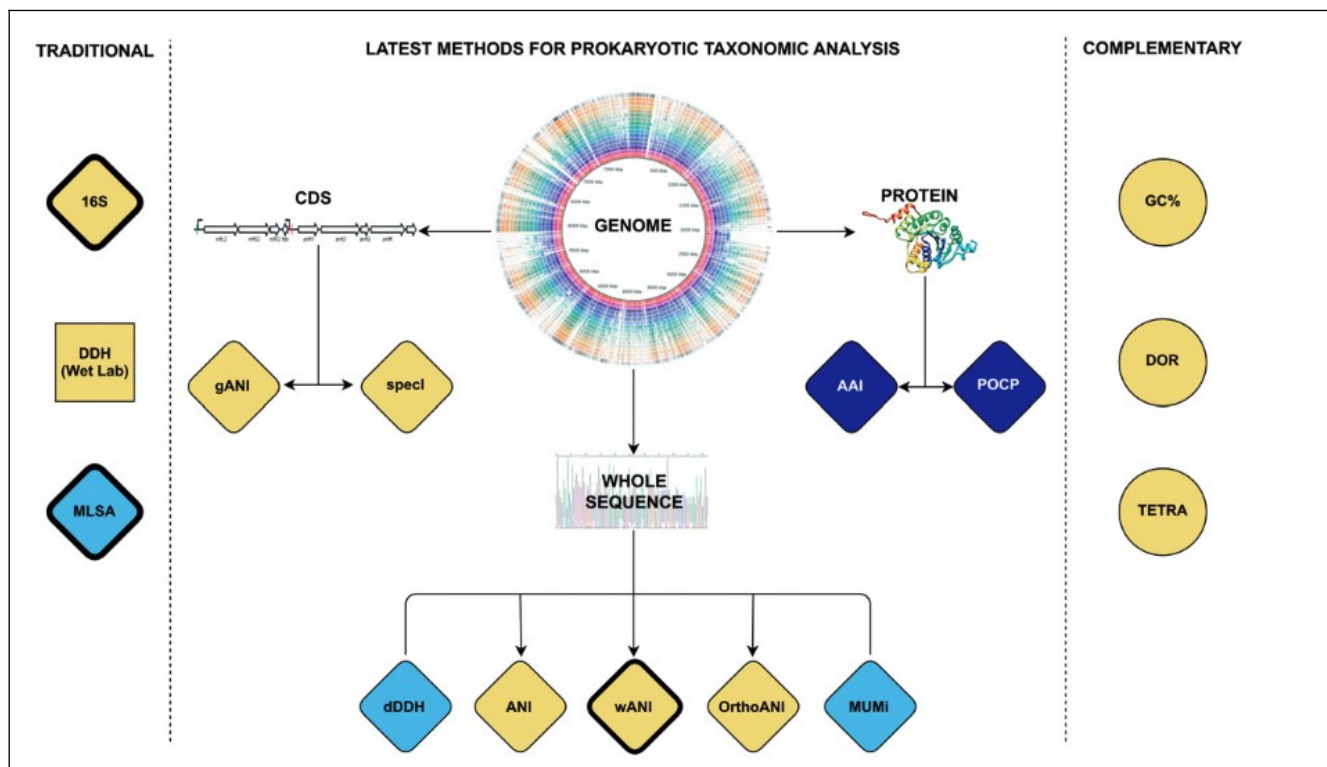
Η μέθοδος ANI (average nucleotide identity) όπως περιγράφει ο ορισμός της αναφέρεται στην μέση τιμή της νουκλεοτιδικής ταύτισης μεταξύ δύο συγκρινόμενων ακολουθιών, και υποδηλώνει πόσο όμοιες είναι δύο ακολουθίες DNA. Σε έρευνα του 2005 ανακτήθηκε με τον αλγόριθμο BLAST η μέτρηση της τιμής ANI σε όλα τα συντηρημένα γονίδια των υπό εξέταση ειδών, με σκοπό τον υπολογισμό της εξελικτικής απόστασης 70 βακτηριακών στελεχών, που είχαν ήδη ταξινομηθεί ως τότε με την χρήση προηγούμενων μεθόδων (16s rRNA, DDH) (Konstantinidis & Tiedje, 2005). Τα αποτελέσματα έδειξαν ότι, όταν υπήρχε ποσοστό υβριδοποίησης 70% και μεγαλύτερο τότε η τιμή ANI αντιστοιχούσε σε τιμές 93-94 %. Επομένως θεωρήθηκε εκ τότε ότι τιμές ANI μεγαλύτερες του 94% κατατάσσει τα υπό διερεύνηση στελέχη στο ίδιο είδος σύμφωνα με τον κανόνα υβριδοποίησης του DNA. Τα πλεονεκτήματα της ANI πέραν της απλότητας, της ταχύτητας και του χαμηλού κόστους που προσφέρει δίνει την δυνατότητα σύγκρισης πολλών γενετικών τόπων ταυτόχρονα, και για αυτό υπερέχει από τις τεχνικές που χρησιμοποιούν ένα ή μερικά μόνο γονίδια (16s rRNA, MLST) ενώ ταυτόχρονα για τον ίδιο λόγο δεν επηρεάζεται σημαντικά από εξελικτικά φαινόμενα που καθιστούν προβληματική την χρήση ενός ή περισσότερων γονιδίων όπως φαινόμενα ομόλογου ανασυνδυασμού (HGT) καθώς αυτός συμβαίνει και μεταξύ διαφορετικών ειδών και η υπό εξέταση ενός ή μερικών γονιδίων ενδέχεται να έχει παραπλανητικά αποτελέσματα στην φυλογενετική ταξινόμηση (Konstantinidis & Tiedje, 2005).

Υπάρχουν πολλές παραλλαγές της μεθόδου ANI που έχουν αναπτυχθεί με την χρήση διαφορετικών εργαλείων κάθε φορά. Για παράδειγμα, έχουν αναπτυχθεί οι τεχνικές wANI, gANI, OrthoANI κ.α.. Η μέθοδος wANI προαπαιτεί την τμηματοποίηση των γονιδιωμάτων και στην συνέχεια την στοίχιση των επιμέρους τμημάτων για την εξαγωγή της τιμής ANI χρησιμοποιώντας τον αλγόριθμο BLAST, έχει δείχθει ότι είναι ικανή τεχνική για τον διαχωρισμό στελεχών του ίδιου είδους (Goris et al., 2007). Η τεχνική gANI είναι παρόμοια της

προαναφερθείσας με την διαφορά ότι δεν απαιτείται ο κατακερματισμός του DNA και χρησιμοποιούνται οι κωδικές περιοχές του γονιδιώματος για τον υπολογισμό της γενετικής ταύτισης ενώ επίσης χρησιμοποιείται άλλος αλγόριθμος αντί του BLAST για την στοίχιση (NsimScAN) ο οποίος μειώνει τον χρόνο υπολογισμού. Η μέθοδος OrthoANI βασίζεται στον υπολογισμό της γενετικής ταύτισης χρησιμοποιώντας μόνο τις περιοχές που έχουν δώσει τα καλύτερα ανταποδοτικά αποτελέσματα όσον αφορά το ποσοστό ομοιότητας. Χρησιμοποιεί δηλαδή μόνο ορθόλογα γονίδια για την εξαγωγή της γενετικής ταύτισης και θεωρείται πιο αξιόπιστη με τις προηγούμενες τεχνικές ενώ επιπλέον προϋποθέτει την τμηματοποίηση και των δυο γονιδιωμάτων που συγκρίνονται κάθε φορά (τμήματα μήκους 1020 bp) (Lee et al., 2016b). Ωστόσο όλες αυτές οι παραλλαγές της μεθόδου αν και μπορούν να χρησιμοποιηθούν όταν εξετάζεται πολύ μεγάλος αριθμός διαφορετικών γονιδιωμάτων, δεν ενδείκνυται η χρήση τους διότι αυξάνεται εκθετικά ο χρόνος υπολογισμού με την προσθήκη περισσότερων ακολουθιών εξαιτίας της εξάρτησης των τεχνικών αυτών από την απαραίτητη στοίχιση που πρέπει να γίνει για την απόδοση των τιμών ANI (Jain et al., 2018).

Η μέθοδος FastANI δεν προαπαιτεί την στοίχιση των συγκρινόμενων αλληλουχιών και έτσι ο υπολογισμός των τιμών ANI γίνεται εξαιρετικά πιο γρήγορα σε σύγκριση με τις άλλες μεθόδους, διατηρώντας ταυτόχρονα την αξιοπιστία των αποτελεσμάτων καθώς τα αποτελέσματα με προηγούμενες μεθόδους είναι σχεδόν πανομοιότυπα. Βασίζεται στον αλγόριθμο Mashmap που χρησιμοποιεί την τεχνική των MiniHash και είναι υπεύθυνος για την απεξαρτητοποίηση της μεθόδου από την στοίχιση. Η μέθοδος αυτή βρίσκει ένα ακόμα πλεονέκτημα καθώς μπορεί να χρησιμοποιεί και μη ολοκληρωμένα γονιδιώματα για τον υπολογισμό της γενετικής ταύτισης (Jain et al., 2018).

Μια παραλλαγή της μεθόδου αποτελεί η AAI (average amino-acid identity) που αναφέρεται στο ποσοστό ομοιότητας των αμινοξέων. Σε σύγκριση με τις μεθόδους ANI η συγκεκριμένη τεχνική προσφέρει καλύτερο διαχωρισμό των στελεχών σε επίπεδο γένους, η γενικότερα σε υψηλότερες ταξινομικές ομάδες (Qin et al., 2014). Παραπλήσια αυτής της τεχνικής είναι η POCF (percentage of conserved proteins) η οποία αποτελεί το αριθμητικό αποτέλεσμα της πράξης : το άθροισμα των συντηρημένων πρωτεϊνών των συγκρινόμενων στελεχών διαιρεμένο με το άθροισμα των συνολικών πρωτεϊνών και των δύο στελεχών (Qin et al., 2014). Η τεχνική αυτή έχει χρησιμοποιηθεί για την ενοποίηση των γενών *Chlamydia* και *Chlamydophila* στο ίδιο γένος *Chlamydia* (Pannekoek et al., 2016) και για την κατηγοριοποίηση του βακτηρίου *Burkholderia andropogonis* στο γένος *Robbsia* (Lopes-Santos et al., 2017). Μια σύνοψη όλων των τεχνικών που έχουν αναφερθεί ως τώρα βρίσκεται στην εικόνα.5.



**Εικόνα 5 :** Συνοπτική απεικόνιση διάφορων ταξινομικών αναλύσεων. Στα αριστερά διακρίνονται προηγούμενες τεχνικές (16s rRNA, DDH, MLSA). Στην μέση διακρίνονται οι σύγχρονες μέθοδοι που βασίζονται είτε στις κωδικές περιοχές (CDS) όπως η gANI, specl είτε σε ολόκληρο το γονιδίωμα (dDDH, ANI, wANI, OrthoANI, MUMi που είναι παραπλήσια τεχνική της ANI) είτε στο επίπεδο πρωτεϊνών (AAI, POCP). Ενώ στα δεξιά παρουσιάζονται τα συμπληρωματικά εργαλεία που βοηθούν στο σύνολο των αναλύσεων (GC%, DOR, TETRA). Υπάρχει κατηγοριοποίηση βάσει του σχήματος όπου με τετράγωνο απεικονίζονται οι τεχνικές που απαιτούν εργαστήριο (wet lab), με ρόμβο οι τεχνικές που προαπαιτούν την στοίχιση και είναι πιο χρονοβόρες ενώ με κύκλο οι τεχνικές που δεν χρειάζονται στοίχιση.



## 1.3. Το γένος *Vibrio*

### 1.3.1. Γενικά χαρακτηριστικά

Το γένος *Vibrio* spp. είναι μια ομάδα κοινών, αρνητικών κατά Gram προαιρετικά αναερόβιων βακτηρίων με ραβδωτό σχήμα, διαθέτουν μαστίγιο και αποτελούν φυσικά συστατικά του γλυκού νερού, των εκβολών ποταμών και του θαλάσσιου περιβάλλοντος και συναντώνται σε μεγέθη 0,5 μm σε πλάτος και 1,5 με 3.0 μm σε μήκος (The Editors of Encyclopaedia Britannica, 2023). Κολυμπούν ελεύθερα σαν ατομικά κύτταρα στο νερό, ζουν σε συνδυασμό με άλλους οργανισμούς και μπορούν να συμβιώσουν με άλλους οργανισμούς. Αυτή η ποικιλία των βιολογικών δυνατοτήτων στο υδάτινο περιβάλλον είναι δυνατή μόνο επειδή τα είδη του γένους *Vibrio* είναι κινητικά βακτήρια και έχουν την ικανότητα να αισθάνονται και να επικοινωνούν χημικά. Σε επίπεδο μικροκλίμακας, η χημειοταξία διασφαλίζει τον προσανατολισμό προς το βέλτιστο ενδιαίτημα με βάση τα χημικά σήματα που εκπέμπονται από διαφορετικές θέσεις (Samraio et al., 2022).



**Εικόνα 6 :** Απεικόνιση του είδους *V. cholerae* με την χρήση ηλεκτρονικού μικροσκοπίου.

Τα είδη του γένους *Vibrio* μοιράζονται πολλά βιολογικά και γονιδιωματικά χαρακτηριστικά. Τα γονιδιώματά τους χωρίζονται σε δύο χρωμοσώματα, τα οποία έχουν

διαμορφωθεί με ανασυνδυασμό και οριζόντια μεταφορά γονιδίων (HGT, δηλαδή την απόκτηση γενετικού υλικού με μεταφορά από άλλους οργανισμούς). Αν και αυτά τα είδη μπορεί να είναι γονιδιωματικά διαφορετικά, όλα προέρχονται από υδάτινα και θαλάσσια περιβάλλοντα: προτιμούν ζεστό, υφάλμυρο (ελαφρώς αλμυρό) νερό και η αφθονία τους στο φυσικό περιβάλλον τείνει να αντικατοπτρίζει τις περιβαλλοντικές θερμοκρασίες. Μερικά από τα συμπεριλαμβανόμενα είδη αποτελούν ευκαιριακά παθογόνα θαλάσσιων οργανισμών αλλά και του ανθρώπου (Baker-Austin et al., 2018). Μάλιστα το 1854 με την επιδημία της χολέρας (υπεύθυνο είδος *Vibrio cholerae*) ο Paccini παρατηρώντας στο μικροσκόπιο δείγματα από μολυσμένους ασθενείς αντιλαμβάνεται την παρουσία μικροοργανισμών τα οποία ονομάζει *vibrions* εξαιτίας της έντονης κινητικότητας τους, δίνοντας την προέλευση του ονόματος του γένους. (Carboni, 2021).

Τα είδη *Vibrio* είναι επίσης ικανοί καθοριστές της ανακύκλωσης της οργανικής ύλης, παίζοντας σημαντικό ρόλο στην ανοργανοποίηση της λόγω των ενζύμων τους, τα οποία τους επιτρέπουν να χρησιμοποιούν μια μεγάλη ποικιλία υποστρωμάτων (Thompson & Polz, 2014). Μπορούν και υδρολύουν υδατάνθρακες, λιπίδια και πρωτεΐνες και έχουν την ικανότητα να διασπών πολυμερή όπως ζελατίνη, κολλαγόνο, άμυλο, χιτίνη, λιγνίνη και υδρογονάνθρακες, συμβάλλοντας στην ανακύκλωση άνθρακα και αζώτου σε υδάτινα περιβάλλοντα. Συγκεκριμένα σε μελέτη του 2018 εντοπίστηκε ότι τα βακτήρια του γένους *Vibrio* κατέχουν πρωταρχικό ρόλο στο κύκλο του άνθρακα στα υδάτινα οικοσυστήματα χρίζοντας έτσι όλους τους σχετικούς αντιπροσώπους υψηλής περιβαλλοντικής σημασίας (Zhang et al., 2018).

### 1.3.2. Ταξινόμηση του γένους *Vibrio*

Το γένος *Vibrio* ταξινομείται στην οικογένεια Vibrionaceae, στην τάξη Vibrionales, στην κλάση γ-πρωτεοβακτηρίων στο φύλο των πρωτεοβακτηρίων (αλλιώς γνωστό και ως pseudomonadota) και ευκόλως εννοούμενα ανήκει στο βασίλειο των Βακτηρίων.

Το φύλο των πρωτεοβακτηρίων (Proteobacteria) περιλαμβάνει μια μεγάλη ποικιλία παθογόνων γενών, όπως *Escherichia*, *Salmonella*, *Vibrio*, *Yersinia*, *Legionella* και πολλά άλλα. Κάποια από αυτά είναι ελεύθερα (μη παρασιτικά) και περιλαμβάνουν πολλά από τα βακτήρια που είναι υπεύθυνα για τη δέσμευση του αζώτου. Ο Carl Woese ίδρυσε αυτή την ομάδα το 1987, αποκαλώντας την ανεπίσημα τα «μωβ βακτήρια και τους συγγενείς τους» (Woese, 1987). Τα βακτήρια του φύλου αυτού είναι κατά κανόνα Gram-αρνητικά βακτήρια αν και μερικά έχουν χαρακτηριστεί και κατά Gram θετικά.

Η κλάση των γ-πρωτεοβακτηρίων είναι η πιο πλούσια σε είδη κλάση των προκαρυωτών και περιέχει πάνω από 250 γένη με αρκετά από αυτά να είναι ιατρικά, οικολογικά και επιστημονικά σημαντικά (“Bergey’s Manual® of Systematic Bacteriology,” 2005). Αυτοί οι μικροοργανισμοί μπορούν να ζουν σε πολλά χερσαία και θαλάσσια περιβάλλοντα, στα οποία παίζουν διάφορους σημαντικούς ρόλους, συμπεριλαμβανομένων των ακραίων περιβαλλόντων όπως οι υδροθερμικοί πίδακες. Γενικά συναντώνται με διαφορετικά σχήματα είτε ραβδωτά, ή με την μορφή κόκκων, ή σε νημάτια και περιλαμβάνουν ελεύθερα ζωντανά βακτήρια, βακτήρια που σχηματίζουν βιοφίλμ, συμβιώτες παθογόνους και μη, μερικά έχουν επίσης το χαρακτηριστικό γνώρισμα της βιοφωταύγειας (Williams et al., 2010).

Η τάξη Vibrionales περιλαμβάνει κυρίως ετερότροφα κατά Gram αρνητικά βακτήρια τα οποία έχουν σημαντικό ρόλο σε διάφορα περιβάλλοντα κυρίως συμμετέχουν στην ισορροπία των υδάτινων οικοσυστημάτων ενώ πολλά από αυτά επίσης ευθύνονται και για την εμφάνιση παθογένειας σε σπονδυλωτά και ασπόνδυλα ζώα (Lydon & Lipp, 2018).

Η Vibrionaceae είναι μια μεγάλη οικογένεια αρνητικών κατά Gram θαλάσσιων, προαιρετικών αναερόβιων βακτηρίων που ανήκουν στα γ-πρωτεοβακτήρια. Αυτή η οικογένεια περιλαμβάνει πολλά γένη από τα οποία τα μεγαλύτερα είναι τα *Vibrio* και *Photobacterium* που αποτελούν παθογόνα για τον άνθρωπο και τα ψάρια όπως το *Vibrio cholerae*, το *V. anguillarum*, και το *V. vulnificus*. Ανάμεσα τους επίσης βρίσκονται και συμβιωτικά βακτήρια με φύκια ή καλαμάρια όπως τα *V. pomeroiy*, *V. aestuarianus* και *Aliivibrio fischeri* (Takemura et al., 2014).

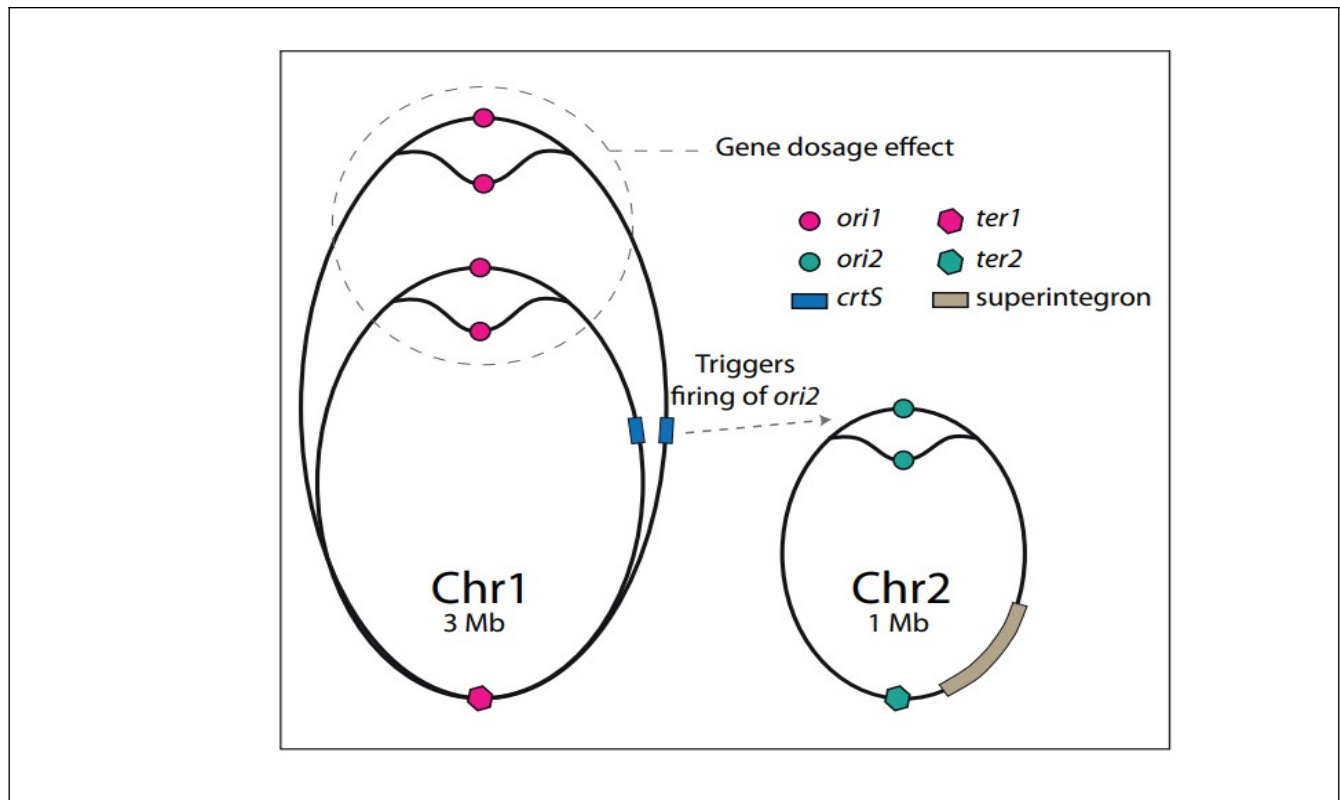
Το γένος *Vibrio* αποτελείται από 584 χαρακτηρισμένα είδη (<https://lpsn.dsmz.de/search?word=vibrio>) και αποτελεί μια από τις πολυπληθέστερες ταξινομικές ομάδες στην οικογένεια Vibrionaceae μάλιστα σε έρευνα του 2022 αναφέρεται ότι το γένος *Vibrio* έχει τα περισσότερα είδη (Sampaio et al., 2022).

### 1.3.3. Το γονιδίωμα του γένους

Αν και τα περισσότερα βακτήρια συναντώνται με ένα κυκλικό χρωμόσωμα το γονιδίωμα του γένους *Vibrio* περιέχει δύο, με το πρώτο να είναι μεγαλύτερο από το δεύτερο ενώ σε ορισμένα είδη υπάρχουν ένα ή περισσότερα πλασμίδια. Το μέγεθος του γονιδιώματος των *Vibrio* κυμαίνεται από 4,09 έως 6,32 Mb, με μέσο μέγεθος 5,14 Mb. Τα *Vibrio* έχουν μεγάλο εύρος από περιεκτικότητα σε G + C, που κυμαίνεται από 38,3% σε *V. fischeri* έως 50,7% σε *V. Furnissii* (Lin et al., 2018). Παρόλο που το συνολικό γονιδίωμα διαχωρίζεται σε δύο χρωμοσώματα έχει δείχθει ότι η αντιγραφή συμβαίνει ταυτόχρονα με το χρωμόσωμα 1 να παρέχει το κατάλληλο σήμα για την εκκίνηση της αντιγραφής στο χρωμόσωμα 2 (Escudero & Mazel, 2017).

Σε μια έρευνα η αλληλούχιση του γονιδιώματος ενός στελέχους του είδους *Vibrio cholerae* έδειξε ότι η συντριπτική πλειονότητα των αναγνωρισθέντων γονιδίων που ευθύνονται για βασικές κυτταρικές λειτουργίες (όπως αντιγραφή DNA, μεταγραφή, μετάφραση και βιοσύνθεση κυτταρικού τοιχώματος) και για την εμφάνιση παθογένειας (για παράδειγμα, γονίδια που κωδικοποιούν τοξίνες, επιφανειακά αντιγόνα) βρίσκονται στο μεγάλο χρωμόσωμα. Αντίθετα, το μικρό χρωμόσωμα περιέχει υψηλότερη αναλογία υποθετικών γονιδίων από το μεγάλο χρωμόσωμα. Το μικρό χρωμόσωμα φέρει επίσης ένα σύστημα σύλληψης γονιδίων (gene capture system) που περιέχει ιντεργκόνια μια κατηγορία μεταθετών στοιχείων και γονίδια που σχετίζονται με τον εθισμό του ξενιστή που βρίσκονται συνήθως σε πλασμίδια. Έτσι, το μικρό χρωμόσωμα μπορεί να ήταν αρχικά ένα μεγάλο πλασμίδιο που ενσωματώθηκε από ένα προγονικό είδος *Vibrio* (Heidelberg et al., 2000). Είναι σημαντικό να αναφερθεί ότι το γονιδίωμα του γένους *Vibrio* διαθέτει μια πολύ μεγάλη πλαστικότητα η οποία αντικατοπτρίζεται στα φαινοτυπικά χαρακτηριστικά του κάθε είδους με το κυριότερο την

ικανότητάς τους να προσαρμόζονται σε όλα τα υδάτινα περιβάλλοντα αλλά και να προσβάλλουν διαφορετικής φυσιολογίας ξενιστές (Escudero & Mazel, 2017). Για παράδειγμα έχουν την ικανότητα να μολύνουν ξενιστές πέραν του ανθρώπου όπως, ψάρια και θαλάσσια ασπόνδυλα.



**Εικόνα 7 :** Αναπαράσταση της δομής του γονιδιώματος του γένους *Vibrio*. Το χρωμόσωμα 1 εμφανίζεται μεγαλύτερο του χρωμοσώματος 2. Η αντιγραφή του ολόκληρου γονιδιώματος ξεκινάει στην θέση ori1 του 1ου χρωμοσώματος και σταματάει συγχρονισμένα στις θέσεις ter1 και ter2. Η αντιγραφή του χρωμοσώματος 2 πυροδοτείται από την αντιγραφή του crtS, ενός μικρού τμήματος DNA, που βρίσκεται στο χρωμόσωμα 1. Σε συνθήκες αυξημένου ρυθμού ανάπτυξης παρατηρείται στην περιοχή ori1 το φαινόμενο gene dosage effect που αντιπροσωπεύει τον αυξημένο αριθμό των γονιδιακών προϊόντων. Επίσης στο μικρό χρωμόσωμα παρατηρείται το superintegron ένα είδος μεταθετού στοιχείου που συναντάται κυρίως σε πλασμίδια, υποθέτοντας την εξελικτική του προέλευση (Escudero & Mazel, 2017)

### 1.3.4. Είδη του γένους *Vibrio* ως παθογόνα στον άνθρωπο

Οι ανθρώπινες ασθένειες που προκαλούνται από παθογόνα βακτήρια του γένους *Vibrio* μπορούν να χωριστούν σε δύο μεγάλες ομάδες: τις λοιμώξεις από χολέρα και τις λοιμώξεις που δεν προκαλούν χολέρα (cholera infections, and non-cholera). Το *V. cholerae* είναι ο αιτιολογικός παράγοντας της χολέρας, μιας σοβαρής νόσου που επιφέρει διάρροια και συνήθως προκαλείται από την κατάποση μολυσμένων τροφών ή νερού, αν και είναι επίσης δυνατή η μετάδοση από άτομο σε άτομο. Αξίζει να σημειωθεί ότι το *V. cholerae* μπορεί να βρεθεί και σε γλυκό νερό. Τα παθογόνα είδη του γένους *Vibrio* που δεν προκαλούν χολέρα., όπως τα *V. parahaemolyticus* και *V. vulnificus*, ευθύνονται για την βιμπρίωση (vibriosis), ένα είδος φλεγμονής με διαφορετικές κλινικές εκδηλώσεις ανάλογα με το είδος του παθογόνου, την οδό μόλυνσης και την ευαισθησία του ξενιστή. Για παράδειγμα, η κατάποση βακτηρίων που είναι Non-Cholera μπορεί να προκαλέσει ήπια γαστρεντερίτιδα ή πρωτογενή σηψαιμία (δηλαδή σηψαιμία μετά από κατάποση ωμής ή μη μαγειρεμένης μολυσμένης τροφής), ενώ η έκθεση των τραυμάτων του δέρματος σε μολυσμένο νερό μπορεί να προκαλέσει μόλυνση του τραύματος που μπορεί να οδηγήσει σε δευτερογενή σηψαιμία. Τα είδη *Vibrio* που δεν προκαλούν χολέρα καταλαμβάνουν ενδιαιτήματα μέτριας ή υψηλής αλατότητας και μπορούν να βρεθούν στο θαλασσινό νερό ως ελεύθερα βακτήρια ή ως συμβιώτες και παθογόνα σε θαλάσσιους οργανισμούς. Αυτά τα βακτήρια είναι τα πιο σημαντικά περιβαλλοντικά ανθρώπινα παθογόνα που προέρχονται από υδρόβιους και θαλάσσιους οικοτόπους (Baker-Austin et al., 2018). Ενώ το *V. parahaemolyticus* είναι το πιο κοινό και το *V. vulnificus* είναι η πιο θανατηφόρα αιτία βιμπρίωσης, άλλα είδη που σχετίζονται με τη θάλασσα μπορεί να προκαλέσουν βιμπρίωση και περιλαμβάνουν τα *V. alginolyticus*, *V. fluvialis*, *V. Hollisae* και *V. Metschnikovii* (Bell & Bott, 2021)

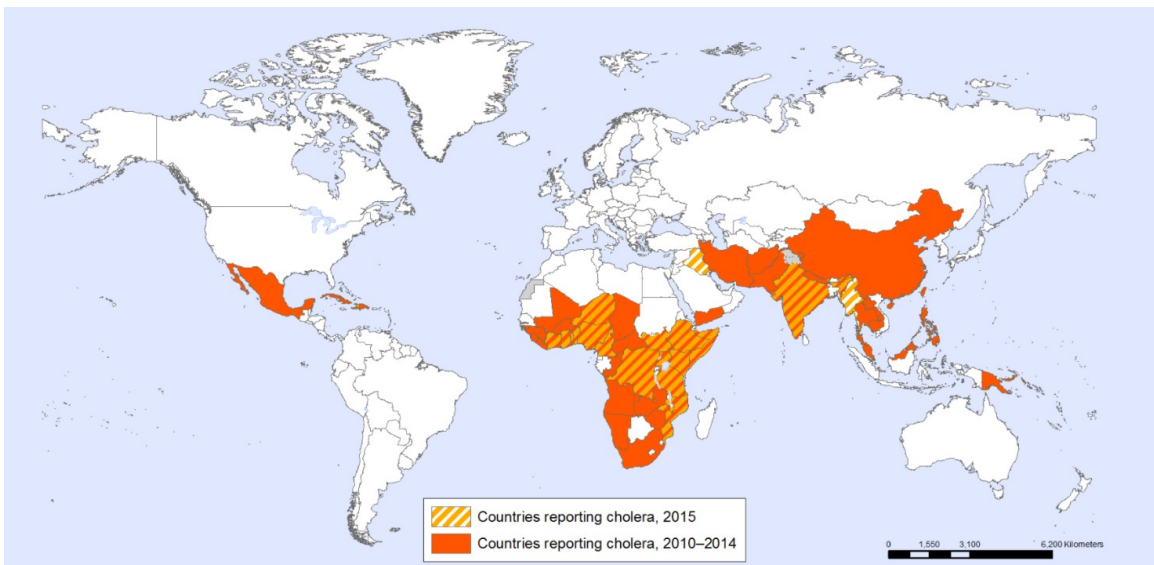
#### 1.3.4.1. Η νόσος της χολέρας

##### Επιδημιολογία

Η χολέρα είναι μια νόσος που χαρακτηρίζεται από έντονη διάρροια και προκαλείται από τα βακτήρια του είδους *Vibrio cholerae*. Χωρίς έγκαιρη θεραπεία, η χολέρα μπορεί να προκαλέσει σοβαρή αφυδάτωση μέχρι και θάνατο. Η χολέρα είναι ενδημική σε πολλές περιοχές της Αφρικής και της Ασίας, όπου εμφανίζονται εποχιακά ή σποραδικά κρούσματα, κυρίως σε χώρες με κακές συνθήκες υγιεινής, όπως η ανοιχτή απόφευση, ο ανθυγιεινός χειρισμός των τροφίμων και η περιορισμένη πρόσβαση σε ασφαλές πόσιμο νερό.

Το *Vibrio cholerae* αποτελεί μείζον κίνδυνο για τη δημόσια υγεία λόγω της δυνατότητάς του να προκαλεί πανδημίες. Από το 1817 έχουν σημειωθεί επτά πανδημίες χολέρας, με την έβδομη να ξεκινά το 1961 και να συνεχίζεται μέχρι σήμερα. Το 2015, η εκτιμώμενη ετήσια εμφάνιση της χολέρας ήταν 1,3–4 εκατομμύρια περιπτώσεις, με αποτέλεσμα 21.000–143.000 θανάτους (Ali et al., 2015). Ωστόσο, η κοινοποίηση κρουσμάτων χολέρας στον WHO (World Health Organization) δεν είναι υποχρεωτική. Ως εκ τούτου, είναι μια ασθένεια που δεν

αναφέρεται σε πολλές χώρες. Εκτιμάται ότι 3–5 εκατομμύρια άνθρωποι προσβάλλονται από χολέρα σε όλο τον κόσμο ετησίως με ~100.000 θανάτους (World Health Organization: WHO & World Health Organization: WHO, 2022). Στις ενδημικές χώρες, περίπου οι μισοί θάνατοι συμβαίνουν σε παιδιά ηλικίας <5 ετών<sup>13</sup>. Η χολέρα έχει την υψηλότερη συχνότητά της σε παιδιά ηλικίας <5 ετών. περίπου τα μισά από όλα τα κρούσματα χολέρας συμβαίνουν σε αυτήν την ηλικιακή ομάδα, αν και η εμφάνιση της ποικίλλει ετησίως, πιθανώς σχετιζόμενη με το κλίμα και τους μέχρι τώρα άγνωστους παράγοντες (Sack et al., 2004).



**Εικόνα 8 :** Χάρτης που αναπαριστά τα περιστατικά χολέρας το 2010 έως και το 2014 με πορτοκαλί ρίγες, και τα περιστατικά του 2015 με πορτοκαλί χρώμα. Διακρίνεται η έξαρση της ασθένειας της χολέρας στην περίοδο του 2015 (World Health Organization: WHO & World Health Organization: WHO, 2016)

### **Το είδος *V. cholerae* : γονιδίωμα και εξέλιξη**

Το *Vibrio cholerae* χωρίζεται σε περισσότερες από 200 οροομάδες (serotypes) που προσδιορίζονται από τη δομή του Ο-αντιγόνου του λιποπολυσακχαριδίου (LPS). Μεταξύ αυτών, ένα υποσύνολο στελεχών που ανήκουν στις οροομάδες O1 και O139 μπορεί να προκαλέσει χολέρα και επιδημίες λόγω της ικανότητάς τους να παράγουν μια τοξίνη cholera toxin (CTX). Οι οροομάδες που δεν είναι O1 και O139, που συλλογικά αναφέρονται ως non-O1/non-O139, στερούνται τυπικά την CTX και προκαλούν μικρές εστίες γαστρεντερίτιδας,

σποραδικές περιπτώσεις βακτηριαμίας και μολύνσεις τραυμάτων, αλλά δεν προκαλούν χολέρα. Σε αντίθεση με το O1 βιότυπο, περισσότερο από το 85% των μη-O1 ορομάδων (συμπεριλαμβανομένου του O139) έχουν μια κάψουλα που είναι κρίσιμη για τη λοιμογόνο δράση σε εξωεντερικές λοιμώξεις (Montero et al., 2023).

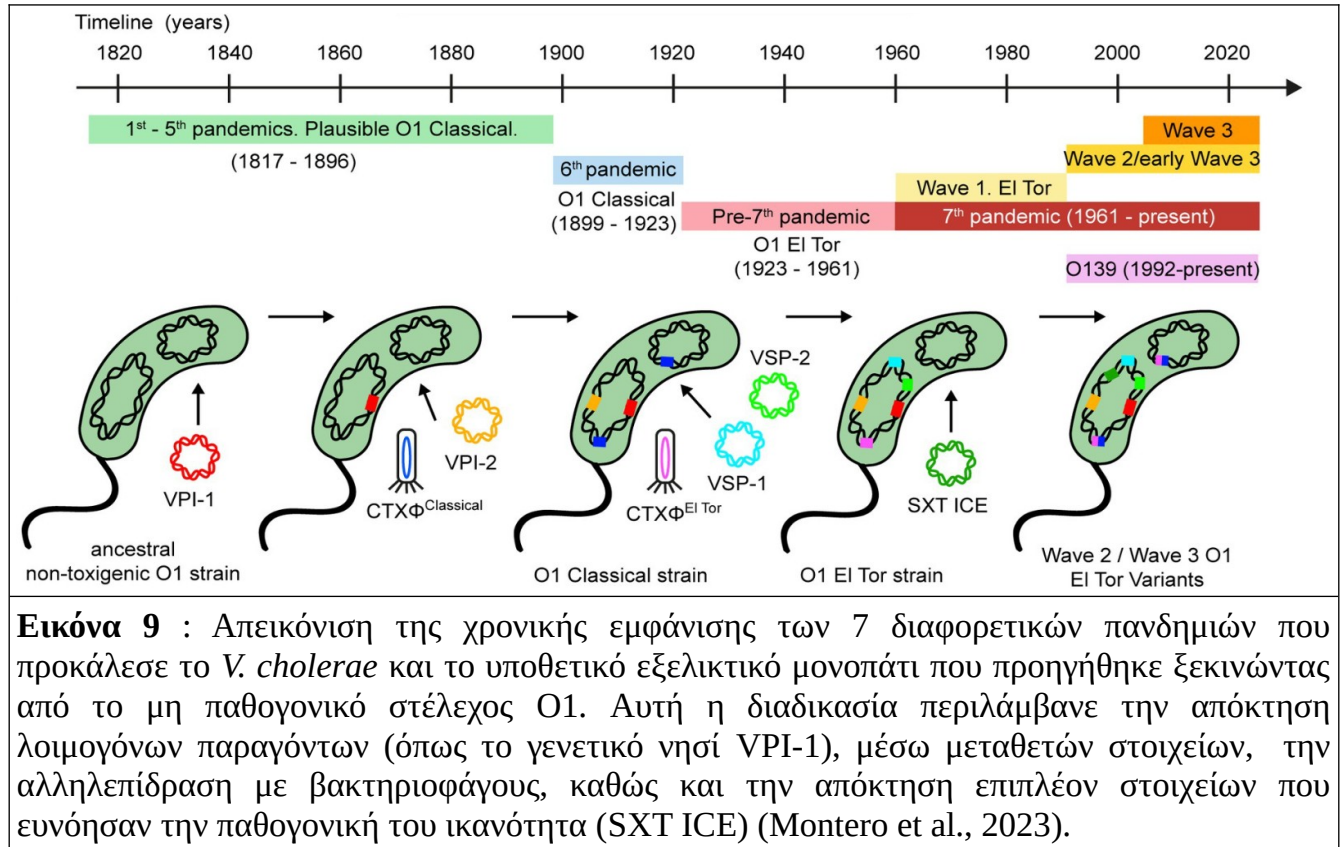
Η γονιδιωματική πλαστικότητα του *V. Cholerae* που είναι χαρακτηριστική σε όλο το γένος και η ικανότητά του να ανταλλάσσει γονίδια μέσω φυσικού μετασχηματισμού, σύζευξης και μεταγωγής είναι η εξελικτική δύναμη αυτού του βακτηρίου. Η εξέλιξή του είναι συνεχής λόγω της απόκτησης ή απώλειας γονιδιωματικών τμημάτων (Pant et al., 2020). Η απόκτηση κινητικών γενετικών στοιχείων (mobile genetic elements MGEs) είναι γνωστό ότι είναι ο κύριος μοχλός για την εξέλιξη της λοιμογόνου δράσης του *V. cholerae* και ένας καθοριστικός παράγοντας της γενετικής απόκλισης μεταξύ περιβαλλοντικών και πανδημικών στελεχών (Faruque et al., 2004). Από αυτή την άποψη, η κατανόηση των εξελικτικών γεγονότων που οδηγούν στην εμφάνιση πανδημικών στελεχών του *V. cholerae* μπορεί να προσφέρει νέες προσεγγίσεις για τον έλεγχο αυτού του παθογόνου.

Ένα υποθετικό εξελικτικό μονοπάτι για την εμφάνιση των στελεχών *V. cholerae* που οφείλεται για τις 7 πανδημίες προτάθηκε σε μελέτη του 2009 (Chun et al., 2009) και διακρίνεται στην εικόνα 9. Σύμφωνα με αυτό το μοντέλο, η διαφοροποίηση ενός κοινού προγονικού στελέχους προέκυψε μέσω της διαδοχικής απόκτησης κινητικών γενετικών στοιχείων MGE, πιθανότατα λόγω περιβαλλοντικών παραγόντων. Μετά την απόκτηση του αντιγόνου O1, ένα προγονικό στέλεχος O1 πιθανότατα απέκτησε το VPI-1 και το VPI-2, που αναφέρονται ως γενετικοί τόποι “νησιά” και είναι πανταχού παρόντα μεταξύ των στελεχών της 6ης (Κλασικός βιότυπος) και της 7ης (Βιότυπος El Tor) πανδημίας (Karaolis et al., 2001). Το νησί VPI-1 κωδικοποιεί το toxic-coregulated-pilus (TCP), το οποίο είναι ο υποδοχέας για τον βακτηριοφάγο CTXΦ. Φαίνεται λοιπόν, η μεταφορά αυτής της γενετικής πληροφορίας από το βακτηριοφάγο CTXΦ να έχει προηγηθεί και έτσι οδήγησε στην απόκτηση του VPI-1. Η απόκλιση μεταξύ των βιοτύπων Classical και El Tor οφειλόταν στην απόκτηση διαφορετικών βακτηριοφάγων CTXΦ και των γενετικών νησιών (VSP-1 και VSP-2) που βρίσκονται μόνο στα στελέχη της 6ης και της 7ης πανδημίας. Αρκετές σειρές αποδεικτικών στοιχείων το υποστηρίζουν αυτό. Για παράδειγμα, οι συγκριτικές αναλύσεις DNA έχουν αποκαλύψει ότι το CTXΦ από τους βιοτύπους Classical και El Tor περιλαμβάνει δύο διακριτές γενεές, υποδεικνύοντας ότι αποκτήθηκαν σε ανεξάρτητα γεγονότα (Boyd et al., 2000). Επιπλέον, τα VSP-1 και VSP-2 βρίσκονται σταθερά στα στελέχη O1 El Tor και O139, αλλά κυρίως απουσιάζουν στα στελέχη O1 Classical που απομονώθηκαν μεταξύ 1817 και 1923.

Επιπλέον γεγονότα οριζόντιας μεταφοράς γονιδίων είναι πιθανόν να έχουν συμβεί μεταξύ των στελεχών από την 6η πανδημία η οποία διακρίνεται σε 3 κύματα. Σε αντίθεση με τα στελέχη του 1ου κύματος (wave 1), τα στελέχη του 2ου και του 3ου κύματος (wave 2 και wave 3), περιέχουν ένα αυτομεταδιδόμενο ενοποιητικό συζευκτικό στοιχείο (self-transmissible integrative conjugative element) που φέρει πολλαπλά γονίδια ανθεκτικότητας στα αντιβιοτικά (SXT ICE). Η απόκτηση του SXT ICE πιθανότατα επηρέασε τη μετατόπιση του πληθυσμού φέρνοντας την κυριαρχία τα στελέχη υπεύθυνα για το δεύτερο και το τρίτο κύμα (Mitreja et al., 2011). Είναι ενδιαφέρον ότι τα στελέχη O139 που εμφανίστηκαν τη δεκαετία του 1990 φιλοξενούν επίσης το SXT ICE. Επιπλέον, τα στελέχη Wave 2 και Wave 3 έχουν υποστεί



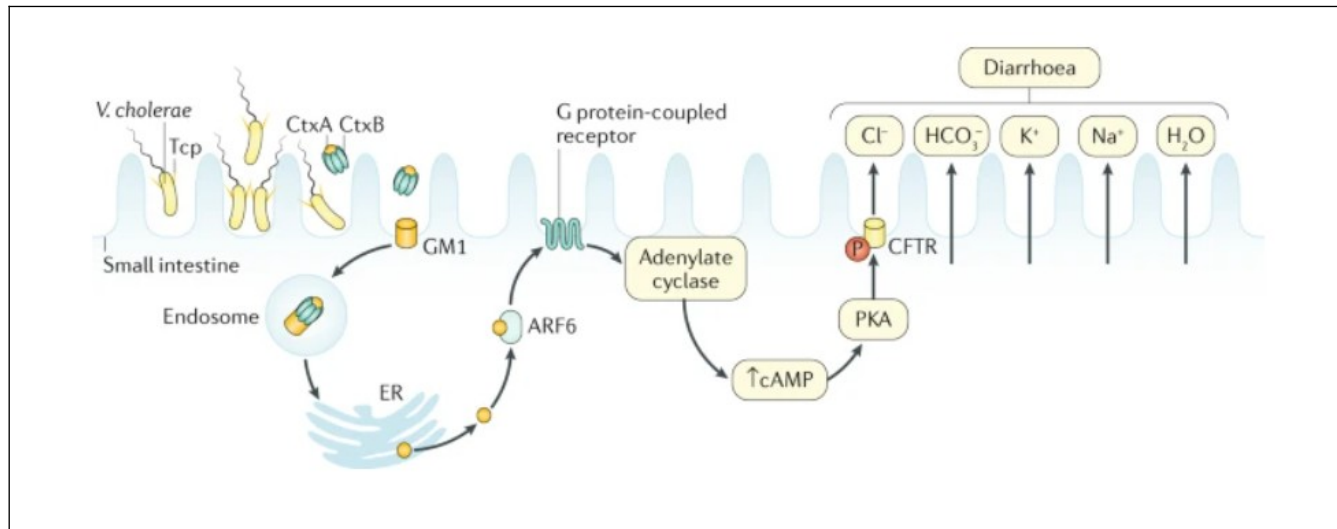
πολλαπλές αντικαταστάσεις και μεταθέσεις CTXΦ, οδηγώντας στην εμφάνιση στελεχών παραλλαγής El Tor. Όλα τα παραπάνω συνοψίζονται στην εικόνα.10.



## Μηχανισμός παθογένεσης

Το *V. cholerae* είναι ένα παράδειγμα μη επεμβατικού παθογόνου του βλεννογόνου (non-invasive pathogen mucosal) : μετά την κατάποση μολυσμένου νερού ή τροφής από τον ξενιστή, το παθογόνο πολλαπλασιάζεται σε υψηλή πυκνότητα κατά μήκος της βλεννογόνου επιφάνειας του λεπτού εντέρου, αλλά δεν διαταράσσει την ακεραιότητα του επιθηλιακού φραγμού ή προκαλεί σημαντική βλάβη στα επιθηλιακά κύτταρα (Gangarosa et al., 1960). Αντίθετα, τα βακτήρια προκαλούν μια έντονη εκκριτική απόκριση, με αποτέλεσμα άφθονη υδαρή διάρροια που εάν δεν αντιμετωπιστεί, οδηγεί συχνά σε θάνατο λόγω αφυδάτωσης εντός 1-2 ημερών. Μελέτες σε μια ποικιλία ζωικών μοντέλων και σε εθελοντές ανθρώπους έχουν δείξει ότι η χολερική διάρροια είναι κυρίως μια απόκριση σε έναν παράγοντα, την τοξίνη της χολέρας (CT). Η απουσία των *ctxA* και *ctxB* (που κωδικοποιούν τις υπομονάδες εντεροτοξίνης χολέρας A και B, αντίστοιχα) από το παθογόνο *V. cholerae* καταργεί την ικανότητα των βακτηρίων να προκαλούν διάρροια σε ζωικά μοντέλα ενώ η χορήγηση καθαρής CT είναι αρκετή για να προκαλέσει διάρροια σε εθελοντές,





**Εικόνα 10 :** Μόλις εισέλθει στον ξενιστή, αφού φτάσει στο λεπτό έντερο, το *Vibrio cholerae* αρχίζει να εκφράζει γονίδια που κωδικοποιούν παράγοντες λοιμογόνου δράσης, όπως το συν-ρυθμιζόμενο από τις τοξίνες pilus (Tcp) και την τοξίνη χολέρας. Η τοξίνη της χολέρας αποτελείται από δύο υπομονάδες, CtxA και CtxB, και συνδέεται με το γαγγλιοσίτιο (ένα σιαλυλιωμένο γλυκοσφιγγολιπίδιο) GM1 στην πλασματική μεμβράνη των εντεροκυττάρων μέσω της πενταμερούς υπομονάδας CtxB. Η δεσμευμένη τοξίνη ενδοκυτταρώνεται και στη συνέχεια μεταφέρεται στο ενδοπλασματικό δίκτυο (ER), όπου οι υπομονάδες διαχωρίζονται. Η απελευθέρωση της ενζυμικής υπομονάδας CtxA από το ER στο κυτταρόπλασμα επιτρέπει την αλλοστερική ενεργοποίησή της από τον παράγοντα ριβοσυλίωσης ADP 6 (ARF6). Η συνδεδεμένη με ARF6, ενεργοποιημένη υπομονάδα CtxA με τη σειρά της ενεργοποιεί την αδενυλική κυκλάση καταλύοντας τη ριβοζυλίωση ADP μέσω ενός GPCR υποδοχέα. Αυξημένα κυτταρικά επίπεδα κυκλικού AMP cAMP οδηγούν σε φωσφορυλίωση από την κινάση PKA, του διαμεμβρανικού υποδοχέα της κυστικής ίνωσης (CFTR), η οποία προκαλεί την εκροή ιόντων και νερού στον αυλό του λεπτού εντέρου το οποίο εν τέλει προκαλεί διάρροια (Baker-Austin et al., 2018).

### Αντίσταση του είδους σε φυσιολογικούς αμυντικούς μηχανισμούς του ξενιστή

Όταν φτάσει στο λεπτό έντερο, το κύριο σημείο αποικισμού, η προσαρμογή σε αντιμικροβιακούς παράγοντες, όπως τα χολικά άλατα και τα αντιμικροβιακά πεπτίδια, είναι απαραίτητη για την επιβίωση του. Για να επιτευχθεί αυτό, το βακτήριο ρυθμίζει το προφίλ των πρωτεϊνών της εξωτερικής μεμβράνης (OMP outer membrane protein) μέσω της ενεργοποίησης μιας στενά ρυθμισμένης οδού σηματοδότησης γνωστής ως ToxR Regulon (Ramamurthy et al., 2020). Παρουσία χολικών αλάτων, το μονοπάτι ToxR ρυθμίζει θετικά την έκφραση του OmpU και μειώνει την έκφραση του OmpT, δύο από τις πιο άφθονες εξωμεμβρανικές πρωτεΐνες (OMP) του *V. cholerae* (Childers & Klose, 2007). Η αλλαγή της σύνθεσης OMP περιλαμβάνει επίσης την αφαίρεση του OmpT με την παραγωγή κυστιδίων εξωτερικής μεμβράνης (OMV outer membrane vesicles). Το OmpU και το OmpT έχουν

διακριτές ιδιότητες όσο αναφορά τον έλεγχο κυτταρικών καναλιών: το OmpU είναι πιο επιλεκτικό ως προς τα κατιόντα από το OmpT. Το δεοξυχολικό οξύ που είναι χολικό άλας εμποδίζει τη δραστηριότητα της πορίνης OmpT αλλά όχι αυτή του OmpU. Ως εκ τούτου, το OmpU προσδίδει αντοχή στα χολικά άλατα και τα αντιμικροβιακά πεπτίδια, παίζοντας καθοριστικό ρόλο στον αποικισμό και την επιβίωση του *V. cholerae* στο λεπτό έντερο (Pennetzdorfer et al., 2021).

Για να αποικίσει με επιτυχία το λεπτό έντερο, το *V. cholerae* πρέπει να διεισδύσει σε ένα πολύ παχύρρευστο στρώμα βλέννας που έχει πάχος περίπου 100–400 μm περίπου 30–130 φορές το μέγεθος του βακτηρίου. Για αυτό, τα βακτήρια του είδους *V. cholerae* χρησιμοποιούν το μαστίγιο τους για να προωθηθούν μέσω του στρώματος της βλέννας και να φτάσουν στην επιφάνεια του επιθηλίου (Butler & Camilli, 2005). Αξίζει να σημειωθεί ότι τα μη κινητικά βακτήρια είναι λιγότερο αποτελεσματικά στον αποικισμό και πολλές φορές ανίκανα να προκαλέσουν λοίμωξη (Attridge & Rowley, 1983). Τα βακτήρια που αποτυγχάνουν να διεισδύσουν στο στρώμα της βλέννας δεν αποικίζουν τον εντερικό βλεννογόνο και αποβάλλονται στα κόπρανα λόγω της συνεχούς παραγωγής και αναπλήρωσης βλέννας.

Παράλληλα, το *V. cholerae* χρειάζεται να υπερνικήσει και τους μηχανισμούς αντίστασης της μικροχλωρίδας του εντέρου (Cho et al., 2021). Από αυτή την άποψη, η βλεννίνη (mucin) ενεργοποιεί το σύστημα έκκρισης τύπου VI *V. cholerae* (T6SS), το οποίο λειτουργεί ως μοριακή σύρριγγα που σκοτώνει βακτηριακούς ανταγωνιστές. Σε ποντίκια, το *V. cholerae* σύστημα T6SS έχει αποδειχθεί ότι επιτίθεται σε μέλη της κοινής μικροχλωρίδας του ξενιστή, διευκολύνοντας τον εντερικό αποικισμό (Zhao et al., 2018). Επιπλέον, το T6SS έχει προταθεί ως βασικός μηχανισμός που προσδίδει ανθεκτικότητα σε στελέχη πανδημίας *V. cholerae*. Ωστόσο, τα δευτερογενή χολικά οξέα που παράγονται από τη μικροχλωρίδα του εντέρου μπορούν να αναστείλουν τη συναρμολόγηση του συστήματος T6SS. Πρόσφατα, οι διαφορές στη μικροχλωρίδα του εντέρου μεταξύ των ατόμων έχουν προταθεί ως πιθανή εξήγηση για την ευαισθησία ή την αντοχή στη χολέρα (Alavi et al., 2020).

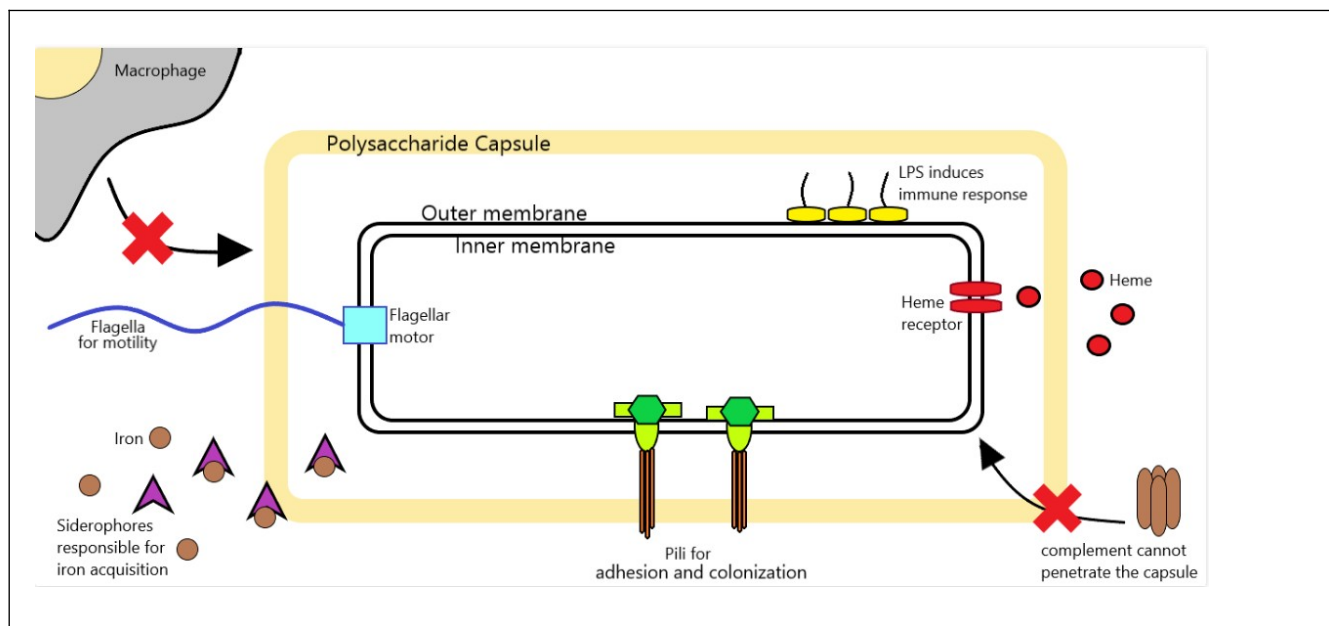
#### **1.3.4.2. Βιμπρίωση (vibriosis) και το *V. vulnificus***

Το *V. vulnificus* είναι ένα ευκαιριακό παθογόνο δηλαδή δεν εμφανίζει μόλυνση σε υγιή άτομα, εκμεταλλεύεται αδυναμίες του ξενιστή όπως η εξασθένηση του ανοσοποιητικού του συστήματος. Οι λοιμώξεις που προκαλούνται από αυτόν τον μικροοργανισμό είναι σοβαρές, συνήθως απαιτούν νοσηλεία και μπορεί να σχετίζονται τόσο με την κατανάλωση μολυσμένων θαλασσινών όσο και με την έκθεση ανοιχτών πληγών στο θαλασσινό νερό. Η κατανάλωση μολυσμένων θαλασσινών, ιδιαίτερα μαλακίων, μπορεί να οδηγήσει σε καταστάσεις που κυμαίνονται από γαστρεντερίτιδα έως σηψαιμία. Τα στρείδια είναι η κύρια τροφή που σχετίζεται με λοιμώξεις από *V. vulnificus*, λόγω της μεγάλης ποσότητας αυτών των μικροοργανισμών σε αυτό το προϊόν (Martins et al., 2022). Τα στελέχη *V. vulnificus* μπορούν να ταξινομηθούν σε τρεις βιότυπους, 1, 2 ή 3, με βάση τα βιοχημικά χαρακτηριστικά τους. Ο βιότυπος 1 σχετίζεται κυρίως με περιπτώσεις γαστρεντερίτιδας και επίσης με τις περισσότερες λοιμώξεις που προέρχονται από τραύματα. Ο βιότυπος 2, από την άλλη πλευρά, σπάνια σχετίζεται με λοιμώξεις στον άνθρωπο. Οι γονιδιωματικές συγκρίσεις υποδηλώνουν ότι ο

βιότυπος 3 είναι ένα υβρίδιο των βιοτύπων 1 και 2, που μπορεί να προκαλέσει σοβαρές λοιμώξεις στον άνθρωπο, ωστόσο με ποσοστά θνησιμότητας μικρότερα από 8% των περιπτώσεων (Heng et al., 2017).

Σε αντίθεση με το *V. cholerae*, υπάρχουν πολλά ακόμη που πρέπει να μάθουμε ως προς το ποια γονίδια είναι ζωτικής σημασίας για την ικανότητα αυτού του παθογόνου να προκαλεί τόσο γρήγορα θανατηφόρες λοιμώξεις. Ένας σημαντικός παράγοντας λοιμογόνου δράσης που έχει αποδειχθεί πειραματικά ότι είναι απαραίτητος για την επιτυχή μόλυνση είναι η βακτηριακή κάψουλα και συγκεκριμένα το λιπολυσασακχαρίδιο LPS, καθώς τα βακτήρια που δεν περιβάλλονται από αυτή φαγοκυτταρώνονται εύκολα από τα μακροφάγα. Το λιποπολυσασακχαρίδιο LPS του βακτηρίου είναι μια ενδοτοξίνη (Bahrani & Oliver, 1990). Οι βακτηριακές ενδοτοξίνες διεγείρουν τη φλεγμονή και την παραγωγή κυτοκίνης από ενεργοποιημένα μακροφάγα και Β κύτταρα, οδηγώντας σε αγγειοδιαστολή, αυξημένη διαπερατότητα τριχοειδών και ενεργοποίηση του συμπληρώματος και των οδών πήξης. Η ενδοτοξίνη του *V. vulnificus* είναι πιθανώς υπεύθυνη για την πρόκληση σημαντικής υπότασης και γενικευμένης ανεπάρκειας οργάνων που μπορεί να είναι θανατηφόρα (Park et al., 2007). Απευθείας ένεση LPS σε αρουραίους και ποντίκια είχε ως αποτέλεσμα μια δραματική πτώση της μέσης αρτηριακής πίεσης και ταχύ θάνατο των ζώων, υποδηλώνοντας το ρόλο του στην ανάπτυξη συμπτωμάτων και τη θνησιμότητα (McPherson et al., 1991).

Μόλις το βακτήριο εισέλθει στην κυκλοφορία του αίματος, το ανοσοποιητικό σύστημα ανταποκρίνεται με ένα σύνολο πρωτεϊνών που είναι γνωστές ως πρωτεΐνες συμπληρώματος. Σε άλλα παθογόνα οι πρωτεΐνες αυτές συνδέονται με το βακτήριο και το σηματοδοτούν για την εξουδετέρωση του από κύτταρα του ανοσοποιητικού γνωστό ως μακροφάγα. Αυτό στην πραγματικότητα αποφεύγεται στο *V. vulnificus* λόγω της κάψουλας που διαθέτει το πολυσασακχαρίδιο (LPS) που δεν επιτρέπει σε αυτές τις πρωτεΐνες να συνδεθούν, αποφεύγοντας έτσι την καταστροφή τους από τα μακροφάγα. Ένας άλλος ουσιαστικός παράγοντας λοιμογόνου δράσης είναι η αφομοίωση σιδήρου. Συγκεκριμένα έχει αποδειχθεί ότι ενισχύει την ανάπτυξη και θέτει σε κίνδυνο την ανοσολογική απόκριση. Αυτό επιτυγχάνεται μέσω ενός μοναδικού μηχανισμού στο *V. vulnificus* που χρησιμοποιεί μια πρωτεΐνη κωδικοποιημένη στο γονιδίωμά του που παίρνει σίδηρο από τα ερυθρά αιμοσφαίρια (Cmangiola, 2019). Το *V. vulnificus* έχει την ικανότητα να ρυθμίζει αυτούς τους παράγοντες λοιμογόνου δράσης με την αίσθηση απαρτίας, που είναι η ρύθμιση της γονιδιακής έκφρασης σε απόκριση στην κυτταρική πυκνότητα και τα σήματα κυτταρικής σηματοδότησης (Ng & Bassler, 2009). Η αίσθηση απαρτίας χρησιμοποιείται για τον έλεγχο της γονιδιακής έκφρασης πρωτεϊνών που αποκτούν σίδηρο από τα ερυθρά αιμοσφαίρια..



**Εικόνα 11** : Αναπαράσταση του κυττάρου του *V. vulnificus* με τους διάφορους λοιμογόνους παράγοντες. Το σύστημα του συμπληρώματος και τα μακροφάγα δεν έχουν καμία επίδραση στα παθογόνα κύτταρα εξαιτίας του λιποπολυσακχαριδίου LPS και της πολυσακχαρικής κάψουλας, επίσης διακρίνονται και τα σιδηροφόρα για την πρόσληψη σιδήρου από την κυκλοφορία του αίματος του ξενιστή, ο υποδοχέας Pili που είναι υπεύθυνος για την προσκόληση και τον αποικισμό του βακτηρίου, το μαστίγιο που δίνει την κινητικότητα στα συγκεκριμένα κύτταρα (Cmangiola, 2019).

### 1.3.5. Προηγούμενες αναλύσεις του γένους

Σε μελέτη του 2018, προσδιορίστηκε η αλληλουχία δύο νέων ολοκληρωμένων *Vibrio* γονιδιωμάτων, και παρουσιάστηκε μια ευρεία συγκριτική ανάλυση αυτών σε όλο το γένος με άλλα 18 πλήρη γονιδιώματα του γένους *Vibrio*, που ήταν διαθέσιμα στη βάση δεδομένων της GenBank. Το σύνολο αυτών των 18 γονιδιωμάτων περιλάμβανε το μικρό data set ενώ δημιουργήθηκε και ένα ακόμα μεγαλύτερο data set με 1582 γονιδιώματα *Vibrio* εξαιρώντας εκείνα που ήταν χαμηλής ποιότητας (Contig N50 <10.000 bp). Στην συγκεκριμένη ανάλυση ανακτήθηκε το core genome για τα 20 γονιδιώματα του γένους χρησιμοποιώντας το μικρό data set, ενώ μεταξύ άλλων χρησιμοποιήθηκε το μεγάλο data set με σκοπό να προσδιοριστούν σε ευρεία κλίμακα τα γενικότερα χαρακτηριστικά του γονιδιώματος του γένους *Vibrio* (Lin et al., 2018).

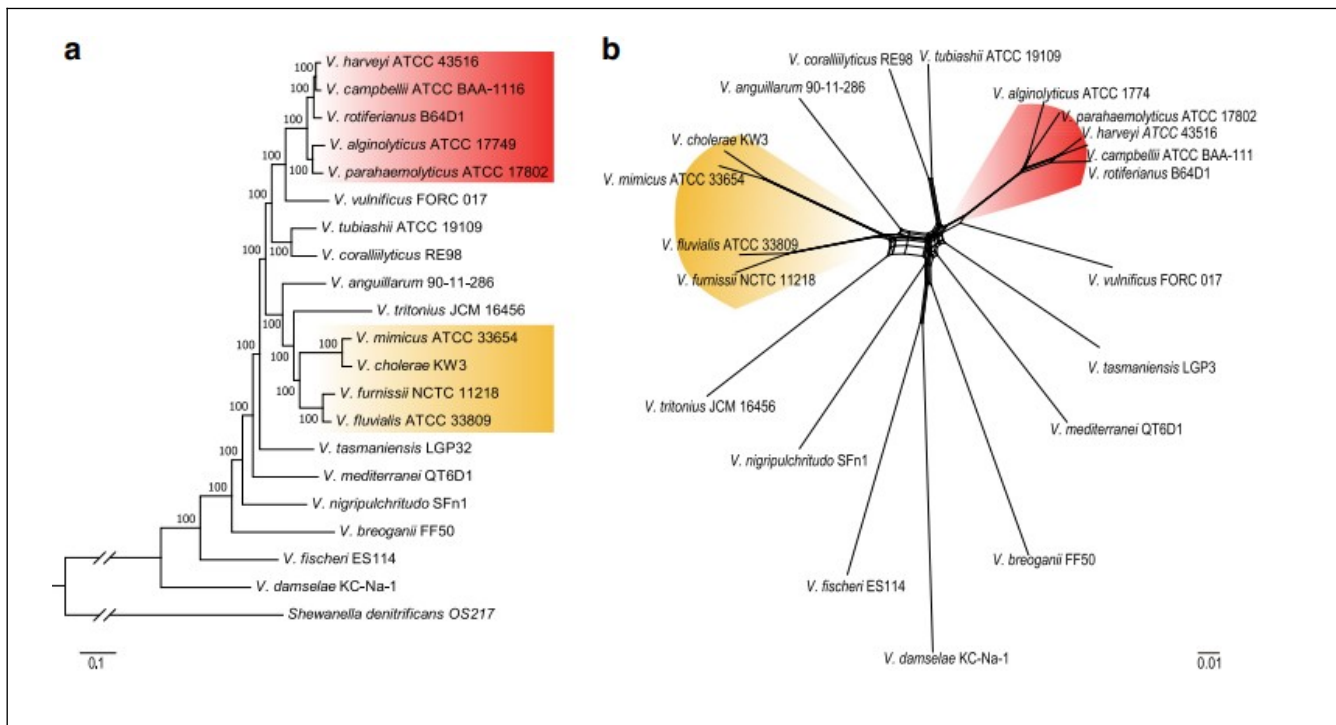
Το πρόγραμμα GET\_HOMOLOGUES v3.0.3, χρησιμοποιώντας δημοφιλείς αλγόριθμους ομαδοποίησης, bidirectional best hit (BDBH), COGtriangles και OrthoMCL, χρησιμοποιήθηκε για την ομαδοποίηση των ορθολόγων γονιδίων και την ταυτοποίηση του γονιδιώματος πυρήνα και του παν-γονιδιώματος εντός του μικρού σύνολο δεδομένων που

αναφέρεται παραπάνω με την χρήση των προεπιλεγμένων κριτηρίων. Οι ακολουθίες των ορθολόγων που ανακτήθηκαν ευθυγραμμίστηκαν και στοιχίστηκαν με το λογισμικό MAFFT χρησιμοποιώντας τον αλγόριθμο «linsi» δημιουργώντας ταυτόχρονα μια υπερακολουθία που περιλάμβανε το σύνολο των ορθολόγων. Το φυλογενετικό δέντρο κατασκευάστηκε με την χρήση του SplitsTree v4.14.6. Τα αποτελέσματα έδειξαν την ύπαρξη 1003 ορθολόγων γονιδίων (single-copy) τα οποία διατηρούνται ανάμεσα στα 20 υπό εξέταση είδη.

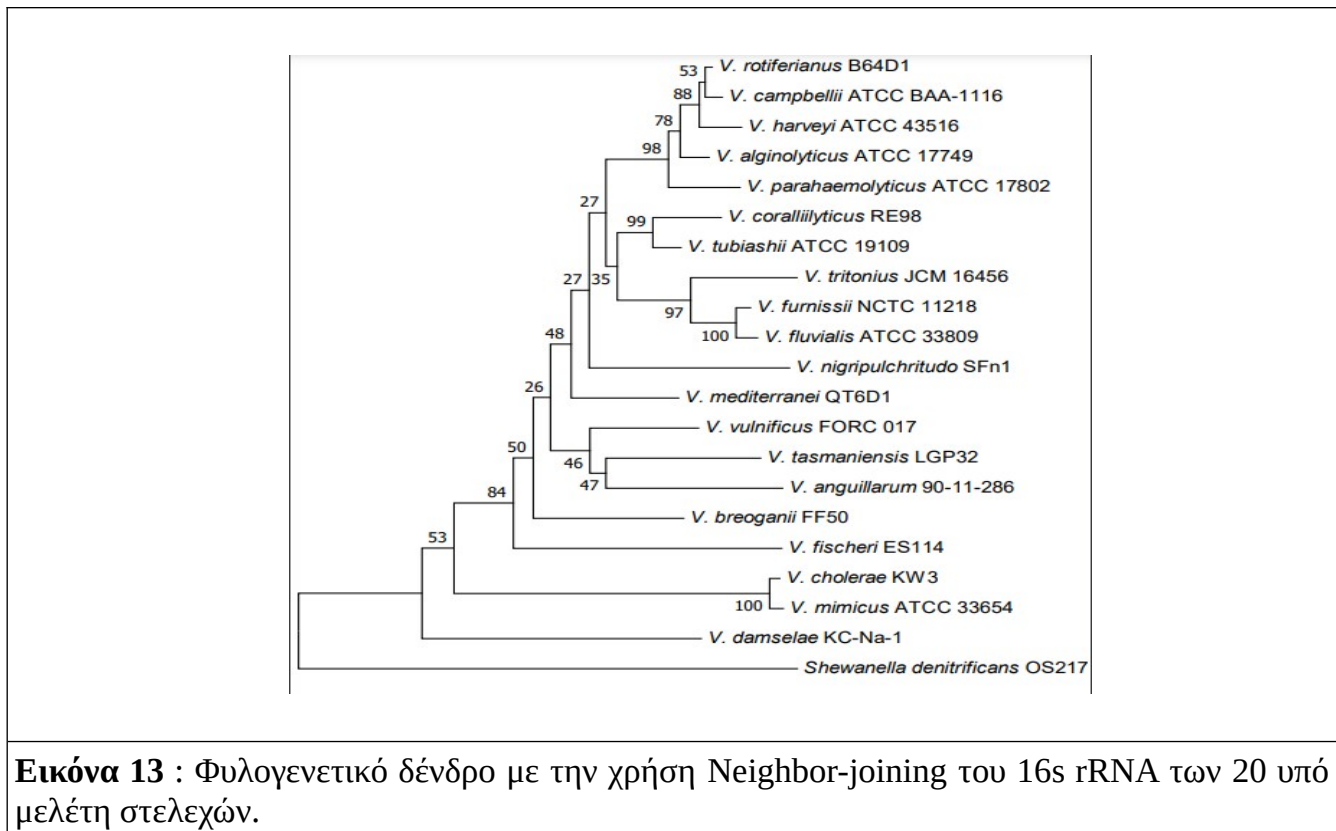
Στο δέντρο διακρίνονται δύο ξεχωριστοί κλάδοι, οι οποίοι αναγνωρίζονται ευκολότερα με την ανάλυση διαχωρισμένου δικτύου (split network analysis) (εικόνα.11). Ο κλάδος I αποτελείται από πέντε είδη: *V. harveyi*, *V. campbellii*, *V. rotiferianus*, *V. alginolyticus* και *V. parahaemolyticus*, που είναι η λεγόμενη ομάδα πυρήνα, *Vibrio core group* ή *Harveyi clade*, και Ο κλάδος II αποτελείται από τα *V. cholerae*, *V. mimicus*, *V. fluvialis* και *V. furnissii*, που σχηματίζουν το *Cholerae clade*.

Στην συγκεκριμένη ανάλυση κατασκευάστηκε και το φυλογενετικό δένδρο με την χρήση του 16s rRNA γονιδίου και τα αποτελέσματα έδειξαν ότι αυτή η ανάλυση ήταν πιο αναξιόπιστη σε σχέση με την προαναφερθείσα καθώς παρουσίαζε εξαιρετικά χαμηλές τιμές στις ρίζες των κόμβων (εικόνα 13). Αξίζει να αναφερθεί ότι τα είδη του γένους *Vibrio* φέρνουν πολλές επαναλήψεις του γονιδίου 16s rRNA το οποίο μπορεί και να εξηγεί την αναξιοπιστία αυτής της τεχνικής. Το γεγονός αυτός αποδεικνύει την σημαντικότητα της ανακάλυψης σύγχρονων τεχνικών που εξυπηρετούν την ποιοτικότερη φυλογενετική ταξινόμηση.

Στα πλαίσια αναζήτησης της γενετικής τους ομοιότητας και με σκοπό την καλύτερη κατανόηση των εξελικτικών τους αποστάσεων πραγματοποιήθηκε η σύγκριση των γονιδιωμάτων τους σε ζευγάρια με την μέθοδο ANI. Τα αποτελέσματα έδειξαν ότι οι τιμές ANI κυμαίνονταν από 70 – 88 % γεγονός που υποδεικνύει ότι κάθε ένα υπό τα υποξεταζόμενα στελέχη όντως ανήκει σε διαφορετικό είδος σύμφωνα με την προτεινόμενη βιβλιογραφία που αφορά την εφαρμογή της μεθόδου. Τα μέλη του κλάδου Clade εμφανίζουν μεγαλύτερες τιμές ANI (80-88%) σε σχέση με τον κλάδο *Cholerae* αποδεικνύοντας εξαιρετική γενετική ομοιότητα στην συγκεκριμένη ομάδα γεγονός που μπορεί να παρατηρηθεί και από το φυλογενετικό δέντρο της εικόνας 12 το οποίο παρουσιάζει στον συγκεκριμένο κλάδο πολύ μικρό μήκος των διακλαδώσεων.



**Εικόνα 12 :** α) Φυλογενετικό δέντρο βασισμένο στο γονιδίωμα-πυρήνα χρησιμοποιώντας το λογισμικό RAxML σαν εξω-ομάδα χρησιμοποιήθηκε το στέλεχος *S. denitrificans* OS217. β) φυλογενετικό δέντρο κατασκευασμένο με το λογισμικό SplitsTree4.



**Εικόνα 13 :** Φυλογενετικό δέντρο με την χρήση Neighbor-joining του 16s rRNA των 20 υπό μελέτη στελεχών.

Τέλος ένα από τα σημαντικά αποτελέσματα που αναδείχθηκαν ήταν ότι από τα στελέχη και του μικρού data set τα 18 από αυτά διέθεταν το ένζυμο χιτινάση ενώ και τα 20 διέθεταν το ένζυμο  $\beta$ -N-acetylhexosaminidase. Παράλληλα παραπάνω από το 99% των στελεχών του μεγάλου dataset διέθεταν και τα δύο από αυτά τα γονίδια γεγονός που υποδεικνύει την σημαντικότητα των υπεύθυνων γονιδίων στην φυσιολογία του γένους. Ο μεταβολισμός της χιτίνης είναι απαραίτητος τον κύκλο του άνθρακα και του αζώτου στα θαλάσσια οικοσυστήματα και το ενζυμικό οπλοστάσιο των βακτηρίων αυτών εξυπηρετεί στην παραγωγή ενέργειας που εξασφαλίζει την επιβίωση τους. Η ικανότητα των βακτηρίων του συγκεκριμένου γένους να χρησιμοποιήσουν τη χιτίνη παρέχει δυνατότητες σχηματισμού βιοφίλμ στην επιφάνεια της χιτίνης και ο σχηματισμός αυτός πιθανότητα ευθύνεται για πολλές άλλες βιολογικές λειτουργίες. Η οικογένεια λοιπόν αυτών των ενζύμων έπαιξε σημαντικό ρόλο στην εξέλιξη του γένους *Vibrio* γεγονός που υποδεικνύεται από την έντονη συντήρηση των σχετικών γονιδίων.

Σκοπός της συγκεκριμένης ανάλυσης είναι ο διαχωρισμός σε είδη του γένους *Vibrio*, η εύρεση του γονιδιώματος πυρήνα του γένους, αλλά και των επιμέρους ειδών καθώς επίσης και η ανάκτηση των πρωτεϊνών αποτυπωμάτων που εμφανίζονται σε μοναδικά είδη κάθε φορά, με σκοπό την καλύτερη κατανόηση της εξέλιξης και της προσαρμογής τους .

## 2. ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΙ

### 2.1. ΛΟΓΙΣΜΙΚΟ

#### 2.1.1. Linux-Ubuntu 20.04 LTS

Το Ubuntu είναι ένα από τα λειτουργικά συστήματα της πλατφόρμας Linux, αποτελεί το πιο πολυχρησιμοποιημένο ανάμεσα σε αυτά και είναι δωρεάν προς χρήση στο κοινό. Χρησιμοποιεί την γλώσσα BASH (Bourne Again Shell) στο περιβάλλον του τερματικού και προσφέρει αξιοπιστία στην διαχείριση αυξημένων όγκων δεδομένων ενώ είναι κατάλληλο για την υποστήριξη διάφορων εργαλείων βιοπληροφορικής (*About the Ubuntu Project | Ubuntu, n.d.*).

#### 2.1.2. PERL5

Η Γλώσσα PERL (Practical Extraction and Report Language) είναι μια γλώσσα που χρησιμοποιεί διερμηνέα (interpreted language), και είναι ιδανική για την σάρωση αρχείων κειμένου την εξαγωγή πληροφοριών από τα διάφορα κείμενα και την εκτύπωση αυτών των πληροφοριών. Τέτοιου είδους αρχεία συναντώνται συνεχώς σε βιοπληροφορικές αναλύσεις, χρίζοντας την κατάλληλη για την διαχείριση παρεμφερών τύπων δεδομένων. Είναι πρακτική εύχρηστη και αποτελεσματική ενώ ταυτόχρονα συνδυάζει τα καλύτερα χαρακτηριστικά των C, sed, awk και sh γλώσσων (Wall, 2011).

#### 2.1.3. Python 3.6

Η Python είναι μια ευέλικτη γλώσσα προγραμματισμού γνωστή για την απλότητα, την αναγνωσιμότητα και τις εκτεταμένες βιβλιοθήκες της. Έχει βρει εφαρμογές σε διάφορους τομείς, συμπεριλαμβανομένης της βιοπληροφορικής, λόγω της ευελιξίας και της ευκολίας χρήσης της. Στη βιοπληροφορική, η Python χρησιμοποιείται εκτενώς για μεθόδους όπως ανάλυση αλληλουχίας DNA, πρόβλεψη δομής πρωτεϊνών, οπτικοποίηση δεδομένων και διαχείριση βάσεων δεδομένων. Εχούν αναπτυχθεί πολυάριθμα εργαλεία βιοπληροφορικής βασισμένα στην συγκεκριμένη γλώσσα όπως η βιβλιοθήκη BioPython (Cock et al., 2009).



#### 2.1.4. Λογισμικό εύρεσης γονιδιώματος πυρήνα

Με σκοπό να προσδιοριστεί το γονιδίωμα πυρήνα χρησιμοποιήθηκε ένα εργαλείο που αναπτύχθηκε από την ερευνητική ομάδα του εργαστηρίου μας (Nikolaidis et al, 2020). Στηρίζεται στον αλγόριθμο DIAMOND, και χρησιμοποιεί 3 διαφορετικά script της γλώσσας Python. Το pipeline αποτελείται από 3 Python scripts, σχεδιασμένα να εκτελούν ανταποδοτικό DIAMOND ανά ζεύγη, με τον καθορισμό κάθε φορά ενός πρωτεύματος αναφοράς. Σε κάθε σύγκριση το πρωτέωμα αναφοράς είναι η βάση για την αναζήτηση ομόλογων στο αντίστοιχο πρωτέωμα. Στην συνέχεια πραγματοποιείται η αντίστροφη διαδικασία, δηλαδή το πρωτέωμα αναφοράς γίνεται πρωτέωμα επερώτησης, και το πρωτέωμα επερώτησης γίνεται πρωτέωμα αναφοράς. Αν σε αυτή την δυαδική σύγκριση οι πρωτεΐνες που συγκρίνονται εμφανιστούν ως οι καλύτερες ανταποδοτικές στοιχίσεις τότε θεωρούνται ορθόλογες.

Σε κάθε σύγκριση μεταξύ δύο πρωτεωμάτων υπολογίζεται το ποσοστό μέσης αμινοξικής ταύτισης (AAI average aminoacid identity) και δεν λαμβάνονται υπόψη όσα αποκλίνουν κατά δύο τυπικές αποκλίσεις από το μέσο όρο. Το επόμενο βήμα είναι η κατασκευή ενός αρχείου με την μορφή πίνακα που εμφανίζει τα καλύτερα ανταποδοτικά χτυπήματα για όλα τα πρωτεώματα επερώτησης, με τις σειρές να αντιπροσωπεύουν τις ορθόλογες πρωτεΐνες και οι στήλες τα διάφορα πρωτεώματα. Ύστερα από την εφαρμογή κατάλληλων φίλτρων στον πίνακα διατηρούνται μόνο οι πρωτεΐνες του πρωτεύματος αναφοράς που έχουν ορθόλογο στα υπόλοιπα πρωτεώματα, οι οποίες είναι εν τελει είναι και οι πρωτεΐνες που περιλαμβάνονται στο πρωτέωμα πυρήνα.

Στο τέλος με τον αλγόριθμο MUSCLE (Edgar, 2004) κάθε ομάδα ορθόλογων στοιχίζεται και ενώνεται με τις υπόλοιπες δημιουργώντας μια υπερακολουθία η οποία και φιλτράρεται με το Gblocks (Castresana, 2000). Με βάση την τελική στοίχιση που προκύπτει δημιουργείται ένα Maximum Likelihood (μέγιστης πιθανοφάνειας) φυλογενωμικό δέντρο, χρησιμοποιώντας το λογισμικό IQ tree2 (Minh et al., 2020), .

#### 2.1.5. IQ-Tree2

Με το IQ-Tree2 είναι δυνατή η κατασκευή φυλογενετικών δέντρων βασισμένη στην εκτίμηση της Μέγιστης Πιθανοφάνειας (Maximum Likelihood). Πρόκειται για ένα γρήγορο και στοχαστικό λογισμικό που επιλέγει το βέλτιστο μοντέλο φυλογενετικού δέντρου, με την δυνατότητα ταυτόχρονα επεξεργασίας από τον χρήστη (Minh et al., 2020). .

#### 2.1.6. FastANI

Το FastANI, είναι ένας αλγόριθμος που υπολογίζει την μέση τιμή νουκλεοτιδικής ταύτισης (Average Nucleotide Identity; ANI) μεταξύ γονιδιωμάτων. Οι τιμές ANI ανάμεσα σε

διαφορετικά στελέχη χρησιμοποιούνται για την ταξινομική τους ομαδοποίηση σε είδη. Ποσοστό ταύτισης άνω του 95% υποδεικνύει ότι οι οργανισμοί που συγκρίνονται ανήκουν στο ίδιο είδος. Η μέθοδος ANI είναι γρήγορη και εύχρηστη ενώ ταυτόχρονα υπερέρχει σε αξιοπιστία συγκρινόμενη με προηγούμενες μεθόδους που αποσκοπούσαν στον διαχωρισμό των ειδών όπως η αλληλούχιση του 16s Rna.

### **2.1.7. Markov Cluster Algorithm (MCL)**

Ο αλγόριθμος Markov, προτάθηκε από τον Stijn van Dongen το 2000, εφαρμόζει αρχές των στοχαστικών πινάκων και χρησιμοποιείται με σκοπό την ομαδοποίηση στοιχείων που μοιράζονται κοινά χαρακτηριστικά σε μεγάλου όγκου δεδομένα. Είναι ιδιαίτερα κατάλληλο για την ανάλυση δικτύων μεγάλης κλίμακας, όπως κοινωνικά δίκτυα, βιολογικά δίκτυα και γραφήματα ιστού (Enright A., 2002). Στην συγκεκριμένη εργασία ο αλγόριθμος MCL χρησιμοποιήθηκε με σκοπό την κατηγοριοποίηση των στελεχών σε είδη με βάση τις τιμές από το FastANI.

### **2.1.8. Interactive Tree of Life (iTOL)**

Το Interactive Tree of Life (<https://itol.embl.de>) είναι ένα διαδικτυακό εργαλείο που χρησιμοποιείται για την απεικόνιση, το χειρισμό και τον σχολιασμό δένδρων φυλογένεσης καθώς και άλλων δέντρων. Είναι δωρεάν διαθέσιμο και ανοιχτό σε όλους. Προσφέρει μεγάλο εύρος επιλογών επεξεργασίας με εύκολο χειροκίνητο τρόπο, όπως η αλλαγή των κλάδων ή των κόμβων, σε ένα φιλικό προς το χρήστη ψηφιακό περιβάλλον. Επιπλέον μια πιο πρόσφατη έκδοση το ItoI v5 παρουσιάζει σημαντική βελτίωση με αναβαθμισμένα γραφικά και λειτουργίες καθιστώντας ένα ποιοτικό και αξιόπιστο εργαλείο στην δημιουργία φυλογενετικών δένδρων σε σχέση με προηγούμενες εκδόσεις (Letunic I. and Bork P., 2021).

### **2.1.9. eggNOG-mapper v2**

Το eggNOG-mapper v2 πρόκειται για ένα λογισμικό που χρησιμοποιείται για την εύρεση συσχέτισης μεταξύ μιας πρωτεϊνικής αλληλουχίας υπερώτησης σε σύγκριση με μια βάση δεδομένων που περιέχει ομάδες ορθόλογων πρωτεϊνών. Η βάση αυτή ονομάζεται eggNOG database και περιλαμβάνει ένα σύνολο από ομάδες ορθόλογων (Orthologous Groups-Ogs) βακτηριακών, αρχαίους και ευκαρυωτικούς οργανισμούς και περιέχει πληροφορίες για την λειτουργία κάθε ομάδας ορθόλογων πρωτεϊνών (Cantalapiedra et al., 2021).

## 2.2. NCBI

Για την διεξαγωγή της εργασίας, χρησιμοποιήθηκαν από τη βάση δεδομένων NCBI Assembly (NCBI Taxonomy) 530 γονιδιώματα που αφορούν βακτηριακά στελέχη του γένους *Vibrio*. Για την λήψη χρησιμοποιήθηκαν τα παρακάτω φίλτρα. Λήφθηκαν μόνο γονιδιώματα με συναρμολόγηση (Assembly level) να έχουν ολοκληρωμένα γονιδιώματα (“Complete genome”) και να εξαιρεθούν τυχόν ανωμαλίες (exclude anomalous). Με την δημιουργία script (Perl) και χειροκίνητα φιλτραρίστηκαν περεταίρω, μόνο όσα είχαν ποσοστό άγνωστων νουκλεοτιδίων (%) < 1%, ποσοστό ψευδογονιδίων (%) < 10% και βάθος αλληλούχισης γονιδιώματος > 10X.). Η επιλεκτική αυτή διάκριση είναι σημαντική καθώς προφυλάσσει από την συμμετοχή κακής ποιότητας γονιδιωμάτων που μπορεί να οδηγήσουν σε ανεπιθύμητα αποτελέσματα όσο αναφορά το γονιδίωμα πυρήνα. Έτσι, απορρίφθηκαν συνολικά 131 γονιδιώματα κι έτσι η ανάλυση πραγματοποιήθηκε με 399 γονιδιώματα του γένους *Vibrio*.

### 3. Αποτελέσματα-Συζήτηση

Στην φυλογενωμική ανάλυση του γένους *Vibrio* χρησιμοποιήθηκαν για την εύρεση του γονιδιώματος πυρήνα 399 πρωτεώματα (και γονιδιώματα) από τα 530, τα οποία συλλέχθηκαν από το NCBI και επιλεκτικά διαχωρίστηκαν ύστερα από την εφαρμογή ειδικών φίλτρων που αναφέρθηκαν προηγουμένως και τοποθετήθηκαν στο pipeline. Εκτός από τα κριτήρια επιλογής που αναφέρθηκαν βασικό πρόβλημα στην επιλογή των γονιδιωμάτων και πρωτεωμάτων έπαιξε η ύπαρξη πολλών πλασμιδιακών γονιδίων. Με την δημιουργία Perl σκριπτ αφαιρέθηκαν όλες οι πλασμιδιακές πρωτεΐνες και χρησιμοποιήθηκαν για την ανάλυση μόνο οι χρωμοσωμικές. Έτσι με την χρήση μόνο των χρωμοσωμικών πρωτεϊνών και την παρουσία ενός πρωτεώματος αναφοράς (GCF\_008369605.1 *V. cholerae*) ανακτήθηκε το πρωτέωμα-πυρήνας.

Το αποτέλεσμα του pipeline έδωσε το σύνολο των ορθολόγων πρωτεϊνών μεταξύ των 399 πρωτεωμάτων ύστερα από την εκτέλεση ανταποδοτικού DIAMOND του πρωτεώματος αναφοράς και όλων των υπολοίπων πρωτεωμάτων. Ανακτήθηκαν λοιπόν 174 πρωτεΐνες με τιμή E-value μικρότερη από  $1 \times 10^{-5}$  τα οποία είναι συντηρημένα σε όλα τα μέλη του γένους, αποτελούν δηλαδή το πρωτέωμα πυρήνα (core-proteome).

Το επόμενο βήμα ήταν η κατασκευή του φυλογενωμικού δένδρου για το οποίο χρησιμοποιήθηκε ο αλγόριθμος MUSCLE. Με την χρήση του αυτού του αλγορίθμου κάθε ομάδα ορθολόγων στοιχίστηκε και όλες μαζί ενώθηκαν δημιουργώντας μια υπερακολουθία που περιλάμβανε το core-proteome. Αυτή η συνολική στοίχιση χρησιμοποιήθηκε για την δημιουργία ενός μέγιστης πιθανοφάνειας Maximum Likelihood φυλογενωμικού δέντρου με την χρήση του λογισμικού IQ Tree 2.

Με σκοπό να προσδιοριστούν τα διάφορα είδη χρησιμοποιήθηκε το λογισμικό FastANI το οποίο και βρίσκει το πόσο ταυτίζονται τα διάφορα γονιδιώματα μεταξύ τους, όπου με βάση την τιμή ANI καθορίζει το ποσοστό ταύτισης της νουκλεοτιδικής ομοιότητας. Στη συνέχεια τα αποτελέσματα που ελήφθησαν μέσω του αλγορίθμου Markov cluster δημιούργησαν ομάδες στελεχών τα οποία έχουν γονιδιώματα με ταύτιση 95% και μεγαλύτερη, κάθε μια τέτοια ομάδα ονομάστηκε cluster το οποίο αντιπροσωπεύει και ένα είδος. Έτσι προέκυψαν 64 διαφορετικά clusters η αλλιώς είδη για το γένος *Vibrio*. Κάθε είδος έχει έναν μοναδικό κωδικό, ο οποίος ξεκινάει με το γράμμα C (από την αγγλική λέξη Cluster) και ακολουθείται από έναν αύξοντα αριθμό με εκκίνηση το 1.

Στην συνέχεια όσα από τα clusters περιείχαν πέντε αντιπροσώπους και πάνω επιλέχθηκαν, με το καθένα από τα cluster να έχει 5 και μόνο πλέον στελέχη, ώστε να γίνει μια λεπτομερέστερη ανάλυση της συντήρησης του γονιδιώματος τόσο εντός του κάθε είδους άλλα και μεταξύ αυτών των ειδών. Πρόκειται για μια διαδικασία λοιπόν κανονικοποίησης που στόχο έχει την αντιπροσωπευτικότερη σύγκριση ανάμεσα σε είδη που έχουν μεγάλες διακυμάνσεις στον αριθμό των αντιπροσώπων που τους αποτελούν. Ο αριθμός 5 έχει πειραματικά δειχθεί ότι είναι ο καταλληλότερος για την βελτιστοποίηση των αποτελεσμάτων. Η επιλογή τους έγινε με βάση το φυλογενωμικό δέντρο με το κριτήριο ότι εντός του είδους οι 5 αυτοί εκπρόσωποι δεν θα πρέπει να βρίσκονται εξελικτικά κοντά να ανήκουν δηλαδή σε όσο πιο απομακρυσμένους κλάδους γίνεται.

Assembly	Species	Cluster
GCF_008369605.1	<i>V. cholerae</i>	C1
GCF_009883875.1	<i>V. parahaemolyticus</i>	C2
GCF_009613065.1	<i>V. alginolyticus</i>	C3
GCF_001675245.1	<i>V. vulnificus</i>	C4
GCF_002310335.1	<i>V. anguillarum</i>	C5
GCF_009764195.1	<i>V. fluvialis</i>	C6
GCF_002906475.1	<i>V. campbellii</i>	C7
GCF_000770115.1	<i>V. harveyi</i>	C8
GCF_000967465.2	<i>V. coralliilyticus</i>	C9
GCF_006364365.1	<i>V. fuminis</i>	C10
GCF_001456255.1	<i>V. natriegens</i>	C11
GCF_002021755.1	<i>V. owensii</i>	C12
GCF_001767355.1	<i>V. mimicus</i>	C13
GCF_002963385.1	<i>V. diabolicus</i>	C14
GCF_019977755.1	<i>V. penaeicida</i>	C15
GCF_023503885.1	<i>V. splendidus</i>	C16
GCF_012274885.1	<i>V. navarensis</i>	C17
GCF_009763885.1	<i>V. cincinnatiensis</i>	C18
GCF_023206055.1	<i>V. cyclitrophicus</i>	C19
GCF_019670485.1	<i>V. alfacensis</i>	C20
GCF_009763805.1	<i>V. metoecus</i>	C21
GCF_001995825.2	<i>V. kanaloae</i>	C22
GCF_023206035.1	<i>V. cyclitrophicus</i>	C23
GCF_007182835.1	<i>V. rotiferianus</i>	C24
GCF_001685465.1	<i>V. scophthalmi</i>	C25
GCF_023277445.1	<i>V. jasicida</i>	C26
GCF_015664285.1	<i>V. europaeus</i>	C27
GCF_905175365.1	<i>Vibrio</i> sp. B1FLJ16	C28
GCF_009938165.1	<i>V. takatomensis</i>	C29
GCF_003726835.1	<i>V. mediterranea</i>	C30
GCF_009763825.1	<i>V. metschnikovii</i>	C31
GCF_015097735.1	<i>V. nigripulchritudo</i>	C32
GCF_000272185.2	<i>V. crassostreae</i>	C33
GCF_023221665.1	<i>V. crassostreae</i>	C34
GCF_022371215.1	<i>V. gigantis</i>	C35
GCF_023206895.1	<i>Vibrio</i> sp. ED004	C36
GCF_013394185.1	<i>Vibrio</i> sp. Scap24	C37
GCF_002119625.1	<i>V. crassostreae</i>	C38
GCF_014879975.1	<i>V. bathopelagicus</i>	C39
GCF_004022545.1	<i>V. chagasii</i>	C40
GCF_009363365.1	<i>Vibrio</i> sp. THAF190c	C41
GCF_001677275.1	<i>V. biocanii</i>	C42
GCF_010587385.1	<i>V. astriarenae</i>	C43
GCF_002906865.1	<i>V. hyugensis</i>	C44
GCF_000772105.1	<i>V. tubashii</i>	C45
GCF_900233005.1	<i>V. tapetis</i>	C46
GCF_023195835.1	<i>V. sinoensis</i>	C47
GCF_009938225.1	<i>V. ponticus</i>	C48
GCF_009938205.1	<i>V. panuliri</i>	C49
GCF_011064285.1	<i>V. ziniensis</i>	C50
GCF_009665255.1	<i>V. metoecus</i>	C51
GCF_019048845.1	<i>V. mimicus</i>	C52
GCF_001547935.1	<i>V. tritonius</i>	C53
GCF_002849855.1	<i>V. azureus</i>	C54
GCF_009363415.1	<i>V. aquimaris</i>	C55
GCF_003716875.1	<i>V. zhugei</i>	C56
GCF_002157735.2	<i>V. aphrogenes</i>	C57
GCF_002218045.2	<i>V. rumoiensis</i>	C58
GCF_002218025.2	<i>V. casei</i>	C59
GCF_009601765.2	<i>V. algicola</i>	C60
GCF_002196515.1	<i>V. gazogenes</i>	C61
GCF_014083825.1	<i>V. spartinae</i>	C62
GCF_015394925.2	<i>Vibrio</i> sp. V.16	C63
GCF_021513135.1	<i>Vibrio</i> sp. SS-MA-C1-2	C64

**Πίνακας 1 :** Το σύνολο των 64 διαφορετικών ειδών, με τον κωδικό του NCBI, και το cluster στο οποίο ανήκουν.

Έτσι λοιπόν από τα 64 clusters πλέον αποσπάστηκαν 15 τα οποία συνολικά διέθεταν 5 η παραπάνω αντιπροσώπους και αναλυθήκαν περαιτέρω για τα ενδο-ειδικά χαρακτηριστικά τους. Το γονιδίωμα αναφοράς για το κάθε είδος λήφθηκε υπόψιν της βιβλιογραφίας από το NCBI ενώ όπου αυτή ήταν ανεπαρκής το γονιδίωμα αναφοράς θα ήταν αυτό με τις περισσότερες πρωτεΐνες. Συνολικά το πλήθος των στελεχών που χρησιμοποιήθηκε απλοποιήθηκε σε ένα δεύτερο σετ οποίο απαρτιζόταν από 324 στελέχη.

Assembly	Species	Cluster
GCF_008369605.1	<i>V. cholerae</i>	C1
GCF_009883875.1	<i>V. parahaemolyticus</i>	C2
GCF_003613065.1	<i>V. alginolyticus</i>	C3
GCF_001675245.1	<i>V. vulnificus</i>	C4
GCF_002310335.1	<i>V. anguillarum</i>	C5
GCF_009764195.1	<i>V. fluvialis</i>	C6
GCF_002906475.1	<i>V. campbellii</i>	C7
GCF_000770115.1	<i>V. harveyi</i>	C8
GCF_000967465.2	<i>V. coralliilyticus</i>	C9
GCF_006364355.1	<i>V. furnissii</i>	C10
GCF_001456255.1	<i>V. natriegens</i>	C11
GCF_002021755.1	<i>V. owensii</i>	C12
GCF_001767355.1	<i>V. mimicus</i>	C13
GCF_002953395.1	<i>V. diabolicus</i>	C14
GCF_019977755.1	<i>V. penaeicida</i>	C15

**Πίνακας 2 :** Τα 15 διαφορετικά είδη (Species), ο κωδικός GCF του NCBI που αντιστοιχεί στα γονιδιώματα αναφοράς του κάθε είδους (Assembly) και το Cluster που ανήκει το κάθε είδος

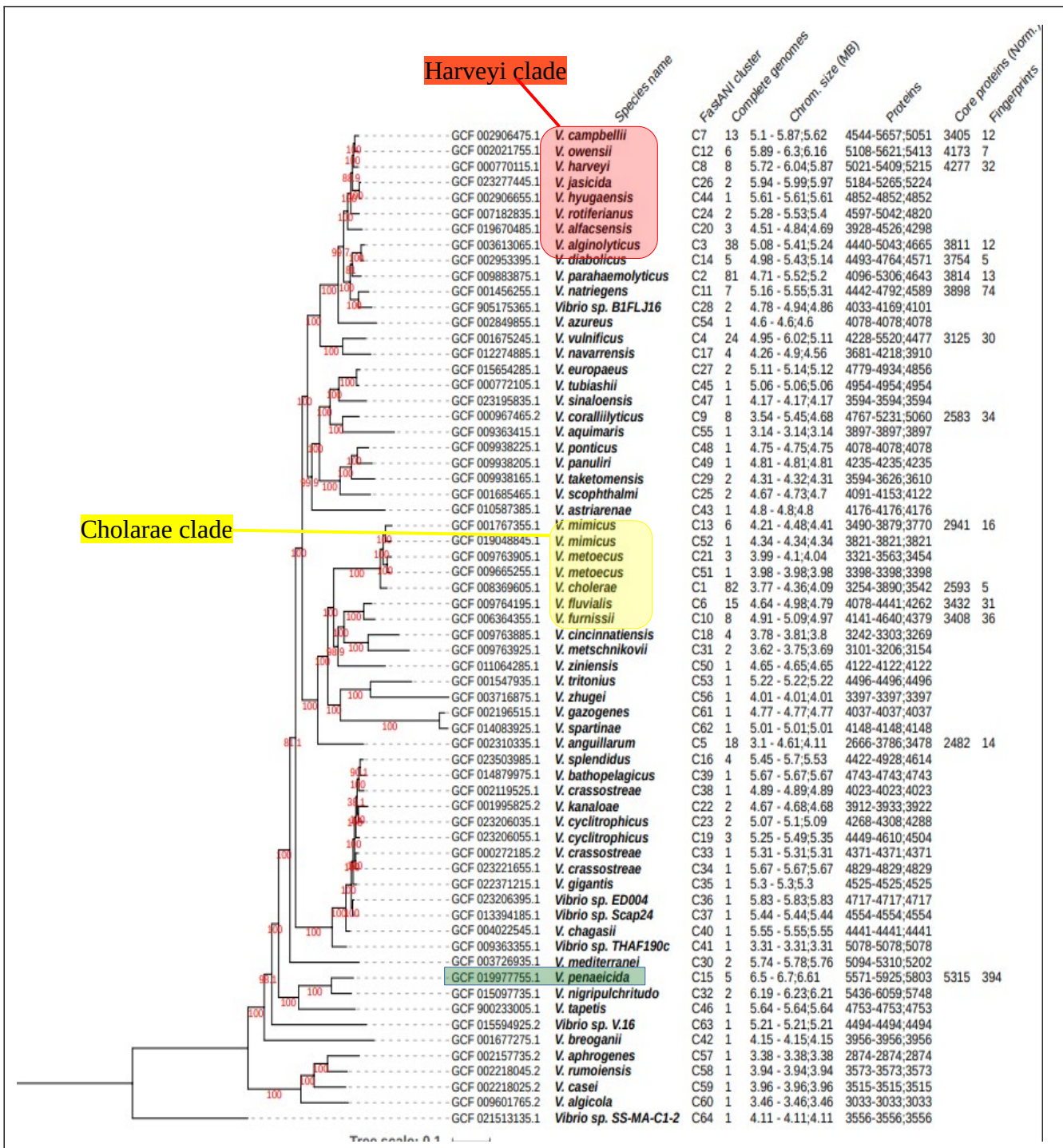
Στη συνέχεια τα 324 γονιδιώματα αυτά χρησιμοποιήθηκαν για ανάλυση στο pipeline που αυτήν την φορά στόχος ήταν η ανάκτηση του core-proteome του κάθε είδους συγκρίνοντας κάθε φορά τους 5 επιλεγμένους αντιπροσώπους από το κάθε cluster και εξάγοντας έτσι το σύνολο των ορθόλογων για το κάθε είδος ξεχωριστά. Ταυτόχρονα υπολογίστηκε και το σύνολο των γονιδίων που εμφανίζονταν και στα 5 στελέχη ενός είδους αλλά σε κανένα άλλο από τα υπόλοιπα είδη. Δίνοντας έτσι τις πρωτεΐνες fingerprints

Reference_assembly	FastANI_cluster	Core_proteins (Norm.)	Fingerprints (Norm.)
GCF_008369605.1	cluster_1	2593	5
GCF_009883875.1	cluster_2	3814	13
GCF_003613065.1	cluster_3	3811	12
GCF_001675245.1	cluster_4	3125	30
GCF_002310335.1	cluster_5	2482	14
GCF_009764195.1	cluster_6	3432	31
GCF_002906475.1	cluster_7	3405	12
GCF_000770115.1	cluster_8	4277	32
GCF_000967465.2	cluster_9	2583	34
GCF_006364355.1	cluster_10	3408	36
GCF_001456255.1	cluster_11	3898	74
GCF_002021755.1	cluster_12	4173	7
GCF_001767355.1	cluster_13	2941	16
GCF_002953395.1	cluster_14	3754	5
GCF_019977755.1	cluster_15	5315	394

**Πίνακας 3 :** Τα γονιδιώματα αναφοράς με το κάθε cluster στο οποίο ανήκουν, ο αριθμός των πρωτεϊνών-πυρήνα αλλά και ο αριθμός των μοναδικών πρωτεϊνών κάθε είδους fingerprint ύστερα της κανονικοποίησης.

Σε επόμενο βήμα έγινε η κατασκευή του φυλογενωμικού δένδρου (πλήρης εικόνα με όλα τα 399 στελέχη βρίσκεται στο τέλος της εργασίας) το οποίο περιλαμβάνει έναν αντιπρόσωπο από το κάθε cluster από τα 64 συνολικά διαφορετικά είδη. Τα αποτελέσματα του δένδρου συμβαδίζουν με αυτά της ομάδας ερευνητών του 2018 (Lin et al., 2018) όπου φαίνονται οι στενές συγγενικές σχέσεις των ειδών που ανήκουν στους κλάδους harveyi και cholerae. Συγκεκριμένα τα είδη *V. harveyi*, *V. campbellii*, *V. rotiferianus*, *V. alginolyticus* και στις δύο έρευνες ομαδοποιούνται μαζί στο κλάδο harveyi. Ενώ τα είδη *V. mimicus*, *V. cholerae*, *V. furnisii*, *V. fluvialis* συναντώνται και στις δύο μελέτες στον κλάδο cholerae.

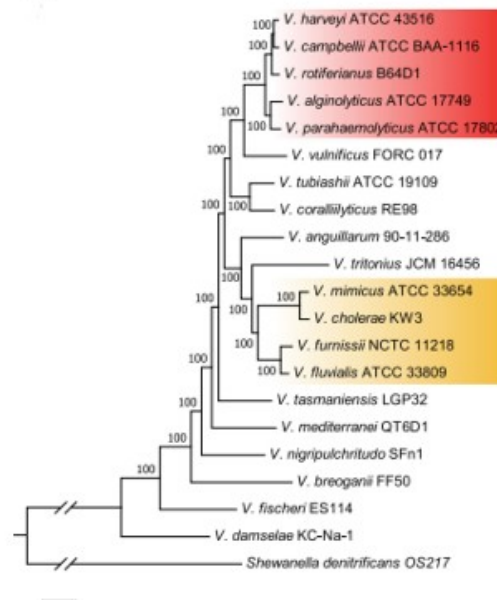




**Εικόνα 14 :** Το φυλογενωμικό δέντρο του γένους *Vibrio*. Χαρακτηριστικά δίπλα από το κωδικό του NCBI το οποίο αφορά τους οργανισμούς αναφοράς βρίσκονται από αριστερά προς δεξιά με την σειρά: 1) Το όνομα του είδους στο οποίο ανήκει (Species name), 2) το cluster που ανήκουν με βάση το FastANI (FastANI cluster), 3) το πλήθος των ολοκληρωμένων γονιδιωμάτων (complete genomes) το μέγεθος του χρωμοσώματος (Chrom. Size; Ελάχιστον-Μέγιστο: Μέσος ορος), 4) ο αριθμός των πρωτεϊνών με το μικρότερο αριθμό που εμφανίζεται



στα αριστερά, το μέσο όρο αλλά και τον μεγαλύτερο αριθμό πρωτεϊνών που συναντάται μεταξύ του είδους 5) το σύνολο των ορθόλογων μεταξύ των 324 γονιδιωμάτων που προαναφέρθηκαν και 6) ο αριθμός των μοναδικών πρωτεϊνών που συναντώνται μόνο στο συγκεκριμένο είδος (fingerprints)



**Εικόνα 15 :** Το φυλογενωμικό δένδρο για τα 20 είδη *Vibrio* (Lin et al., 2018).

Το επόμενο βήμα ήταν ο χαρακτηρισμός των COGS (cluster orthologous genes). Από την δεξαμενή των ορθόλογων που προέκυψε ήταν ενδιαφέρον να διαπιστώσουμε αν υπήρχε κάποιο σημαντικό εξελικτικό πλεονέκτημα, είτε αυτό προέκυπτε από την γονιδιακή λειτουργία των core-genes η οποία ευνοούσε το είδος στην φυσιολογία και την βιωσιμότητα του είτε εάν σε κάθε είδος οι μοναδικές πρωτεΐνες fingerprints ήταν αυτές οι οποίες πρόσδιδαν κάποιο ιδιαίτερο φαινοτυπικό χαρακτηριστικό το οποίο μπορεί να ήταν απαραίτητο για την βιολογική δραστηριότητα και επομένως την επιβίωση του κάθε είδους ξεχωριστά.

Για αυτό το σκοπό χρησιμοποιήθηκε το eggNOG-mapper για την ομαδοποίηση των core γονιδίων ανάλογα την λειτουργία τους όπως και των fingerprint πρωτεϊνών κάθε είδους από τα 15 του γένους *Vibrio*. Ύστερα από αυτό για να διακρίνουμε ποιες από αυτές τις κατηγορίες ήταν over-represented (δηλαδή εμφανίζονταν πολλές φορές ανάμεσα στο πλήθος των ορθόλογων εντός του είδους) η under-represented (εμφανίζονταν στατιστικά πολύ λιγότερες

φορές ανάμεσα στο πλήθος των ορθόλογων), πραγματοποιήθηκε το υπεργεωμετρικό τεστ για όλες τις ορθόλογες πρωτεΐνες που προέκυψαν στο κάθε είδος.

Τα αποτελέσματα έδειξαν ότι οι κατηγορίες που εμφανίζονταν ως over-represented στα περισσότερα είδη ήταν οι εξής : C (energy production and dynamics) 12/15 80%, E (Amino acid transport and metabolism) 10/15 66.66%, F (Nucleotide transport and metabolism) 14/15 93.33%, H (Coenzyme transport and metabolism) 14/15 93.33%, J (translation, ribosomal structure and biogenesis) 14/15 93.33%, O (Post-translational modification, protein turnover and chaperones) 12/15 80% και P (Inorganic ion transport and metabolism) 10/15 66.66%

Ενώ οι κατηγορίες που εμφανίζονται ως under-represented σε παραπάνω από τα μισά είδη, είναι οι εξής: L(Replication, recombination and repair) 9/15 60% και S(Function unknown)..Είναι σημαντικό να αναφερθεί ότι η κατηγορία S εμφανίστηκε under-represented και στα 15 είδη. Αυτό είναι λογικό διότι η μελέτη γύρω από το πρωτέωμα-πυρήνα του γένους *Vibrio* είναι εκτεταμένη, επομένως το μεγαλύτερο πλήθος των ορθόλογων που έχουν μελετηθεί από τους επιστήμονες είναι χαρακτηρισμένο με αποτέλεσμα να είναι πολύ λίγα τα γονίδια εκείνα που δεν γνωρίζουμε ακόμα την κατηγορία στην οποία ανήκουν.

Επομένως διακρίνεται ότι τα περισσότερα είδη έχουν συντηρημένα γονίδια υπεύθυνα για βασικές κυτταρικές λειτουργίες όπως η παραγωγή ενέργειας, ο μεταβολισμός των αμινοξέων, των νουκλεοτιδίων, η μετάφραση, η βιογένεση των ριβοσωμάτων, οι μετα-μεταφραστικές τροποποιήσεις, ο μεταβολισμός και η μεταφορά ανόργανων ιόντων.

Cluster	Fingerprints (Norm.)	Hypothetical proteins	% hypothetical	Proteins with S category	% with S category
cluster_1	5	3	60.00	4	80.00
cluster_2	13	9	69.23	11	84.62
cluster_3	12	7	58.33	8	66.67
cluster_4	30	11	36.67	16	53.33
cluster_5	14	7	50.00	9	64.29
cluster_6	31	14	45.16	21	67.74
cluster_7	12	6	50.00	8	66.67
cluster_8	32	23	71.88	30	93.75
cluster_9	34	21	61.76	25	73.53
cluster_10	36	15	41.67	21	58.33
cluster_11	74	29	39.19	37	50.00
cluster_12	7	5	71.43	5	71.43
cluster_13	16	6	37.50	8	50.00
cluster_14	5	0	0.00	2	40.00
cluster_15	394	205	52.03	248	62.94

**Πίνακας 4 :** Το σύνολο των μοριακών αποτυπωμάτων fingerprints για το κάθε είδος από τα 15 είδη του γένους *Vibrio* που μελετήθηκαν, φαίνεται ο αριθμός των fingerprints του κάθε είδους, ο αριθμός και το ποσοστό των υποθετικών πρωτεϊνών, ο αριθμός και το ποσοστό των πρωτεϊνών που είναι άγνωστης λειτουργίας (S κατηγορία)..

Για τον χαρακτηρισμό των fingerprints πραγματοποιήθηκε με την ίδια διαδικασία το υπεργεωμετρικό τεστ και τα αποτελέσματα έδειξαν ότι στα 14 από τα 15 είδη η κατηγορία S

ήταν over-represented, άρα σημαίνει ότι οι χαρακτηριστικές πρωτεΐνες fingerprints του κάθε είδους είναι άγνωστης λειτουργίας και επομένως χρήζεται η ανάγκη για την καλύτερη μελέτη αυτών των πρωτεΐνων. Στον πίνακα.3 διακρίνεται μια σύνοψη των αποτελεσμάτων. Συνολικά βρέθηκαν 715 πρωτεΐνες αποτυπώματα σε όλα τα διαφορετικά είδη και προσδιορίστηκε η οικογένεια στην οποία ανήκουν σύμφωνα με το NCBI, με 25 από αυτά να είναι πρωτεΐνες τύπου ABC transporters οι οποίες μοιράζονται στα διάφορα είδη, αλλά χρειάζεται περισσότερη μελέτη για τον πιθανό ρόλο του που έχουν στην βιολογία του εκάστοτε είδους.

Η μελέτη των πρωτεΐνων fingerprint είναι σημαντική γιατί μπορεί να υποδείξει πρωτεΐνες οι οποίες προσφέρουν ένα σημαντικό εξελικτικό πλεονέκτημα το οποίο προσφέρει την ικανότητα στα στελέχη του συγκεκριμένου είδους να ευδοκίμουν στο αντίστοιχο περιβάλλον. Εκτός όμως από αυτό η μελέτη αυτών των αποτυπωμάτων μπορεί να αξιοποιηθεί για την δημιουργία διαγνωστικών τεχνικών, καθώς αν γνωρίζουμε την μοριακή και την γενετική βάση αυτών των πρωτεΐνων μπορεί να αναπτυχθούν κατάλληλα εργαλεία για την ταξινόμηση άγνωστων στελεχών εντός μιας ταξινομικής ομάδας.

Εντύπωση προκαλεί ο μεγάλος αριθμός πρωτεΐνων fingerprint του Cluster 15 που αντιστοιχεί στο είδος *V. parvula*, όπου διακρίνονται 394 διαφορετικές μοναδικές πρωτεΐνες. Ένας τέτοιος μεγάλος αριθμός μοναδικών πρωτεΐνων υποδηλώνει μια στενή συγγενική σχέση ανάμεσα στα στελέχη αυτής της ταξινομικής ομάδας. Φαίνεται δηλαδή να υπάρχει μια υψηλότερη συντήρηση μεταξύ των στελεχών σε σχέση με τα υπόλοιπα 14 είδη και αυτό παρατηρείται και στο δέντρο της εικόνας 14 όπου το συγκεκριμένο είδος διακρίνεται ως πιο απομακρυσμένο από τα υπόλοιπα.

Για το *V. cholerae* το οποίο αντιστοιχεί στο cluster 1 εντοπίστηκαν 5 πρωτεΐνες fingerprint με 3 από αυτές να είναι υποθετικές ενώ 4 από αυτές ήταν άγνωστης λειτουργίας. Μία από αυτές κατηγοριοποιήθηκε στην ομάδα T των COG annotation, που αντιπροσωπεύει πρωτεΐνες οι οποίες ευθύνονται για μηχανισμούς κυτταρικής σηματοδότησης. Η πρωτεΐνη αυτή με βάση το NCBI είναι η Hpt domain-containing protein (histidine phosphotransferase domain), και αποτελεί ένα τμήμα της κινάσης ιστιδίνης. Η λειτουργία αυτού του τμήματος είναι να πραγματοποιεί την μεταφορά μιας φωσφορικής ομάδας από ένα κατάλοιπο ιστιδίνης, στο οποίο έχει προηγηθεί αυτοφωσφωρυλίωση από την κινάση, σε ένα κατάλοιπο ασπαρτικού μιας άλλης πρωτεΐνης τελεστή. Αυτό το σύστημα είναι γνωστό ως two-component system όπου οι δύο συμμετέχοντες στην σηματοδότηση είναι η κινάση ιστιδίνης και ένας ρυθμιστής απόκρισης (response regulator) (Stock et al., 2000).

Γενικότερα η λειτουργία αυτής της πρωτεΐνης έχει αποδειχθεί ότι σχετίζεται με την ρύθμιση της ανάπτυξης των βακτηρίων ως ένα συνεργαζόμενο σύνολο στο περιβάλλον. Συγκεκριμένα σχετίζονται με την οσμωρύθμιση, την χειμειόταξια και ειδικότερα την δημιουργία βιοφίλμ, διεργασίες που αφορούν την αίσθηση απαρτίας (quorum sensing) των βακτηρίων (InterPro, n.d.).

Όσον αφορά τα βιοφίλμ είναι γνωστό ότι τα βακτήρια του γένους *Vibrio* ευθύνονται για την δημιουργία τέτοιων δομών οι οποίες είναι ευνοϊκές για την βιωσιμότητα και τον βιολογικό κύκλο των συγκεκριμένων βακτηρίων. Πιο συγκεκριμένα όμως το είδος *V. cholerae* έχει αποδειχθεί ότι είναι ικανό για την δημιουργία βιοφίλμ (Shikuma et al., 2009). Σε ενίσχυση αυτής της πληροφορίας, συνεισφέρει η εύρεση μιας ακόμα πρωτεΐνης fingerprint του συγκεκριμένου είδους η οποία χαρακτηρίζεται σύμφωνα με το NCBI ως VpsP family

polysaccharide biosynthesis protein κατατάσσεται ως πρωτεΐνη με άγνωστη λειτουργία (S κατηγορία). Η οικογένεια αυτών των πρωτεϊνών είναι μεμβρανικές πρωτεΐνες οι οποίες είναι υπεύθυνες για την παραγωγή πολυσακχαριτών, που εξωκυτταρικά δρουν ως συνεκτικές δομές ανάμεσα στα διαφορετικά κύτταρα συμβάλλοντας σημαντικά στην δημιουργία του βιοφίλμ. Μάλιστα σε άρθρο του 2009 φαίνεται ότι υπάρχει ένα two-component σύστημα που λειτουργεί στο *V. cholerae* στο οποίο συμμετέχουν μια Hpt domain-containing protein (LUX) που δρα ως το σήμα για την ενεργοποίηση του VpsS μιας πρωτεΐνης που ρυθμίζει θετικά την παραγωγή πολυσακχαριτών και επομένως την δημιουργία βιοφίλμ (Shikuma et al., 2009). Η ίδια οικογένεια πρωτεϊνών παρατηρήθηκε ως πρωτεϊνικό αποτύπωμα στα βακτήρια του είδους *V. anguillarum* που επίσης δημιουργούν βιοφίλμ .

Επιπλέον στα βακτήρια του είδους *V. vulnificus* διακρίνεται μια πρωτεΐνη που κατατάσσεται ως O-antigen ligase family protein. Το O-αντιγόνο αποτελεί τμήμα του λιποπολυσακχαριδίου (LPS) που όπως προαναφέρθηκε και στο κομμάτι της εισαγωγής αποτελεί λοιμογόνο παράγοντα που ευνοεί την προσβολή του ξενιστή και την ολοκλήρωση του βιολογικού κύκλου του συγκεκριμένου είδους. Γενικότερα η συγκεκριμένη οικογένεια πρωτεϊνών είναι συνδεδεμένη με την εμφάνιση λοιμογόνου δράσης και σε άλλα παθογόνα είδη.

Ενδιαφέρον προκαλεί η εμφάνιση μιας ομάδας πεπτιδίων στο είδος *V. peneicida* που ανήκει στην οικογένεια, oniridin/marinostatin family tricyclic proteinase inhibitor. Αυτή η οικογένεια πεπτιδίων είναι γνωστές στα κυανοβακτήρια, παράγονται από τα ριβοσώματα και είναι αναστολείς πρωτεασών. Τα συγκεκριμένα πεπτιδία συνήθως δρουν αμυντικά σε φυσικούς θηρευτές με το να αναστέλλουν την λειτουργία υδρολυτικών ενζύμων εμποδίζοντας την πέψη της τροφής και εξασφαλίζουν έτσι την κυριαρχία τους στα διάφορα ενδιαιτήματα. Ένας από αυτούς τους θηρευτές των κυανοβακτηρίων είναι η *Daphnia microcrustacean* που αποτελεί ένα είδος μικρού αρθροποδού. Αξίζει να σημειωθεί ότι τα βακτήρια του είδους *V. peneicida* είναι παθογόνα αρθρόποδων μεταξύ αυτών είναι διάφορα είδη γαρίδας, υποδηλώνοντας πιθανόν μια παρόμοια λειτουργία.

Στα βακτήρια του είδους *V. furnissii* εμφανίζονται δύο πρωτεϊνικά αποτυπώματα fingerprint που ανήκουν στην οικογένεια πρωτεϊνών OmpA family protein (Omp, outer membrane proteins). Οι έξω μεμβρανικές πρωτεΐνες Omp έχει δειχθεί ότι παίζουν σημαντικό ρόλο στον μηχανισμό παθογένεσης διάφορων βακτηριακών ειδών, όπως *Pseudomonas aeruginosa*, *Haemophilus influenzae*, *Klebsiella pneumoniae* και *Chlamydia trachomatis* και συνήθως ευθύνονται για την κυτταρική προσκόλληση, την είσβολή στα κύτταρα του ξενιστή, την επιβίωση στο ενδοκυτταρικό περιβάλλον, ακόμα και την αποφυγή αμυντικών μηχανισμών του ξενιστή όπως ο καταρράκτης κυτοκινών (Confer & Ayalew, 2013). Στο τελευταίο όπως αναφέρθηκε στο κομμάτι της εισαγωγής τα βακτήρια του είδους *V. cholerae* χρησιμοποιούν την κατανομή των διαφορών εξωμεμβρανικών πρωτεϊνών Omp με στόχο την επιβίωση στο περιβάλλον του ξενιστή. Τα βακτήρια του είδους *V. furnissii* ευθύνονται για την πρόκληση γαστρεντερίτιδας, γεγονός το οποίο μπορεί να συνδέεται με την λειτουργία αυτής της πρωτεΐνης. Τέλος τέτοιες ομάδες πρωτεϊνών θεωρούνται πολλές φορές πιθανοί φαρμακευτικοί στόχοι για την ανάπτυξη εμβολίων (Confer & Ayalew, 2013).

Συμπερασματικά τα αποτελέσματα μας συμβαδίζουν με τις μελέτες του 2018, με τη διαφορά ότι η χρήση του μεγάλου αριθμού γονιδιωμάτων που χρησιμοποιήσαμε χαρακτηρίζει την μελέτη μας αρκετά ακριβής και αποδοτική διαδικασία, ο μεγάλος αριθμός στελεχών που χρησιμοποιήσαμε ευθύνεται και για το πλήθος των fingerprint που αναδείχθηκαν τα οποία είχαν πιθανούς λοιμογόνους μηχανισμούς, ή μηχανισμούς προστασίας των ειδών *Vibrio* από τους περιβαλλοντικούς θηρευτές του. Αξίζει η λεπτομερέστερη ανάλυση των αποτελεσμάτων και η εξακρίβωση των μυστικών των πρωτεϊνών αποτυπωμάτων καθώς μπορεί να αποτελούν πιθανούς φαρμακευτικούς στόχους.

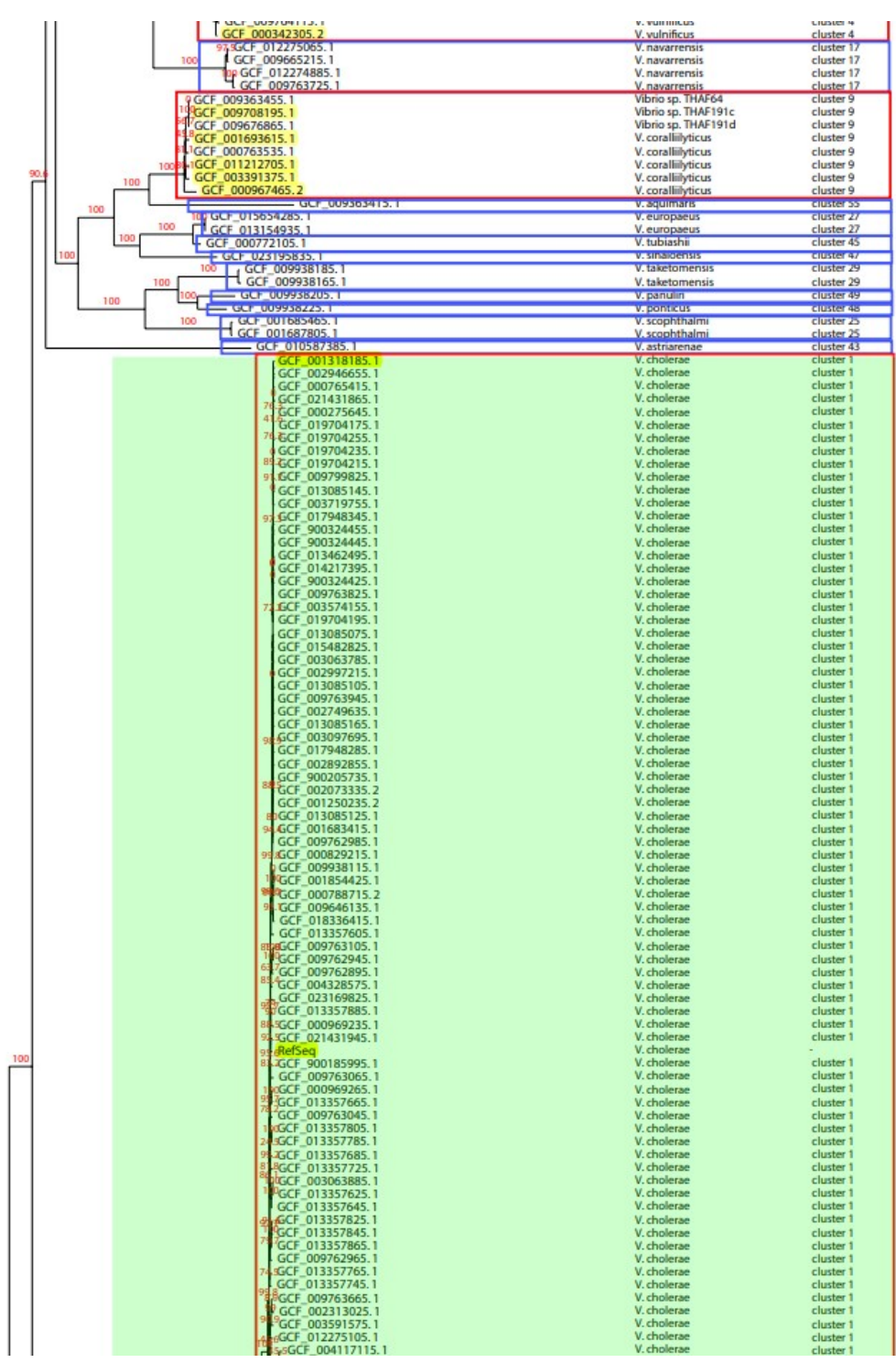
Category	Annotation
A	RNA processing and modification
B	Chromatin structure and dynamics
C	Energy production and conversion
D	Cell cycle control, cell division, chromosome partitioning
E	Amino acid transport and metabolism
F	Nucleotide transport and metabolism
G	Carbohydrate transport and metabolism
H	Coenzyme transport and metabolism
I	Lipid transport and metabolism
J	Translation, ribosomal structure and biogenesis
K	Transcription
L	Replication, recombination and repair
M	Cell wall/membrane/envelope biogenesis
N	Cell motility
O	Post-translational modification, protein turnover, and chaperones
P	Inorganic ion transport and metabolism
Q	Secondary metabolites biosynthesis, transport, and catabolism
S	Function unknown
T	Signal transduction mechanisms
U	Intracellular trafficking, secretion, and vesicular transport
V	Defense mechanisms
W	Extracellular structures
Z	Cytoskeleton

**Πίνακας 5 :** Οι λειτουργικές κατηγορίες των ορθόλογων γονιδίων από το eggNOG-mapper v2.

84	GCF_016132485.1	V. parahaemolyticus	cluster 2
74	GCF_013393865.1	V. parahaemolyticus	cluster 2
74	GCF_003047085.1	V. parahaemolyticus	cluster 2
84	GCF_018135645.1	V. parahaemolyticus	cluster 2
	GCF_001758605.1	V. parahaemolyticus	cluster 2
84	GCF_004194515.1	V. parahaemolyticus	cluster 2
	GCF_020041945.1	V. parahaemolyticus	cluster 2
99	GCF_015779285.1	V. parahaemolyticus	cluster 2
99	GCF_009734325.1	V. parahaemolyticus	cluster 2
	GCF_023205895.1	V. parahaemolyticus	cluster 2
74	GCF_009763025.1	V. parahaemolyticus	cluster 2
	GCF_021729985.1	V. parahaemolyticus	cluster 2
	GCF_002209725.2	V. parahaemolyticus	cluster 2
	GCF_009763565.1	V. parahaemolyticus	cluster 2
80	GCF_001682175.1	V. parahaemolyticus	cluster 2
3	GCF_002220985.3	V. parahaemolyticus	cluster 2
84	GCF_001951115.1	V. parahaemolyticus	cluster 2
99	GCF_003612695.1	V. parahaemolyticus	cluster 2
74	GCF_000430425.1	V. parahaemolyticus	cluster 2
74	GCF_009763425.1	V. parahaemolyticus	cluster 2
74	GCF_014217295.1	V. parahaemolyticus	cluster 2
	GCF_013393845.1	V. parahaemolyticus	cluster 2
64	GCF_013393885.1	V. parahaemolyticus	cluster 2
34	GCF_012274865.1	V. parahaemolyticus	cluster 2
	GCF_009764055.1	V. parahaemolyticus	cluster 2
89	GCF_016132465.1	V. parahaemolyticus	cluster 2
1	GCF_009763545.1	V. parahaemolyticus	cluster 2
74	GCF_009763405.1	V. parahaemolyticus	cluster 2
74	GCF_009883815.1	V. parahaemolyticus	cluster 2
89	GCF_009883835.1	V. parahaemolyticus	cluster 2
	GCF_009883855.1	V. parahaemolyticus	cluster 2
	GCF_016403045.1	V. parahaemolyticus	cluster 2
14	GCF_009883895.1	V. parahaemolyticus	cluster 2
24	GCF_003085735.1	V. parahaemolyticus	cluster 2
74	GCF_003612715.1	V. parahaemolyticus	cluster 2
84	GCF_001636035.1	V. parahaemolyticus	cluster 2
	GCF_001996365.2	V. parahaemolyticus	cluster 2
	GCF_003691525.1	V. parahaemolyticus	cluster 2
64	GCF_001244315.1	V. parahaemolyticus	cluster 2
64	GCF_001887055.1	V. parahaemolyticus	cluster 2
64	GCF_004006515.1	V. parahaemolyticus	cluster 2
	GCF_002073775.2	V. parahaemolyticus	cluster 2
84	GCF_016834555.1	V. parahaemolyticus	cluster 2
	GCF_008693745.1	V. parahaemolyticus	cluster 2
	GCF_009763645.1	V. parahaemolyticus	cluster 2
	GCF_009764075.1	V. parahaemolyticus	cluster 2
	GCF_009763505.1	V. parahaemolyticus	cluster 2
	GCF_009763525.1	V. parahaemolyticus	cluster 2
14	GCF_009763165.1	V. parahaemolyticus	cluster 2
64	GCF_001188185.2	V. parahaemolyticus	cluster 2
94	GCF_009649015.1	V. parahaemolyticus	cluster 2
1	GCF_021729965.1	V. parahaemolyticus	cluster 2
84	GCF_021730025.1	V. parahaemolyticus	cluster 2
1	GCF_009665495.1	V. parahaemolyticus	cluster 2
74	GCF_009763445.1	V. parahaemolyticus	cluster 2
1	GCF_003076895.1	V. parahaemolyticus	cluster 2
94	GCF_003119375.1	V. parahaemolyticus	cluster 2
34	GCF_001433415.1	V. parahaemolyticus	cluster 2
94	GCF_023205915.1	V. parahaemolyticus	cluster 2
84	GCF_006517795.1	V. parahaemolyticus	cluster 2
84	GCF_001700835.1	V. parahaemolyticus	cluster 2
74	GCF_021730085.1	V. parahaemolyticus	cluster 2
1	GCF_023206515.1	V. parahaemolyticus	cluster 2
94	GCF_001879585.1	V. parahaemolyticus	cluster 2
14	GCF_012275005.1	V. parahaemolyticus	cluster 2
84	GCF_009763625.1	V. parahaemolyticus	cluster 2
44	GCF_000568495.1	V. parahaemolyticus	cluster 2
	GCF_002504185.1	V. parahaemolyticus	cluster 2
84	GCF_023205955.1	V. parahaemolyticus	cluster 2
84	GCF_023205935.1	V. parahaemolyticus	cluster 2
84	GCF_012274985.1	V. parahaemolyticus	cluster 2
84	GCF_008693625.1	V. parahaemolyticus	cluster 2
	GCF_000430405.1	V. parahaemolyticus	cluster 2
	GCF_009763605.1	V. parahaemolyticus	cluster 2
00	GCF_021730005.1	V. parahaemolyticus	cluster 2
	GCF_001558495.2	V. parahaemolyticus	cluster 2
	GCF_021730065.1	V. parahaemolyticus	cluster 2
	GCF_019321785.1	V. parahaemolyticus	cluster 2
34	GCF_003351885.1	V. parahaemolyticus	cluster 2
1	GCF_021730045.1	V. parahaemolyticus	cluster 2
	GCF_009883875.1	V. parahaemolyticus	cluster 2
	GCF_002149065.1	V. alginolyticus	cluster 3
	GCF_002119565.1	V. alginolyticus	cluster 3
	GCF_002119545.1	V. alginolyticus	cluster 3
	GCF_002119585.1	V. alginolyticus	cluster 3
	GCF_002149105.1	V. alginolyticus	cluster 3
	GCF_002149085.1	V. alginolyticus	cluster 3
	GCF_003613035.1	V. alginolyticus	cluster 3
	GCF_002119505.2	V. alginolyticus	cluster 3
	GCF_003613065.1	V. alginolyticus	cluster 3
	GCF_022343165.1	V. alginolyticus	cluster 3
	GCF_002953535.1	V. alginolyticus	cluster 3
	GCF_023169625.1	V. alginolyticus	cluster 3
	GCF_009763745.1	V. alginolyticus	cluster 3
	GCF_022343085.1	V. alginolyticus	cluster 3
	GCF_022343105.1	V. alginolyticus	cluster 3
	GCF_009763845.1	V. alginolyticus	cluster 3
	GCF_002953515.1	V. alginolyticus	cluster 3
	GCF_023650755.1	Vibrio sp. SC5IO 43097	cluster 3

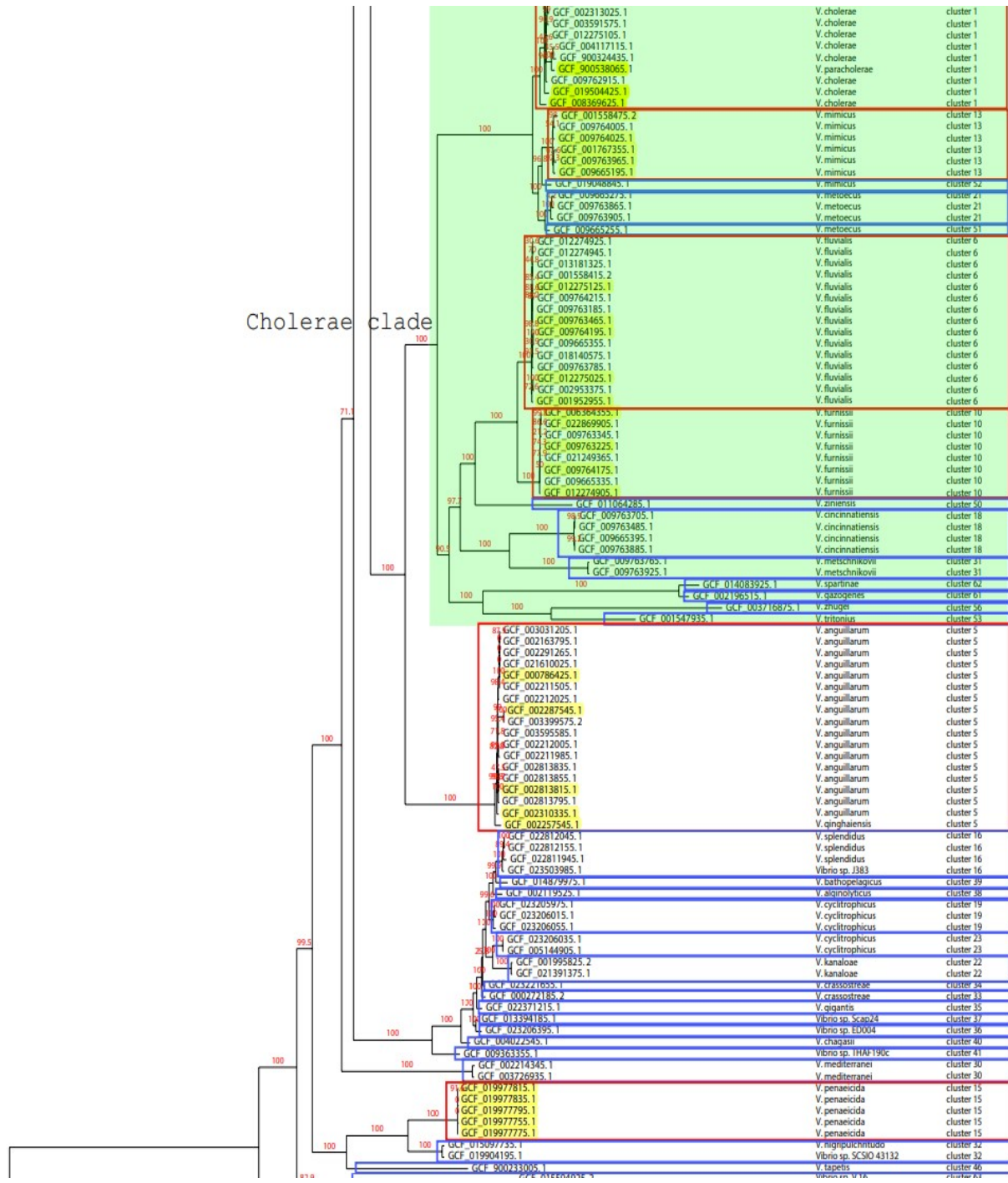


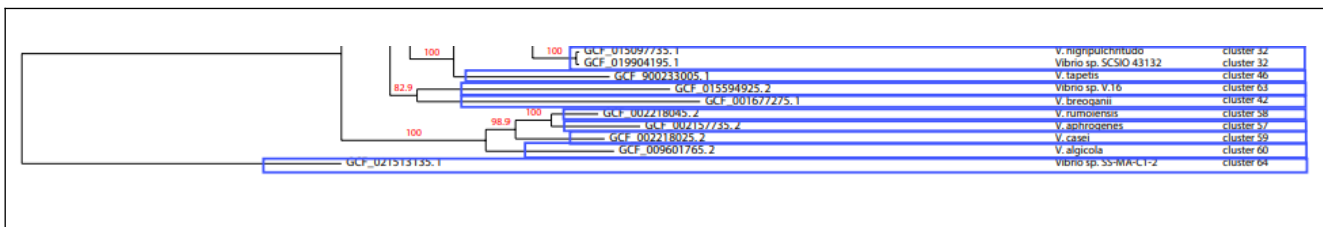






Cholerae clade





**Εικόνα 16 :** Το Φυλογενωμικό δέντρο του γένους *Vibrio* που έγινε με το IQ\_Tree 2 χρησιμοποιώντας μοντέλο αντικατάστασης LG+I+F+G4 με υπερστοίχιση 174 πρωτεϊνών πυρήνα του γένους (36.158 sites). Με κόκκινο σημειώνονται τα είδη με παραπάνω από 5 αντιπροσώπους ενώ με μπλε τα είδη με λιγότερα από 5 στελέχη. Το κίτρινο Highlight σηματοδοτεί τους πέντε αντιπροσώπους από κάθε είδος που επιλέχθηκαν για την ανάλυση των πρωτεωμάτων πυρήνα και την εύρεση των πρωτεϊνών fingerprint.

## 4. Βιβλιογραφία

- Aanensen, D. M., & Spratt, B. G. (2005). The multilocus sequence typing network: mlst.net. *Nucleic Acids Research*, 33(Web Server), W728–W733. <https://doi.org/10.1093/nar/gki415>
- About the Ubuntu project | Ubuntu*. (n.d.). Ubuntu. <https://ubuntu.com/about>
- Alavi, S., Mitchell, J., Cho, J. Y., Li, R., Macbeth, J., & Hsiao, A. (2020). Interpersonal gut microbiome variation drives susceptibility and resistance to cholera infection. *Cell*, 181(7), 1533-1546.e13. <https://doi.org/10.1016/j.cell.2020.05.036>
- Ali, M., Nelson, A. R., Lopez, A. L., & Sack, D. A. (2015). Updated Global burden of cholera in endemic countries. *PLOS Neglected Tropical Diseases*, 9(6), e0003832. <https://doi.org/10.1371/journal.pntd.0003832>
- Attridge, S. R., & Rowley, D. (1983). The Role of the Flagellum in the Adherence of *Vibrio cholerae*. *The Journal of Infectious Diseases*, 147(5), 864–872. <https://doi.org/10.1093/infdis/147.5.864>
- Bahrani, K., & Oliver, J. D. (1990). Studies on the lipopolysaccharide of a virulent and an avirulent strain of *Vibrio vulnificus*. *Biochemistry and Cell Biology*, 68(2), 547–551. <https://doi.org/10.1139/o90-078>
- Baker-Austin, C., Oliver, J. D., Alam, M., Ali, A., Waldor, M. K., Qadri, F., & Martínez-Urtaza, J. (2018). *Vibrio* spp. infections. *Nature Reviews Disease Primers*, 4(1), 1–19. <https://doi.org/10.1038/s41572-018-0005-8>
- Bell, A., & Bott, M. (2021). Vibriosis:: What You and Your Patients Need To Know. *Delaware Journal of Public Health*, 7(1), 14–21. <https://doi.org/10.32481/djph.2021.001.005>
- Bergey's Manual® of Systematic Bacteriology. (2005). In *Springer eBooks*. <https://doi.org/10.1007/0-387-28022-7>

- Bogdanchikov, A., Zhaparov, M., & Suliyev, R. (2013). Python to learn programming. *Journal of Physics*, 423, 012027. <https://doi.org/10.1088/1742-6596/423/1/012027>
- Bosi, E., Fani, R., & Fondi, M. (2015). Defining orthologs and pangenome size metrics. In *Methods in molecular biology* (pp. 191–202). [https://doi.org/10.1007/978-1-4939-1720-4\\_13](https://doi.org/10.1007/978-1-4939-1720-4_13)
- Boyd, E. F., Heilpern, A. J., & Waldor, M. K. (2000). Molecular Analyses of a Putative CTX $\phi$  Precursor and Evidence for Independent Acquisition of Distinct CTX $\phi$ s by Toxigenic *Vibrio cholerae*. *Journal of Bacteriology*, 182(19), 5530–5538. <https://doi.org/10.1128/jb.182.19.5530-5538.2000>
- Butler, S. M., & Camilli, A. (2005). Going against the grain: chemotaxis and infection in *Vibrio cholerae*. *Nature Reviews Microbiology*, 3(8), 611–620. <https://doi.org/10.1038/nrmicro1207>
- Carboni, G. P. (2021). The enigma of Pacini's *Vibrio cholerae* discovery. *Journal of Medical Microbiology*, 70(11). <https://doi.org/10.1099/jmm.0.001450>
- Childers, B. M., & Klose, K. E. (2007). Regulation of virulence in *Vibrio cholerae*: the ToxR regulon. *Future Microbiology*, 2(3), 335–344. <https://doi.org/10.2217/17460913.2.3.335>
- Cho, J. Y., Li, R., Macbeth, J., & Hsiao, A. (2021). The Interface of *Vibrio cholerae* and the Gut Microbiome. *Gut Microbes*, 13(1). <https://doi.org/10.1080/19490976.2021.1937015>
- Chun, J., Grim, C. J., Hasan, N. A., Lee, J., Choi, S. Y., Haley, B. J., Taviani, E., Jeon, Y., Kim, D. W., Lee, J., Brettin, T., Bruce, D., Challacombe, J. F., Detter, J. C., Han, C., Munk, A. C., Chertkov, O., Meincke, L., Saunders, E., . . . Colwell, R. R. (2009). Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae*. *Proceedings of the National Academy of Sciences of the United States of America*, 106(36), 15442–15447. <https://doi.org/10.1073/pnas.0907787106>

- Chun, J., & Rainey, F. A. (2014). Integrating genomics into the taxonomy and systematics of the Bacteria and Archaea. *International Journal of Systematic and Evolutionary Microbiology*, 64(Pt\_2), 316–324. <https://doi.org/10.1099/ij.s.0.054171-0>
- Cmangiola. (2019, December 5). *Vibrio vulnificus*. Mechanisms of Pathogenicity. <https://mechpath.com/2019/12/05/vibrio-vulnificus/>
- Cock, P. J. A., Antão, T., Chang, J. T., Chapman, B., Cox, C. J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczyński, B., & De Hoon, M. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*, 25(11), 1422–1423. <https://doi.org/10.1093/bioinformatics/btp163>
- Confer, A. W., & Ayalew, S. (2013). The OmpA family of proteins: Roles in bacterial pathogenesis and immunity. *Veterinary Microbiology*, 163(3–4), 207–222. <https://doi.org/10.1016/j.vetmic.2012.08.019>
- Drancourt, M., & Raoult, D. (1994). Taxonomic position of the Rickettsiae: Current knowledge. *Fems Microbiology Reviews*, 13(1), 13–24. <https://doi.org/10.1111/j.1574-6976.1994.tb00032.x>
- Drews, G. (2000). The roots of microbiology and the influence of Ferdinand Cohn on microbiology of the 19th century. *Fems Microbiology Reviews*, 24(3), 225–249. <https://doi.org/10.1111/j.1574-6976.2000.tb00540.x>
- Escudero, J. A., & Mazel, D. (2017). Genomic Plasticity of *Vibrio cholerae*. *PubMed*, 20(3), 138–148. <https://doi.org/10.2436/20.1501.01.295>
- Fang, Y., Wang, Y., Liu, Z., Dai, H., Cai, H., Li, Z., Du, Z., Wang, X., Jing, H., Wei, Q., Kan, B., & Wang, D. (2019). Multilocus Sequence Analysis, a Rapid and Accurate Tool for Taxonomic Classification, Evolutionary Relationship Determination, and Population Biology Studies of the Genus *Shewanella*. *Applied and Environmental Microbiology*, 85(11). <https://doi.org/10.1128/aem.03126-18>

- Faruque, S. M., Chowdhury, N., Kamruzzaman, M., Dziejman, M., Rahman, M., Sack, D. A., Nair, G. B., & Mekalanos, J. J. (2004). Genetic diversity and virulence potential of environmental *Vibrio cholerae* population in a cholera-endemic area. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(7), 2123–2128. <https://doi.org/10.1073/pnas.0308485100>
- Fu, X., Gong, L., Liu, Y., Lai, Q., Li, G., & Shao, Z. (2021). Bacillus pumilus Group Comparative Genomics: Toward Pangenome Features, Diversity, and Marine Environmental Adaptation. *Frontiers in Microbiology*, *12*. <https://doi.org/10.3389/fmicb.2021.571212>
- Gangarosa, E. J., Beisel, W. R., Benyajati, C., Sprinz, H., & Piyaratn, P. (1960). The nature of the gastrointestinal lesion in Asiatic cholera and its relation to pathogenesis: a biopsy study. *American Journal of Tropical Medicine and Hygiene*, *9*(2), 125–135. <https://doi.org/10.4269/ajtmh.1960.9.125>
- Gevers, D., Cohan, F. M., Lawrence, J. G., Spratt, B. G., Coenye, T., Feil, E. J., Stackebrandt, E., Van De Peer, Y., Vandamme, P., Thompson, F. L., & Swings, J. (2005). Re-evaluating prokaryotic species. *Nature Reviews Microbiology*, *3*(9), 733–739. <https://doi.org/10.1038/nrmicro1236>
- Glaeser, S. P., & Kämpfer, P. (2015). Multilocus sequence analysis (MLSA) in prokaryotic taxonomy. *Systematic and Applied Microbiology*, *38*(4), 237–245. <https://doi.org/10.1016/j.syapm.2015.03.007>
- Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P., & Tiedje, J. M. (2007). DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *International Journal of Systematic and Evolutionary Microbiology*, *57*(1), 81–91. <https://doi.org/10.1099/ijms.0.64483-0>
- Heidelberg, J. F., Eisen, J. A., Nelson, W., Clayton, R. A., Gwinn, M. L., Dodson, R. J., Haft, D. H., Hickey, E. K., Peterson, J., Umayam, L., Gill, S. R., Nelson, K. E., Read, T. D., Tettelin, H.,

- Richardson, D., Ermolaeva, M. D., Vamathevan, J., Bass, S., Qin, H., . . . Fraser, C. M. (2000). DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature*, *406*(6795), 477–483. <https://doi.org/10.1038/35020000>
- Heng, S. P., Letchumanan, V., Deng, C. Y., Mutalib, N. S. A., Khan, T. M., Chuah, L. H., Chan, K., Goh, B. H., Pusparajah, P., & Lee, L. (2017). *Vibrio vulnificus*: An Environmental and Clinical Burden. *Frontiers in Microbiology*, *8*. <https://doi.org/10.3389/fmicb.2017.00997>
- InterPro*. (n.d.). InterPro. <https://www.ebi.ac.uk/interpro/entry/InterPro/IPR036641/>
- Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., & Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nature Communications*, *9*(1). <https://doi.org/10.1038/s41467-018-07641-9>
- Kämpfer, P., & Glaeser, S. P. (2011). Prokaryotic taxonomy in the sequencing era - the polyphasic approach revisited. *Environmental Microbiology*, *14*(2), 291–317. <https://doi.org/10.1111/j.1462-2920.2011.02615.x>
- Karaolis, D. K., Lan, R., Kaper, J. B., & Reeves, P. R. (2001). Comparison of *Vibrio cholerae* Pathogenicity Islands in Sixth and Seventh Pandemic Strains. *Infection and Immunity*, *69*(3), 1947–1952. <https://doi.org/10.1128/iai.69.3.1947-1952.2001>
- Konstantinidis, K. T., & Tiedje, J. M. (2005). Genomic insights that advance the species definition for prokaryotes. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(7), 2567–2572. <https://doi.org/10.1073/pnas.0409727102>
- Lee, I., Kim, Y. O., Park, S. C., & Chun, J. (2016a). OrthoANI: An improved algorithm and software for calculating average nucleotide identity. *International Journal of Systematic and Evolutionary Microbiology*, *66*(2), 1100–1103. <https://doi.org/10.1099/ijsem.0.000760>

- Lee, I., Kim, Y. O., Park, S. C., & Chun, J. (2016b). OrthoANI: An improved algorithm and software for calculating average nucleotide identity. *International Journal of Systematic and Evolutionary Microbiology*, *66*(2), 1100–1103. <https://doi.org/10.1099/ijsem.0.000760>
- Lin, H., Yu, M., Wang, X., & Zhang, X. (2018). Comparative genomic analysis reveals the evolution and environmental adaptation strategies of vibrios. *BMC Genomics*, *19*(1). <https://doi.org/10.1186/s12864-018-4531-2>
- Lopes-Santos, L., Castro, D. B. A., Ferreira-Tonin, M., Corrêa, D. B. A., Weir, B. S., Park, D., Ottoboni, L. M. M., Neto, J. R., & Destéfano, S. a. L. (2017). Reassessment of the taxonomic position of *Burkholderia andropogonis* and description of *Robbsia andropogonis* gen. nov., comb. nov. *Antonie Van Leeuwenhoek*, *110*(6), 727–736. <https://doi.org/10.1007/s10482-017-0842-6>
- López-Aladid, R., Fernández-Barat, L., Alcaraz-Serrano, V., Bueno-Freire, L., Vázquez, N., Pastor-Ibáñez, R., Palomeque, A., Oscanoa, P., & Torres, A. (2023). Determining the most accurate 16S rRNA hypervariable region for taxonomic identification from respiratory samples. *Scientific Reports*, *13*(1). <https://doi.org/10.1038/s41598-023-30764-z>
- Lydon, K. A., & Lipp, E. K. (2018). Taxonomic annotation errors incorrectly assign the family Pseudoalteromonadaceae to the order Vibrionales in Greengenes: implications for microbial community assessments. *PeerJ*, *6*, e5248. <https://doi.org/10.7717/peerj.5248>
- Mahato, N. K., Gupta, V., Singh, P., Kumari, R., Verma, H., Tripathi, C., Rani, P., Sharma, A., Singhvi, N. R., Sood, U., Hira, P., Kohli, P., Nayyar, N., Puri, A., Bajaj, A., Kumar, R., Negi, V., Talwar, C., Khurana, H., . . . Lal, R. (2017). Microbial taxonomy in the era of OMICS: application of DNA sequences, computational tools and techniques. *Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology*, *110*(10), 1357–1371. <https://doi.org/10.1007/s10482-017-0928-1>



- Martins, V. G. P., Nascimento, J. D. S., Da Silva Martins, F. M., & Vigoder, H. C. (2022). Vibriosis and its impact on microbiological food safety. *Food Science and Technology*, 42. <https://doi.org/10.1590/fst.65321>
- McPherson, V., Watts, J. A., Simpson, L. M., & Oliver, J. D. (1991). Physiological effects of the lipopolysaccharide of *Vibrio vulnificus* on mice and rats. *PubMed*, 67(272–273), 141–149. <https://pubmed.ncbi.nlm.nih.gov/1779875>
- Medini, D., Donati, C., Tettelin, H., Massignani, V., & Rappuoli, R. (2005). The microbial pan-genome. *Current Opinion in Genetics & Development*, 15(6), 589–594. <https://doi.org/10.1016/j.gde.2005.09.006>
- Montero, D. A., Vidal, R., Velasco, J., George, S., Lucero, Y., Gómez, L. A., Carreño, L. J., García-Betancourt, R., & O’Ryan, M. (2023). *Vibrio cholerae*, classification, pathogenesis, immune response, and trends in vaccine development. *Frontiers in Medicine*, 10. <https://doi.org/10.3389/fmed.2023.1155751>
- Mutreja, A., Kim, D. W., Thomson, N. R., Connor, T. R., Lee, J., Kariuki, S., Croucher, N. J., Choi, S. Y., Harris, S. R., Lebens, M., Niyogi, S. K., Kim, E. J., Ramamurthy, T., Chun, J., Wood, J., Clemens, J. D., Czerkinsky, C., Nair, G. B., Holmgren, J., . . . Dougan, G. (2011). Evidence for several waves of global transmission in the seventh cholera pandemic. *Nature*, 477(7365), 462–465. <https://doi.org/10.1038/nature10392>
- Ng, W. L., & Bassler, B. L. (2009). Bacterial Quorum-Sensing Network Architectures. *Annual Review of Genetics*, 43(1), 197–222. <https://doi.org/10.1146/annurev-genet-102108-134304>
- Pannekoek, Y., Qin, Q., Zhang, Y., & Van Der Ende, A. (2016). Genus delineation of *Chlamydiales* by analysis of the percentage of conserved proteins justifies the reunifying of the genera *Chlamydia* and *Chlamydophilum* into one single genus *Chlamydia*. *Pathogens and Disease*, 74(6), ftw071. <https://doi.org/10.1093/femspd/ftw071>

- Pant, A., Bag, S., Saha, B., Verma, J., Kumar, P., Banerjee, S., Kumar, B., Kumar, Y., Desigamani, A., Maiti, S., Maiti, T. K., Banerjee, S. K., Bhadra, R. K., Koley, H., Dutta, S., Nair, G. B., Ramamurthy, T., & Das, B. (2020). Molecular insights into the genome dynamics and interactions between core and acquired genomes of *Vibrio cholerae*. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(38), 23762–23773. <https://doi.org/10.1073/pnas.2006283117>
- Park, K. H., Kim, J., Lee, Y. R., Moon, Y. J., Hur, H., Choi, Y. H., Kim, C. H., Kim, U., Song, E. K., Yoo, W. H., Lee, C. S., Kim, B. S., Lee, S. H., Ryu, P. Y., & Han, M. (2007). Low-density lipoprotein protects *Vibrio vulnificus*-induced lethality through blocking lipopolysaccharide action. *Experimental and Molecular Medicine*, *39*(5), 673–678. <https://doi.org/10.1038/emm.2007.73>
- Pennetzdorfer, N., Höfler, T., Wölflingseder, M., Tutz, S., Schild, S., & Reidl, J. (2021).  $\sigma^E$  controlled regulation of porin OmpU in *Vibrio cholerae*. *Molecular Microbiology*, *115*(6), 1244–1261. <https://doi.org/10.1111/mmi.14669>
- Physiological effects of the lipopolysaccharide of Vibrio vulnificus on mice and rats*. (1991). PubMed. <https://pubmed.ncbi.nlm.nih.gov/1779875/>
- Qin, Q., Xie, B., Zhang, X., Chen, X., Zhou, B., Zhou, J., Oren, A., & Zhang, Y. (2014). A proposed genus boundary for the prokaryotes based on genomic insights. *Journal of Bacteriology*, *196*(12), 2210–2215. <https://doi.org/10.1128/jb.01688-14>
- Ramamurthy, T., Nandy, R. K., Mukhopadhyay, A. K., Dutta, S., Mutreja, A., Okamoto, K., Miyoshi, S. I., Nair, G. B., & Ghosh, A. (2020). Virulence Regulation and Innate Host Response in the Pathogenicity of *Vibrio cholerae*. *Frontiers in Cellular and Infection Microbiology*, *10*. <https://doi.org/10.3389/fcimb.2020.572096>

- Rosselló-Móra, R. (2006). DNA-DNA reassociation methods applied to microbial taxonomy and their critical evaluation. In *Springer eBooks* (pp. 23–50). [https://doi.org/10.1007/978-3-540-31292-5\\_2](https://doi.org/10.1007/978-3-540-31292-5_2)
- Sack, D. A., Sack, R. B., Nair, G. B., & Srivastava, A. K. (2004). Cholera. *The Lancet*, 363(9404), 223–233. [https://doi.org/10.1016/s0140-6736\(03\)15328-7](https://doi.org/10.1016/s0140-6736(03)15328-7)
- Sampaio, A., Silva, V., Poeta, P., & Aonofriesei, F. (2022). *Vibrio* spp.: Life Strategies, Ecology, and Risks in a Changing Environment. *Diversity*, 14(2), 97. <https://doi.org/10.3390/d14020097>
- Shikuma, N. J., Fong, J. C. N., Odell, L. S., Perchuk, B. S., Laub, M. T., & Yildiz, F. H. (2009). Overexpression of VpsS, a Hybrid Sensor Kinase, Enhances Biofilm Formation in *Vibrio cholerae*. *Journal of Bacteriology*, 191(16), 5147–5158. <https://doi.org/10.1128/jb.00401-09>
- Stock, A., Robinson, V., & Goudreau, P. N. (2000). Two-Component signal transduction. *Annual Review of Biochemistry*, 69(1), 183–215. <https://doi.org/10.1146/annurev.biochem.69.1.183>
- Takemura, A. F., Chien, D. M., & Polz, M. F. (2014). Associations and dynamics of Vibrionaceae in the environment, from the genus to the population level. *Frontiers in Microbiology*, 5. <https://doi.org/10.3389/fmicb.2014.00038>
- Tettelin, H., Massignani, V., Cieslewicz, M. J., Donati, C., Medini, D., Ward, N., Angiuoli, S. V., Crabtree, J., Jones, A., Durkin, A. S., DeBoy, R. T., Davidsen, T. M., Mora, M., Scarselli, M., Ros, I. M. Y., Peterson, J., Hauser, C. R., Sundaram, J. P., Nelson, K. E., . . . Fraser, C. M. (2005). Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial “pan-genome.” *Proceedings of the National Academy of Sciences of the United States of America*, 102(39), 13950–13955. <https://doi.org/10.1073/pnas.0506758102>
- The Editors of Encyclopaedia Britannica. (2023, September 25). *Vibrio* | Marine, pathogenic, infectious. Encyclopedia Britannica. <https://www.britannica.com/science/vibrio>

- Thompson, J. R., & Polz, M. F. (2014). Dynamics of *Vibrio* Populations and Their Role in Environmental Nutrient Cycling. In *ASM Press eBooks* (pp. 190–203). <https://doi.org/10.1128/9781555815714.ch13>
- Tsai, M., Liu, Y., Soo, V., & Chen, C. (2019). A new Genome-to-Genome Comparison Approach for Large-Scale Revisiting of current microbial taxonomy. *Microorganisms*, 7(6), 161. <https://doi.org/10.3390/microorganisms7060161>
- Wall, L. (2011). The PERL programming language. *ResearchGate*. [https://www.researchgate.net/publication/265628667\\_The\\_PERL\\_Programming\\_Language](https://www.researchgate.net/publication/265628667_The_PERL_Programming_Language)
- Wayne, L. G., Brenner, D. J., Colwell, R. R., Grimont, P. a. D., Kandler, O., Krichevsky, M. I., Moore, L., Moore, W. E. C., Murray, R. G. E., Stackebrandt, E., Starr, M. P., & Trüper, H. G. (1987). Report of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics. *International Journal of Systematic and Evolutionary Microbiology*, 37(4), 463–464. <https://doi.org/10.1099/00207713-37-4-463>
- Williams, K. P., Gillespie, J. J., Sobral, B. W. S., Nordberg, E. K., Snyder, E. E., Shallom, J. M., & Dickerman, A. W. (2010). Phylogeny of gammaproteobacteria. *Journal of Bacteriology*, 192(9), 2305–2314. <https://doi.org/10.1128/jb.01480-09>
- Woese, C. R. (1987). Bacterial evolution. *Microbiological Reviews*, 51(2), 221–271. <https://doi.org/10.1128/mr.51.2.221-271.1987>
- World Health Organization: WHO & World Health Organization: WHO. (2022). Cholera. [www.who.int](http://www.who.int). <https://www.who.int/en/news-room/fact-sheets/detail/cholera>
- Zhang, X., Lin, H., Wang, X., & Austin, B. (2018). Significance of *Vibrio* species in the marine organic carbon cycle—A review. *Science China-earth Sciences*, 61(10), 1357–1368. <https://doi.org/10.1007/s11430-017-9229-x>

Zhao, W., Caro, F., Robins, W., & Mekalanos, J. J. (2018). Antagonism toward the intestinal microbiota and its effect on *Vibrio cholerae* virulence. *Science*, 359(6372), 210–213.  
<https://doi.org/10.1126/science.aap8775>