

ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΥΓΕΙΑΣ
ΤΜΗΜΑ ΒΙΟΧΗΜΕΙΑΣ ΚΑΙ ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ

ΕΡΓΑΣΤΗΡΙΟ ΠΕΡΙΒΑΛΛΟΝΤΙΚΗΣ ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ
ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

«ΣΥΓΚΡΙΤΙΚΗ ΓΟΝΙΔΙΩΜΑΤΙΚΗ ΜΕΛΕΤΗ ΣΤΕΛΕΧΩΝ
ΡΑΕΝΑΡΤΗΡΟΒΑΚΤΕΡ ΠΟΥ ΜΠΟΡΟΥΝ ΝΑ ΚΑΤΑΒΟΛΙΣΟΥΝ ΤΟ
ΦΥΤΟΦΑΡΜΑΚΟ ΙΠΡΟΔΙΟΝΕ.»

«COMPARATIVE GENOMIC ANALYSIS OF PAENARTHROBACTER
STRAINS INVOLVED IN THE DEGRADATION OF IPRADIONE»



Όνοματεπώνυμο: Ζαφρανάς Αλέξανδρος
Πατρώνυμο: Νικόλαος
Ακαδημαϊκό έτος: 2021
Επιβλέπων καθηγητής: Καρπούζας Δημήτριος

ΤΡΙΜΕΛΗΣ ΕΠΙΤΡΟΠΗ

Καρπούζας Δημήτριος (Υπεύθυνος καθηγητής), Καθηγητής Περιβαλλοντικής Μικροβιολογίας και Βιοτεχνολογίας, Τμήμα Βιοχημείας και Βιοτεχνολογίας, Πανεπιστήμιο Θεσσαλίας

Βασιλειάδης Σωτήριος, Επίκουρος Καθηγητής Μοριακής Μικροβιακής Οικολογίας – Γονιδιωματικής, Τμήμα Βιοχημείας και Βιοτεχνολογίας, Πανεπιστήμιο Θεσσαλίας

Παναγιώτης Καράς, Διδάσκων, Τμήμα Βιοχημείας και Βιοτεχνολογίας, Πανεπιστήμιο Θεσσαλίας

ΕΥΧΑΡΙΣΤΙΕΣ

Η παρούσα πτυχιακή εργασία εκπονήθηκε στο εργαστήριο Περιβαλλοντικής Βιοτεχνολογίας του Τμήματος Βιοχημείας και Βιοτεχνολογίας, υπό την επίβλεψη του καθηγητή κ. Δημήτριου Καρπούζα. Θα ήθελα αρχικά, να τον ευχαριστήσω για την ανάθεση μίας ενδιαφέρουσας πτυχιακής εργασίας, για την εμπιστοσύνη που έδειξε στο πρόσωπο μου καθώς και για την βοήθεια και καθοδήγηση του.

Επιπροσθέτως, θα ήθελα να ευχαριστήσω τον Επίκουρο Καθηγητή κ. Βασιλειάδη Σωτήριο και τον κ. Παναγιώτη Καραά που δέχτηκαν να συμμετέχουν στην τριμελή συμβουλευτική επιτροπή. Μάλιστα, ιδιαίτερες ευχαριστίες οφείλω, στον Επίκουρο Καθηγητή κ. Βασιλειάδη Σωτήριο, ο οποίος με τη συνεχή καθοδήγηση, συμβουλές και μετάδοση γνώσεων, συνέβαλε καθοριστικά στην διεκπεραίωση της παρούσας πτυχιακής εργασίας.

Θα ήθελα επίσης να ευχαριστήσω όλη την ομάδα του εργαστηρίου για το φιλικό κλίμα και την προθυμία να απαντήσουν σε οποιαδήποτε απορία μου.

Τέλος, θέλω να ευχαριστήσω την οικογένεια μου και τους κοντινούς μου φίλους, η συμβολή και στήριξη των οποίων ήταν καθοριστική, στο σύνολο της διάρκειας των σπουδών μου.

ΠΕΡΙΛΗΨΗ

Το *iprodione* είναι ένα μυκητοκτόνο ευρείας χρήσης με αρνητικές επιπτώσεις στη βιολογία οργανισμών μη στόχων και στο περιβάλλον κατά την αλόγιστη εφαρμογή του, και για το οποίο έχει παρατηρηθεί επιταχυνόμενη βιο-αποδόμηση από μικροοργανισμούς σε αξενικές καλλιέργειες. Φαινότυποι αποδόμησης του *iprodione* παρατηρήθηκαν σε τρεις (3) απόπειρες απομόνωσης μικροοργανισμών του εργαστηρίου μας, που κατετάγησαν στο είδος *Raenarthrobacter* με τη χρήση φυλογενετικής ανάλυσης του δείκτη 16S rRNA γονιδίου. Από τις καλλιέργειες αυτές εκχυλίστηκε το DNA των μικροοργανισμών και αλληλουχήθηκε με μεθόδους 2^{ης} και 3^{ης} γενιάς αλληλούχησης. Έπειτα, συναρμολογήθηκαν τα γονιδιώματα και τα πλασμίδια των καλλιεργειών και ταυτοποιήθηκαν. Ακολούθησε δομικός και λειτουργικός χαρακτηρισμός των γονιδιωμάτων, και ταυτόχρονα, βάσει γνωστού μεταβολικού μονοπατιού, έγινε έλεγχος των γονιδιωμάτων (δυνητικών χρωμοσωμάτων και πλασμιδίων) για σχετικές αλληλουχίες. Συστηματική ταξινόμηση των συνεχειών ακολουθιών DNA του γονιδιώματος έδειξε ότι στις δύο από τις τρεις καλλιέργειες (T1.6, T1.8) βρέθηκαν χρωμοσωμικές αλληλουχίες που ανήκαν σε δύο μικροοργανισμούς που ταυτοποιήθηκαν ως *Arthrobacter* και *Microbacterium* και επομένως οι καλλιέργειες δεν ήταν αξενικές. Επιπλέον, βρέθηκε ότι οι δύο αυτές καλλιέργειες είναι γενετικά πανομοιότυπες. Οι αλληλουχίες ανοιχτών πλαισίων ανάγνωσης υποθετικά υπεύθυνες για τα γονίδια υπεύθυνα για το μεταβολικό μονοπάτι του *iprodione* βρέθηκαν σε αυτά τα γονιδιώματα μετά από σύγκριση με γονίδια αναφοράς. Συγκριτική γονιδιωματική ανάλυση των γονιδιακών τόπων όπου βρέθηκαν τα ομόλογα γονίδια καθώς και ανάλυση για μεταθετά γενετικά στοιχεία έδειξαν ότι υπάρχουν επανεμφανιζόμενα γενετικά μοτίβα στις περιοχές των γονιδίων υπεύθυνων για τον καταβολισμό του *iprodione* και παρατηρούνται μεταθετά στοιχεία στα χρωμοσώματα και στα πλασμίδια των υπό μελέτη μικροοργανισμών παραπέμποντας σε φαινόμενα οριζόντιας γονιδιακής μεταφοράς.

ABSTRACT

Iprodione is a widely used fungicide with negative effects on the biology of non-target organisms and the environment during its reckless application, and for which accelerated biodegradation by microorganisms has been observed in valuable crops. Iprodione degradation phenotypes were observed in three (3) attempts to isolate microorganisms in our laboratory, classified as *Paenarthrobacter* using phylogenetic analysis of the 16S rRNA gene marker. The DNA of the microorganisms was extracted from these cultures and sequenced by 2nd and 3rd generation sequencing methods. The genomes and plasmids of the cultures were then assembled and identified. Structural and functional characterization of the genomes followed, and at the same time, based on a known metabolic pathway, the genomes (potential chromosomes and plasmids) were tested for relevant DNA sequences. Systematic classification of genome sequences showed that in two of the three cultures (T1.6, T1.8) chromosomal sequences belonging to two microorganisms were identified that were identified as *Arthrobacter* and *Microbacterium* and therefore the cultures were not valuable. In addition, these two cultures were found to be genetically identical. The open reading frames hypothetically responsible for the genes responsible for the Iprodione metabolic pathway were found in these genomes after comparison with reference genes. Comparative genomic analysis of the gene sites where the homologous genes were found as well as analysis for transposable genetic elements showed that there are recurring genetic patterns in the regions of the genes responsible for iprodione catabolism. Also, transposable elements were observed in the chromosomes and plasmids of the studied microorganisms, hinting towards lateral gene transfer.

ΠΕΡΙΕΧΟΜΕΝΑ

1 Εισαγωγή

- 1.1 Φυτοφάρμακα
- 1.2 Iprodione
- 1.3 Περιβαλλοντικές επιπτώσεις και απορρύπανση
- 1.4 Γένη *Arthrobacter*, *Paenarthrobacter* και *Microbacterium*
- 1.5 Προτεινόμενο μεταβολικό μονοπάτι καταβολισμού του Iprodione
- 1.6 Υδρολάσες και Αμιδάσες
- 1.7 Στοιχεία μεθοδολογικής προσέγγισης συγκριτικής γονιδιοματικής
 - 1.7.1 Χαρακτηρισμός γονιδιώματος
 - Δομικός χαρακτηρισμός γονιδίων/μη-κωδικευμένων στοιχείων
 - Αλγόριθμοι και μέθοδοι χαρακτηρισμού λειτουργικών γονιδίων
→ Αλγόριθμοι αναζήτησης BLAST και HMMER
 - 1.7.2 Ανάλυση οριζόντια μεταθετών στοιχείων γονιδιώματος
 - 1.7.3 Ανάλυση synteny
 - 1.7.4 Ανάλυση του πανγονιδιώματος
- 1.8 Προηγούμενη εργασία και στόχοι

2 Μέθοδοι

- 2.1 Χαρακτηρισμός γονιδιώματος
 - 2.1.1 Λογισμικό και περιβάλλον
 - 2.1.2 Δομικός και λειτουργικός χαρακτηρισμός
 - PROKKA
 - BLAST
 - HMMER
- 2.2 Στατιστική ανάλυση και παρουσίαση δεδομένων
- 2.3 Γονίδια αναφοράς και περιγραφή γονιδιωμάτων των δειγμάτων

3 Αποτελέσματα

- 3.1 Γονίδια υπεύθυνα για τον μεταβολισμό του Iprodione στα υπό μελέτη στελέχη
- 3.2 Μεταθετά στοιχεία
- 3.3 Γονιδιακοί τόποι των υπεύθυνων γονιδίων για τον καταβολισμό του Iprodione
- 3.4 Σύγκριση γονιδιακών τόπων των γονιδίων υπεύθυνων για τον καταβολισμό του Iprodione
- 3.5 Σύγκριση γονιδιωμάτων *Paenarthrobacter*
- 3.6 Ανάλυση του πανγονιδιώματος του γένους *Paenarthrobacter*

4 Συζήτηση

5 Συμπεράσματα-μελλοντική έρευνα

6 Βιβλιογραφία

7 Παραρτήματα

- 7.1 παράρτημα I

Αναλυμένα αποτελέσματα για τις στοιχίσεις IpaH στα ανοιχτά πλαίσια ανάγνωσης των βασικών χρωμοσωμάτων του δείγματος T1.6 (*Paenarthrobacter* και *Microbacterium*)

7.2 παραρτημα II

Αναλυμένα αποτελέσματα για τις στοιχίσεις DdaH στα ανοιχτά πλαίσια ανάγνωσης των βασικού χρωμοσώματος του *Paenarthrobacter* του δείγματος T1.6

7.3 παραρτημα III

Αναλυμένα αποτελέσματα για τις στοιχίσεις DuaH στα ανοιχτά πλαίσια ανάγνωσης των βασικού χρωμοσώματος του *Microbacterium* του δείγματος T1.6

7.4 παραρτημα IV

Αποτελέσματα στοιχίσεων BLAST για το δείγμα T1.8

7.5 παραρτημα V

Κατανομή γονιδίων τρανσποζάσης στο πλασμίδιο 2 του δείγματος T1.6

7.6 παραρτημα VI

Κατανομή γονιδίων τρανσποζάσης στο πλασμίδιο 1 του δείγματος T1.6

7.7 παραρτημα VII

Τμήματα DNA που αναλύθηκαν, τα οποία περιέχουν τα υπό μελέτη ανοιχτά πλαίσια ανάγνωσης και τμήματα DNA που χρησιμοποιήθηκαν για τον έλεγχο synteny.

7.8 παραρτημα VIII

Συμπληρωματικό υλικό της ανάλυσης του πανγονιδιώματος

1. Εισαγωγή

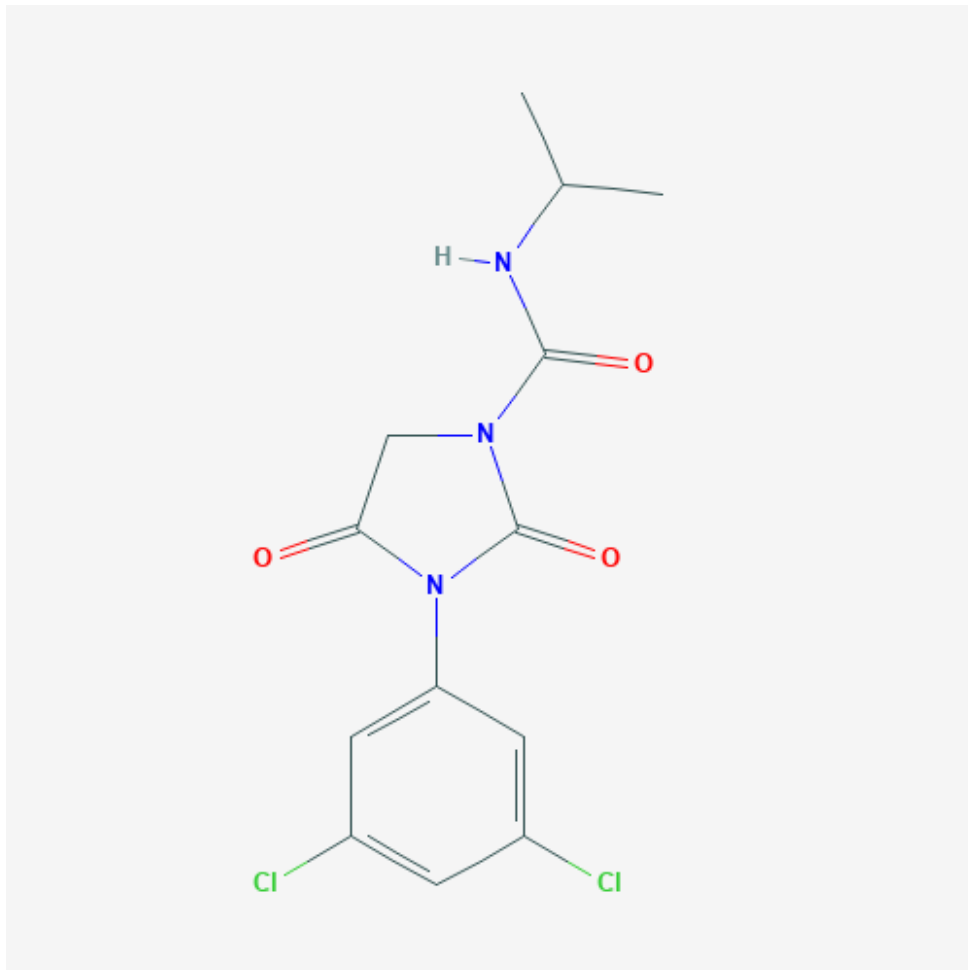
1.1 Φυτοφάρμακα

Τα φυτοφάρμακα είναι ουσίες που προορίζονται για τον έλεγχο των παρασίτων, συμπεριλαμβανομένων των ζιζανίων (*Al-Saleh IA Journal of Environmental Pathology, Toxicology and Oncology 1994*). Ο όρος φυτοφάρμακα περιλαμβάνει όλα από τα ακόλουθα: ζιζανιοκτόνα, εντομοκτόνα (όπως ρυθμιστές ανάπτυξης εντόμων, τερμιτοκτόνα), νηματωδοκτόνα, μαλακιοκτόνα, ιχθυοκτόνα, πτηνοκτόνα, τρωκτικοκτόνα, βακτηριοκτόνα, εντομοαπωθητικά, απωθητικά ζώων, αντιμικροβιακά, μυκητοκτόνα, απολυμαντικά, και αποστειρωτικά. Τα ζιζανιοκτόνα αντιπροσωπεύουν το 80% της χρήσης φυτοφαρμάκων, και προστατεύουν τα φυτά (και τις καλλιέργειες) από ζιζάνια, μύκητες ή έντομα. Γενικά, ένα φυτοφάρμακο είναι μια χημική ουσία ή βιολογικός παράγοντας (όπως ιός, βακτήριο ή μύκητας) που καταστέλλει, εξουδετερώνει, σκοτώνει ή απλά αποθαρρύνει τα παράσιτα. Τα φυτοφάρμακα έχουν οφέλη και ορισμένα μειονεκτήματα, όπως πιθανή τοξικότητα για τον άνθρωπο και άλλα είδη.

1.2 Iprodione

Το iprodione [3-(3,5-dichlorophenyl) N-isopropyl-2,4- dioxoimidazolidine-1-carboxamide] είναι ένα δικαρβοξαμίδικό μυκητοκτόνο το οποίο χρησιμοποιείται ευρέως για την προστασία καλλιεργειών από διάφορους μύκητες που προκαλούν προβλήματα (*Miñambres et al.2010 · Grabke et al.2014*).

Το iprodione χρησιμοποιείται σε καλλιέργειες που προσβάλλονται από μύκητες του γένους *Botrytis*, του γένους *Sclerotinia* και άλλες μυκητολογικές ασθένειες στα φυτά (*FAO, 2004*). Το iprodione εφαρμόζεται σε μια ποικιλία καλλιεργειών: φρούτα, λαχανικά, καλλωπιστικά δέντρα και θάμνους και σε γκαζόν. Είναι ένα μυκητοκτόνο επαφής που εμποδίζει τη βλάστηση των σπορίων των μυκήτων και εμποδίζει την ανάπτυξη του μυκηλίου τους.



Εικόνα 1 : Η χημική δομή του φυτοφαρμάκου iprodione (<https://pubchem.ncbi.nlm.nih.gov/compound/Iprodione>).

1.3 Περιβαλλοντικές επιπτώσεις και απορρύπανση

Το iprodione έχει χαρακτηριστεί ως μερικώς τοξικό για μικρά ζώα (*Blystone et al. 2007*) και ως πιθανό καρκινογόνο για τον άνθρωπο (*U.S. Environmental Protection Agency (USEPA), 1998*). Το iprodione παρουσιάζει μεγάλη κινητικότητα στο έδαφος. Το εδαφικό pH αποτελεί τον βασικό παράγοντα που επηρεάζει το ποσοστό αποδόμησης του iprodione, με τα αλκαλικά εδάφη να παρουσιάζουν τα μεγαλύτερα ποσοστά αποδόμησης (*Walker et al. 1987*). Ωστόσο, έχει βρεθεί ότι επαναλαμβανόμενες δόσεις iprodione σε εδάφη είναι ικανές να προκαλέσουν επιταχυνόμενη αποικοδόμηση του iprodione (*Martin et al. 1990 · Slade et al. 1992 · Mitchel et al. 1996*). Απομονώσεις από τέτοια εδάφη σε εκλεκτικά θρεπτικά μέσα οδήγησαν σε επιλογή στελεχών με ικανότητα επιταχυνόμενης αποδόμησης του iprodione στο παρελθόν, με στελέχη που ανήκουν στα γένη *Paenarthrobacter* και *Arthrobacter* να κυριαρχούν την ομάδα αυτών των στελεχών (*Campos et al. 2015 · Campos et al. 2017 · Zhang et al. 2021*).

1.4 Γένη *Arthrobacter*, *Paenarthrobacter* και *Microbacterium*

Το γένος *Arthrobacter* ανήκει στην οικογένεια των *Micrococcaceae*, στην τάξη των *Micrococcales* στην κλάση των *Actinobacteria*, στο φύλο των *Actinobacteria*, στην επικράτεια των βακτηρίων (*Conn & Dimmick 1947*).

Τα *Arthrobacteria* χαρακτηρίζονται από πλεομορφισμό (μεταβλητό σχήμα) και μεταβλητότητα Gram (χρώση θετικό ή αρνητικό) αν και γενετικά προέρχονται από τα θετικά κατά Gram *Actinobacteria*. Τα *Arthrobacteria* είναι διατροφικά ευπροσάρμοστα, χρησιμοποιώντας μια ποικιλία υποστρωμάτων στον οξειδωτικό μεταβολισμό τους, συμπεριλαμβανομένης της νικοτίνης, νουκλεϊκών οξέων και διαφόρων ζιζανιοκτόνων και φυτοφαρμάκων. Τα περισσότερα είδη *Arthrobacter* είναι υποχρεωτικά αερόβια, αλλά έρευνες έδειξαν ότι τουλάχιστον δύο είδη *Arthrobacter*, τα *A. globiformis* και *A. nicotianae*, παρουσιάζουν αναερόβιο μεταβολισμό. Στα ανώτερα στρώματα του εδάφους που κατοικούνται από *Arthrobacteria* οι αλλαγές στις συγκεντρώσεις οξυγόνου είναι συχνές και κάποια είδη *Arthrobacter* έχουν προσαρμόσει στρατηγικές ανάπτυξης ανεξάρτητες οξυγόνου προκειμένου να επιβιώσουν σε περιόδους που αυτό είναι περιορισμένο. Αυτά τα είδη χρησιμοποιούν νιτρικά ως δέκτη ηλεκτρονίων στο τέλος της αναπνευστικής τους αλυσίδας, ανάγοντας τα σε αμμωνία μέσω νιτρώδων (*Eschbach et al. 2003*).

Το γένος *Paenarthrobacter* ανήκει στην οικογένεια των *Micrococcaceae*, στην τάξη των *Micrococcales* στην κλάση των *Actinobacteria*, στο φύλο των *Actinobacteria*, στην επικράτεια των βακτηρίων (*Busse et al. 2016*).

Τα στελέχη του γένους *Paenarthrobacter* μοιράζονται τις παρακάτω ιδιότητες:

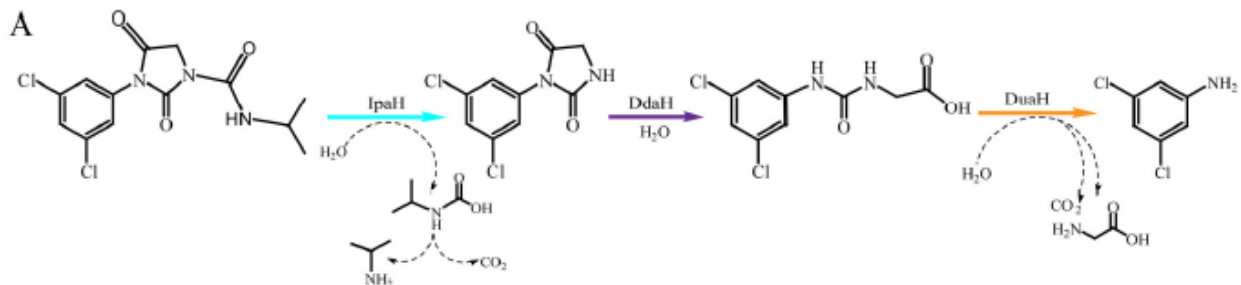
- Πεπτιδογλυκάνη τύπου A3a
- Menaquinone MK-9(H2) (βιταμίνη K2)
- Προφίλ πολικών λιπιδίων το οποίο συνίσταται από diphosphatidylglycerol,

- phosphatidylglycerol, phosphatidylinositol, dimannosylglyceride, και monogalactosyldiacylglycerol, και σε μικρότερο βαθμό trimannosyldiacylglycerol.
- Σύσταση λιπαρών οξέων κυρίως από: anteiso-C15:0, και σε μικρότερο βαθμό iso-C15:0, iso-C16:0, anteiso-C17:0 και iso-C14.
 - Χρωμοσωμική περιεκτικότητα GC σε εύρος 61.3–62.5%.

Το γένος *Microbacterium* ανήκει στην οικογένεια των *Microbacteriaceae*, στην τάξη των *Micrococcales* στην κλάση των *Actinobacteria*, στο φύλο των *Actinobacteria*, στην επικράτεια των βακτηρίων (*Orla-Jensen et al. 1919*).

1.5 Προτεινόμενο μεταβολικό μονοπάτι καταβολισμού του Iprodione

Απομόνωση και συγκριτική γονιδιωματική ανάλυση βακτηρίων που είναι ικανά να διασπούν το iprodione οδήγησε στην διευκρίνιση του μονοπατιού μεταβολισμού του iprodione και των γονιδίων που είναι υπεύθυνα για κάθε στάδιο του μονοπατιού (*Zhang et al. 2021*).

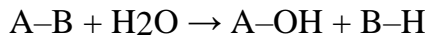


Εικόνα 2: Πιθανό μονοπάτι καταβολισμού του Iprodione από το στέλεχος *Paenarthrobacter YJN-D* και ένζυμα που συμμετέχουν.

Το πρώτο στάδιο του μεταβολισμού καταλύεται από το ένζυμο IpaH, μία αμιδάση, η οποία είναι υπεύθυνη για την διάσπαση του αμιδικού δεσμού μεταξύ του N-1 του ιμιδαζολικού δακτυλίου και του ισοπροπυλαμιδίου. Το δεύτερο στάδιο του μεταβολισμού καταλύεται από το ένζυμο DdaH, μία διακετυλάση του N-(3,5 – διχλωροφαινύλ)-2,4-διόξιμιδαζολίου, η οποία υδρολύει τον N-3 αμιδικό δεσμό του ιμιδαζολίου. Τέλος, το τρίτο στάδιο του μεταβολισμού καταλύεται από το ένζυμο DuaH, την υδρολάση του (3,5 – διχλωροφαινυλουρία) – οξικού οξέος, η οποία υδρολύει τον N-1 αμιδικό δεσμό του διχλωροφαινυλίου (*Zhang et al. 2021*).

1.6 Υδρολάσες και Αμιδάσες

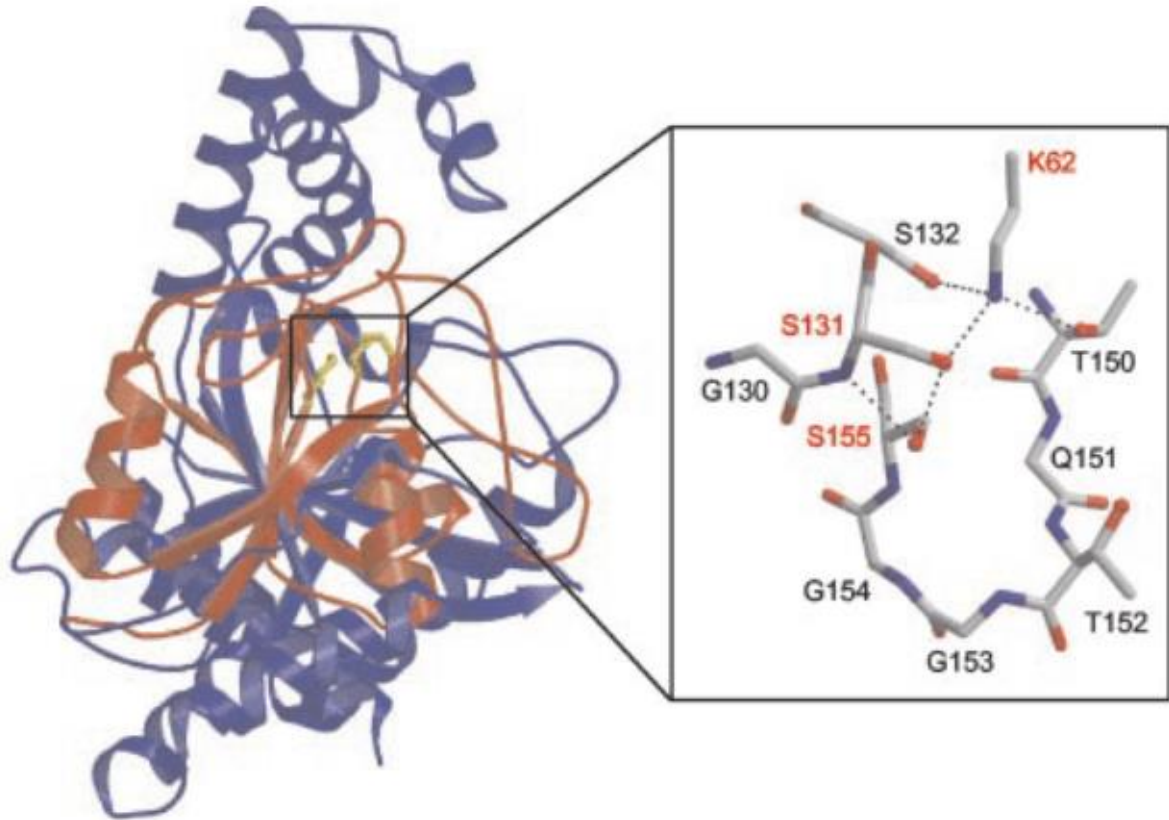
Οι υδρολάσες είναι μια κατηγορία ενζύμων που λειτουργούν ως βιοχημικοί καταλύτες που χρησιμοποιούν νερό (H₂O) για να σπάσουν έναν χημικό δεσμό, κάτι που έχει ως αποτέλεσμα τη διαίρεση ενός μεγαλύτερου μορίου σε μικρότερα μόρια. Για παράδειγμα, κάθε ένζυμο που καταλύει την ακόλουθη αντίδραση είναι υδρολάση:



όπου η ένωση A-B αποτελεί χημικό δεσμό μεταξύ μη συγκεκριμένων μορίων. Οι υδρολάσες ταξινομούνται ως EC 3 στην EC (Enzyme Commission number (EC number)) κατάταξη των ενζύμων. Μπορούν να ταξινομηθούν επιπλέον, ανάλογα με τον τύπο δεσμού που υδρολύουν (*Encyclopedia Britannica, Accessed 27 September 2021*). Ένας από τους τύπους δεσμών που υδρολύουν είναι οι δεσμοί αζώτου-άνθρακα. Οι υδρολάσες που δρουν επί τέτοιου τύπου δεσμών χωρίζονται σε πεπτιδάσες (EC 3.4), οι οποίες δρουν σε πεπτιδικούς δεσμούς και υδρολάσες μη πεπτιδικών δεσμών αζώτου-άνθρακα (EC 3.5). Οι αμιδάσες αποτελούν υποκατηγορία των EC 3.5 υδρολασών.

Στην ενζυμολογία, οι αμιδάσες (EC 3.5.1, acylamidase, acylase, amidohydrolase, deaminase, fatty acylamidase, N-acetylaminohydrolase) είναι ένα ένζυμο το οποίο καταλύει την υδρόλυση ενός αμιδικού δεσμού (*Valiña et al. 2004*).

Οι αμιδάσες (AS - Amidase signature) είναι μια μεγάλη ομάδα υδρολυτικών ενζύμων που περιέχουν μια συντηρημένη περιοχή περίπου 130 αμινοξέων γνωστών ως η αλληλουχία AS. Τα ένζυμα αυτά βρίσκονται τόσο σε προκαρυώτες όσο και σε ευκαρυώτες. Τα ένζυμα AS καταλύουν την υδρόλυση αμιδικών δεσμών (CO-NH₂), αν και η οικογένεια των ενζύμων αυτών είναι ευρεία όσον αφορά την ειδικότητα προς το υπόστρωμα και τη λειτουργία του κάθε ενζύμου. Ωστόσο, αυτά τα ένζυμα διατηρούν μια βασική δομή άλφα / βήτα / άλφα, όπου οι τοπολογίες των N- και C-τελικών άκρων είναι παρόμοιες. Τα ένζυμα AS έχουν χαρακτηριστικά μια ιδιαίτερα διατηρημένη καρβοξυτελική περιοχή πλούσια σε κατάλοιπα σερίνης και γλυκίνης, αλλά στερούνται κατάλοιπα ασπαρτικού οξέος και ιστιδίνης, επομένως διαφέρουν από τις κλασικές υδρολάσες σερίνης. Αυτά τα ένζυμα διαθέτουν μια μοναδική, πολύ συντηρημένη καταλυτική τριάδα Ser-Ser-Lys που χρησιμοποιείται για υδρόλυση αμιδίου, αν και ο καταλυτικός μηχανισμός για τον ενδιάμεσο σχηματισμό ακυλο-ενζύμου μπορεί να διαφέρει μεταξύ των ενζύμων.



Εικόνα 3: Τρισδιάστατη δομή αμιδάσης. Με κόκκινο είναι υπογραμμισμένη η συντηρημένη AS αλληλουχία, ενώ με μπλε απεικονίζεται η μη συντηρημένη περιοχή της αμιδάσης. Η καταλυτική τριάδα Ser-Ser-Lys απεικονίζεται σε μεγέθυνση και παρίσταται το ενεργό κέντρο του ενζύμου (*Shin et al. 2003*).

1.7 Στοιχεία μεθοδολογικής προσέγγισης συγκριτικής γονιδιωματικής

1.7.1 Χαρακτηρισμός γονιδιώματος

Ο σχολιασμός γονιδιώματος (Genome annotation) είναι η διαδικασία αναγνώρισης λειτουργικών στοιχείων κατά μήκος της ακολουθίας ενός γονιδιώματος (*Josep et al. 2019*). Τις τελευταίες τρεις δεκαετίες, ο σχολιασμός γονιδιώματος εξελίχθηκε από τον υπολογισμό σχολιασμού μερικών γονιδίων που κωδικοποιούν πρωτεΐνες σε κάποια λίγα γονιδιώματα και τον πειραματικό σχολιασμό σύντομων ρυθμιστικών στοιχείων σε μικρό αριθμό από αυτά, σε σχολιασμό πληθυσμού μεμονωμένων νουκλεοτιδίων σε χιλιάδες γονιδιώματα (πολλά ανά είδος). Αυτή η αυξημένη ανάλυση και η περιεκτικότητα των σχολιασμών γονιδιώματος (από γονότυπους έως φαινότυπους) οδηγεί σε ακριβείς γνώσεις για τη βιολογία των ειδών, των πληθυσμών και των ατόμων.

Ο σχολιασμός γονιδιώματος αποτελείται από τρία βασικά βήματα:

1. Προσδιορισμός τμημάτων του γονιδιώματος που δεν κωδικοποιούν πρωτεΐνες

2. Προσδιορισμός δομικών στοιχείων στο γονιδίωμα (gene prediction)(δομικός σχολιασμός)
3. Χαρακτηρισμός βιολογικών πληροφοριών σε αυτά τα στοιχεία

Τα αυτόματα εργαλεία σχολιασμού προσπαθούν να εκτελέσουν αυτά τα βήματα με την αξιοποίηση ανάλυσης υπολογιστή (*Lincoln Stein. 2001*).

1.7.1.1 Δομικός χαρακτηρισμός γονιδίων/μη-κωδικευμένων στοιχείων

Ο δομικός σχολιασμός συνίσταται στην αναγνώριση γονιδιωματικών στοιχείων.

- Ανοιχτά πλαίσια ανάγνωσης
- Δομή γονιδίων
- Περιοχές κωδικοποίησης
- Θέση ρυθμιστικών μοτίβων

Υπάρχει πληθώρα αυτοματοποιημένων μεθόδων δομικού σχολιασμού αλληλουχιών ή και ολόκληρων γονιδιωμάτων. Τα προγράμματα αυτά ποικίλουν ως προς την ταχύτητα, την αξιοπιστία και την ευχρηστία τους, όμως βασίζονται σε κοινά αξιώματα και μεθοδολογίες. Το πρώτο βήμα στον δομικό χαρακτηρισμό ενός γονιδιώματος είναι η αναγνώριση των ανοιχτών πλαισίων ανάγνωσης. Έπειτα ακολουθεί η ταυτοποίηση των γονιδίων που δεν μεταφράζονται, κυρίως των tRNA και rRNA γονιδίων. Επιπλέον, μπορεί να αξιοποιηθούν προγράμματα τα οποία εντοπίζουν περιοχές που αναγνωρίζονται από ένζυμα περιορισμού ή και προγράμματα που εντοπίζουν υποκινητές ή άλλες ρυθμιστικές αλληλουχίες.

1.7.1.2 Αλγόριθμοι και μέθοδοι χαρακτηρισμού λειτουργιών γονιδίων

Ο λειτουργικός σχολιασμός συνίσταται στη σύνδεση βιολογικών πληροφοριών σε γονιδιωματικά στοιχεία.

- Βιοχημική λειτουργία
- Βιολογική λειτουργία
- Ρύθμιση και αλληλεπιδράσεις
- Έκφραση

Τα προγράμματα που εκπονούν τον λειτουργικό χαρακτηρισμό μιας αλληλουχίας ή ενός γονιδιώματος συγκρίνουν τα γενετικά στοιχεία του δομικού χαρακτηρισμού με γνωστά στοιχεία αναφοράς και αποδίδουν νόημα σύμφωνα με την ομολογία ή και την συγγένεια και ομοιότητα που έχουν τα γενετικά στοιχεία υπό μελέτη με τα αντιπαραβαλλόμενα στοιχεία αναφοράς. Μια απλή μέθοδος σχολιασμού γονιδίων βασίζεται σε εργαλεία αναζήτησης που βασίζονται σε ομολογία για την αναζήτηση ομόλογων γονιδίων σε συγκεκριμένες βάσεις δεδομένων. Οι πληροφορίες που προκύπτουν από τις συγκρίσεις αυτές, χρησιμοποιούνται για τον σχολιασμό νέων γονιδίων και γονιδιωμάτων. Ορισμένες βάσεις δεδομένων

χρησιμοποιούν πληροφορίες του περιβάλλοντος και των συμφραζόμενων γενετικών στοιχείων στο σημείο σχολιασμού, βαθμολογίες ομοιότητας, πειραματικά δεδομένα και ενσωματώσεις άλλων πόρων για να παρέχουν σχολιασμούς γονιδιώματος μέσω της προσέγγισής τους. Άλλες βάσεις δεδομένων βασίζονται σε επιμελημένες πηγές δεδομένων, καθώς και σε μια σειρά από διαφορετικά εργαλεία λογισμικού (*Jonathan Pevsner. 2009*).

Αλγόριθμοι αναζήτησης BLAST και HMMER

Για την ανάλυση των δεδομένων χρησιμοποιούνται κυρίως δύο αλγόριθμοι, το βασικό εργαλείο αναζήτησης τοπικής στοίχισης (basic local alignment search tool, BLAST) (*Altschul et al. 1990*).

Το BLAST είναι ένας ευρετικός αλγόριθμος σύγκρισης αλληλουχιών ανά ζεύγη όπου μπορεί να επιτύχει ταχύες αναζητήσεων αλληλουχιών έναντι μεγάλων σχετικά βάσεων δεδομένων. Βασίζεται στον κατακερματισμό της πληροφορίας κάθε αλληλουχίας αναζήτησης ή αναφοράς σε μικρότερα τμήματα (υπό-αλληλουχίες ή «λέξεις» νουκλεϊκών οξέων ή αμινοξέων), πολλά από τα οποία είναι κοινά μεταξύ διαφορετικών αλληλουχιών. Με την ομαδοποίηση των κοινών αυτών τμημάτων στις βάσεις δεδομένων μειώνεται δραματικά η απαραίτητη υπολογιστική ισχύς αναζήτησής τους και συνεπώς επιταχύνεται η διαδικασία αυτή. Έτσι, μία αναζήτηση βασίζεται σε αυτές τις κοινές «λέξεις» οι οποίες μόλις αναγνωριστούν οδηγούν στον αναθεωρητικό περιορισμό της προς αναζήτηση βάσης δεδομένων στο τμήμα που φέρει μονάχα τις κοινές αυτές «λέξεις». Μόλις αναγνωριστούν αλληλουχίες που φέρουν αρκετές όμοιες λέξεις με την εκάστοτε αλληλουχία αναζήτησης, τότε εφαρμόζεται πιο ακριβής αλγόριθμος για την στοίχιση των αλληλουχιών στο σύνολό τους ο οποίος βασίζεται στον τοπικής στοίχισης αλγόριθμο των Smith-Waterman (*Smith et al. 1981*). Το HMMER βασίζεται σε πιθανολογικά μοντέλα εμφάνισης καταλοίπων αμινοξέων σε αλληλουχίες πρωτεϊνών, τα οποία μοντέλα προέρχονται από στοιχίσεις πολλαπλών αλληλουχιών της εκάστοτε οικογένειας πρωτεϊνών. Με τον τρόπο αυτόν αναζητούνται μονάχα τα λειτουργικά απαραίτητα για την κάθε οικογένεια πρωτεΐνης συντηρημένα κατάλοιπα μέσω της σύγκρισης των μοντέλων της βάσης δεδομένων με την κάθε πρωτεΐνη αναζήτησης, αντί ολόκληρων των σετ αλληλουχιών που χρησιμοποιήθηκαν για τη δημιουργία των μοντέλων (*Durbin et al. 1998*).

1.7.2 Ανάλυση οριζόντια μεταθετών στοιχείων γονιδιώματος

Για την ανάλυση των μεταθετών στοιχείων σε προκαρυωτικό γενετικό υλικό υπάρχουν πληθώρα αυτοματοποιημένων προγραμμάτων. Η βασική φιλοσοφία των προγραμμάτων αυτών είναι στοίχιση ομολογίας με βάσεις δεδομένων που περιέχουν αλληλουχίες γνωστών μεταθετών στοιχείων. Η εύρεση συσχέτισης σε επίπεδο οικογένειας του μεταθετού στοιχείου συνήθως οδηγεί στην επαλήθευση της ύπαρξης του γενετικού στοιχείου στον τόπο της στοίχισης. Εκτός από τα αυτοματοποιημένα προγράμματα μπορούν να

αξιοποιηθούν στοχευμένες έρευνες για αλληλουχίες οι οποίες θα συγκριθούν με συγκεκριμένες βάσεις δεδομένων.

1.7.3 Ανάλυση synteny

Η γονιδιωματική αλληλούχηση και χαρτογράφηση επέτρεψαν τη σύγκριση των γενικών δομών των γονιδιωμάτων πολλών διαφορετικών ειδών. Το γενικό εύρημα είναι ότι οι οργανισμοί με μεγάλη ταξινομική συνάφεια εμφανίζουν παρόμοιες γονιδιακές δομές στις ίδιες σχετικές θέσεις στο γονιδίωμα. Αυτή η κατάσταση ονομάζεται synteny (heredity (genetics): Microevolution - Britannica Online Encyclopedia). Η synteny που μοιράζονται μεταξύ τους διαφορετικά είδη μπορεί να συναχθεί από τις γονιδιωματικές αλληλουχίες τους. Αυτό γίνεται συνήθως χρησιμοποιώντας μια έκδοση του αλγορίθμου MCScan, ο οποίος βρίσκει συνθετικά μπλοκ μεταξύ των ειδών συγκρίνοντας τα ομόλογα γονίδια τους και αναζητώντας κοινά μοτίβα σε χρωμοσωμική κλίμακα. Οι ομολογίες καθορίζονται συνήθως με βάση τις επιτυχίες αλγορίθμων στοίχισης. Synteny μπορεί να παρατηρηθεί σε μεγέθη εύρους από γενετικών τόπων έως χρωμοσωμάτων (*Wang et al. 2012*).

1.7.4 Ανάλυση του παν-γονιδιώματος

Η αύξηση της προσβασιμότητας στις μεθόδους αλληλούχησης νέας γενιάς λόγω του χαμηλού τους κόστους είχε ως αποτέλεσμα την έκρηξη του αριθμού των διαθέσιμων μικροβιακών γονιδιωμάτων. Μία από τις αναλύσεις συγκριτικής γονιδιωματικής που μπορεί να εκμεταλλευτεί το σύνολο της πληροφορίας των γονιδιωμάτων είναι αυτή του παν-γονιδιώματος. Κατά την ανάλυση αυτήν, το σύνολο των στελεχών ή και αντιπροσωπευτικά στελέχη μίας ταξινομικής ομάδας ενδιαφέροντος υπόκεινται σε σύγκριση προς αναζήτηση ομολόγων γονιδίων εντός και μεταξύ των γονιδιωμάτων τους. Στη συνέχεια γίνεται αναγνώριση του πυρηνικό σετ γονιδίων και των στελεχο-ειδικών γονιδίων για κάθε υπό-ομάδα στελεχών. Με τον τρόπο αυτόν μπορούν να αναγνωριστούν τα γονίδια που είναι υπεύθυνα για την οικοθέση και το φαινότυπο της εν λόγω ταξινομικής ομάδας, ενώ επίσης είναι δυνατή η αναγνώριση γονιδίων (και συνεπώς λειτουργιών) που διαφοροποιούν ένα στέλεχος από τα υπόλοιπα ως προς την οικοθέση του.

1.8 Προηγούμενη εργασία και στόχοι

Πρόσφατα, έγινε απόπειρα απομόνωσης μικροοργανισμών από ριζόσφαιρα εκτεθειμένη σε iprodione αλλά και από έδαφος ριζόσφαιρας χωρίς προηγούμενη έκθεση σε φυτοφάρμακα, οι οποίοι είχαν την ικανότητα να βιοδιασπούν το iprodione. Δεδομένων των παρατηρήσεων αυτών των φαινοτύπων προβήκαμε στην εκχύλιση DNA από τις απόπειρες απομόνωσης. Τα δείγματα T1.6 και T1.8 προερχόταν από ριζόσφαιρα εκτεθειμένη σε iprodione (*Katsoula et al. 2020*) · το δείγμα επονομαζόμενο στην παρούσα εργασία ως “Marco” προέρχονταν από

ριζόσφαιρα μη εκτεθειμένη σε φυτοφάρμακα (*Campos et al. 2015*). Τα δείγματα DNA αλληλουχήθηκαν με μεθόδους Illumina και Oxford Nanopore και έγινε παραγωγή των συναρμολογήσεων των γονιδιωμάτων των μικροοργανισμών. Ταξινομική ταυτοποίηση των contigs που προέκυψαν έδειξε ότι τα T1.6 και T1.8 ήταν σχεδόν πανομοιότυπα και περιείχαν γενετικό υλικό από δύο χρωμοσώματα, ένα που ανήκε σε ένα *Arthrobacter* ή *Paenarthrobacter* και ένα που ανήκε σε ένα *Microbacterium*, ενώ περιείχαν και 3 κυκλικά πλασμίδια. Στην περίπτωση του δείγματος Marco προέκυψαν 4 contigs, εκ των οποίων το ένα ήταν μεγέθους χρωμοσώματος και κατετάγη στο γένος *Paenarthrobacter* βάσει φυλογεντικών δεικτών, ενώ τα υπόλοιπα είχαν μεγέθη πλασμιδίων (μεσαία ως μεγάλα), με τα δύο να είναι κυκλικά και το μεγαλύτερο γραμμικό. Τα στοιχεία της συναρμολόγησης και της ταξινομικής προέκυψαν από την πτυχιακή εργασία του Μυρισιώτη Χρήστου .

Η εκπόνηση της συγκεκριμένης εργασίας αποσκοπεί στην διερεύνηση και επιβεβαίωση ή μη του προτεινόμενου από *Zhang et al. (Zhang et al. 2021)* μηχανισμού μεταβολισμού του iprodione από απομονώσεις και απόπειρες απομονώσεων αυτών. Για την επίτευξη του στόχου αξιοποιήθηκαν τα δεδομένα συναρμολόγησης των εν λόγω γονιδιωμάτων όπου και πραγματοποιήθηκε λειτουργικός χαρακτηρισμός και συγκριτική γονιδιωματική ανάλυση προκειμένου να αναγνωριστούν πιθανά γονίδια που εμπλέκονται στο μεταβολικό μονοπάτι καταβολισμού του iprodione. Ολοκληρώνοντάς, έγινε ανάλυση του πανγονιδιώματος (pangenome) για το γένος *Paenarthrobacter* προς αναζήτηση της διάδοσης των σχετικών γονιδίων μεταξύ μελών του γένους.

2. Μέθοδοι

2.1 Χαρακτηρισμός γονιδιώματος

Για την πραγμάτωση των αναλύσεων της διπλωματικής εργασίας ήταν απαραίτητη η χρήση εξειδικευμένου λογισμικού και γλωσσών προγραμματισμού τα οποία και παρατίθενται. Στη συγκεκριμένη μελέτη αξιοποιήθηκαν τμήματα των προγραμμάτων ή εξ'ολοκλήρου τα προγράμματα και ιδιομορφές των αλγορίθμων των: Prokka, Ugene, BLAST, HMMER, ISfinder, Genoplots, micropan.

2.1.1 Λογισμικό και περιβάλλον

Linux Ubuntu

Τα Linux Ubuntu αποτελούν ένα λειτουργικό σύστημα, ανοιχτού λογισμικού για υπολογιστές που διατίθενται δωρεάν για τους χρήστες. Το βασικό περιβάλλον αλληλεπίδρασης του Ubuntu με τον χρήστη χρησιμοποιεί τον φλοιό BASH και ενδείκνυται για τη διαχείριση μεγάλου όγκου δεδομένων (*Canonical, n.d.*).

R και R-studio

Η γλώσσα προγραμματισμού R και μία από τις πλατφόρμες χρήσης/υποστήριξης της, το R-studio, ενδείκνυται για μεταχείριση μεγάλου όγκου δεδομένων. Το περιβάλλον περιέχει ενσωματωμένη τη δυνατότητα χρήσης προγράμματος φλοιού του BASH εκτός από τερματικό της R. Η γλώσσα προγραμματισμού όπως και το περιβάλλον διατίθενται δωρεάν για τους χρήστες και είναι ανοιχτού λογισμικού (*RStudio Team 2020*).

2.1.2 Δομικός και λειτουργικός χαρακτηρισμός

PROKKA

Για τον χαρακτηρισμό χρησιμοποιήθηκε ως πρώτη προσέγγιση το πρόγραμμα Prokka v1.14.6 (*Seemann et al. 2014*). Το Prokka για να πραγματοποιήσει τον σχολιασμό πρώτα προβαίνει στο δομικό χαρακτηρισμό των μη κωδικευόντων γονιδίων και γενετικών στοιχείων και έπειτα χαρακτηρίζει τα ανοιχτά πλαίσια ανάγνωσης. Στη συνέχεια συγκρίνει την κάθε δομικά χαρακτηρισμένη αλληλουχία με σειρά βάσεων δεδομένων με τη βοήθεια των αλγορίθμων BLAST και HMMER3 προκειμένου να αποδώσει τον λειτουργικό χαρακτηρισμό τους.

Οι βάσεις δεδομένων που χρησιμοποιεί το PROKKA (BLAST +)

Οι τρεις βασικές βάσεις δεδομένων, που εφαρμόζονται με τη σειρά, είναι:

1. ISfinder: Μόνο οι αλληλουχίες τρανσποζάσης (πρωτεΐνη). ολόκληρο το τρανσποζόνιο δεν ανιχνεύεται.

2. NCBI Βάση δεδομένων γονιδίων αναφοράς για την αντιμικροβιακή αντοχή σε βακτήρια: Γονίδια ανθεκτικότητας σε μικροβιακή επιμέλεια από το NCBI

3. UniProtKB (SwissProt): Πρωτεΐνες με στοιχεία (i) από βακτήρια (ή Archaea ή ιούς) (ii) τα οποία να μην είναι καταχωρήσεις "Fragment". και (iii) τα οποία έχουν ένα επίπεδο απόδειξης ("PE") 2 ή χαμηλότερο, το οποίο αντιστοιχεί σε πειραματικά στοιχεία mRNA ή πρωτεωμικής.

Οι βάσεις δεδομένων HMM.

Το Prokka χρησιμοποιεί βάσεις δεδομένων πιθανολογικών HMM προφίλ για το HMMER3. Είναι ως επί το πλείστον εξειδικευμένες σε βακτήρια (*Seemann et al. 2014*).

Επιπλέον χαρακτηρισμοί γονιδίων οπουδήποτε απαιτήθηκε πραγματοποιήθηκαν μέσω διαδικτυακής πλατφόρμας του Εθνικού αμερικάνικου Κέντρου Βιοτεχνολογικής Πληροφορίας (National Center for Biotechnology Information – NCBI) και πιο ενδελεχείς βάσεις δεδομένων σε σύγκριση με το PROKKA, καθώς και έναντι της πλατφόρμας βάσης δεδομένων HMM της οικογένειας πρωτεϊνών.

BLAST

Για το BLAST χρησιμοποιήθηκε η τοπική εκδοχή του v2.11.0+ καθώς και ο webserver του NCBI όπως και το πακέτο BLAST+ suite. Στον διαδουκτυακό χώρο πραγματοποιήθηκαν στοιχίσεις μεταξύ αλληλουχιών DNA-DNA, DNA-πρωτεϊνών και πρωτεϊνών-πρωτεϊνών, αξιοποιώντας τα προγράμματα BLASTn, BLASTp, BLASTx, tBLASTn και του αλγορίθμους MegaBLAST, discontinuous MegaBLAST, BLASTn, BLASTp, PSI-BLAST, BLASTx και tBLASTn (*Altschul et al. 1990 · NCBI Resource Coordinators 2016*). Το πακέτο προγραμμάτων BLAST+ suite BLAST v2.11.0+ αξιοποιήθηκε για την ανάλυση

synteny μεταξύ περιοχών DNA και μεταξύ γονιδιωμάτων ως υποπρόγραμμα για το GenPlotR (*Guy et al. 2010*) και για την ανάλυση του πανγονιδιώματος ως υποπρόγραμμα για το micropan (*Snipen et al. 2015*).

HMMER

Για την διερεύνηση δομικών μοτίβων για συγκεκριμένες πρωτεΐνες χρησιμοποιήθηκε ο webserver (hidden Markov model profiler (HMMER) HmmerWeb version 2.41.1 (*Finn et al. 2011*), στον οποίο έγιναν έλεγχοι ομολογίας έναντι της βάσης δεδομένων Pfam-A (*Mistry et al. 2021*).

2.2 Στατιστική ανάλυση και παρουσίαση δεδομένων

Ugene

Για οπτικοποίηση των συναρμολογημάτων και διαχείριση αλληλουχιών DNA χρησιμοποιήθηκε το πρόγραμμα Ugene v39 (*Okonechnikov et al. 2012*).

ISfinder

Για την ανάλυση μεταθετών στοιχείων χρησιμοποιήθηκε ο webserver του Isfinder (*Siguiet et al. 2006*) στον οποίο διερευνήθηκε η ύπαρξη ή μη μεταθετών στοιχείων σε αλληλουχίες DNA.

GenoPlotR

Το πακέτο προγραμμάτων και αλγορίθμων GenoPlotR v0.8.11 (*Guy et al. 2010*) χρησιμοποιήθηκε για την παρουσίαση των αποτελεσμάτων συγκριτικής γονιδιωματικής μεταξύ γονιδιακών τόπων και την παραγωγή γραφημάτων ανάλυσης synteny.

micropan

Το πακέτο προγραμμάτων και αλγορίθμων micropan v2.1 (*Snipen et al. 2015*) αξιοποιήθηκε για την δημιουργία και ανάλυση του πυρηνικού γονιδιώματος (σύνολο κοινών λειτουργικά γονιδίων μεταξύ του συνόλου των δοκιμασμένων γονιδιωμάτων) και του πανγονιδιώματος (στελέχο-ειδικών ομογύλων), όπως επίσης και για την απεικόνιση της ανάλυσης αυτής.

2.3 Γονίδια αναφοράς και περιγραφή γονιδιωμάτων των δειγμάτων

Ως γονίδια αναφοράς για την αναζήτηση του μεταβολικού μονοπατιού του iprodione στους μικροοργανισμούς της παρούσας εργασίας χρησιμοποιήθηκαν τα γονίδια ipaH, ddaH και duaH, τα οποία χαρακτηρίστηκαν στην μελέτη των *Zhang et al.* (*Zhang et al. 2021*). Οι αμινοξικές αλληλουχίες τους ανακτήθηκαν από την GenBank (*Clark et al. 2016*) και παρατίθενται στον Πίνακα 1. Τα γονίδια αυτά πραγματοποιούν τα 3 στάδια καταβολισμού του iprodione 1, 2 και 3 αντίστοιχα όπως παρατίθενται στην Εικόνα 2 της εισαγωγής της

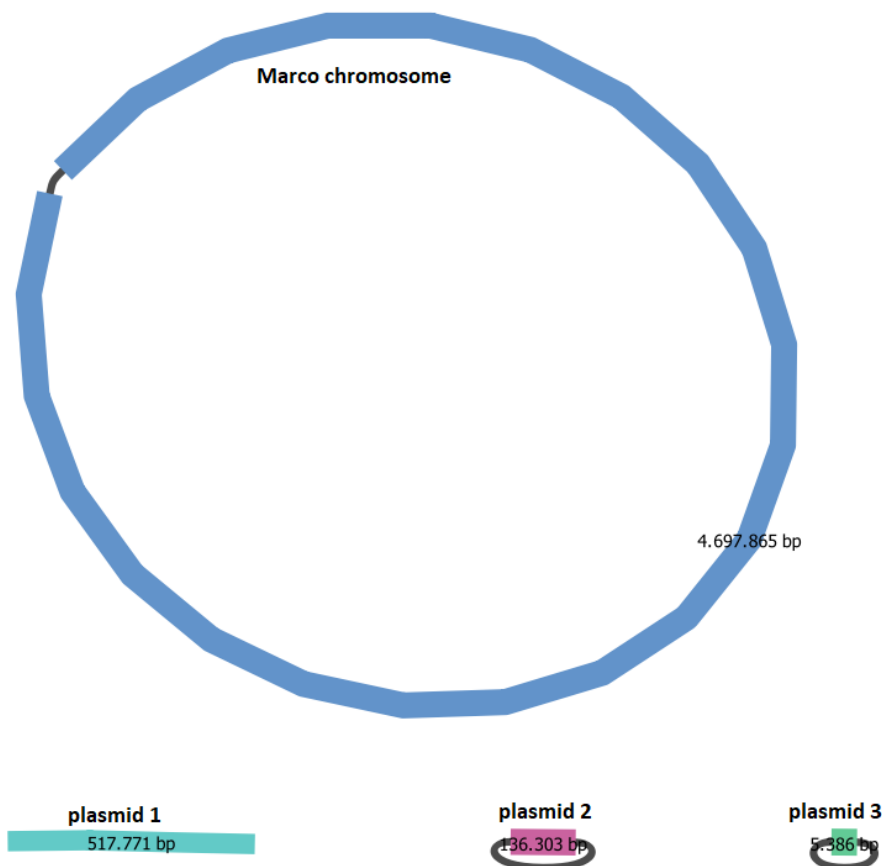
παρούσας εργασίας.

Γονίδιο	Στάδιο μεταβολισμού του Iprodione	Κωδικός στην βάση δεδομένων Genbank του NCBI
ipaH	1 ^ο	QOU11568.1
ddaH	2 ^ο	QOU11570.1
duaH	3 ^ο	QOU11572.1

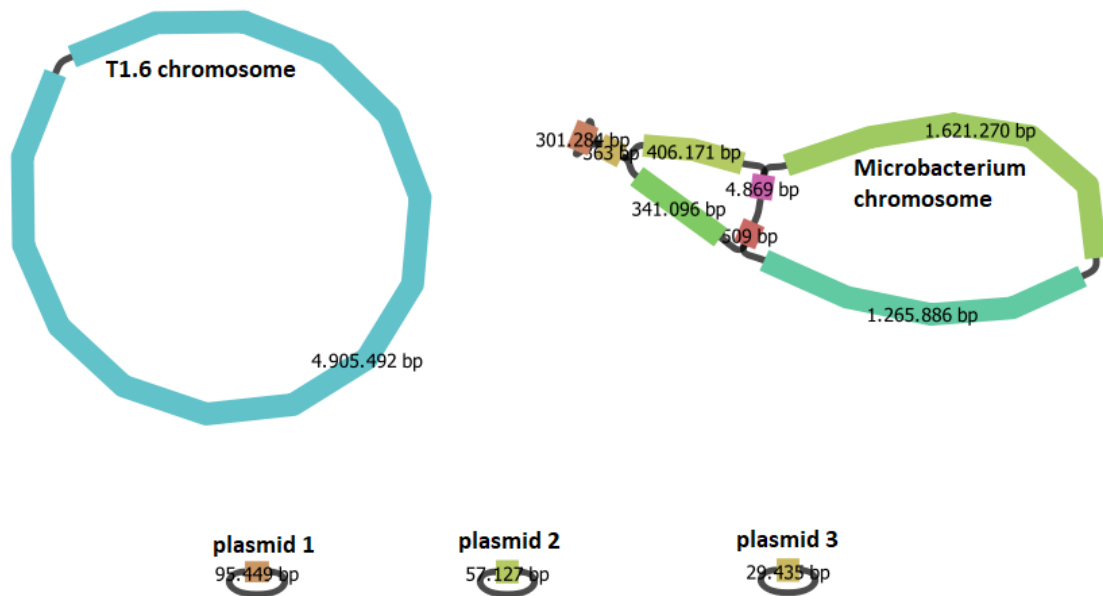
Πινάκας 1: Αρ. πρόσβασης αλληλουχιών αναφοράς για την ανεύρεση του καταβολικού μονοπατιού του Iprodione στους μικροοργανισμούς της παρούσας εργασίας.

Οι συναρμολόγησεις των δειγμάτων απεικονίζονται με την χρήση του προγράμματος Bandage v0.8.1 (Wick et al. 2015) παρακάτω:

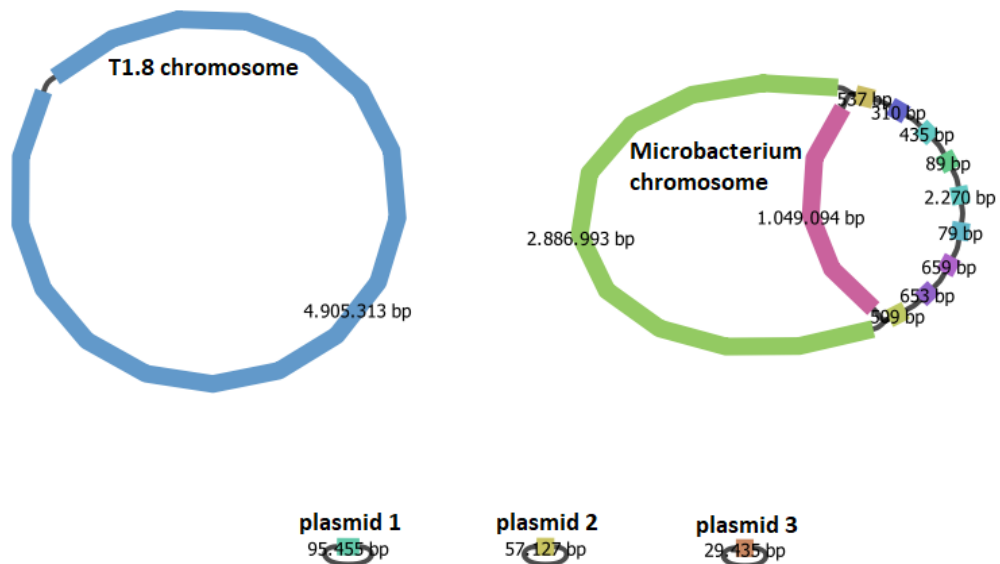
Στην Εικόνα 4 απεικονίζεται η συναρμολόγηση για το δείγμα Marco, Στην Εικόνα 5 για το δείγμα T1.6 και στην Εικόνα 6 για το δείγμα T1.8. Οι συναρμολογήσεις προέκυψαν στην πτυχιακή εργασία του Χρήστου Μυρισιώτη.



Εικόνα 4: Απεικόνιση της συναρμολόγησης για το δείγμα Marco



Εικόνα 5: Απεικόνιση της συναρμολόγησης για το δείγμα T1.6



Εικόνα 6: Απεικόνιση της συναρμολόγησης για το δείγμα T1.8

3. Αποτελέσματα

3.1 Γονίδια υπεύθυνα για τον μεταβολισμό του Iprodione στα υπό μελέτη στελέχη

Δείγμα T1.6

1)IraH

A

ρ

χ

ι

κ

ή

α

ν

Η καλύτερη στοίχιση ανά ζεύγη του BLAST που παρατηρήθηκε για την πρωτεϊνική αλληλουχία αναφοράς IraH στα μεταφρασμένα ανοιχτά πλαίσια ανάγνωσης του T1.6, ήταν στο πλασμίδιο 2, έναντι της αλληλουχίας αμινοξέων με αριθμό πρόσβασης (accession number) PHPLPGGG_00034, με στατιστική σημαντικότητα τοπικής στοίχισης (expectation value, ή E-value) 0, bit score ίσο με 921 και ποσοστό ταυτόσημων και θετικών αμινοξέων 97% (Εικόνα 7). Το καλύτερο αποτέλεσμα BLAST φαίνεται στην Εικόνα 7 που ενέχει ανοιχτό πλαίσιο ανάγνωσης που έχει χαρακτηριστεί από το πρόγραμμα Prokka ως ακυλαμιδάση, ένα ένζυμο με παραπλήσια δράση αυτής της IraH.

τ

ο

υ

γ

ο

ν

ι

δ

ί

ο

υ

ι

ρ

α

H

σ

τ

ο

ν

λ

PHPLPGGG_00034 Acylamidase

Sequence ID: Query_60786 Length: 469 Number of Matches: 1

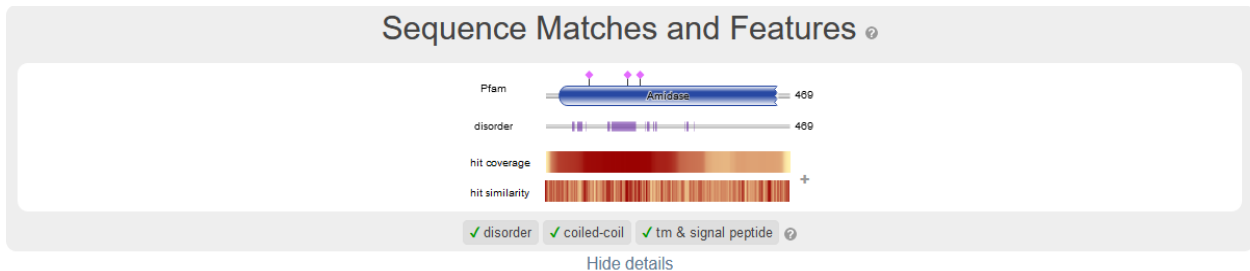
Range 1: 1 to 469 [Graphics](#)

[▼ Next Match](#) [▲ Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
921 bits(2381)	0.0	Compositional matrix adjust.	456/469(97%)	458/469(97%)	0/469(0%)
Query 1		MSDQLWSKSATELAGLVRSKVVSATEVVEAHLQRIEDVNPENAIIVVLADDARQSARIA			60
Sbjct 1		MSDQLWSKSATELAGLVRSKVVSATEVVEAHLQRIEDVNPENAIIVVLADDARQSARIA			60
Query 61		DEKTRTEPDQLGRLHGVPITLKVNIIDLVGSAVPAFKDFYPPMNTPLVDRLLGEGAI			120
Sbjct 61		DEKTRTEPDQLGRLHGVPITLKVNIIDLVGSAVPAFKDFYPPMNTPLVDRLLGEGAI			120
Query 121		VVGRTNMSDMGMRMTTDSLHLGLTRNPWHGPRTAGGSSGGEGAALASGMSALGVGNDLVG			180
Sbjct 121		VVGRTNMS+MGMR TTDSTLHLGLTRNPWHGPRTAGGSSGGEGAALASGMSALGVGNDL G			180
Query 181		SLRNPAHCCGISTLKPTPGRIPWADSI LPPDGPLSFQMMLVHGPMARQVADVRLGMEIMS			240
Sbjct 181		SLRNPAHCCGISTLKPTPGRIPWADSI LPPDGPLSFQMM VHGPMARQVADVRLGMEIMS			240
Query 241		GAHPRDPYSVDVPLYRADSNQPFKVAVMAAPPGAPTDPVSSVVRKAGEAVAAAGYEVDE			300
Sbjct 241		GAHPRDPYSVDVPLYRADSNQPFKVAVMAAPPGAPTDPVSSVVRKAGEAVAAAGYEVDE			300
Query 301		IDAPEYLETRQIWLDFLMTEVNVLQDHITEAMGPAGRQFLRGFLDHAGPLDLPGYIDL FV			360
Sbjct 301		IDAPEYLETRQIWLDFLMTEVNVLQD ITEAMGPAGRQFL DHAGPLDLPGYIDL FV			360
Query 361		RRRMSRVWNDFQYEPV I IAPTWLDVAFEHDWDLKSVADTVDNCGPVVPANVLGLPAAV			420
Sbjct 361		RRRMSRVWNDFQYEPV I IAPTWLDVAFEHDWDLKSVADTVD GP+ PANVLGLPAAV			420
Query 421		TFGGLAKGMPVGVQCI AA AFHDDQAMQVAEVI EAAVGATRPVSPFINAG			469
Sbjct 421		TFGGLAKGMPVGVQCI AA AFHDDQAMQVAEVI EAAVGATRPVSPFINAG			469

Εικόνα 7: Στοιχισμός και στατιστικά του καλύτερου αποτελέσματος της αναζήτησης BLAST βάσει bit score της πρωτεϊνικής αλληλουχίας του *iraH* έναντι της PHPLPGGG_00034.

Στη συνέχεια, σύμφωνα με την αναζήτηση μέσω του HMMER-Web έναντι της Pfam-A φάνηκε ότι παρουσιάζονται πιθανά ενεργά κέντρα τα οποία προσδίδουν δράση αμιδάσης (Εικόνα 8). Η PHPLPGGG_00034 παρουσιάζει 3 motifs ενεργού κέντρου αμιδάσης, δηλαδή υδρολάσης αμιδικού δεσμού (τα motifs αυτά απεικονίζονται ως ροζ ρόμβοι σχηματικά). Το HMMER δίνει μέγιστη ομοιότητα για την αλληλουχία, με την πρωτεΐνη A0A1P8B760 (AMI4G_ARATH), μια “υποθετική” αμιδάση του *Arabidopsis thaliana* (UniProt) (*The UniProt Consortium 2021*). Στην Pfam-A επίσης κατατάσσεται η πρωτεϊνική αλληλουχία ως αμιδάση, bit score 294.66 και κάλυψη 418AA/469AA δηλαδή 89,12%.



Pfam Matches														Standard	
Family	Accession	Clan	Description	Cross-references	Alignment		Model			Bit Score	Domain E-values				
					Start	End	Start	End	Length		Ind.	Cond.			
> Amidase	PF01425.21	n/a	Amidase		26	446	26	444	1	449	451	294.66	1.3e-87	7.3e-92	

Εικόνα 8: Αποτελέσματα αναζήτησης της πρωτεϊνικής αλληλουχίας του μεταφρασμένου ανοιχτού πλάσιου IpaH.

Ολοκληρώνοντας, πραγματοποιήσαμε αναζήτηση στοίχισης της αλληλουχίας RHPLPGGG_00034 (ακυλαμιδάση) στην βάση δεδομένων μονογραφών non redundant protein sequences (nr) του National Center for Biotechnology Information (NCBI). Τα αποτελέσματα της αναζήτησης φαίνονται στην Εικόνα 9 με μεγαλύτερη ομοιότητα της πρωτεΐνης RHPLPGGG_00034 με την IpaH στελέχους *Microbacterium*.

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> IpaH [<i>Microbacterium paraoxydans</i>]	Microbacterium paraoxydans	950	950	100%	0.0	99.79%	469	AXF48700.1
<input checked="" type="checkbox"/> indole acetamide hydrolase [<i>Paenarthrobacter</i> sp. YJN-5]	Paenarthrobacter sp. YJN-5	922	922	100%	0.0	97.44%	469	WP_193593892.1
<input checked="" type="checkbox"/> MULTISPECIES: indole acetamide hydrolase [<i>Micrococcales</i>]	Micrococcales	921	921	100%	0.0	97.23%	469	WP_193591308.1
<input checked="" type="checkbox"/> indole acetamide hydrolase [<i>Geodermatophilus siccatus</i>]	Geodermatophilus siccatus	550	550	100%	0.0	57.20%	472	WP_091224217.1

Εικόνα 9: Τα πρώτα 4 αποτελέσματα αναζήτησης της μεταφρασμένης αλληλουχίας ανοιχτού πλάσιου IpaH.

Στα πλασμίδια 1 και 3 δεν εμφανίστηκε πρωτεϊνική αλληλουχία με ομολογία ως προς τις καταβολικές πρωτεΐνες με την βιβλιογραφία όμως εμφανίστηκε στα βασικά χρωμοσώματα και των δύο βακτηριακών ειδών του *Paenarthrobacter* και *Microbacterium*. Η στοίχιση των ανοιχτών πλαισίων ανάγνωσης των βακτηριακών χρωμοσωμάτων παρουσίασε αποτελέσματα πολύ χαμηλότερου ποσοστού ταύτισης (27.7% και 28.1% με αντίστοιχη κάλυψη 97% και 92% για τα *Arthrobacter/Paenarthrobacter* και *Microbacterium* αντίστοιχα) συγκριτικά με την στοίχιση του ανοιχτού πλαισίου ανάγνωσης του πλασμιδίου 2 (97% με αντίστοιχη κάλυψη 100%). Τα αποτελέσματα των στοίχισεων για τις περιπτώσεις των πιθανών χρωμοσωμάτων με την IpaH βρίσκονται στο παράρτημα I.

2)DdaH

Αρχική αναζήτηση του γονιδίου *ddaH* στον λειτουργικό χαρακτηρισμό του προγράμματος Prokka, όπως και στην περίπτωση του *ipaH*, δεν έδωσε τον αναμενόμενο χαρακτηρισμό, οπότε προχωρήσαμε σε περαιτέρω αναζήτηση με BLAST, ενώ σε δεύτερη προσέγγιση αναζητήσαμε motifs με το HMMER-Web έναντι της Pfam-A βάσης προφίλ πρωτεϊνών.

Η καλύτερη στοίχιση ανά ζεύγη του τοπικού BLAST που έναντι της αλληλουχίας αναφοράς DdaH στα μεταφρασμένα ανοιχτά πλαίσια ανάγνωσης των contigs του T1.6, ήταν στο πλασμίδιο 1 με σχεδόν πλήρη ταύτιση (99%) και θετικά (99% στοίχιση αμινοξέων με παραπλήσιες χημικές ιδιότητες), bit score ίσο με 571, e-value ίσο με 0 και ποσοστό κάλυψης 99% (Εικόνα 10). Το καλύτερο αποτέλεσμα BLAST φαίνεται στην Εικόνα 10 που ενέχει ανοιχτό πλαίσιο ανάγνωσης που έχει χαρακτηριστεί από το πρόγραμμα Prokka ως αποακετυλάση, ένα ένζυμο (PCAODCGP_00013) με παραπλήσια δράση αυτής της DdaH.

PCAODCGP_00013 Peptidoglycan deacetylase

Sequence ID: Query_21330 Length: 279 Number of Matches: 1

Range 1: 1 to 279 [Graphics](#)

[▼ Next Match](#) [▲ Previous Match](#)

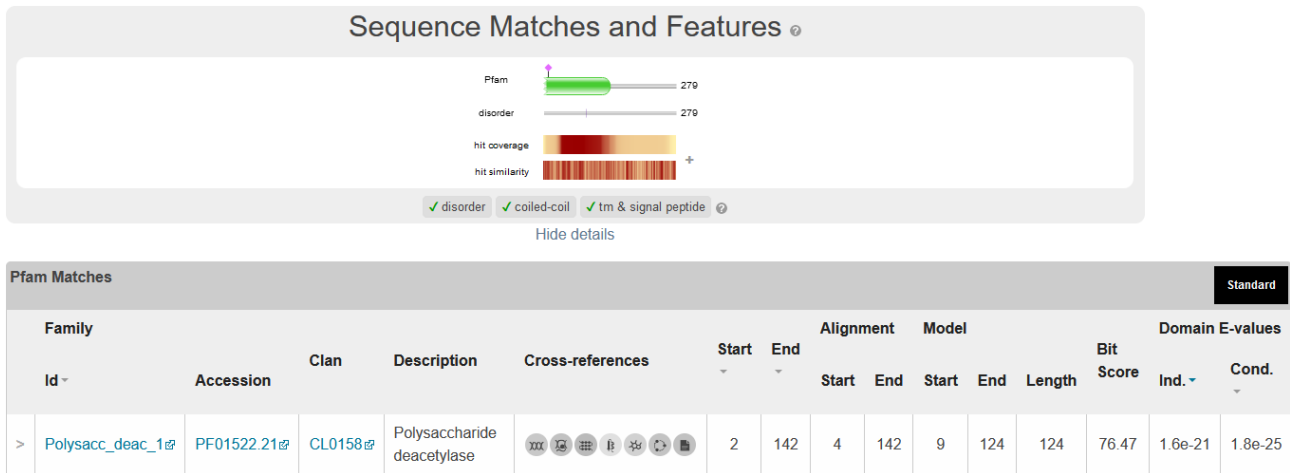
Score	Expect	Method	Identities	Positives	Gaps
571 bits(1472)	0.0	Compositional matrix adjust.	276/279(99%)	278/279(99%)	0/279(0%)
Query 1		MATICLTFDFDAISAWAGTFGLTTPQYVSRGEYGANVATPRILDLLEREGVKSTWYIPGM			60
Sbjct 1		MATICLTFDFDAISAWAGTFGLTTPQYVSRGEYGANVATPRILDLLEREGVKSTWYIPGM			60
Query 61		DVDTYPEVCRRIRDAGHEIGHHGYAHEGPTALSEQAERDVLERGLDALDRVLGVPYGYR			120
Sbjct 61		DVDTYPEVCRRIRDAGHEIGHHGYAHEGPTALSEQAERDVLERGLDALDRVLGVPYGYR			120
Query 121		SPAFDLSPNSTRLLHEYGFAYDSSMMANDFELYKCRGTGDI IHTDKAVEFGPTLDLVEVPV			180
Sbjct 121		SPAFDLSPNSTRLLHEYGFAYDSSMMANDFELYKCRGTGDI IHTDKAVEFGPTLDLVEVPV			180
Query 181		TWTLDDFPFTEFVLAGMLVPGSTDVEAMAQRWMADLDFMIEEYPTGVFTQTFHPQSIGR			240
Sbjct 181		TWTLDDFPFTEFVLAGMLVPGSTDVEAMAQRWMADLDFMIEEYPTGVFTQTFHPQSIGR			240
Query 241		GSRLRILERIIRRAKDHGAEFSTVNDVAVKVGARQGTLR		279	
Sbjct 241		GSRLRILERIIRRAKDHGAEFSTVNDVAVKVGARQGTLR		279	

Εικόνα 10: Στοίχιση και στατιστικά του καλύτερου αποτελέσματος της αναζήτησης BLAST βάσει bit score της πρωτεϊνικής αλληλουχίας του *DdaH* έναντι της μεταφρασμένης αλληλουχίας ανοιχτού πλαισίου ανάγνωσης του PCAOTCGP_00013 του δείγματος T1.6.

Στη συνέχεια, σύμφωνα με την αναζήτηση μέσω του HMMER-Web έναντι της Pfam-A φάνηκε ότι η αλληλουχία PCAOTCGP_00013 ενέχει πιθανά ενεργά κέντρα τα οποία φέρουν ενεργότητα αποακετυλάσης (Εικόνα 11).

Η αποακετυλάση του ανοιχτού πλαισίου ανάγνωσης PCAOTCGP_00013 παρουσιάζει 1 motif ενεργού κέντρου αμιδάσης, δηλαδή υδρολάσης αμιδικού δεσμού, όπως φαίνεται στην Εικόνα 11. Το HMMER δίνει μέγιστη ομοιότητα για την αλληλουχία, με την πρωτεΐνη

Q8DP63 (PGDA_STRR6), μια αποακετυλάση της πεπτιδογλυκάνης του *Streptococcus pneumoniae* (strain ATCC BAA-255 / R6) (UniProt). Στην Pfam-A επίσης κατατάσσεται η πρωτεϊνική αλληλουχία ως αποακετυλάση πολυσακχαριτών, με bit score 76.47 και κάλυψη 138AA/279AA δηλαδή 49,46%.



Εικόνα 11: Αποτέλεσμα αναζήτησης της πρωτεϊνικής αλληλουχίας του μεταφρασμένου ανοιχτού πλαισίου ανάγνωσης του PCAOTCGP_00013 (αποακετυλάση) με την nr βάση δεδομένων μη πλεοναζόντων αλληλουχιών του NCBI. Τα αποτελέσματα της αναζήτησης έδειξαν πλήρη ταύτιση με αποακετυλάση από *Paenarthrobacter* και αποακετυλάσες του ίδιου και άλλων ειδών (Εικόνα 12).

Ολοκληρώνοντας, πραγματοποιήσαμε αναζήτηση στοίχισης της αλληλουχίας του ανοιχτού πλαισίου ανάγνωσης του PCAOTCGP_00013 (αποακετυλάση) με την nr βάση δεδομένων μη πλεοναζόντων αλληλουχιών του NCBI. Τα αποτελέσματα της αναζήτησης έδειξαν πλήρη ταύτιση με αποακετυλάση από *Paenarthrobacter* και αποακετυλάσες του ίδιου και άλλων ειδών (Εικόνα 12).

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> polysaccharide deacetylase [Paenarthrobacter sp. YJN-5]	Paenarthrobacter sp. YJN-5	578	578	100%	0.0	100.00%	279	WP_193594614.1
<input checked="" type="checkbox"/> polysaccharide deacetylase [Paenarthrobacter sp. YJN-D]	Paenarthrobacter sp. YJN-D	571	571	100%	0.0	98.92%	279	WP_193591300.1
<input checked="" type="checkbox"/> polysaccharide deacetylase [Pseudonocardia asaccharolytica]	Pseudonocardia asaccharolytica	468	468	97%	7e-165	80.15%	279	WP_028931167.1
<input checked="" type="checkbox"/> polysaccharide deacetylase [Rhodococcus sp. JVH1]	Rhodococcus sp. JVH1	455	455	96%	1e-159	78.81%	288	WP_009476013.1
<input checked="" type="checkbox"/> polysaccharide deacetylase [Streptomyces sp. SID8352]	Streptomyces sp. SID8352	442	442	99%	2e-154	80.14%	282	WP_161227658.1
<input checked="" type="checkbox"/> polysaccharide deacetylase [Microbacterium sp. 10M-3C3]	Microbacterium sp. 10M-3C3	401	401	97%	2e-138	72.53%	280	WP_125130039.1

Εικόνα 12: Τα πρώτα 6 αποτελέσματα αναζήτησης της πρωτεϊνικής αλληλουχίας της μ ε τ

Στα πλασμίδια 2 και 3, όπως επίσης στο βακτηριακό χρωμόσωμα του *Microbacterium*, δεν εμφανίστηκαν παρόμοιες πρωτεϊνικές αλληλουχίες με αυτές των Zhang et al. (Zhang et al. 2021), όμως εμφανίστηκαν στο βασικό χρωμόσωμα του *Arthrobacter/Paenarthrobacter* (38,4% ταύτιση στο 91% της συνολικής αλληλουχίας, Παράρτημα II). Τα αναλυμένα αποτελέσματα για τις στοίχισεις ανοιχτού πλαισίου ανάγνωσης για το βασικό χρωμόσωμα του *Paenarthrobacter* με την DdaH βρίσκεται στο παράρτημα II.

3)DuaH

Αρχική αναζήτηση του γονιδίου *duaH* στον λειτουργικό χαρακτηρισμό του προγράμματος Prokka δεν έδωσε αποτελέσματα υψηλού βαθμού εμπιστοσύνης οπότε προχωρήσαμε σε περαιτέρω αναζήτηση. Προκειμένου να βρούμε πιθανές ομολογίες μεταξύ χαρακτηρισμένων αλληλουχιών αμινοξέων DuaH και των μεταφρασμένων ανοιχτών πλαισίων ανάγνωσης του γενετικού υλικού του δείγματος T1.6 χρησιμοποιήσαμε τοπικό BLAST έναντι των αλληλουχιών αναφοράς των Zhang et al. (Zhang et al. 2021), ενώ σε δεύτερη προσέγγιση αναζητήσαμε motifs με το HMMER έναντι της Pfam-A βάσης προφίλ πρωτεϊνών.

Η καλύτερη στοίχιση ανά ζεύγη του BLAST που παρατηρήθηκε για την πρωτεϊνική αλληλουχία αναφοράς DuaH στα μεταφρασμένα ανοιχτά πλαίσια ανάγνωσης του T1.6, ήταν η AHEJCLAN_02683 στο βακτηριακό χρωμόσωμα του *Arthrobacter/Paenarthrobacter* με bit score ίσο με 728 και ποσοστό ταυτόσημων αμινοξέων 88% με κάλυψη 93% (Εικόνα 13). Στο βασικό χρωμόσωμα του *Microbacterium* ταυτοποιήθηκαν επίσης δυο αλληλουχίες (51,3% ταύτιση με 93% κάλυψη και 28,4% ταύτιση με 21% κάλυψη· Παράρτημα III). Το καλύτερο αποτέλεσμα BLAST φαίνεται στην Εικόνα 13 που ενέχει ανοιχτό πλαίσιο ανάγνωσης που έχει χαρακτηριστεί από το πρόγραμμα Prokka ως υδρολάση του ιππουρονικού οξέως, ένα ένζυμο (AHEJCLAN_02683) με παραπλήσια δράση αυτής της DuaH.

AHEJLAN_02683 Hippurate hydrolase

Sequence ID: Query_66510 Length: 409 Number of Matches: 1

Range 1: 1 to 409 [Graphics](#)

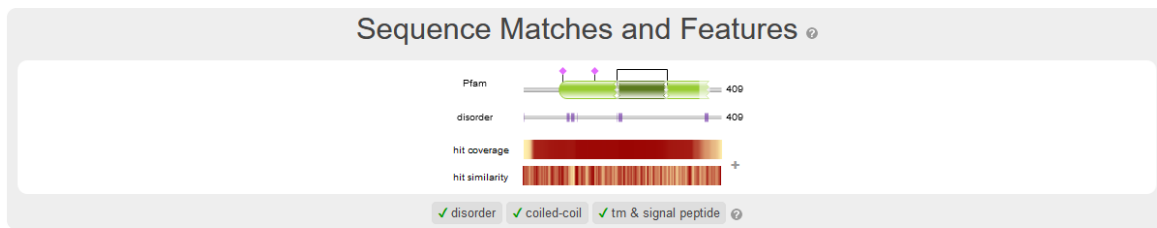
[▼ Next Match](#) [▲ Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
728 bits(1878)	0.0	Compositional matrix adjust.	361/409(88%)	381/409(93%)	1/409(0%)
Query 1		MTQTATVLQLDTEQTTALHELYRRLHAHPELSMQEHATAALIEEQLDRLGADHFRCGRG			60
Sbjct 1		MTSTVTALRLDSDQTTKLHALYRHLHAHPELSMQEFATAKLIEEQLDGLGIEHFRCGGTG			60
Query 61		VVGILRNGEGPVVAFRADTDGLPIEEATGLDYASTDTGML-DGSEVPVMHGCGHDTHVAS			119
Sbjct 61		VVGILRNG GPVVAFRAD+DGLPIEEATGL+YASTDTG L DG+ VPVMHGCGHDTHVAS			120
Query 120		LLGAAEILAGNQDAWSTLVLI FQPGEETAAGALAMLEDGLWDRAPRPEVVFQHVMPRP			179
Sbjct 121		LLGAAEVLAGNLNAWSTLVLI FQPGEETAAGALAMLDDGLWDKAPRPEVVFQHVMPRP			180
Query 180		VGTVSISSGPVAAMADSLRVTVHGGQSHGSQPQDSIDPIVMAAHMVTRLQTVSRELDPR			239
Sbjct 181		VGTVDISSGPVAAMADSLRVTVHGGQSHGSQPQDSIDPIVMAAHMVTRLQGVISRELDPR			240
Query 240		KSAVVTVGTFFHAGLKENIIPASAEFTLNIRTFDEEVRGQVLTAVRRIIEAEAAASGAPGP			299
Sbjct 241		KSAVVTVGTFFHAGLKENIIPASAEFTLNIRTFDE+VR QVL AVRRII AEAAASGAP P			300
Query 300		KIEELYRFPQCYNPDPAVPAVLEALRDALGAESVEVTPPMMGSEDFGHLGTAIGVPSVFW			359
Sbjct 301		IEE+YRFPQC+NDPDAVP+V+EALR ALG E+VE+TPPMMGSEDFGHLGTAIGVPSVFW			360
Query 360		MFGGTPGAVFDGPGPVFNHSPFFAPELEGLAGGISAAVAVLLSRLGN			408
Sbjct 361		MFGGTP A +GPGPVFNHSPFFAPELEGLAGG+SAAVAVLLSRLGN			409

Εικόνα 13: Στοιχισμός και στατιστικά του καλύτερου αποτελέσματος της αναζήτησης BLAST βάσει bit score της πρωτεϊνικής αλληλουχίας του *DuaH* έναντι της μεταφρασμένης αλληλουχίας ανοιχτού πλαισίου ανάγνωσης AHEJLAN_02683 του πιθανού βακτηριακού χρωμοσώματος του *Raenarthrobacter* του δείγματος T1.6.

Στη συνέχεια, σύμφωνα με την αναζήτηση μέσω του HMMER έναντι της Pfam-A φάνηκε ότι παρουσιάζονται πιθανά ενεργά κέντρα τα οποία φέρουν ενεργότητα πεπτιδάσης/καρβόξυ-υδρολάσης (Εικόνα 14).

Η “πεπτιδάση” του ανοιχτού πλαισίου ανάγνωσης του βακτηριακού χρωμοσώματος του *Raenarthrobacter* παρουσιάζει 2 motif ενεργού κέντρου πεπτιδάσης, δηλαδή υδρολάσης αμιδικού δεσμού, όπως φαίνεται στην Εικόνα 14 (τα motifs αυτά απεικονίζονται ως ροζ ρόμβοι). Το HMMER δίνει μέγιστη ομοιότητα για την αλληλουχία, με την πρωτεΐνη P44514 (DAPE_HAEIN), μια καρβόξυ-υδρολάση του *Haemophilus influenzae* (strain ATCC 51907 / DSM 11121 / KW20 / Rd). Στην Pfam-A κατατάσσεται η πρωτεϊνική αλληλουχία ως πεπτιδάση, στην οικογένεια M20 πεπτιδασών, με bit score 82.81 και ταύτιση 291AA/409AA δηλαδή 71,15%. Επίσης, σύμφωνα με την αναζήτηση στην Pfam-A, η αλληλουχία του ανοιχτού πλαισίου ανάγνωσης φέρει έναν γονιδιακό τόπο διμερισμού η οποία συχνά εμφανίζεται σε πεπτιδάσες, ειδικά σε αυτές της οικογένειας M20 (*Hani et al. 1995*).



Pfam Matches														Standard	
Family	Accession	Clan	Description	Cross-references	Start	End	Alignment		Model			Bit Score	Domain E-values		
							Start	End	Start	End	Length		Ind.	Cond.	
> Peptidase_M20	PF01546.28	CL0035	Peptidase family M20/M25/M40		74	385	74	365	1	177	207	82.81	2.6e-23	2.9e-27	
> M20_dimer	PF07687.14	n/a	Peptidase dimerisation domain		194	296	197	293	8	106	109	49.43	3.5e-13	3.9e-17	

Εικόνα 14: Αποτέλεσμα αναζήτησης της πρωτεϊνικής αλληλουχίας του μεταφρασμένου ανοιχτού πλαισίου ανάγνωσης του AHEJLAN_02683 του δείγματος T1.6 που έφερε την καλύτερη στοίχιση στο BLAST με την *DuaH*, με το HMMER έναντι της Pfam A. Οι ροζ ρόμβοι αντιστοιχούν σε εξελικτικά συντηρημένες αλληλουχίες (motifs).

Ολοκληρώνοντας, πραγματοποιήσαμε αναζήτηση στοίχισης της αλληλουχίας του ανοιχτού πλαισίου ανάγνωσης του βακτηριακού χρωμοσώματος του *Arthrobacter/Paenarthrobacter* (πεπτιδάση/καρβοξυδρολάση) με την βάση δεδομένων NCBI-nr. Τα αποτελέσματα της αναζήτησης φαίνονται στην εικόνα 15: εμφανίζονται μόνο οι πρώτες 6 αντιστοιχίσεις με τα υψηλότερα bit scores.

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> amidohydrolase [Paenarthrobacter nitroguajacolicus]	Paenarthrobacter nitroguajacolicus	790	790	100%	0.0	95.35%	409	WP_144651748.1
<input checked="" type="checkbox"/> amidohydrolase [Arthrobacter sp. Rue61a]	Arthrobacter sp. Rue61a	780	780	100%	0.0	93.89%	409	WP_014920817.1
<input checked="" type="checkbox"/> amidohydrolase [Arthrobacter sp. 31Y]	Arthrobacter sp. 31Y	780	780	99%	0.0	94.36%	408	WP_024819642.1
<input checked="" type="checkbox"/> amidohydrolase [Pseudarthrobacter sp. PvP004]	Pseudarthrobacter sp. PvP004	779	779	100%	0.0	94.38%	409	WP_210064555.1
<input checked="" type="checkbox"/> amidohydrolase [Arthrobacter sp. yr096]	Arthrobacter sp. yr096	778	778	99%	0.0	94.36%	409	WP_090824954.1
<input checked="" type="checkbox"/> amidohydrolase [Paenarthrobacter aurescens]	Paenarthrobacter aurescens	777	777	100%	0.0	93.89%	409	WP_011773600.1

Εικόνα 15: Τα πρώτα 6 αποτελέσματα αναζήτησης της πρωτεϊνικής αλληλουχίας της μεταφρασμένης αλληλουχίας ανοιχτού πλαισίου ανάγνωσης του βακτηριακού χρωμοσώματος του *Paenarthrobacter* του δείγματος T1.6 που έφερε την καλύτερη στοίχιση BLAST με την *DuaH*, με την βάση δεδομένων του NCBI.

Στα πλασμίδια 1, 2 και 3, δεν εμφανίστηκαν παρόμοιες πρωτεϊνικές αλληλουχίες με την βιβλιογραφία όμως εμφανίστηκαν στο βασικό χρωμόσωμα του *Microbacterium* (παράρτημα III).

Δείγμα T1.6	Τοποθεσία γενετικού τόπου καλύτερου αποτελέσματος BLAST	Εναλλακτική τοποθεσία γενετικού τόπου αποτελέσματος BLAST
Γονίδια αναφοράς		
iraH	Πλασμίδιο 2	Βακτηριακό χρωμόσωμα <i>Arthrobacter/Paenarthrobacter</i> και βακτηριακό χρωμόσωμα <i>Microbacterium</i>
ddaH	Πλασμίδιο 1	Βακτηριακό χρωμόσωμα <i>Arthrobacter/Paenarthrobacter</i>
duaH	Βακτηριακό χρωμόσωμα <i>Arthrobacter/Paenarthrobacter</i>	βακτηριακό χρωμόσωμα <i>Microbacterium</i>

Πίνακας 2: Τοποθεσία γενετικών τόπων ανοιχτών πλαισίων ανάγνωσης ομολόγων με τα γονίδια αναφοράς για το δείγμα T1.6.

T1.8

Εξαιτίας της γενετικής ομοιότητας των δειγμάτων T1.6 και T1.8, όλα τα αποτελέσματα για το δείγμα T1.8 είναι όμοια με αυτά του T1.6 αναφορικά με τα γονίδια που βρέθηκαν στα ανοιχτά πλαίσια ανάγνωσης και τις συγκρίσεις BLAST. Τα αποτελέσματα βρίσκονται στο παράρτημα IV.

Δείγμα Marco

1)IraH

Κατά την αρχική αναζήτηση του γονιδίου *IraH* στον λειτουργικό χαρακτηρισμό του

π

ρ

ο

γ

ρ

ά

Η καλύτερη στοίχιση ανά ζεύγη του BLAST που παρατηρήθηκε για την πρωτεϊνική μλληλουχία αναφοράς *IraH* στα μεταφρασμένα ανοιχτά πλαίσια ανάγνωσης του δείγματος 29% και 49% ομοιότητας (Εικόνα 16) με αριθμό πρόσβασης GMNKJGKN_00801. Το GMNKJGKN_00801 είχε χαρακτηριστεί από το πρόγραμμα Prokka ως ακυλαμιδάση, ένζυμο με παραπλήσια δράση αυτής της *IraH*.

ς

P

r

ο

k

a

δεν προέκυψε ο συγκεκριμένος χαρακτηρισμός. Προκειμένου να βρούμε πιθανές ομολογίες μεταξύ χαρακτηρισμένων αλληλουγιών αμινοξέων *IraH* και των μεταφρασμένων ανοιχτών

GMNKJGKN_00801 Acylamidase

Sequence ID: Query_90195 Length: 482 Number of Matches: 1

Range 1: 1 to 472 [Graphics](#)

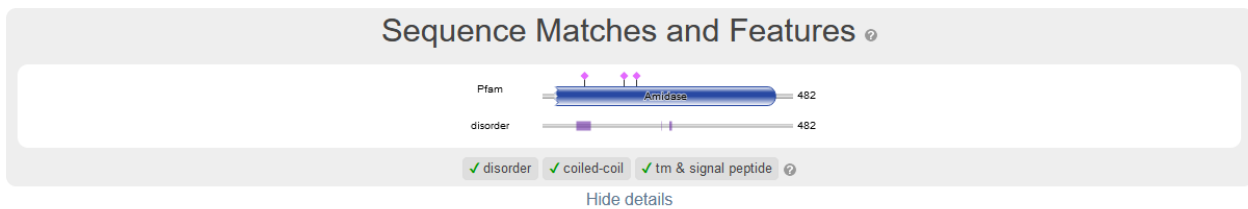
[▼ Next Match](#) [▲ Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
172 bits(437)	3e-49	Compositional matrix adjust.	138/481(29%)	224/481(46%)	22/481(4%)
Query 1	MSDQLWSKSATELAGLVRKSVVSATEVVEAHLQRIEDVNPENAIIVVLADDARQSARIA	60			
Sbjct 1	MSSTIGELSAVELTAAIRSKEISAREALDDHLQRI SDVNPVINA VVTLDAEGARALALNA	60			
Query 61	DEKTRTEPDQLGRLHGVFPIITLKVNI DLVGSATSDAVPAFKDFYPPMNTPLVDRLLEGAI	120			
Sbjct 61	D+ T + +L LHG+P+T K + G T+ A +DF P + ++ RL G +				
Query 121	VVGRTNMSDMGMRMTTDSLHLGLTRNPWHFPGRTAGGSSGGEGAALASGMSALGVGNDLVG	180			
Sbjct 120	G++N+ + G T + L G T NP+ P +AGGSSGG A +AS + A+G G+D+ G				
Query 181	SLRNPAAHCCGISTLKPTPGRI PWADSI LPPDGPLSFQMLLVHGPMARQVADVRLGMEIMS	240			
Sbjct 180	SLR PA C + +P+ IP P ++ + GP MAR V D+ L M + +				
Query 241	GAHPRDPY--SVDVPLYRADSNQPFKVAVMAAPP----GAPTDPEVSSVVRKAGEAVAAA	294			
Sbjct 234	G+ PR P+ +D ++R+ + + G P +PE+ + + A				
Query 295	GVEVDEIDAPEYLETRQIWLDFLMTVEVNLQDHITEAMGPAGRQFLRGFLDHAGPLDLPG	354			
Sbjct 294	G V+E P++ E ++ + + I E G + + ++ L				
Query 355	YIDL FVRRRRMSRVWNDFQOEYPVIIAPTW----LDVA FEHDWDLKSVADT--VDMCGPV	408			
Sbjct 353	I+ R R+ R +FF + + +P D + D+ VA +D				
Query 409	VPANVLGLPA-AVTFGGLAKGMPVGVQCIAAAFHDDQAMQVAEVIEAAVGATRPVSPFIN	467			
Sbjct 413	+ GLP +V G + G+P+G+Q + + D Q ++ E G V+P +				
Query 468	A 468				
Sbjct 472	A 472				

Εικόνα 16: Στατιστικά του καλύτερου αποτελέσματος της αναζήτησης BLAST βάσει bit score της πρωτεϊνικής αλληλουχίας του *IraH* έναντι της μεταφρασμένης αλληλουχίας ανοιχτού πλαισίου ανάγνωσης του βασικού βακτηριακού χρωμοσώματος του δείγματος *Marco*

Στη συνέχεια, σύμφωνα με την αναζήτηση μέσω του HMMER-Web έναντι της Pfam-A φάνηκε ότι παρουσιάζονται πιθανά ενεργά κέντρα τα οποία φέρουν ενεργότητα αμιδάσης (Εικόνα 17).

Η GMNKJGKN_00801 του βακτηριακού χρωμοσώματος παρουσιάζει 3 motifs ενεργού κέντρου αμιδάσης, δηλαδή υδρολάσης αμιδικού δεσμού (τα motifs αυτά απεικονίζονται ως ροζ ρόμβοι σχηματικά). Το HMMER δίνει μέγιστη ομοιότητα για την αλληλουχία, με την πρωτεΐνη A0A0C1BT09 (A0A0C1BT09_9MICC), μια αμιδάση του *Arthrobacter sp.* MWB30 (UniProt). Στην Pfam-A επίσης κατατάσσεται η πρωτεϊνική αλληλουχία ως αμιδάση, με bit score 317.24 και ταύτιση 424AA/482AA δηλαδή 87,96%.



Pfam Matches														Standard	
Family	Accession	Clan	Description	Cross-references	Start	End	Alignment		Model		Bit Score	Domain E-values			
							Start	End	Start	End		Length	Ind.	Cond.	
> Amidase	PF01425.21	n/a	Amidase		26	451	27	451	2	451	451	317.24	1.8e-94	1.0e-98	

Εικόνα 17: Αποτέλεσμα αναζήτησης της πρωτεϊνικής αλληλουχίας του μεταφρασμένου

α
ν
ο
ι

όλοκληρώνοντας, πραγματοποιήσαμε αναζήτηση στοίχισης της αλληλουχίας του ανοιχτού πλαισίου ανάγνωσης του βακτηριακού χρωμοσώματος (ακυλαμιδάση) στην NCBI-nr με BLAST. Τα αποτελέσματα της αναζήτησης φαίνονται στην Εικόνα 18: εμφανίζονται μόνο οι πρώτες 6 αντιστοιχίσεις.

Η δράση των ενζύμων που αντιστοιχούν στα πιο συγγενικές αλληλουχίες είναι αυτή του πρώτο στάδιο καταβολισμού του iprodione (της υδρόλυσης του αμιδικού μεταξύ δεσμού N-ημιδαζολικού δακτυλίου και ισοπροπυλαμιδίου). Στα πλασμίδια δεν εμφανίστηκε παρόμοια πρωτεϊνική αλληλουχία.

σ
ί

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> MULTISPECIES: amidase [Paenarthrobacter]	Paenarthrobacter	970	970	100%	0.0	99.79%	482	WP_110505063.1
<input checked="" type="checkbox"/> amidase [Paenarthrobacter sp. 1092]	Paenarthrobacter sp. 1092	967	967	100%	0.0	99.38%	482	WP_210077503.1
<input checked="" type="checkbox"/> amidase [Arthrobacter sp. 161MFSHa2.1]	Arthrobacter sp. 161MFSHa2.1	967	967	100%	0.0	99.59%	482	WP_018777913.1
<input checked="" type="checkbox"/> amidase [Paenarthrobacter nicotinovorans]	Paenarthrobacter nicotinovorans	966	966	100%	0.0	99.38%	482	WP_064722645.1
<input checked="" type="checkbox"/> amidase [Arthrobacter sp. 31Cvi3.1E]	Arthrobacter sp. 31Cvi3.1E	963	963	100%	0.0	99.38%	482	WP_079581513.1
<input checked="" type="checkbox"/> amidase [Arthrobacter sp. Leaf145]	Arthrobacter sp. Leaf145	963	963	100%	0.0	99.17%	482	WP_055973278.1

Εικόνα 18: Τα πρώτα 6 αποτελέσματα αναζήτησης της πρωτεϊνικής αλληλουχίας της μεταφρασμένης αλληλουχίας ανοιχτού πλαισίου ανάγνωσης του βακτηριακού χρωμοσώματος του δείγματος Marco που έφερε την καλύτερη στοίχιση στο BLAST με την IraH, σε δοκιμή θανάτι της NCBI-nr.

ς

τ
ο
υ

π
λ
α
σ
μ
ι
δ
ί

2)DdaH

Αρχική αναζήτηση του γονιδίου *DdaH* στον λειτουργικό χαρακτηρισμό του Prokka δεν έδωσε αποτελέσματα, οπότε προχωρήσαμε σε περαιτέρω αναζήτηση έναντι της σχετικής αλληλουχίας από την εργασία των Zhang et al. (Zhang et al. 2021) με BLAST, ενώ σε δεύτερη προσέγγιση αναζητήσαμε motifs με το HMMER-Web έναντι της Pfam-A βάσης προφίλ πρωτεϊνών των πιθανών συγγενικών αλληλουχιών.

Η καλύτερη στοίχιση BLAST που παρατηρήθηκε για την πρωτεϊνική αλληλουχία αναφοράς DdaH από Zhang et al στα μεταφρασμένα ανοιχτά πλαίσια ανάγνωσης του γονιδιώματος Marco, ήταν στο βακτηριακό χρωμόσωμα (αλληλουχία GMNKJGKN_04118) με bit score ίσο με 153 και ποσοστό ταύτισης 39%. Το αποτέλεσμα BLAST φαίνεται στην Εικόνα 19.

GMNKJGKN_04118 hypothetical protein

Sequence ID: Query_111105 Length: 331 Number of Matches: 1

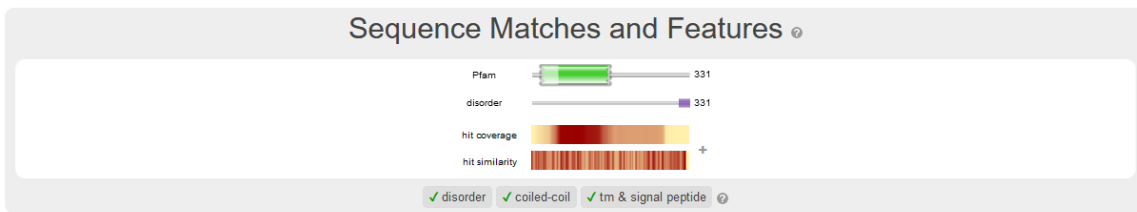
Range 1: 21 to 262 [Graphics](#)

▼ [Next Match](#) ▲ [Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
153 bits(386)	2e-45	Compositional matrix adjust.	99/257(39%)	137/257(53%)	18/257(7%)
Query 2	ATICLTFDFDAISAWAGTFGLTTPQ--YVSRGEYGANVATPRILDLLEREGVKSTWYIPG	59	A TFD DA S +T + +S YG +A PR+L +L+R+ ++ T++IPG		
Sbjct 21	AAASFTFDVDAESCITIAHDPSSSTRMSLSHQSYGPKIAVPRLLQILDRQDIQGTFFIFPG	80			
Query 60	MDVDTYPEVCCRIRDAGHEIGHHGYAHEGPTALSEQAERDVLERGLDALDRVLGVPYGY	119	++YP+V RRI DAGHEI HHGY HE + E L+RGL+AL V G++P GY		
Sbjct 81	FTAESYPDVVRRRIADAGHEIAHHGYLHEP MQGIDAATEASYLDRGLEALANVAGIRPVGY	140			
Query 120	RSPAFDLSPNSTRLLHEYGFAYDSSMMANDFELYKCRGTGDIHTDKAVEFGPTLDLVEVP	179	R+P ++L+ S LL + GF YDSS++ GD + E P D+VE+P		
Sbjct 141	RAPWWELNWQSPALLADRGFLYDSSLL-----DGDAPYRMAVAEDDPR-DIVEIP	189			
Query 180	VTWTLDDFPFTEFVFLAPGMLVPGSTDVEAMAQRW-MADLDFMIEEYPTGVFTQTFHPQSI	238	V W LDD + ++ PG V GS +E+ A+ M L+ F T HP		
Sbjct 190	VDWALDD--WEQYGFYPG--VTGSGVIESPAKALEMWTLEAEAHHSQGSFVLTNHPFIS	245			
Query 239	GRGSRRLRILERIIRRAK	255			
	GR SR LER+I R K				
Sbjct 246	GRPSRAVALERLIERVK	262			

Εικόνα 19: Στατιστικά στοιχεία του πιο συγγενικού αποτελέσματος της αναζήτησης BLAST βάση bit score της πρωτεϊνικής αλληλουχίας της υδρολάσης έναντι της μεταφρασμένης αλληλουχίας του ανοιχτού πλαισίου ανάγνωσης του κύριου βακτηριακού χρωμοσώματος του δείγματος Marco.

Στη συνέχεια, έγινε αναζήτηση μέσω του HMMER-Web έναντι της Pfam-A (Εικόνα 20). Το μεταφρασμένο ανοιχτό πλαίσιο ανάγνωσης GMNKJGKN_04118 του βακτηριακού χρωμοσώματος δεν παρουσιάζει motifs ενεργού κέντρου αμιδάσης. Το HMMER δίνει μέγιστη ομοιότητα για την αλληλουχία, με την πρωτεΐνη A0A0C1DNU9 (A0A0C1DNU9_9MICC), μια δεακετυλάση πολυσακχαριτών του *Arthrobacter sp. MWB30* (UniProt). Στην Pfam-A επίσης κατατάσσεται η πρωτεϊνική αλληλουχία ως δεακετυλάση πολυσακχαριτών, με bit score 59.67 και ταύτιση 105AA/331AA δηλαδή 31,72%.



Pfam Matches														Standard	
Family				Alignment		Model			Domain		E-values				
Id	Accession	Clan	Description	Start	End	Start	End	Start	End	Length	Bit Score	Ind.	Cond.		
> Polysacc_deac_1	PF01522.21	CL0158	Polysaccharide deacetylase	22	163	55	160	16	121	124	59.67	2.6e-16	1.4e-20		

Εικόνα 20: Αποτέλεσμα αναζήτησης της πρωτεϊνικής αλληλουχίας του μεταφρασμένου ανοιχτού πλαισίου ανάγνωσης του βακτηριακού χρωμοσώματος του δείγματος Macro που έφερε την καλύτερη στοίχιση στο BLAST με την DdaH, με το HMMER έναντι της Pfam A.

Ολοκληρώνοντας, πραγματοποιήσαμε αναζήτηση στοίχισης BLAST της αλληλουχίας του ανοιχτού πλαισίου ανάγνωσης του βακτηριακού χρωμοσώματος (δεακετυλάση πολυσακχαριτών) με τη βάση δεδομένων NCBI-nr. Τα αποτελέσματα της αναζήτησης φαίνονται στην Εικόνα 21: εμφανίζονται οι πρώτες 5 αντιστοιχίσεις.

Η δράση των ενζύμων που αντιστοιχούν στις πιο συγγενικές αλληλουχίες της αναζήτησης είναι σημαντικά παραπλήσια με την δράση DdaH η οποία ελέγχει το δεύτερο στάδιο καταβολισμού του iprodione – την υδρόλυση του N-3 αμιδικό δεσμό του ιμιδαζολίου και διάνοιξη του δακτυλίου. Στα πλασμίδια δεν εμφανίστηκε παρόμοια πρωτεϊνική αλληλουχία με την βιβλιογραφία.

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
✓ polysaccharide deacetylase [Paenarthrobacter nicotinovorans]	Paenarthrobacter nicotinovorans	670	670	100%	0.0	99.09%	331	WP_110506185.1
✓ polysaccharide deacetylase [Paenarthrobacter nicotinovorans]	Paenarthrobacter nicotinovorans	660	660	98%	0.0	99.38%	327	WP_189076915.1
✓ MULTISPECIES: polysaccharide deacetylase [unclassified Arthrobacter]	unclassified Arthrobacter	659	659	97%	0.0	99.38%	327	WP_039243273.1
✓ polysaccharide deacetylase [Arthrobacter sp. 31Cv3.1E]	Arthrobacter sp. 31Cv3.1E	659	659	97%	0.0	99.38%	327	WP_079583006.1
✓ MULTISPECIES: polysaccharide deacetylase [Micrococcaceae]	Micrococcaceae	652	652	96%	0.0	99.38%	323	WP_018780326.1

Εικόνα 21: Τα πρώτα 5 αποτελέσματα αναζήτησης της πρωτεϊνικής αλληλουχίας της μεταφρασμένης αλληλουχίας ανοιχτού πλαισίου ανάγνωσης του βακτηριακού χρωμοσώματος του δείγματος Marco που έφερε την καλύτερη στοίχιση στο BLAST με την DdaH, με τις βάσεις δεδομένων του BLAST (NCBI-nr).

3)DuaH

Αρχική αναζήτηση του γονιδίου *duaH* στον λειτουργικό χαρακτηρισμό του προγράμματος Prokka δεν έδωσε αποτελέσματα. Προκειμένου να βρούμε ομολογίες μεταξύ χαρακτηρισμένων αλληλουχιών αμινοξέων DuaH και των μεταφρασμένων ανοιχτών πλαισίων ανάγνωσης του γενετικού υλικού του δείγματος Marco χρησιμοποιήσαμε το BLAST, ενώ σε δεύτερη προσέγγιση αναζητήσαμε motifs με το HMMER-Web έναντι της Pfam-A βάσης προφίλ πρωτεϊνών.

Η καλύτερη στοίχιση ανά ζεύγη του BLAST που παρατηρήθηκε για την πρωτεϊνική αλληλουχία αναφοράς DuaH στα μεταφρασμένα ανοιχτά πλαίσια ανάγνωσης του γονιδιώματος Marco, ήταν στο βακτηριακό χρωμόσωμα του *Paenarthrobacter* με bit score ίσο με 812 και ποσοστό ταύτισης 99%. Το αποτέλεσμα BLAST φαίνεται στην Εικόνα 22 που αφορά γονίδιο που έχει χαρακτηριστεί από το πρόγραμμα Prokka ως υδρολάση του ιππουρονικού οξέως, ένα ένζυμο με παραπλήσια δράση αυτής της DuaH.

GMNKJGKN_00802 Hippurate hydrolase

Sequence ID: Query_443992 Length: 408 Number of Matches: 1

Range 1: 1 to 408 [Graphics](#)

[▼ Next Match](#) [▲ Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
812 bits(2097)	0.0	Compositional matrix adjust.	402/408(99%)	404/408(99%)	0/408(0%)
Query 1	MTQTATVQLDTEQTTALHELYRRLHAHPELSMQEHATAALIEEQDLRLGADHFRCGRTG	60			
Sbjct 1	MTQTATVL LDTEQTTALHELYRRLHAHPELSMQEHATAALIEEQDLRLG DHFRCG TG	60			
Query 61	VVGILRNNGEGPVVAFRADTDGLPIEEATGLDYASTDTGMLDGEVFPVMHGCCHDTHVASL	120			
Sbjct 61	VVGILRNNGEGPVVAFRADTDGLPIEEATGLDYASTDTGMLDGEVFPVMHGCCHDTHVASL	120			
Query 121	LGAAEILAGNQDAWSGTLVLI FQPGEEAAGALAMLEDGLWDRAPRPEVVFQGHVMPRPV	180			
Sbjct 121	LGAAEILAGNQDAWSGTLVLI FQPGEEAAGALAMLEDGLWDRAPRPEVVFQGHVMPRPV	180			
Query 181	GTVSISSGPVAAMADSLRVTVHGQQSHGSQPQDSIDPIVMAAHMVTRLQTIVSRELDPRK	240			
Sbjct 181	GTVSISSGPVAAMADSLRVTVHGQQSHGSQPQDSIDPIVMAAHMVTRLQTIVSRELDPRK	240			
Query 241	SAVVTVGTFFHAGLKENIIPASAEFTLNIRTFDEEVRGQVLTAVRRIIEAEAAAASGAPGPK	300			
Sbjct 241	SAVVTVGTFFHAGLKENIIPASAEFTLNIRTFDEEVRGQVLTAVRRIIEAEAAAASGAPGPK	300			
Query 301	IEELYRFPQCYNDFDAVPVLEALRDALGAESVEVTPPMMGSEDFGHLGTAIGVPSVFWM	360			
Sbjct 301	IEELYRFPQCYNDFDAVPVLEALRDALGAESVEVTPPMMGSEDFGHLGTAIGVPSVFWM	360			
Query 361	FGGTPGAVFDGPGVPVNHSPFFAPELEGSLAGGISAAVAVLLSRLGN	408			
Sbjct 361	FGGTPGAVFDGPGVPVNHSPFFAPELEGSL GG+SAAVAVLLSRLGN	408			

Εικόνα 22: Στοίχιση και στατιστικά του καλύτερου αποτελέσματος της αναζήτησης BLAST βάσει bit score της πρωτεϊνικής αλληλουχίας του *DuaH* έναντι της μεταφρασμένης αλληλουχίας ανοιχτού πλαισίου ανάγνωσης του βακτηριακού χρωμοσώματος του *Paenarthrobacter* του δείγματος Marco.

Η δράση των ενζύμων που αντιστοιχούν πλέον συγγενικές αλληλουχίες σύμφωνα με τα αποτελέσματα, είναι παραπλήσια με την δράση της υποθετικής πρωτεΐνης-στόχου του βακτηριακού χρωμοσώματος του *Paenarthrobacter*, η οποία ελέγχει το τρίτο στάδιο καταβολισμού του iprodione – την υδρόλυση του N-1 αμιδικού δεσμού του διχλωροφαινυλίου. Στα πλασμίδια 1, 2 και 3, δεν εμφανίστηκαν παρόμοιες πρωτεϊνικές αλληλουχίες με την βιβλιογραφία.

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> MULTISPECIES: amidohydrolase [Micrococcaceae]	Micrococcaceae	818	818	100%	0.0	99.26%	408	WP_018777910.1
<input checked="" type="checkbox"/> amidohydrolase [Arthrobacter sp. 31Cv3.1E]	Arthrobacter sp. 31Cv3.1E	816	816	100%	0.0	99.26%	408	WP_079581512.1
<input checked="" type="checkbox"/> amidohydrolase [Paenarthrobacter nicotinovorans]	Paenarthrobacter nicotinovorans	815	815	100%	0.0	99.02%	408	WP_110505064.1
<input checked="" type="checkbox"/> amidohydrolase [Paenarthrobacter nicotinovorans]	Paenarthrobacter nicotinovorans	815	815	100%	0.0	98.77%	408	WP_189076110.1
<input checked="" type="checkbox"/> amidohydrolase [Paenarthrobacter nicotinovorans]	Paenarthrobacter nicotinovorans	815	815	100%	0.0	98.77%	408	WP_064722647.1
<input checked="" type="checkbox"/> amidohydrolase [Paenarthrobacter sp. YJN-D]	Paenarthrobacter sp. YJN-D	812	812	100%	0.0	98.53%	408	WP_193590919.1
<input checked="" type="checkbox"/> amidohydrolase [Arthrobacter sp. MWB30]	Arthrobacter sp. MWB30	810	810	100%	0.0	98.04%	408	WP_039239768.1
<input checked="" type="checkbox"/> amidohydrolase [Arthrobacter sp. Leaf145]	Arthrobacter sp. Leaf145	810	810	100%	0.0	98.28%	408	WP_055973281.1

Εικόνα 24: Τα πρώτα 6 αποτελέσματα αναζήτησης της πρωτεϊνικής αλληλουχίας της μεταφρασμένης αλληλουχίας ανοιχτού πλαισίου ανάγνωσης του βακτηριακού χρωμοσώματος του *Paenarthrobacter* του δείγματος Marco που έφερε την καλύτερη στοίχιση στο BLAST με την *DuaH*, με την *nr* βάση δεδομένων του NCBI.

Δείγμα Marco	Τοποθεσία γενετικού τόπου καλύτερου αποτελέσματος BLAST	Εναλλακτική τοποθεσία γενετικού τόπου αποτελέσματος BLAST
Γονίδια αναφοράς		
<i>ipaH</i>	Βακτηριακό χρωμόσωμα <i>Arthrobacter/Paenarthrobacter</i>	-
<i>ddaH</i>	Βακτηριακό χρωμόσωμα <i>Arthrobacter/Paenarthrobacter</i>	-
<i>duaH</i>	Βακτηριακό χρωμόσωμα <i>Arthrobacter/Paenarthrobacter</i>	-

Πίνακας 3: Τοποθεσία γενετικών τόπων ανοιχτών πλαισίων ανάγνωσης ομολόγων με τα γονίδια αναφοράς για το δείγμα Marco.

3.2 Μεταθετά στοιχεία

T1.6

IraH/amidase

Με το πρόγραμμα Ugene v39 (*Okonechnikov et al. 2012*) χειρίστηκαν τα πλασμίδια 1, 2, 3 και τα βασικά χρωμοσώματα των μικροοργανισμών των δειγμάτων. Έγινε έλεγχος για ανεστραμμένες επαναλήψεις, όταν βρίσκονταν κοντά και εκατέρωθεν σε αλληλουχίες ανοιχτού πλαισίου ανάγνωσης που αντιστοιχούν στα υπό μελέτη γονίδια, για το αν αποτελούν μεταθετά στοιχεία με τον Webserver IS finder (*Siguier et al. 2006*). Δεν υπήρξαν αποτελέσματα για τον έλεγχο αυτό. Έτσι, χρησιμοποιήθηκε ο IS finder για να ανιχνευθούν τρανσποζάσες, ανεστραμμένες επαναλήψεις και γενικότερα μεταθετά στοιχεία στο σύνολο των πλασμιδίων και βακτηριακών χρωμοσωμάτων χειροκίνητα. Επίσης, μέσω του λειτουργικού χαρακτηρισμού του προγράμματος Prokka, ανιχνεύθηκαν αλληλουχίες γονιδίων τρανσποζάσης, οι οποίες οπτικοποιήθηκαν με την βοήθεια του προγράμματος Ugene και συγκρίθηκαν με τις αλληλουχίες των γονιδίων τρανσποζάσης στα μεταθετά που ανιχνεύθηκαν χειροκίνητα.

Για τον έλεγχο του πλασμιδίου 2, που φέρει την αλληλουχία PHPLPGGG_00034 με πιθανή λειτουργία IraH το καλύτερο αποτέλεσμα με 274 bit score αποτελεί το μεταθετό στοιχείο ISMlu14 της οικογένειας IS21 από *Micrococcus luteus* (Εικόνα 25).

Οι ανεστραμμένες περιοχές του μεταθετού που βρέθηκε από το IS finder δεν βρέθηκαν στο πλασμίδιο 2. Η θέση του μεταθετού τοποθετείται περίπου 4.500 - 7.000 ζ.β. καθοδικά της αλληλουχίας-στόχου της PHPLPGGG_00034.

NODE_8+_length_57127_cov_1.0865

Sequence ID: Query_68079 Length: 57127 Number of Matches: 1

Range 1: 33128 to 34591 [Graphics](#)

[▼ Next Match](#) [▲ Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps	Frame
593 bits(1529)	0.0	Compositional matrix adjust.	301/490(61%)	358/490(73%)	4/490(0%)	+2
Query 1		MRGAAELAGCSHHTVARLVRARDAGQTPGSGASRPKVTDVWLPKIEEWVEASTGRIRADV				60
Sbjct 33128		+R A ELAGCSHHTV V R AG +RP++ D +LPK+EEW+E S G++RAD LRDAGELAGCSHHTVKHYVDRRAAGGELDKAVTRPQLIDEYLPKVEEWIERSKGVVADR				33307
Query 61		VHAKLTAMGYTGSESTRRAVAEVKAAWRAGKRRVHRPWITEPGAWLQYDFGDGPAIDGA				120
Sbjct 33308		H KL AMGY GSER+TRRAVA VK ++R+G RVHREPW+TEPG WLQYD+GDGP +DG AHEKLLAMGYKGSERTRRAVASVKKSYRSRSHVHRVHRPWVTEPGMWLQYDYGDGPPVVDGV				33487
Query 121		KTVLFVAVLAWSRVRIVIALRDRTPSPVFAALDRCFRLIGGAPTYVLTDNEKTVTVSHVA				180
Sbjct 33488		KTVLFVAVLAWSR+R+VIALRD+T PSVFAALD+ FR +GG PTYVLTDNEKTVT H+A KTVLFVAVLAWSRFRVVIALLRDKTMPSPVFAALDQTFRRLGGVPTYVLTDNEKTVTTEHIA				33667
Query 181		GVPVRNQATVAFARHYGVEVLTCQPADPAKGGVENAVKLAADLVPKDNTNLRDQYASFA				240
Sbjct 33668		G+PVRN V FA HY V V TC PADPA+KGG E++VK++KADLVPKDNTNLR++YASF+ GIPVRNGQLVTFAEHYSVVVHTCVPADPASKGGTESSVKISKADLVPKDNTNLRREYASFS				33847
Query 241		ELEAACAGFMAMVNSREHRVTRRRRPHMLAEETSALHRI PATAHTVAYGVGRKVPENTPM				300
Sbjct 33848		ELE AC F VN+R HR T+R P MLAE E + LH +P HTVA+G R VP NTPM ELEEACEAFQKQVNTTRAHRTTKRPPVEMLAEEERTRLHPVPTQPHTVAFGTTRVVPNTPM				34027
Query 301		VSFENAQYSVPAHLLGAEVFRHHGTGPDSMVVIMHAGAQPVEVARHRVARPGSPAIDD				360
Sbjct 34028		V FE+ QYSVP LLGA V+VR G G VVI+H G GP EVARH A PG+P I+D VMFESGQYSVPHTLLGATVWVRAQGLGDGEEVVIVHVGEDGPAEVARHARATPGTTPRIND				34207
Query 361		AHFPGHETDKVPRDYTPVPRSAEAEFLAIGAGARTWLVEAAAAGTSRIGQKMAEAVTLA				420
Sbjct 34208		HFP + P + P ++ AEAFL +G GAR WLVEAAAAGT R+ KMAEA++LA EHFP--PQPEGPLNRQPRAKNPAEAEFLDLGEGARLWLVEAAAAGTTPRMRVKMAEALSIA				34381
Query 421		KLAGTDQVDRALGVAHVQRFAHGDLLASLLTAAGHR--TGMHTATEERSLTQGTAGWAGL				478
Sbjct 34382		KL +VD ALG AAVH RFA DL+S+L + G H A EERSLTQGTAGWA L KLFDPEVDWALGHAHVHRFAEADLSSILDHHRQPVAEGEHRAGEERSLTQGTAGWARL				34561
Query 479		GTSNTEGAAR 488				
Sbjct 34562		G +++ +R GQHDSQHDSR 34591				

Εικόνα 26: Στοιχισμός και στατιστικά του αποτελέσματος για τον έλεγχο της τρανσποζάσης του στοιχείου ISMlu14 στο πλασμιδίου 2 του δείγματος T1.6.

Επίσης, το πρόγραμμα Prokka, στον λειτουργικό χαρακτηρισμό του πλασμιδίου 2 του δείγματος T1.6, ανίχνευσε μια αλληλουχία τρανσποζάσης του μεταθετού στοιχείου IS1415(2580 ζ.β.). Η τρανσποζάση αυτή ανήκει επίσης στην οικογένεια IS21 και ανιχνεύθηκε στο ίδιο σημείο με την τρανσποζάση της χειροκίνητης αναζήτησης. Τα μεταθετά στοιχεία IS1415 και ISMlu14 είναι σχεδόν πανομοιότυπα.

Στο παράρτημα V φαίνεται η κατανομή γονιδίων τρανσποζάσης στο πλασμιδίου 2 του δείγματος T1.6.

DdaH/deacetylase

Χρησιμοποιήθηκε ο webserver IS finder για τον έλεγχο του πλασμιδίου 1, που φέρει την αλληλουχία ανοιχτού πλαισίου ανάγνωσης που παρομοιάζει με την DdaH. Το καλύτερο αποτέλεσμα με 293 bit score αποτελεί το μεταθετό στοιχείο ISArsp14 της οικογένειας ISNCY το οποίο βρέθηκε σε *Arthrobacter sp* (Εικόνα 27).

Η αλληλουχία του μεταθετού στοιχείου στην πλήρη έκταση της (17.533 ζβ), παρουσιάζει απόκλιση από την αλληλουχία του πλασμιδίου. Έλεγχος με κριτήρια μικρής συγγένειας για τον αλγόριθμο του BLAST μεταξύ του μεταθετού στοιχείου και του πλασμιδίου επίσης έδειξαν σχετικά μικρή ταύτιση.

```
> ISArsp14 Family: ISNCY
Length=17533

Score = 293 bits (148), Expect = 3e-77
Identities = 346/412 (84%), Gaps = 0/412 (0%)
Strand=Plus/Plus

Query 6958 TGTTCGTCGGCTACATGCGCGTCTCGAAAGCCGACGGCTCCCAAACAACCGACCTGCAGC 7017
      ||||| ||||| ||||| ||||| || || ||||| ||||| || || || ||||| ||
Sbjct 3835 TGTGGTCGGATACATGCGCGTGTCCAAGGCCGACGGCTCACAGACGACAGACCTGCAAC 3894

Query 7018 GCGACACACTCCTTGCCGCCGGCATCGGCGCAGACTCACTCTACGAAGACCGTGGCTCCG 7077
      ||||| | ||| ||||| ||||| ||||| | ||||| ||||| ||||| | || |||||
Sbjct 3895 GCGACGCCCTCATTGCTGCCGGCATGGGCCCGGACTCCCTCTACGAAGACAGGGCTTCCG 3954

Query 7078 GCAAGAAAGACGACCGTCCCCAGCTGGCCTCCTGCCTCAAAGCGCTACGCCACGGGGATA 7137
      | ||||| || ||||| ||||| ||||| ||||| | ||||| ||||| || || || || || |
Sbjct 3955 GGAAGAAGGATGACCGACCCAGTTGGCTGCATGCCTCAAAGCCCTCCGTGAGGGCGACA 4014

Query 7138 CCCTGGTTCGTCGGAAACTCGACCGACTGGGCCGGGACCTGCGCCACCTGGTCAACATCG 7197
      | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct 4015 CTCTGGTTCGTCGGAAACTCGACCGGCTGGGCCGGGACCTGCGCCACTTGGTGAACATCG 4074

Query 7198 TCCACACCCTCACTGAACGCGGCATCGGACTCAAAGTCCTCACCGGCCAAGGCGCGGCCA 7257
      ||||| ||||| ||||| ||||| || || ||||| ||||| || || |||||
Sbjct 4075 TCCACGACCTCACCGAACGCGGAATCGGGCTGAGGGTCCTCACCGGCCAGGGGGCTGCCA 4134

Query 7258 TCGACACCACCACCGCCTCCGGCAAACCTGGTCTTTGGAATCTTCGCCGCCCTCGCGGAAT 7317
      ||||| ||||| || ||||| || ||||| ||||| ||||| ||||| || || |||||
Sbjct 4135 TCGACACCACCACCGGCTTCCGGGAAGCTGGTCTTCGGAACTCTTCGCAGCGTTGGCCGAAT 4194

Query 7318 TCGAACGCGAACTGATCTCCGAACGCACCCTCGCCGGCCTCGCCGCCGCCCG 7369
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct 4195 TCGAACGGGAACTCATCTCGGAACGAACCTTGGCCGGCCTCGCCTCCGCCCG 4246
```

Εικόνα 27: Στοιχισμός και στατιστικά του καλύτερου αποτελέσματος για τον έλεγχο μεταθετών στοιχείων του πλασμιδίου 1 του δείγματος T1.6.

Για περαιτέρω ταυτοποίηση έγινε BLAST της αμινοξικής αλληλουχίας της τραπεζοζάσης στα ανοιχτά πλαίσια ανάγνωσης του πλασμιδίου 1, το αποτέλεσμα φαίνεται στην Εικόνα 28. Οι ανεστραμμένες περιοχές του μεταθετού που βρέθηκε από το IS finder δεν βρέθηκαν στο πλασμίδιο 1. Η θέση του ανοιχτού πλαισίου ανάγνωσης της τραπεζοζάσης του μεταθετού τοποθετείται περίπου 7.000 ζ.β. ανοδικά της αλληλουχίας-στόχου της “δεακετυλάσης”.

NODE_7+_length_95449_cov_0.711475

Sequence ID: Query_55995 Length: 95449 Number of Matches: 1

Range 1: 4642 to 5850 [Graphics](#)

[▼ Next Match](#) [▲ Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps	Frame
402 bits(1033)	1e-127	Compositional matrix adjust.	207/410(50%)	268/410(65%)	9/410(2%)	+1
Query 17		DDGVFWTRIAAESGVPLRTLARWSARYRTDPTSHGLQRRADRGV--RRIPVELIEAIE				74
Sbjct 4642		+DGVP T +A G+ LRTL RW A ++ + G+ A+RG RR EL+ +E				
EDGVPLTALAEHEGIGLRLTLQRWHAHAKRN----GIAGLASANRGTTRRRSTSELVALVE						4809
Query 75		ALALRRPEPTAAFIHRRVGDIDTRGLSPPSYSSVRAIIGGIDPGLRTLAAHGDAAAYRDK				134
Sbjct 4810		LAL P + A I R+ +A D G P SYS+VR+II G+DPG++TLA G AAYRD				
GLALVNPPLSVAATARKANRVAADHGWPPLSYSTVRSIIAGVDPGMKTLAQLGTAAYRDT						4989
Query 135		FELVLRRTAARPNEQWQADHTLLDIIVLGKKGAVVRPWLTVVLDHDSRAIAGYTLFAGAP				194
Sbjct 4990		+EL R A PN WQADHT LDI+VL RPWLT ++DD+SRAI GYT+F GAP				
YELAWRHRAEGPNAIWQADHTELDILVLNANMKPARPWLTTIVDDYSRAICGYTIFLGAP						5169
Query 195		SAEQTALALHQAINTKTNPWFVYGLPDVLYSDHGSDFTSERLKQVCLDTHIRLIHSQVG				254
Sbjct 5170		SA TALAL QAI K WP+ G+PDVLY DHGSDFTS L Q D H + +S V				
SALNTALALRQAIWPKAGTGWPMCGIPDVLYVDHGSDFTSNHLTQTAKDLHFEITYSTVA						5349
Query 255		VPQGRGKIERFYGTITTELLPHLPHGHIPHGTHGTPATAPALTEQLDQILEQFIVTDYNA				314
Sbjct 5350		PQGRGKIERF+GT+ TELL LPHG+ T G AP LTL++LD + +FI+ DY+				
RPQGRGKIERFFGTVNTTELLAELPHGL---TRGYTPAPVLTLELDAAIGRFIIEDYHQ						5520
Query 315		RPHSQTHQIPVQRWAESGFIPRAPAHPELDLDDLLLTATSPRKVQRDGIKFAATRYVSPIL				374
Sbjct 5521		R H + P Q W G++PR PA ++L+LLLLT PR V RDG+ F RYVSP+L				
RKHPEIKATPHQAWVGDGWLPRLPATMDELNLLLLTVAKPRIVHRDGVHFQGLRYVSPLL						5700
Query 375		AAVVGHEVTVRFNPRDVGEIRVYFNDEYLCRAIAPELAADSISLKDLOAA			424	
Sbjct 5701		AAV E V VR++PRD+ EIR++ ++++C+A+ P+ +++LKD+QAA				
AAVREPVIVRYDPRDITEIRIFHKNQFICKAVDPDHETSTLTTLKDIQAA					5850	

Εικόνα 28: Στοιχισμός και στατιστικά του αποτελέσματος για τον έλεγχο της τραπεζοζάσης του στοιχείου ISArsp14 στο πλασμίδιο 1 του δείγματος T1.6.

Με την βοήθεια του Ugene ανασύρθηκε από το συναρμολόγημα η αλληλουχία του ανοιχτού πλαισίου ανάγνωσης η οποία αντιστοιχεί στην θέση της τραπεζοζάσης και η αλληλουχία αυτή υπέστη BLAST αναζήτηση έναντι της NCBI-nr. Το καλύτερο αποτέλεσμα με bit score 890 για την αλληλουχία αυτή είναι η αλληλουχία μιας τραπεζοζάσης/ιντεγκράσης/ρεκομπινάσης τύπου DDE από *Paenarthrobacter sp.* YJN-5.

Επίσης, το πρόγραμμα Prokka, στον λειτουργικό χαρακτηρισμό του πλασμιδίου 1 του δείγματος T1.6, ανίχνευσε μια αλληλουχία τραπεζοζάσης του μεταθετού στοιχείου ISbli29 (4278 ζ.β.) που αντιστοιχεί στις ίδιες συντεταγμένες. Η τραπεζοζάση αυτή ανήκει επίσης στην οικογένεια ISNCY. Παρότι τα μεταθετά στοιχεία ISArsp14 και ISbli29 διαφέρουν

σημαντικά, μοιράζονται δύο ανοιχτά πλαίσια ανάγνωσης: μια ATP εξαρτώμενη πρωτεΐνη και μια ρεσολβάση. Επίσης, οι τρανσποζάσες τους σε σύγκριση BLAST παρουσιάζουν bit score ίσο με 739. Στο παράρτημα VI φαίνεται η κατανομή γονιδίων τρανσποζάσης στο πλασμίδιο 1 του δείγματος T1.6

DuaH/Hippurate Hydrolase

Χρησιμοποιήθηκε ο webserver του προγράμματος IS finder για τον έλεγχο του βακτηριακού χρωμοσώματος του *Paenarthrobacter* του δείγματος T1.6, που φέρει την αλληλουχία AHEJCLAN_02683 που έδειξε ομολογία με την DuaH. Το καλύτερο αποτέλεσμα με 1362 bit score αποτελεί το μεταθετό στοιχείο IS Ac1 της οικογένειας IS3 από *Arthrobacter chlorophenolicus* (Εικόνα 29).

Η αλληλουχία αυτή παρουσιάζει καλή στοίχιση με το βακτηριακό χρωμόσωμα. Η αλληλουχία είναι ένα απλό μεταθετό στοιχείο το οποίο αποτελείται μόνο από το γονίδιο της τρανσποζάσης και τις ανεστραμμένες επαναλήψεις και επικαλύπτεται σχεδόν εξ'ολοκλήρου στο βακτηριακό χρωμόσωμα. Παρόλαυτα, βρέθηκαν πολλές αλληλουχίες μεταθετών στοιχείων με bit score μεγαλύτερο του 50 και ακόμα περισσότερες με bit score μεγαλύτερο του 40 στο βακτηριακό χρωμόσωμα του *Arthrobacter/Paenarthrobacter* του δείγματος T1.6 (Εικόνα 30).

```

> ISAc12 Family: IS3
Length=1243

Score = 1362 bits (687), Expect = 0.0
Identities = 1104/1243 (89%), Gaps = 0/1243 (0%)
Strand=Plus/Plus

Query 2154414 GGGACTGCCCCCGGTTTTGTTGACTCCTTGACCTAGCGAGCTTGTGCTCGTGAAGGATG 2154473
|
Sbjct 1 GGGACTGCCCCCGGTTTTGTTGACTCCTTGACCTAGCGAGCTTGTGCTCGTGAAGGATG 60

Query 2154474 TTGACCATGCCACGGCTTATGGGGCGGAGTTCGCCAGGATGTTATTGATGTGCCCGT 2154533
|
Sbjct 61 TTGACCATGCCACGGCTTATGGGGCGGAGTTCGCCAGGATGTTATTGATGTGCCCGC 120

Query 2154534 AAGGGCGAAGCGCCGCTGGCGCAGATTGCCAAGGATTTCCGACTGTCCGTACGACGCTG 2154593
|
Sbjct 121 AAGGGCGAGGCGCCGCTGGCGCAGATTGCCAAGGATTTGGGCTTTCCGTACGACGTTG 180

Query 2154594 AAGCGTTGGATTGCCATAGCTGAGCGTAAGGATTCGGAGCCGGCCCGGCAGGCGATGAT 2154653
|
Sbjct 181 AAGCGCTGGATCGCTATCGCTGAGCGTAAGGACTCCGGGACCAGCCAGCGCGCGGGAG 240

Query 2154654 TCTGCCGAGATGCGGGAGCTGAAGAAGCGCAACCGCCTGCTGGAGCAGGAGAACGAGATT 2154713
|
Sbjct 241 TCGGCTGAGGTGCGGGAGTTGAAGAAGCGGAACCGCCTGTTGGAGCAGGAGAACGAGATT 300

Query 2154714 CTGCGCCGCGCTGCTGCCTATTTGGCCAGGGACATCAACCCAAAATGATCTACCCGCTGG 2154773
|
Sbjct 301 CTGCGCCGGGCGGCTGCCTATCTGCGCAGGGACATCAACCCAAAATGATCTACCCGCTGG 360

Query 2154774 TCACGGATCTTGCCGCCGACGGTGTTCCTAGCGGTGACCTGCAGGGTGTGGGATTCA 2154833
|
Sbjct 361 TCACGGATCTTGCCGCCGACGGTGTTCCTAGCGGTGACCTGCAGGGTGTGGGATTIA 420

Query 2154834 GCAAGCAAGGCTATTACCGGTGGAGGGCCAGCCCGGTAAACGGAACGGGACTGGGTAGATG 2154893
|
Sbjct 421 GCAAGCAAGGCTATTACCGCTGGAAGGCGAACCCTGACGGAACGGGACTGGATCGATG 480

Query 2154894 CGCACTTGGTCAACGCTGCCTTGGATATCCATACCGACGACCCGGCATTCCGGTACCGTT 2154953
|
Sbjct 481 CGCACCTGGTCAACGCGCCCTTGACATCCACTCCGACGATCCGGCGTTCGGGTACCGGT 540

Query 2154954 TTATCGCCGACGAACCTCCGGGAGAAGGGCATCACAGCGGGTGAGAACAGGGTCCAGCGCT 2155013
|
Sbjct 541 TCATCGCCGACGAGCTCCCGGAGAAGGGCATCACCGGGCGGAGAACAGGGTCCAGCGCT 600

Query 2155014 TGTGCAGGGACCACGGTATCTGGTCGGTGTCTCGAAGAAGCGTGGCCTGAACCGCAGAC 2155073
|
Sbjct 601 TATGCAGGGACCACGGCATCTGGTCAGTGTCTCGAAGAAACGGGGTCTGAACCGGAAAC 660

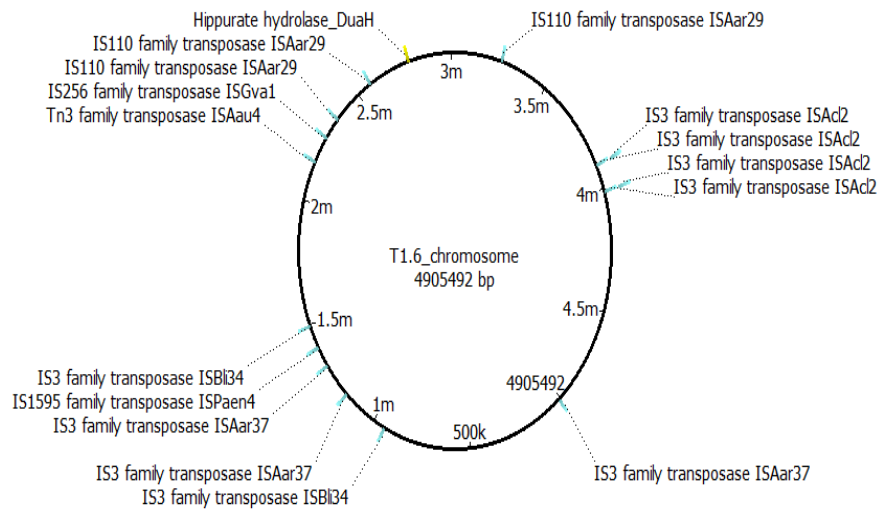
Query 2155074 CCGGGCCCGCGGTTTCATGACGACCTGGTGGAGAGGGACTTCACCGCTGCGGCTCCGAACG 2155133
|
Sbjct 661 CCGGGCCACCGGTCCACGACGACCTGGTTCGAGCGGGATTTCACCGCCGCGGCGCGAACG 720

Query 2155134 AGCTGTGGCTGACGGACATCACCGAACATCCACCGCCGAGGGCAAGCTCTATCTCTGTG 2155193
|
Sbjct 721 AACTGTGGCTGACAGACATCACCGAGCACCCACCGAAGAAGGGAAAGCTCTACCTCTGCG 780

Query 2155194 CTGTGAAGGACGTCTACTCCGGGAGGATCGTTGGGTATTGATGGACTCGCGGATGAAGT 2155253
|
Sbjct 781 CGGTGAAGGACGTCTACTCCGGCAGGATCGTTCGGCTACTCAATGGACGCCCGGATGAAGT 840

```

Εικόνα 29: Απόκομμα στοίχισης και στατιστικά του καλύτερου αποτελέσματος για τον έλεγχο μεταθετών στοιχείων του βακτηριακού χρωμοσώματος του *Raenarthrobacter* του δείγματος T1.6.



Εικόνα 30: Κατανομή γονιδίων τρανσποζασών στο βακτηριακό χρωμόσωμα του *Arthrobacter/Paenarthrobacter* του δείγματος T1.6. Επίσης τοποθετείται το ανοιχτό πλαίσιο ανάγνωσης που αντιστοιχεί στο γονίδιο αναφοράς DuaH.

Marco

Για τον έλεγχο του βακτηριακού χρωμοσώματος, που φέρει όλες τις αλληλουχίες ανοιχτού πλαισίου ανάγνωσης που παρουσιάζουν ομολογία με τα υπό μελέτη ένζυμα, πραγματοποιήθηκε αναζήτηση με το IS-Finder. Ομοιότητα 387 bit score (ταύτιση 83%) βρέθηκε για το μεταθετό στοιχείο ISArsp6 της οικογένειας Tn3 που ανήκει σε *Arthrobacter* sp. (Εικόνα 31).

Παρότι το μεταθετό στοιχείο ISArsp6 έχει μέγεθος 26.712 ζ.β. , αποτελείται μόνο από το γονίδιο της τρανσποζάσης και μιας Tn3 ρεσολβάσης.


```

> ISArsp14 Family: ISNCY
Length=17533

Score = 188 bits (95), Expect = 6e-44
Identities = 194/227 (85%), Gaps = 0/227 (0%)
Strand=Plus/Plus

Query 1479587 ACCTGGTAGTCCATGGCCAGGCCGAACAGCACGCCGATCAGGATGATGGGCAGGAAGCTC 1479646
          |||
Sbjct 8533 ACCTGGTAATCCATGGCCAGGCCGAACAGCACGCCGATGAGGATGATCGGCAGGAAGCTC 8592

Query 1479647 AGGACGGCACCCTGGATTGGCGACGTCGAAGATTCTCCAGCCAGCCCCACTGGTAAACG 1479706
          |||
Sbjct 8593 AGGACCGCCCGGGTTGTGGACCCGAAGATTTCGCCGAGCCAGCCCCACTGGTAGACG 8652

Query 1479707 GCGACCACTGCGCCGAAGGCGGCGGAGAGAGCAGGAATCCACCCGTGGCCAGCAGC 1479766
          |||
Sbjct 8653 GCGACGATGGCGCCGAAGGTCGCTGCCATGGAGAGCAGGAACCCGCCCGTTGCCAGCAGC 8712

Query 1479767 GGCACCACGATGGAGCGGAACACCAGGAGCAGCAGGATCAGGGAGAG 1479813
          |||
Sbjct 8713 GGGACCAGGATGGAGCGGAAGACGAGCAGCAGCAGCACCAGGGAGAG 8759

Score = 119 bits (60), Expect = 4e-23
Identities = 147/176 (84%), Gaps = 0/176 (0%)
Strand=Plus/Plus

Query 56746 CCGGAGCCCTTCACGTTGATCATGGTCTGCCCGATCACCGCGCAGCCGCCAGGCCACCA 56805
          |||
Sbjct 775 CCGGAGGCCCTTGACGTTGATCATGGTCTGCCCGATCATCGCGCAGCCGCCATGCCGCCG 834

Query 56806 AGGAAACCGGTGACGATGTTGGCCACACCTTGCCCCAAGACACCCGGGTCTTATCGGAG 56865
          |||
Sbjct 835 AAGAAGCCGGTGATGATGTTTGGCGGCCCTGGCCCCAGGATTCGCGGGTCTTGTGGAA 894

Query 56866 CGGGTGTCGGTGAATGTCGTCACCAGTTTGGCAGTCATGAGTGATTCCAGCAGGCC 56921
          |||
Sbjct 895 CGGGTGTCGGTGAATGTCGTCACCAAGCTTGGCGGTCAICAGGGATTCAAGCAGGCC 950

Score = 93.7 bits (47), Expect = 2e-15
Identities = 104/123 (85%), Gaps = 0/123 (0%)
Strand=Plus/Plus

Query 1480828 GGCCATGAGCAATGGCAGGCCGGCGGCGATAAGGGTGCCAGCATGAGGATCAGGACAAG 1480887
          |||
Sbjct 9801 GGCCATCAGCAGTGGCAGGCCGGCAGCGATCAGGGTGCCGAGCATGAGGATCAGGACCAG 9860

Query 1480888 GGCCGCAACCGCAACGCCGACGATTTCCAGCGATGCCGAAGATTTCCGAGACGTCCTCGGT 1480947
          |||
Sbjct 9861 GGCCGCGACGCCACACCGATGATTTCCGCGGTGCCGAGGACTTCGGAGACATCTGGGT 9920

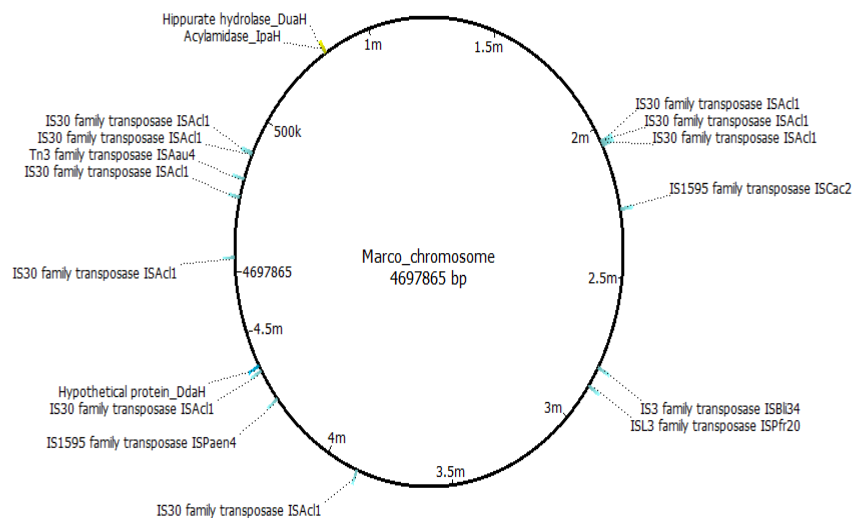
Query 1480948 GAT 1480950
          |||
Sbjct 9921 GAT 9923

```

Εικόνα 32: Απόκομμα στοίχισης και στατιστικών του δεύτερου καλύτερου αποτελέσματος για τον έλεγχο μεταθετών στοιχείων του βακτηριακού χρωμοσώματος του *Raenarthrobacter* του δείγματος Marco.

Η αλληλουχία του μεταθετού στοιχείου στην πλήρη έκταση της (17.533 ζβ), παρουσιάζει μεγάλη απόκλιση από την αλληλουχία του βακτηριακού χρωμοσώματος του δείγματος Marco. Έλεγχος με κριτήρια μικρής συγγένειας για τον αλγόριθμο του BLAST μεταξύ του μεταθετού στοιχείου και του πλασμιδίου επίσης έδειξαν σχετικά μικρή ταύτιση.

Για περαιτέρω ταυτοποίηση έγινε BLAST της αμινοξικής αλληλουχίας της τρανσποζάσης στα ανοιχτά πλαίσια ανάγνωσης του βακτηριακού χρωμοσώματος, όμως δεν υπήρξαν αποτελέσματα. Παρόλαυτα, βρέθηκαν πολλές αλληλουχίες μεταθετών με bit score μεγαλύτερο του 50 και ακόμα περισσότερες με bit score μεγαλύτερο του 40. Με τη βοήθεια του προγράμματος Ugene, οπτικοποιήθηκαν οι αλληλουχίες γονιδίων τρανσποζάσης και η κατανομή τους στο βακτηριακό χρωμόσωμα του *Paenarthrobacter* (Εικόνα 33).



Εικόνα 33: Κατανομή γονιδίων τρανσποζασών στο βακτηριακό χρωμόσωμα του *Paenarthrobacter* του δείγματος Marco. Επίσης τοποθετούνται τα ανοιχτά πλαίσια ανάγνωσης που αντιστοιχούν στα γονίδια αναφοράς *ipaH*, *ddaH* και *duaH*.

3.3 Γενετικοί τόποι των υπεύθυνων γονιδίων για τον καταβολισμό του Iprodione

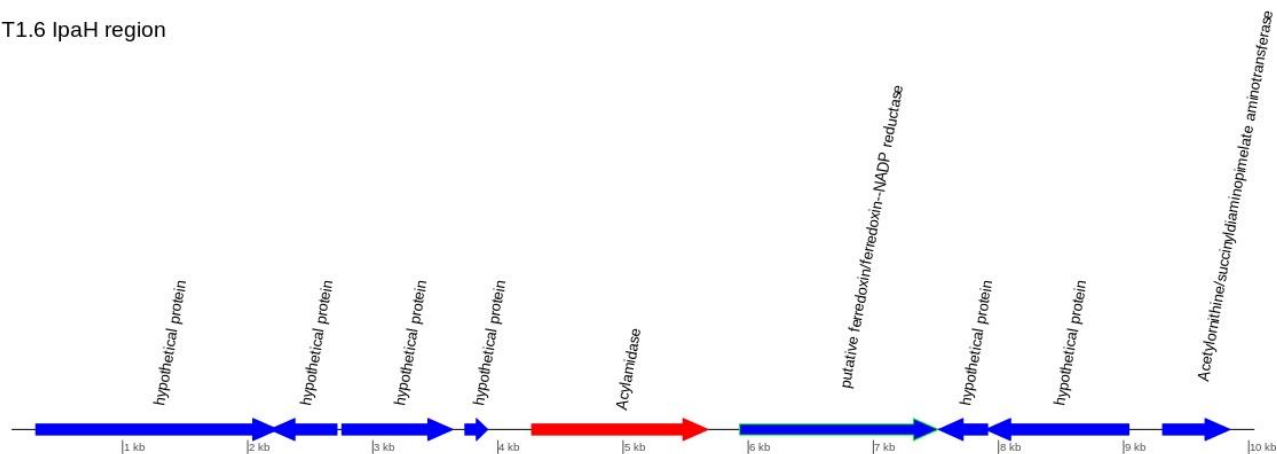
Τα τμήματα DNA που μελετούνται, στα οποία βρίσκονται τα loci των ανοιχτών πλαισίων ανάγνωσης για τις αλληλουχίες που ομοιάζουν με τα γονίδια αναφοράς, φαίνονται αναλυμένα στο παράρτημα VII. Η παρουσίαση των σχετικών γονιδιακών τόπων έγινε με χρήση του πακέτο GenoPlotR του προγράμματος R (Guy *et al.* 2010).

T1.6

IpaH/Acylamidase

Καθοδικά του γονιδίου της αμιδάσης (PHPLPGGG_00034, βρέθηκε στο υποτιθέμενο πλασμίδιο 2), σύμφωνα με τον λειτουργικό χαρακτηρισμό του προγράμματος Prokka, εντοπίστηκαν (Εικόνα 34): ένα ανοιχτό πλαίσιο ανάγνωσης οξειδοαναγωγάσης, δυο ανοιχτά πλαίσια ανάγνωσης υποθετικών πρωτεϊνών και ένα ανοιχτό πλαίσιο ανάγνωσης αμινοτρανσφεράσης. Ανοδικά του γονιδίου της “αμιδάσης”, σύμφωνα με τον λειτουργικό χαρακτηρισμό του προγράμματος Prokka, εντοπίστηκαν τέσσερα(4) ανοιχτά πλαίσια ανάγνωσης υποθετικών πρωτεϊνών.

T1.6 IpaH region



Εικόνα 34: Σχηματική απεικόνιση του τμήματος και των στοιχείων του, του πλασμιδίου 2 του δείγματος T1.6 που περιέχει την αλληλουχία της “ακυλαμιδάσης”.

Οι αλληλουχίες των ανοιχτών πλαισίων ανάγνωσης για υποθετικές πρωτεΐνες υποβλήθηκαν σε έλεγχο ομολογίας μέσω του προγράμματος BLAST με την NCBI-nr ως βάση δεδομένων αναφοράς. Τα αποτελέσματα του ελέγχου αυτού ήταν:

Για την πρώτη υποθετική πρωτεΐνη ανοδικά, μεγαλύτερη ομοιότητα εμφανίστηκε με μια υποθετική πρωτεΐνη που εντοπίστηκε σε *Glutamicibacter protophormiae*. Η δεύτερη αλληλουχία παρουσίασε μέγιστη ομολογία με bit score 599 με μια πρωτεΐνη που δεσμεύει μονόκλωνο DNA, η οποία εντοπίστηκε σε *Glutamicibacter protophormiae*. Για την τρίτη αλληλουχία δεν υπήρχαν αποτελέσματα, και η τέταρτη αλληλουχία παρουσίαζε ομολογία με μια υποθετική πρωτεΐνη που εμφανίστηκε σε *Glutamicibacter protophormiae*.

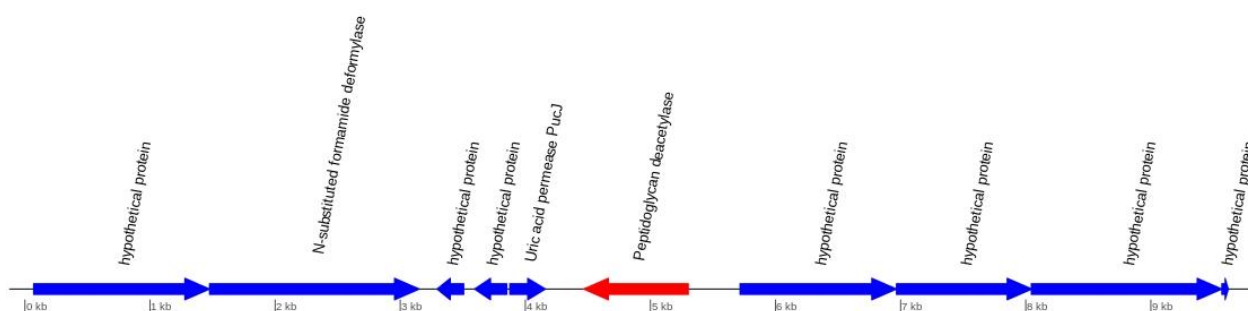
Για την πρώτη αλληλουχία καθοδικά μέγιστη ομοιότητα εμφανίστηκε με μια υποθετική πρωτεΐνη που εντοπίστηκε σε *Nocardioides sp.* OAS665 και για την δεύτερη αλληλουχία, με bit score 461, μια δεϋδρογενάση Glu/Leu/Phe/Val η οποία εντοπίστηκε επίσης σε

Nocardioides sp. OAS665.

DdaH

Ανοδικά του γονιδίου της “αποακετυλάσης(PCAOTCGP_00013, βρέθηκε στο υποτιθέμενο πλασμίδιο 1)”, σύμφωνα με τον λειτουργικό χαρακτηρισμό του προγράμματος Prokka, εντοπίστηκαν (Εικόνα 35): τέσσερα(4) ανοιχτά πλαίσια ανάγνωσης υποθετικών πρωτεϊνών. Καθοδικά του γονιδίου, σύμφωνα με το Prokka, εντοπίστηκαν: ένα ανοιχτό πλαίσιο ανάγνωσης για μια πρωτεΐνη-μεταφορέα ουρικού οξέως, δύο(2) ανοιχτά πλαίσια ανάγνωσης υποθετικών πρωτεϊνών, ένα ανοιχτό πλαίσιο ανάγνωσης μιας υδρολάσης και ένα ανοιχτό πλαίσιο ανάγνωσης μιας υποθετικής πρωτεΐνης.

T1.6 DdaH region



Εικόνα 35: Σχηματική απεικόνιση του τμήματος και των στοιχείων του, του πλασμιδίου 1 του δειγματος T1.6 που περιέχει την αλληλουχία της “αποακετυλάσης”.

Οι αλληλουχίες των ανοιχτών πλαισίων ανάγνωσης για υποθετικές πρωτεΐνες υποβλήθηκαν σε έλεγχο ομολογίας μέσω του προγράμματος BLAST με την NCBI-nr ως βάση δεδομένων αναφοράς. Τα αποτελέσματα της δοκιμής αυτής ήταν:

Για την πρώτη υποθετική πρωτεΐνη καθοδικά, μεγαλύτερη ομοιότητα εμφανίστηκε με μια πρωτεΐνη της οικογένειας RbsD/FucU που εντοπίστηκε σε *Clostridium*. Για την δεύτερη υποθετική πρωτεΐνη καθοδικά, μεγαλύτερη ομοιότητα εμφανίστηκε με μια υποθετική πρωτεΐνη που εντοπίστηκε σε *Paenarthrobacter* sp. YJN-5. Για την τρίτη υποθετική πρωτεΐνη καθοδικά, μεγαλύτερη ομολογία εμφανίστηκε, με bit score 926, με έναν διαμεμβρανικό μεταφορέα κυτοσίνης που εντοπίστηκε σε *Paenarthrobacter* sp. YJN-5.

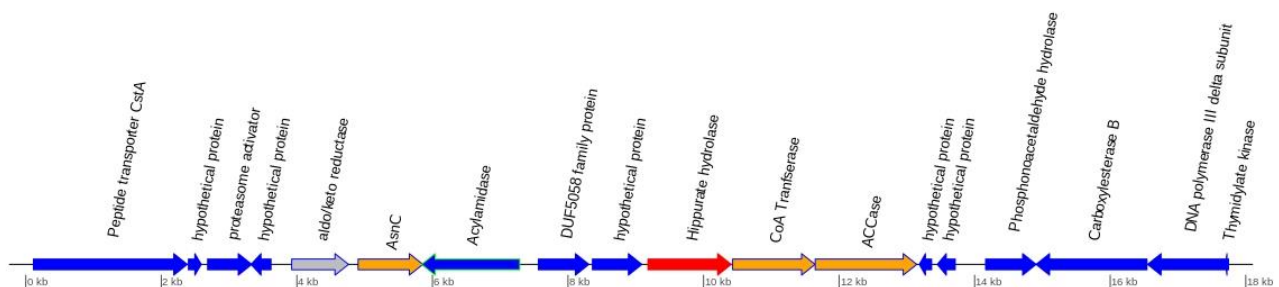
Για την πρώτη υποθετική πρωτεΐνη ανοδικά, μεγαλύτερη ομοιότητα εμφανίστηκε, με bit score 824, με μια N-methylhydantoinase (ATP εξαρτώμενη υδρολάση) που εντοπίστηκε σε *Paenarthrobacter* sp. YJN-5. Για την δεύτερη υποθετική πρωτεΐνη ανοδικά, μεγαλύτερη ομοιότητα εμφανίστηκε, με bit score 727, με μια πρωτεΐνη που περιέχει επικράτεια DUF917

η οποία εντοπίστηκε σε *Paenarthrobacter* sp. YJN-5. Για την τρίτη υποθετική πρωτεΐνη ανοδικά, μεγαλύτερη ομοιότητα εμφανίστηκε, με bit score 994, με πρωτεΐνη που περιέχει δομή (α-έλικα)-(φουρκέτα)-(α-έλικα) η οποία εντοπίστηκε σε *Paenarthrobacter* sp. YJN-5. Για την τέταρτη υποθετική πρωτεΐνη ανοδικά, δεν παρήχθησαν αποτελέσματα.

DuaH/Hippurate Hydrolase

Καθοδικά του γονιδίου της υδρολάσης του ιππουρονικού οξέος (AHEJLAN_02683, βρέθηκε στο χρωμόσωμα του *Arthrobacter/Paenarthrobacter*), σύμφωνα με τον λειτουργικό χαρακτηρισμό του προγράμματος Prokka, εντοπίστηκαν ανοιχτά πλαίσια ανάγνωσης για (Εικόνα 36): μια τρανσφεράση του συνενζύμου-A, μια καρβοξυλάση του ακέτυλο-συνενζύμου-A, δυο υποθετικές πρωτεΐνες, μια υδρολάση, μια καρβοξυλεστεράση, την δέλτα υπομονάδα της DNA πολυμεράσης III και μια κινάση. Ανοδικά του γονιδίου της “υδρολάσης του ιππουρονικού”, σύμφωνα με τον λειτουργικό χαρακτηρισμό του προγράμματος Prokka, εντοπίστηκαν ανοιχτά πλαίσια ανάγνωσης για: μια υποθετική πρωτεΐνη, μια πρωτεΐνη της οικογένειας DUF5058, μια ακυλαμιδάση, έναν μεταγραφικό παράγοντα, μια ρεδουκτάση, μια υποθετική πρωτεΐνη, έναν ενεργοποιητή του πρωτεασώματος, μια υποθετική πρωτεΐνη και έναν μεταφορέα πυροσταφυλικού.

T1.6 DuaH region



Εικόνα 36: Σχηματική απεικόνιση του τμήματος και των στοιχείων του, του βακτηριακού χρωμοσώματος του δείγματος T1.6 που περιέχει την αλληλουχία της “υδρολάσης του ιππουρονικού”.

Οι αλληλουχίες των ανοιχτών πλαισίων ανάγνωσης για υποθετικές πρωτεΐνες υποβλήθηκαν σε έλεγχο ομολογίας μέσω του προγράμματος BLAST με την NCBI-nr ως βάση δεδομένων αναφοράς. Τα αποτελέσματα της δοκιμής αυτής ήταν:

Για την πρώτη υποθετική πρωτεΐνη ανοδικά, μεγαλύτερη ομοιότητα εμφανίστηκε, με μια πρωτεΐνη η οποία περιέχει επικράτεια HipA και ομοιάζει με πρωτεΐνες που έχουν εντοπιστεί σε πολλά και ποικίλα είδη (MULTISPECIES). Για την δεύτερη υποθετική πρωτεΐνη ανοδικά, μεγαλύτερη ομοιότητα εμφανίστηκε, με πρωτεΐνη που περιέχει δομή (α-έλικα)-

(φουρκέτα)-(α-έλικα) η οποία εντοπίστηκε σε πολλά και ποικίλα είδη (MULTISPECIES).

Για την πρώτη υποθετική πρωτεΐνη καθοδικά, μεγαλύτερη ομοιότητα εμφανίστηκε, με bit score 415, με μια υποθετική πρωτεΐνη που εντοπίστηκε σε *Paenarthrobacter sp. CM16*. Για την δεύτερη υποθετική πρωτεΐνη καθοδικά, μεγαλύτερη ομοιότητα εμφανίστηκε, με μια πρωτεΐνη της οικογένειας YbdD/YjiX που εντοπίστηκε σε *Paenarthrobacter nitroguajacolicus*. Για την τρίτη υποθετική πρωτεΐνη καθοδικά, μεγαλύτερη ομοιότητα εμφανίστηκε, με μια υποθετική πρωτεΐνη που εντοπίστηκε σε *Arthrobacter sp. Hiyo4*.

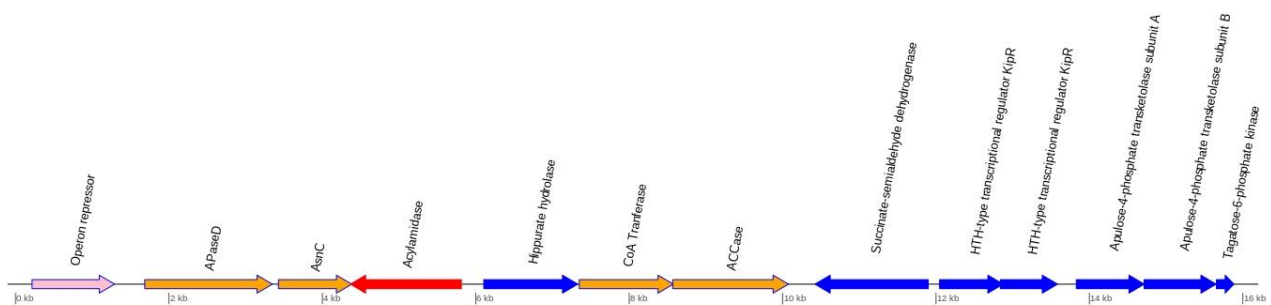
Marco

IpaH/Acylamidase-DuaH/Hippurate Hydrolase

Οι γονιδιακοί τόποι στους οποίους εμφανίζονται τα ανοιχτά πλαίσια ανάγνωσης που μοιάζουν με την IpaH (GMNKJGKN_00801) και την DuaH (GMNKJGKN_00802) στο βακτηριακό χρωμόσωμα του δείγματος Marco αλληλεπικαλύπτονται καθώς τα ανοιχτά πλαίσια ανάγνωσης αυτά απέχουν περίπου 200 ζ.β. μεταξύ τους.

Στα δεξιά των γονιδίων της ακυλαμιδάσης και υδρολάσης του ιππουρονικού, σύμφωνα με τον λειτουργικό χαρακτηρισμό του προγράμματος Prokka, εντοπίστηκαν ανοιχτά πλαίσια ανάγνωσης για (Εικόνα 37): μια τρανσφεράση του συνενζύμου-A, μια καρβοξυλάση του ακέτυλο-συνένζυμου-A, μια δευδρογονάση της ηλεκτρυλο-ημιαλδεΐδης, δύο μεταγραφικούς παράγοντες KipR τύπου HTH, τις α και β υπομονάδες μιας τρανσφεράσης κετονικής ομάδας και μια κινάση. Στα αριστερά εντοπίστηκαν ανοιχτά πλαίσια ανάγνωσης για: έναν μεταγραφικό παράγοντα, μια φωσφατάση και έναν αναστολέα οπερονίου.

Marco IpaH and DuaH region

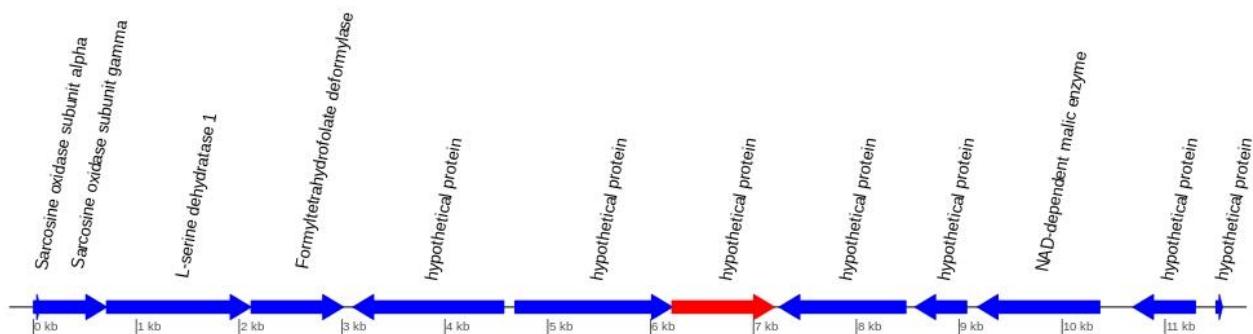


Εικόνα 37: Σχηματική απεικόνιση του τμήματος και των στοιχείων του, του βακτηριακού χρωμοσώματος του δείγματος Marco, που περιέχει τις αλληλουχίες της “ακυλαμιδάσης” και “υδρολάσης του ιππουρονικού”.

DdaH

Καθοδικά του γονιδίου της “υποθετικής πρωτεΐνης” (GMNKJGKN_04118), σύμφωνα με τον λειτουργικό χαρακτηρισμό του προγράμματος Prokka και η οποία χαρακτηρίστηκε ως DdaH σύμφωνα με δοκιμή NCBI-nr BLAST, εντοπίστηκαν: δύο(2) ανοιχτά πλαίσια ανάγνωσης υποθετικών πρωτεϊνών, ένα ανοιχτό πλαίσιο ανάγνωσης μιας NAD-εξαρτώμενης αποκαρβοξυλάσης και δύο(2) ανοιχτά πλαίσια ανάγνωσης υποθετικών πρωτεϊνών. Ανοδικά του της DdaH, εντοπίστηκαν: δύο(2) ανοιχτά πλαίσια ανάγνωσης υποθετικών πρωτεϊνών, ένα ανοιχτό πλαίσιο ανάγνωσης για μια αποφορμυλάση του φορμυλ-τετραυδροφολικού οξέως, ένα ανοιχτό πλαίσιο ανάγνωσης για μια λυάση και δύο ανοιχτά πλαίσια ανάγνωσης για τις α και γ υπομονάδες μιας οξειδάσης.

Marco DdaH region



Εικόνα 38: Σχηματική απεικόνιση του τμήματος και των στοιχείων του, του βακτηριακού χρωμοσώματος του δείγματος Marco, που περιέχει την αλληλουχία της “υποθετικής πρωτεΐνης” η οποία ομοιάζει με το γονίδιο αναφοράς *ddaH*.

Οι αλληλουχίες των ανοιχτών πλαισίων ανάγνωσης για υποθετικές πρωτεΐνες υποβλήθηκαν σε BLAST με την NCBI-nr ως βάση δεδομένων αναφοράς. Τα αποτελέσματα του ελέγχου αυτού ήταν:

Για την πρώτη υποθετική πρωτεΐνη ανοδικά, μεγαλύτερη ομοιότητα εμφανίστηκε, με bit score 1005, με μια πρωτεΐνη διαμεμβρανικού μεταφορέα κυτοσίνης που εντοπίστηκε σε πολλά και ποικίλα είδη (MULTISPECIES). Για την δεύτερη υποθετική πρωτεΐνη ανοδικά, μεγαλύτερη ομοιότητα εμφανίστηκε, με bit score 943, με πρωτεΐνη μεταγραφικού παράγοντα οικογένειας PucR που εντοπίστηκε σε πολλά είδη του γένους *Paenarthrobacter*.

Για την πρώτη υποθετική πρωτεΐνη καθοδικά, μεγαλύτερη ομοιότητα εμφανίστηκε, με bit score 842, με μια τρανσποζάση οικογένειας IS30 που εντοπίστηκε σε *Paenarthrobacter*

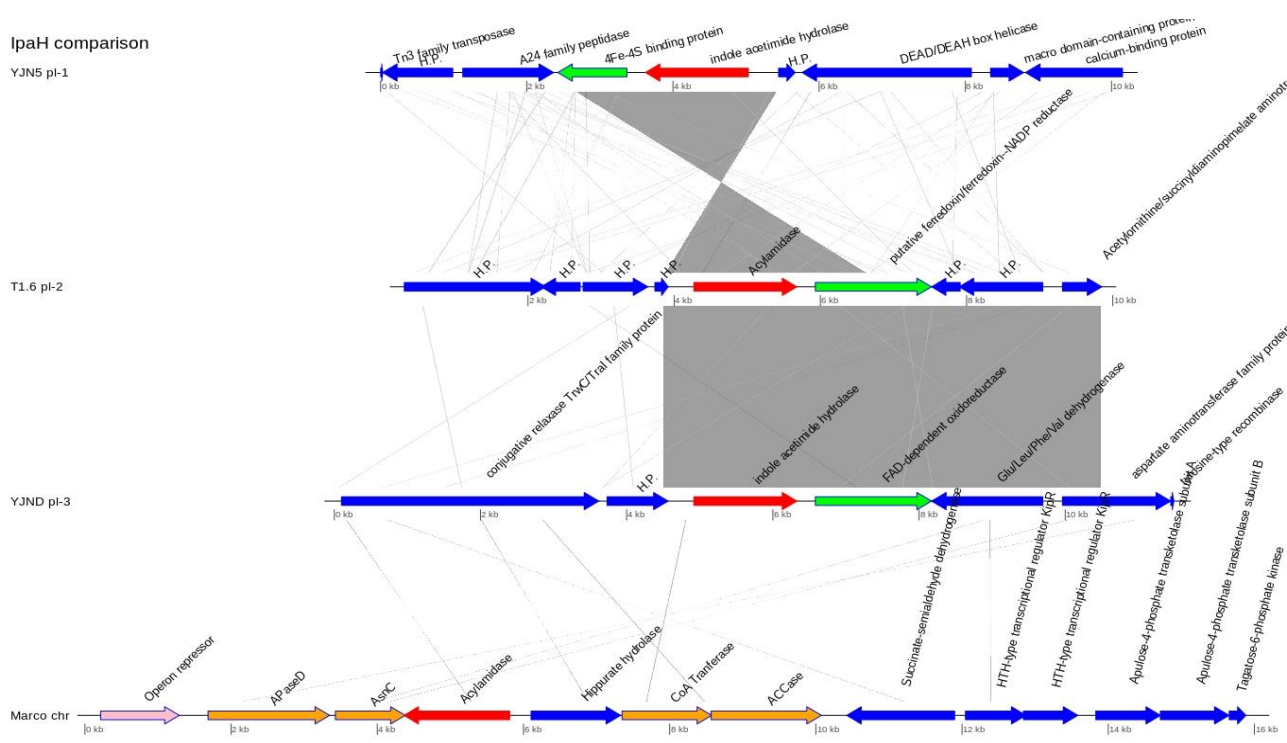
nicotinovorans. Για την δεύτερη υποθετική πρωτεΐνη καθοδικά, μεγαλύτερη ομοιότητα εμφανίστηκε, με bit score 310, με μια πρωτεΐνη που περιέχει επικράτεια *DUF456* που εντοπίστηκε σε *Arthrobacter sp. MWB30*. Για την τρίτη υποθετική πρωτεΐνη καθοδικά, μεγαλύτερη ομοιότητα εμφανίστηκε, με bit score 410, με μεταγραφικό παράγοντα οικογένειας TetR/AcrR που εντοπίστηκε σε πολλά είδη (MULTISPECIES). Για την τέταρτη υποθετική πρωτεΐνη ανοδικά, μεγαλύτερη ομοιότητα εμφανίστηκε, με την υποθετική πρωτεΐνη ASF74_19955 που εντοπίστηκε σε *Arthrobacter sp. Leaf145*.

3.4 Σύγκριση γονιδιακών τόπων των υπεύθυνων για τον καταβολισμό του iprodione γονιδίων

Αφού εξετάστηκε ο γονιδιακός τόπος των γονιδίων που ομοιάζουν με τα γονίδια αναφοράς στα δείγματα T1.6/1.8 και Marco, ακολούθησε σύγκριση των γονιδιακών τόπων αυτών με τους γενετικούς τόπους στα γονιδιώματα δυο στελεχών *Paenarthrobacter*, των YJN-D και YJN-5, τα οποία παρουσιάζουν φαινότυπο καταβολισμού του Iprodione και απομονώθηκαν για παλαιότερη εργασία των Zhang et al. (2018).

Οι συντεταγμένες των τμημάτων DNA στα οποία βρίσκονται τα ανοιχτά πλαίσια ανάγνωσης ενδιαφέροντος για τα δείγματα T1.6 και Marco όπως επίσης και για τα βακτήρια *Paenarthrobacter sp. YJN-D* και *YJN-5*, φαίνονται στο παράρτημα VII.

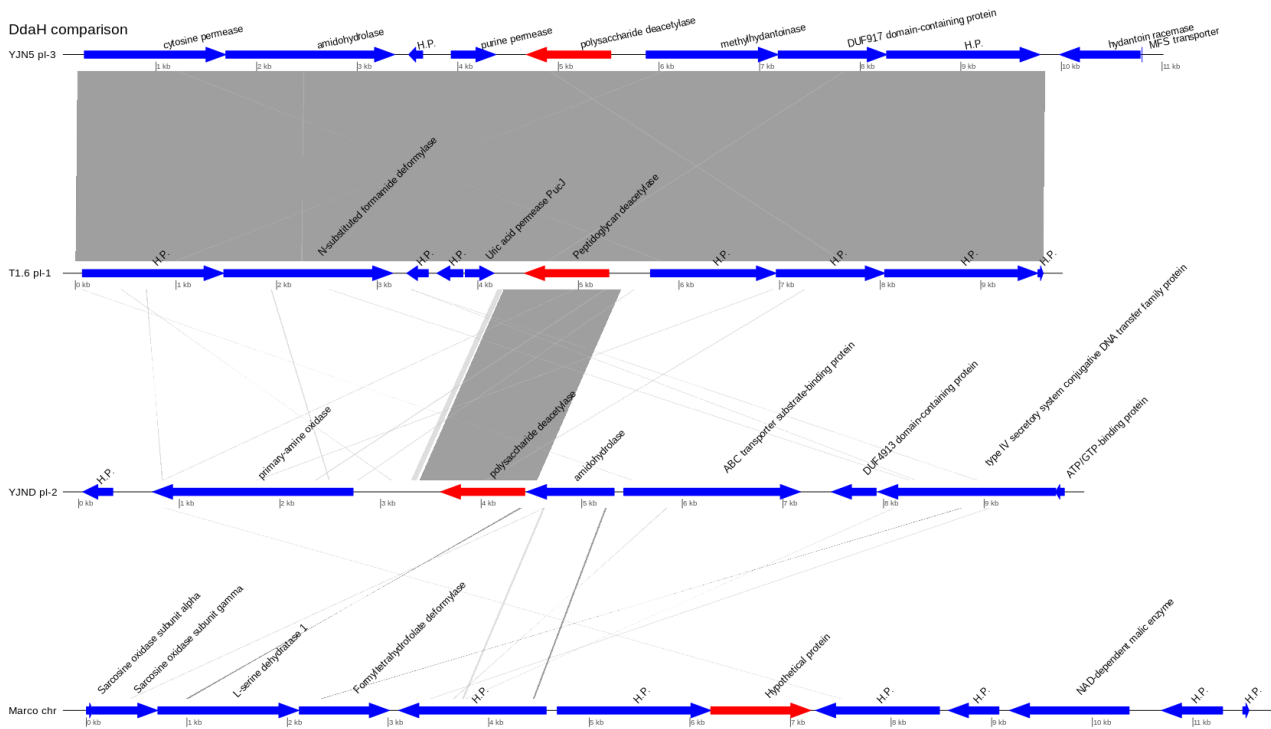
Στην Εικόνα 39 φαίνεται η σύγκριση των γενετικών τόπων του γονιδίου που κωδικεύει την υποτιθέμενη IpaH (indole acetimide hydrolase) μεταξύ των στελεχών *Paenarthrobacter* YJN-5 και YJN-D, και των δειγμάτων T1.6 και Marco. Στο πλασμίδιο του δείγματος YJN-5 το γονίδιο κωδικοποιεί πρωτεΐνη που μεταφέρει ηλεκτρόνια με την βοήθεια ιόντος σιδήρου, ενώ στα δείγματα T1.6 και YJN-D οι μεταφορείς είναι NADP και FAD εξαρτώμενοι. Στο δείγμα Marco, τα γονίδια που ταυτοποιήθηκαν ως υπεύθυνα για το πρώτο βήμα (ακυλαμιδάση) και το τρίτο βήμα (υδρολάση του ιππουρονικού) του μεταβολισμού του iprodione βρίσκονται κοντά μεταξύ τους, απέχουν περίπου 200 ζεύγη βάσεων.



Εικόνα 39: Σύγκριση γονιδιακών τόπων του γονιδίου που κωδικεύει την IpaH, το ένζυμο που καταλύει το πρώτο βήμα καταβολισμού του Iprodione. Τα γονίδια χρωματισμένα με κόκκινο αναπαριστούν τις αμιδάσες-στόχους, τα γκρί πλαίσια και οι γραμμές αναπαριστούν όμοια μπλοκ σύμφωνα με σύγκριση BLAST. Με πράσινο χρώμα είναι χρωματισμένα τα ανοιχτά πλαίσια ανάγνωσης που βάση λειτουργικού χαρακτηρισμού ομοιάζουν σε λειτουργία ως οξειδοαναγωγάσες (μεταφορείς ηλεκτρονίων) τα οποία εντοπίζονται στα πλασμίδια των δειγμάτων YJN-5, T1.6, και YJN-D. Τα χρωματισμένα γονίδια (πορτοκαλί και ροζ) στο δείγμα Marco αναπαριστούν δομή που εμφανίζεται και στα υπόλοιπα δείγματα, αλλά ομοιάζει όταν εξετάζεται ο γενετικός τόπος της αμιδάσης DdaH, του γονιδίου υπεύθυνου για το τρίτο βήμα μεταβολισμού του Iprodione, ο γενετικός αυτός τόπος εξετάζεται και αναλύεται παρακάτω.

Σύγκριση γονιδιακών τόπων που βρίσκονται τα ανοιχτά πλαίσια ανάγνωσης που ομοιάζουν με την DdaH.

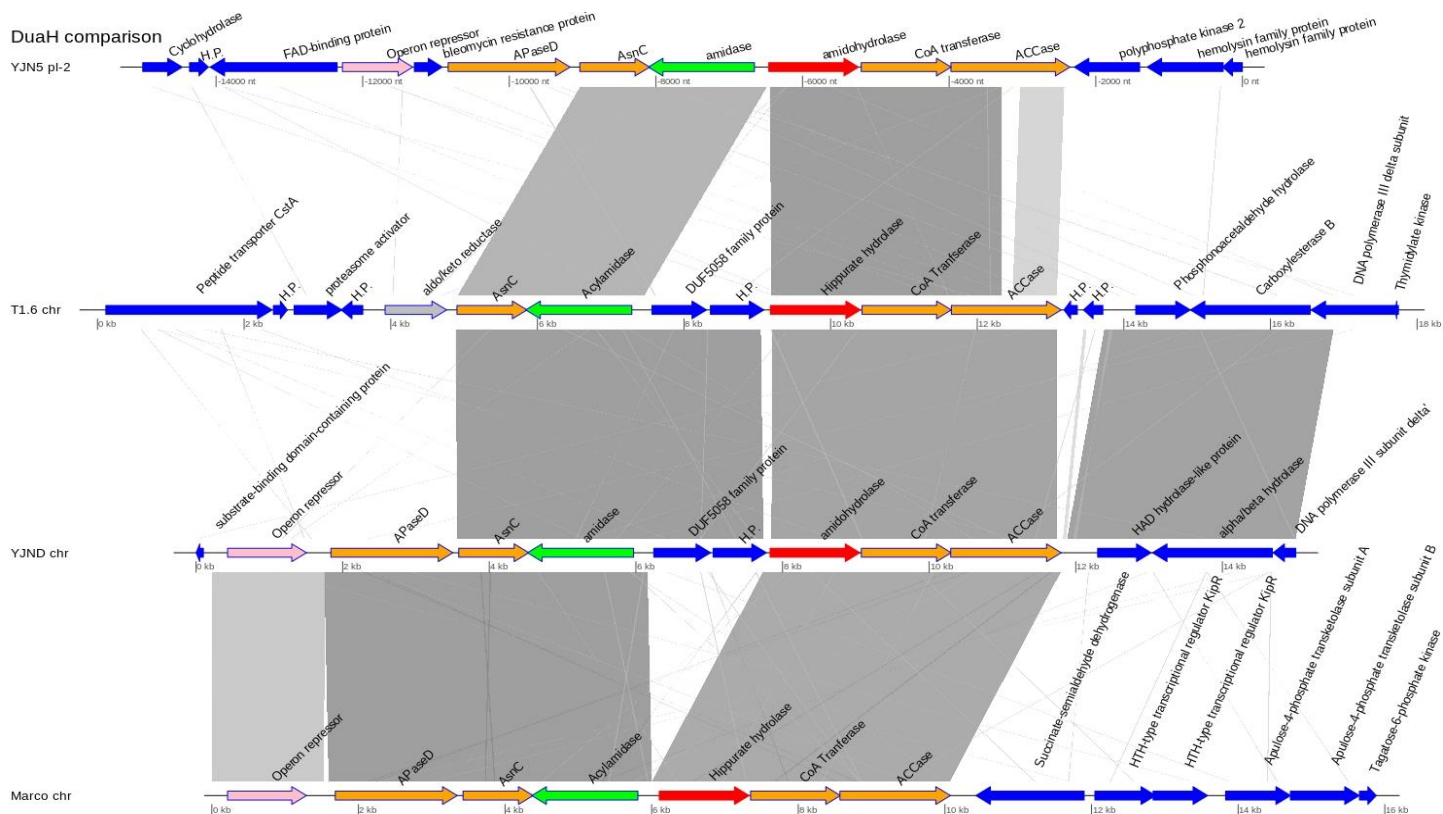
Στην εικόνα 40 φαίνεται η σύγκριση για τους γενετικούς τόπους όπου εντοπίζεται το γονίδιο υπεύθυνο για την κωδικοποίηση της DdaH (polysaccharide/peptidoglycan deacetylase) μεταξύ των στελεχών *Paenarthrobacter* YJN-5 και YJN-D, και των δειγμάτων T1.6 και Marco. Τα γονίδια χρωματισμένα με κόκκινο αναπαριστούν τις αμιδάσες-στόχους, τα γκρί πλαίσια και οι γραμμές αναπαριστούν τα αποτελέσματα στοίχισης που προέκυψαν από την σύγκριση των αλληλουχιών που έγινε με BLAST.



Εικόνα 40: Σύγκριση γονιδιακών τόπων όπου εντοπίζεται το γονίδιο που κωδικοποιεί την *DdaH*, το ένζυμο που καταλύει το δεύτερο βήμα καταβολισμού του *Iprodione*. Τα γονίδια χρωματισμένα με κόκκινο αναπαριστούν τις αμιδάσες-στόχους, τα γκρι πλαίσια και οι γραμμές αναπαριστούν όμοια μπλοκ σύμφωνα με τα αποτελέσματα στοίχισης με BLAST.

Σύγκριση γονιδιακών τόπων που βρίσκονται τα ανοιχτά πλαίσια ανάγνωσης που ομοιάζουν με την *DuaH*.

Στην Εικόνα 41 φαίνεται η σύγκριση για τους γονιδιακούς τόπους όπου εντοπίζεται το γονίδιο υπεύθυνο για την κωδικοποίηση της *DuaH* (Hippurate hydrolase/amidohydrolase) μεταξύ των στελεχών *Paenarthrobacter* YJN-5 και YJN-D, και των δειγμάτων T1.6 και Marco. Οι κοινές αυτές πρωτεΐνες είναι -όπως απεικονίζονται στο δείγμα Marco από αριστερά προς δεξιά- ένας καταστολέας οπερονίου (μεγαλύτερη ομοιότητα έφερε με τον καταστολέα οπερονίου λακτόζης), που απεικονίζεται με ροζ χρώμα, μια αλκαλική φωσφατάση D (APaseD) και ένας μεταγραφικός παράγοντας (*AsnC*) που απεικονίζονται με πορτοκαλί χρώμα. Έπειτα, ακολουθεί ένα γονίδιο ακυλαμιδάσης, το οποίο σημαδεύεται με πράσινο χρώμα, το οποίο για το δείγμα Marco είναι το γονίδιο που φέρει την μεγαλύτερη ομολογία με την *IraH*. Ακολουθούν, ένα γονίδιο Τρανσφεράσης Συνενζύμου Α (*CoA Transferase*) και ένα γονίδιο καρβοξυλάσης του ακέτυλο-συνενζύμου Α (*Acetyl CoA Carboxylase* (ACCase)) τα οποία επίσης είναι χρωματισμένα με πορτοκαλί χρώμα. Επίσης, στο δείγμα T1.6, δεν εντοπίζονται τα γονίδια του καταστολέα οπερονίου και της αλκαλικής φωσφατάσης D. Εντοπίζεται ένα γονίδιο ρεδουκτάσης, το οποίο μερικώς ομοιάζει σε περιοχές του με τα γονίδια του καταστολέα οπερονίου, του μεταγραφικού παράγοντα (*AsnC*) και της καρβοξυλάσης του ακέτυλο-συνενζύμου Α.

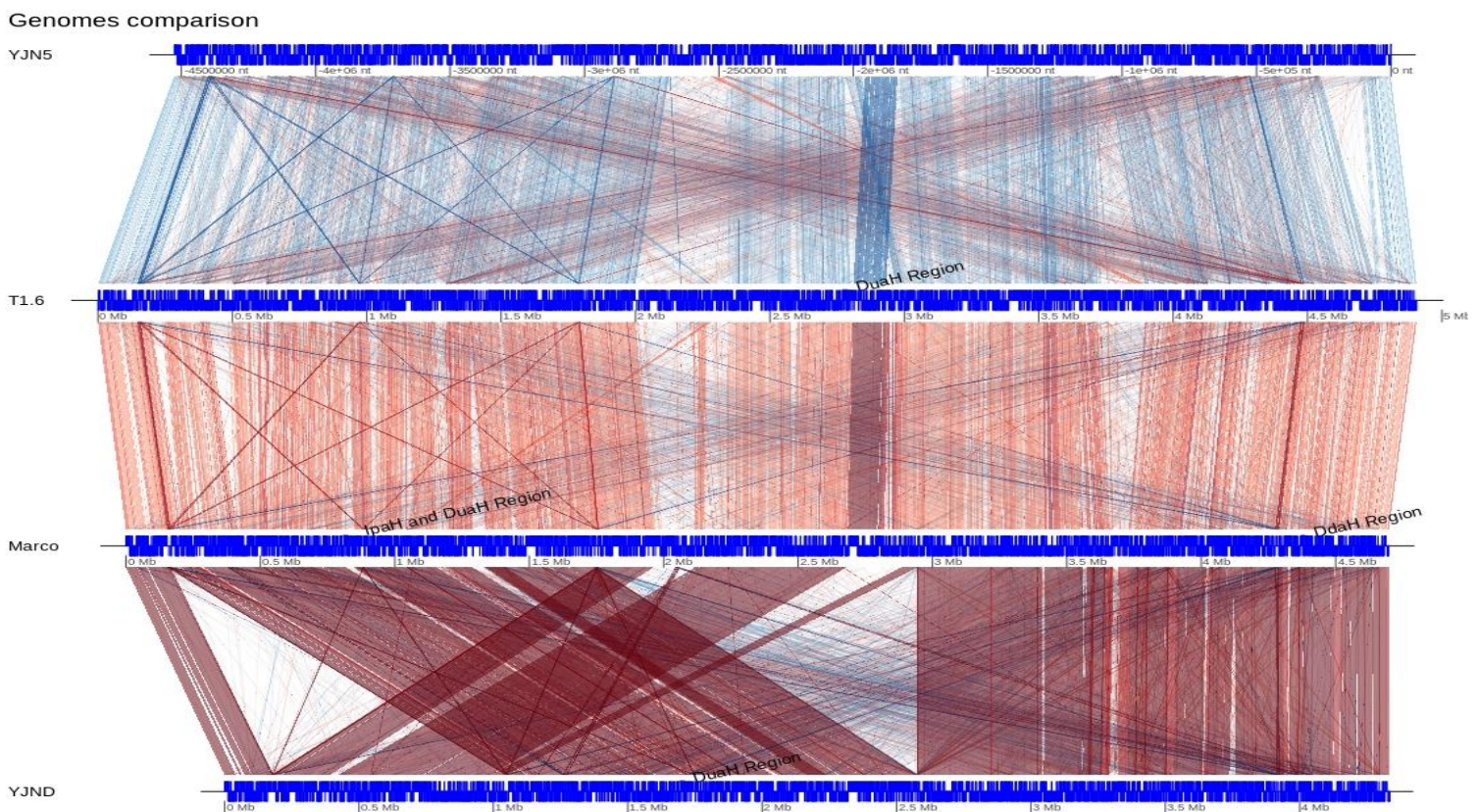


Εικόνα 41: Σύγκριση γονιδιακών τόπων όπου εντοπίζεται το γονίδιο που κωδικοποιεί την *DnaH*, το ένζυμο που καταλύει το τρίτο βήμα καταβολισμού του *Iprodione*. Τα γονίδια χρωματισμένα με κόκκινο αναπαριστούν τις αμιδάσες-στόχους, τα γκρί πλαίσια και οι γραμμές αναπαριστούν τα αποτελέσματα στοίχισης που προέκυψαν από την σύγκριση των αλληλουχιών που έγινε με *BLAST*. Με πορτοκαλί, πράσινο και ροζ χρώμα σημαδεύτηκαν τα γονίδια που κωδικοποιούν παρόμοιες πρωτεΐνες και εντοπίζονται σε όλα τα δείγματα.

3.5 Σύγκριση γονιδιωμάτων *Paenarthrobacter*

Έγινε σύγκριση επί του συνόλου των γονιδιωμάτων των στελεχών *Paenarthrobacter* YJN-5 και YJN-D, και των *Arthrobacter/Paenarthrobacter* των T1.6 και Marco για τα βακτηριακά χρωμοσώματα. Η σύγκριση και απεικόνιση της αναπαράστασης αυτής, πραγματοποιήθηκαν με χρήση του GenoPlotR.

Στην Εικόνα 42 δίδεται η σύγκριση των χρωμοσωμάτων των στελεχών *Paenarthrobacter* YJN-5 και YJN-D, και των δειγμάτων T1.6 και Marco.

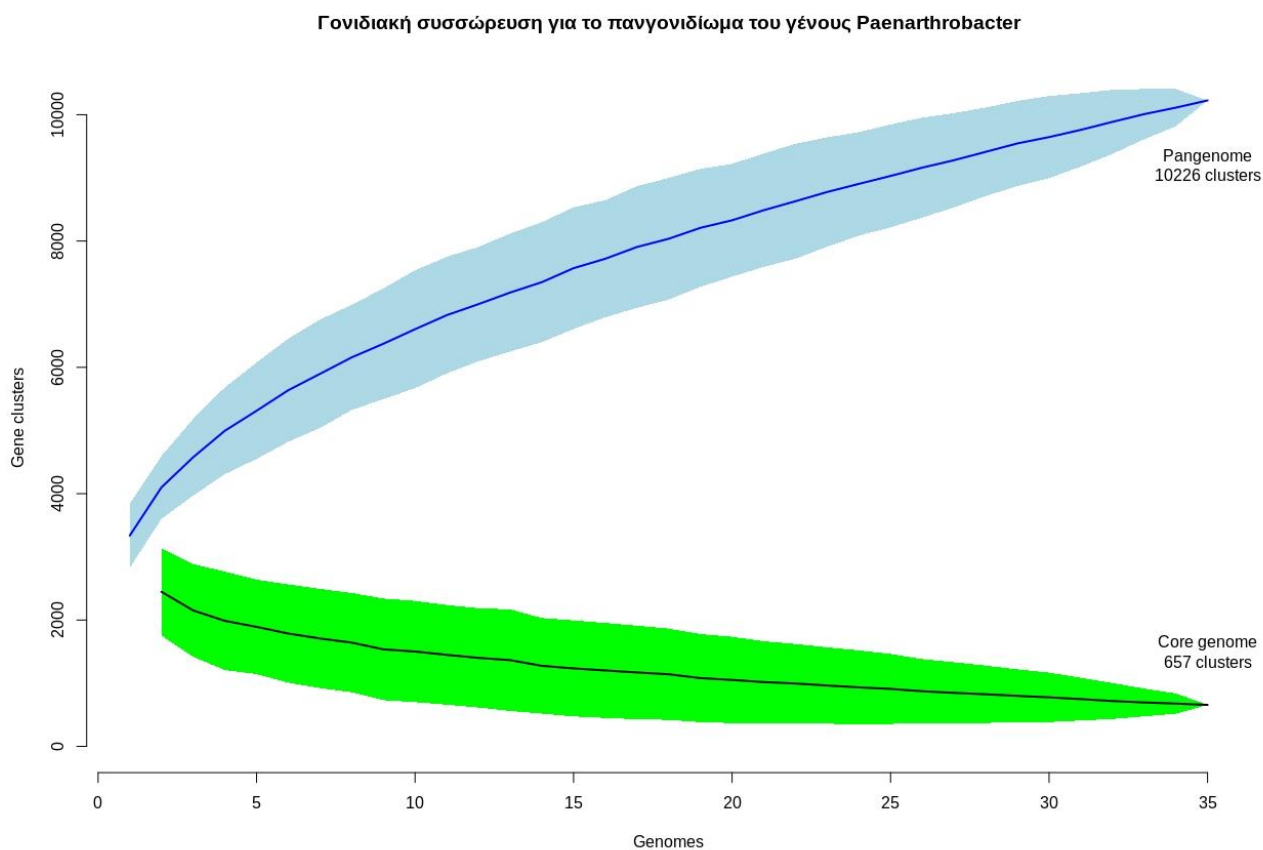


Εικόνα 42: Σύγκριση βακτηριακών χρωμοσωμάτων των δειγμάτων YJN-5, T1.6, YJN-D και Marco. Οι γραμμές που ενώνουν τμήματα των χρωμοσωμάτων με κόκκινο ή μπλε χρώμα αναπαριστούν ταύτιση των αλληλουχιών σε εκείνο το σημείο. Με κόκκινο χρώμα απεικονίζεται η ταύτιση στοίχισης αλληλουχιών σε ίδια φορά ($5' \rightarrow 3'$ και οι δύο αλληλουχίες ενώ βρίσκονται στον κλώνο θετικής πολικότητας), ενώ με μπλε σημαδεύεται η ταύτιση ανεστραμμένων αλληλουχιών ($5' \rightarrow 3'$ και οι δύο αλληλουχίες ενώ βρίσκονται σε κλώνους διαφορετικής μεταξύ τους πολικότητας). Εντοπίζονται στο γράφημα οι γενετικοί τόποι όπου απαντούν τα γονίδια-στόχοι καταβολισμού του *iprodione*, όσοι βρίσκονται στα βακτηριακά χρωμοσώματα των δειγμάτων. Για τις γραμμές που αναπαριστούν την ταύτιση μεταξύ των γονιδιωμάτων των YJN-5 και T1.6 ο χρωματικός κώδικας κόκκινου-μπλε είναι αντίθετος.

3.6 Ανάλυση του πανγονιδιώματος του γένους *Paenarthrobacter*

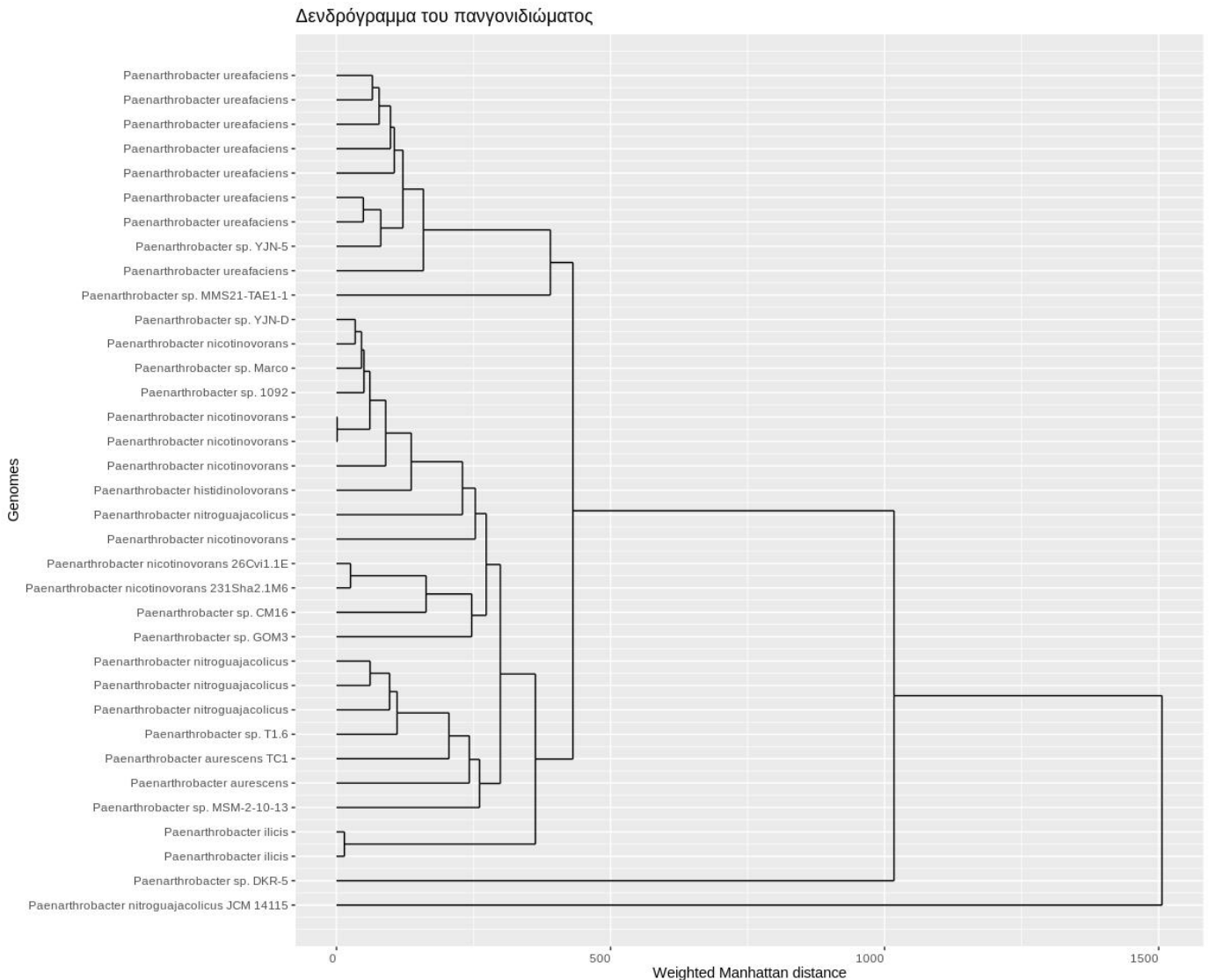
Για την ανάλυση του πανγονιδιώματος για το γένος *Paenarthrobacter* χρησιμοποιήθηκαν 33 γονιδιώματα του γένους από την πλατφόρμα του πλατφόρμα GenBank και υπολογίστηκε το πανγονιδίωμα με την προσθήκη των δειγμάτων T1.6 και Marco ως είδη του γένους *Paenarthrobacter*. Η δημιουργία του πανγονιδιώματος και η ανάλυση του έγιναν με αξιοποίηση του εργαλείου/πακέτου προγραμμάτων micrوران του προγράμματος R.

Ομαδοποίηση των πρωτεϊνών σε συστάδες (clusters) ορθολόγων βάση της ομοιότητας τους, οδήγησε στον υπολογισμό ~657 ορθολόγων που αποτελούσαν το πυρηνικό γονιδίωμα και 10.226 ορθολόγων που αποτελούσαν το παν-γονιδίωμα (Εικόνα 43).



Εικόνα 43: Οι καμπύλες γονιδιακής συσσώρευσης για το πυρηνικό γονιδίωμα και το πανγονιδίωμα. Οι κεντρικές γραμμές αναπαριστούν τις καμπύλες συσσώρευσης και οι χρωματισμένες περιοχές εκατέρωθεν των καμπυλών αναπαριστούν τις τυπικές αποκλίσεις.

Έπειτα, η ανάλυση παν-γονιδιώματος οδήγησε στην φυλογονιδιωματική απεικόνιση των στελεχών που χρησιμοποιήθηκαν στην ανάλυση (Εικόνα 44), τοποθετώντας και τα T1.6/Marco χρωμοσώματα που δοκιμάστηκαν στο γένος *Paenarthrobacter*.



Εικόνα 44: Δενδρόγραμμα των σχέσεων των γονιδιωμάτων του πανγονιδιώματος μεταξύ τους, βασισμένο στον εντοπισμό ομοίων γονιδιακών συστάδων σε έναστο γονιδίωμα.

Τέλος, πραγματοποιήθηκε έλεγχος για την τοποθέτηση των ανοιχτών πλαισίων ανάγνωσης που ομοιάζουν με τα γονίδια αναφοράς στο πυρηνικό γονιδίωμα. Αρχικά, εντοπίστηκαν οι συστάδες γονιδίων που συγκροτούν το πυρηνικό γονιδίωμα. Ταυτοποιήθηκαν τα γονίδια που αντιστοιχούν σε αυτές τις συστάδες για το γονιδίωμα του T1.6 (εφόσον πρόκειται για πυρηνικό γονιδίωμα, ομόλογα γονίδια υπάρχουν σε όλα τα γονιδιώματα που μελετήθηκαν). Απομονώθηκαν οι πρωτεϊνικές αλληλουχίες που αντιστοιχούν στα γονίδια αυτά και έγινε BLAST των αλληλουχιών των IpaH, DdaH και DuaH σε βάση δεδομένων των υποθετικών μεταγράφων του πυρηνικού γονιδιώματος για το γονιδίωμα του T1.6. Στον έλεγχο αυτό δεν παρατηρήθηκε υψηλή ομολογία με την DdaH (το καλύτερο αποτέλεσμα είχε bit score 23,9

και e-value 2,8), όμως παρατηρήθηκε αξιόλογη ομολογία με την IraH και με την DuaH. Το καλύτερο αποτέλεσμα για την IraH φαίνεται στην Εικόνα 45 και το καλύτερο αποτέλεσμα για την DuaH φαίνεται στην Εικόνα 46.

```
> GID34_seq1217 Seqid=CP000000.1;Start=1582613;End=1584166;Strand=+
Length=517

Score = 114 bits (285), Expect = 6e-29, Method: Compositional matrix adjust.
Identities = 135/492 (27%), Positives = 219/492 (45%), Gaps = 44/492 (9%)

Query 1  MSDQLWSKSATELAGLVRSKVVSATEVVEAHLQRIEDVNPENAIIVVLADDARQSARIA 60
+ ++L SA +LA + ++ V+A EV +AHL RI DV+ ++NA V ++A A
Sbjct 4  LKNELIHHSAADLAAKLAAREVTAVEVTQAHLDRADVDGQVNAFLHVNTEEALAVAAEV 63

Query 61 DEKTRTEPDQLGRLH---GVPITLKVNIIDLVGSAVDAVPAFKDFYPPMNTPLVDRLLGE 117
D LH GVPI +K I +G T+ + ++ P + +V +L
Sbjct 64 DAARAAGGSAAEELHALAGVPIAVKDLIVTIGQPTTAGSKILEGWHSPYDATVVKKLRAA 123

Query 118 GAIIVGRTNMSDMGMRMTTDDSTLHGLTRNPWHPGRTAGGSSGGEGAALASGMSALGVGND 177
++G+TN+ + M +T+ + +G TRNPW R GGS GG AA+A+ + L +G D
Sbjct 124 KMPILGKTNLDEFAMGSSTEHSAYGPTRNPDLDRIIPGGSGGSAVAFAEAPLALGTD 183

Query 178 LVGSLRNPAAHCCGISTLKPTPGRIWADSIPLPDGPLSFQMMLVHGPMARQVADVRLGME 237
GS+R P G +KPT G + +I S + GP++R V D L E
Sbjct 184 TGGSIQPGAVTGTGVGKPTYGAVSRYGAIAMAS---SLDQI---GPNVSRVLDLALLQE 237

Query 238 IMSGAHP-----RDPYSVDVPLYRADSNQPFKVAVMAAPPGAPTDPEVSSVVRKAGEAV 291
++ G P DP++ V R + K+ ++ G V + ++ + +
Sbjct 238 VIGGHDPFDSTSLTDPFNDLVAAARVGNVAGMKIGIVKELHGEGYQAGVENRFNESLQLL 297

Query 292 AAAGYEVEIDAPEYLETRQIWLDFLMTEV--NVLQ-DHITEAM-----GPAGRQFLRG 342
AG E+ E+ P + + +EV N+ + D + M GP + + G
Sbjct 298 KDAGAEIVEVSCPNLKYALGAYYLIMPSEVSSNLAKFDGVRFGMRVLPKDGPMTIERVVMG 357

Query 343 FLDHAGPLD-----LPGYIDLFVR-----RRRMSRVWINDFFQEYPVIIAPTW 384
AG D GY D + R + R ++ F + V+I+PT
Sbjct 358 ATRAAGFGDEVKRIILGTYALSAGYYDAYYGSAQKVRTLIQRDFDAAFKADVLISPTA 417

Query 385 LDVAFEHDWDLKSVADTVNCGPVVPANVGLPAAVTFGGLA--KGMPVGVQCIAAAFHD 442
AF+ L N +PAN+ G+P GGLA G+PVG+Q +A A D
Sbjct 418 PTTAFKLGKLDPLAMYLNDVATIPANLAGIPGLSLPGGLADEDGLPVGIIQLLAPARQD 477

Query 443 DQAMQVAEVEIA 454
+ +V V+E+
Sbjct 478 ARLYRVGAVLES 489
```

Εικόνα 45: Στοιχισμός και στατιστικά του καλύτερου αποτελέσματος της αναζήτησης BLAST βάσει bit score της πρωτεϊνικής αλληλουχίας του IraH έναντι των υποθετικών μεταγραφών του πυρηνικού γονιδιώματος για το γονιδίωμα του T1.6.

```
> GID34_seq845 Seqid=CP000000.1;Start=1075087;End=1076310;Strand=-
Length=407
```

```
Score = 362 bits (929), Expect = 3e-123, Method: Compositional matrix adjust.
Identities = 194/395 (49%), Positives = 246/395 (62%), Gaps = 10/395 (3%)
```

```
Query 15 TTALHELYRRLHAHPELSMQEHATAALIEEQLDRLGAD-HFRCGRTGVVGI LRNGEGPVV 73
T L ELY+ LHAHPELS QE TA + L+ LG H G+TGWVG+L NG GPVV
Sbjct 2 TIDLEELYKDLHAHPELSFQETRTAGIAAGYLEDLGFTVHRNIGKTGVVGVLDNGPGPVV 61

Query 74 AFRADTDGLPIEEATGLDYASTDTGM-LDGSEVPVMHGC GHDTHVASLLGAAEILAGNQD 132
RAD D LP++EATGLDYAST G+ +G ++PVMH CGHD HV+ L+GA ++LA +
Sbjct 62 MLRADMDALPVKEATGLDYASTAMGVDHEGKDPVMHACGHDVHV SCLVGAVDVLAQRH 121

Query 133 AWSGTLVLIFQPGEETAAGALAMLEDGLWDRAPRPEVVFQHVMPRPVGTVSISSGPVAA 192
W GTL+ +FQP EE GA M+ DGL+D P+P+VV GQHV P P G I G A
Sbjct 122 EWQGT LIALFQPAEEWGGGAETMVADGLYDAVPKPDVVLGQHVAPFPAGWFGIRPGVTMA 181

Query 193 MADSLRVTVHGQQSHG SQPQDSIDPIVMAAHMVTRLQTIVSRELDPRKSAVVTVGTFHAG 252
ADSL +T+HGQ HGS+P+ ++DP++MAA RLQ IVSREL SAVVTVG H+G
Sbjct 182 SADSLNITMHGQGGHGSRPETTVPVLM AAATTVRLQGIVSRELAASDSAVVTVGQIHSG 241

Query 253 LKENIIPASAEFTLNIRTFDEEVRGQVLTAVRRIIEAEAAAASGAPG-PKIEELYRFPQCY 311
K NIIP +A L++RTF + +R ++L ++ RI + E+ ASGAP P+I FP
Sbjct 242 TKNNIIPETATLGLSVRTFSDAIRDKILASIERIAKGESTASGAPKDPEIHFEHFPLTV 301

Query 312 NDPDAVPAVLEALRDALGAESVEVTPPMGSEDFGHLGTAIGVPSVFWFMFGGTPGAVFDG 371
NDP A V A R G + P+ GSED G L A G P VFW GG ++F
Sbjct 302 NDPAATERVSAAFRAKFGETQIIDPGPVSGSEVDGALAKAAGAPLVFWFLGGAEP SLFKE 361

Query 372 -----PGPVPVNHSPPFAPELEGLAGGISA AV 399
P +P NHSP+FAP ++ +LA G A V
Sbjct 362 WATTGRLPEDIPSNHSPYFAPVIQPTLARGTEALV 396
```

Εικόνα 46: Στοιχισή και στατιστικά του καλύτερου αποτελέσματος της αναζήτησης BLAST βάσει bit score της πρωτεϊνικής αλληλουχίας του DuaH έναντι των υποθετικών μεταγράφων του πυρηνικού γονιδιώματος για το γονιδίωμα του T1.6.

4. Συζήτηση

Βακτήρια που αποικοδομούν φυτοφάρμακα έχουν απομονωθεί από παρθένα περιβάλλοντα χωρίς προηγούμενη έκθεση σε φυτοφάρμακα (*Kitagawa et al. 2002 · Sanchez et al. 2004 · Nautiyal et al. 2008*). Αυτό αποδόθηκε στην τακτική έκθεση μικροβίων της ριζόσφαιρας σε φυσικά αλειφατικά και αρωματικά αλογονωμένα μόρια που παράγονται από φυτά και μικροοργανισμούς και μοιάζουν με τη χημική δομή των συνθετικών φυτοφαρμάκων (*Gribble et al. 1992 · Copping et al. 2007*), ασκώντας έτσι εξελικτική πίεση για την ανάπτυξη νέων καταβολικών μηχανισμών ξενοβιοτικών ουσιών (*Fetzner et al. 1998*). Αναφορικά με το iprodione, το τμήμα ιμιδαζολιδινιδιόνης εντοπίζεται σε φυσικές ουσίες του εδάφους όπως η αλλαντοΐνη και το ουρικό οξύ. Επομένως, αναμένεται για τις αμιδάσες που θα μπορούσαν να διασπάσουν αυτό το τμήμα, να προϋπάρχουν μεταξύ των μικροοργανισμών του εδάφους (*Burns et al. 1986*) και να μπορούν να λειτουργήσουν ως πρόγονοι νέων καταβολικών οδών. Η εκπόνηση της συγκεκριμένης εργασίας αποσκοπεί στην διερεύνηση και επιβεβαίωση ή μη του μηχανισμού μεταβολισμού του iprodione από βακτήρια του γένους *Paenarthrobacter* σύμφωνα με προηγούμενη έρευνα των *Zhang et al. 2021*. Ειδικότερα, μας ενδιαφέρει η σύγκριση μεταξύ στελεχών του είδους σε γενετικό επίπεδο.

Ο φαινότυπος καταβολισμού του iprodione παρουσιάζεται και στο δείγμα T1.6/T1.8 αλλά και στο δείγμα Marco. Επιβεβαιώθηκε ότι υπάρχουν αλληλουχίες ανοιχτών πλαισίων ανάγνωσης οι οποίες ομοιάζουν με το προτεινόμενο μεταβολικό μονοπάτι για το iprodione. Τα ανοιχτά πλαίσια ανάγνωσης που κωδικοποιούν πρωτεΐνες υπεύθυνες για τον καταβολισμό του Iprodione μπορούν να θεωρηθούν υπάρχοντα στα δείγματα της παρούσας μελέτης (T1.6/T1.8 και Marco), σύμφωνα με την ομολογία που παρουσίασαν με τις επιβεβαιωμένες πρωτεΐνες των στελεχών YJN-D και YJN-5 σε επίπεδο αμινοξικής αλληλουχίας και σύμφωνα με τον εντοπισμό ενεργών κέντρων και επικρατειών ενζυμικής δραστηριότητας που απαιτείται, μετά από έλεγχο με το Hmmer. Παρ' ολ' αυτά, οι θέσεις των ανοιχτών πλαισίων ανάγνωσης αυτών στα 2 στελέχη και τα 2 δείγματα που εξετάστηκαν δεν ήταν ίδιες. Τα ομόλογα των γονιδίων αναφοράς βρέθηκαν τόσο σε πλασμίδια όσο και σε βακτηριακά χρωμοσώματα (Πίνακας 2 · Πίνακας 3).

Πρέπει να σημειωθεί ότι το πλασμίδιο 2 μπορεί να ανήκει είτε στο *Microbacterium* είτε στο *Paenarthrobacter*. Η αμφιβολία αυτή δεν μπορεί να επιλυθεί και τα καλύτερα αποτελέσματα της στοίχισης με την nr βάση δεδομένων του NCBI για την υποθετική ακυλαμιδάση υποδεικνύουν ότι μπορεί να ανήκει εξίσου το πλασμίδιο 2 σε οποιοδήποτε από τα δύο γένη, ενισχύοντας την αμφιβολία προέλευσης του πλασμιδίου 2. Επίσης, σημαντικό είναι να σημειωθεί ότι και το πλασμίδιο 1 μπορεί να ανήκει είτε στο *Microbacterium* είτε στο *Paenarthrobacter*. Τα καλύτερα αποτελέσματα της στοίχισης με την nr βάση δεδομένων του NCBI για την υποθετική αποακετυλάση επίσης ενισχύουν την αμφιβολία προέλευσης για το πλασμίδιο 1.

Σύμφωνα με το προτεινόμενο καταβολικό μονοπάτι, τα τρία γονίδια μπορούν να δράσουν ανεξάρτητα μεταξύ τους, κάτι το οποίο ενισχύεται από τον εντοπισμό τους τόσο σε πλασμίδια όσο και σε βακτηριακά χρωμοσώματα.

Στο δείγμα Marco, τα γονίδια που συμμετέχουν στο πρώτο (*ipaH*) και τρίτο (*duaH*) στάδιο του μεταβολισμού του Iprodione απέχουν περίπου 200 ζ.β. και στον γονιδιακό τόπο εντοπίζονται στοιχεία οπερονίου. Επίσης, για τον γενετικό τόπο των ομολόγων του *duaH*, τόσο στα δυο στελέχη όσο και στα δύο δείγματα, παρατηρήθηκε μεγάλη ομοιότητα. Τα γονίδια υπεύθυνα για το τρίτο βήμα καταβολισμού, όπως και γονίδια “ακυλαμιδάσης” (πρώτο βήμα καταβολισμού για το δείγμα marco), παρατηρήθηκαν σε όλα τα δείγματα και στελέχη στο βακτηριακό χρωμόσωμα. Μάλιστα, εντοπίστηκαν συνοδευόμενα από ένα γονίδιο τρανσφεράσης του ακέτυλο-συνεζύμου A και μια καρβοξυλάση του ακετυλο-συνεζύμου A, σε ίδιες αποστάσεις, για όλα τα στελέχη και δείγματα (Εικόνα 41). Επίσης, σε όλα τα στελέχη και δείγματα, ανοδικά του γονιδίου της “ακυλαμιδάσης”, εντοπίζεται ένας μεταγραφικός παράγοντας, ο ίδιος σε όλα τα στελέχη και δείγματα (Εικόνα 41). Επιπλέον, σε όλα εκτός του δείγματος T1.6 εμφανίζεται ένα γονίδιο/ανοιχτό πλαίσιο ανάγνωσης που κωδικοποιεί αλληλουχία καταστολέα οπερονίου η οποία φέρνει σε έλεγχο BLAST μεγαλύτερη ομολογία με τον καταστολέα οπερονίου της λακτόζης (Εικόνα 41). Συνολικά, τα γονίδια/ανοιχτά πλαίσια ανάγνωσης *duaH*, *coA-transferase* και *aCCase* θα μπορούσαν να συνιστούν μέρος οπερονίου ή οπερόνιο εξ' ολοκλήρου.

Στο δείγμα Marco όλα τα γονίδια βρίσκονται στο χρωμόσωμα ενώ στο δείγμα T1.6/T1.8 τα ομόλογα των *ipaH* και *ddaH* βρίσκονται σε δύο διαφορετικά πλασμίδια. Στα γονιδιώματα άλλων *Paenarthrobacter* που διασπών το iprodione τα γονίδια του μονοπατιού είναι κατανομημένα τόσο στο βασικό χρωμόσωμα όσο και στα πλασμίδια (Yang et al. 2018 · Zhang et al. 2021), με εξαίρεση το γονίδιο του *duaH*. Το γονίδιο του *duaH* εντοπίζεται πάντα στο βασικό χρωμόσωμα. Επίσης, το γονίδιο του *duaH*, όπως και το γονίδιο του *ipaH* του δείγματος Marco (ακυλαμιδάση) ανήκουν στο πυρηνικό γονιδίωμα για το γένος *Paenarthrobacter* (Εικόνα 45 · Εικόνα 46). Η ομολογία που παρουσιάζουν τα δύο στελέχη και δύο δείγματα στον γενετικό τόπο του *duaH*, όπως φαίνεται στην Εικόνα 41, επιβεβαιώνει την πρόταση αυτή. Η ταυτοποίηση του γονιδίου αυτού ως κομμάτι του πυρηνικού γονιδιώματος των *Paenarthrobacter* μπορεί μερικώς να αιτιολογήσει την ικανότητα αρκετών ειδών του γένους να καταβολίσουν iprodione.

Η οπερονιακή δομή που εμφανίζεται έντονα στην γενετικό τόπο του *duaH* στο δείγμα Marco μπορεί να παραπέμπει σε «ωρίμανση» του μονοπατιού στο συγκεκριμένο βακτήριο/δείγμα. Αν ισχύει κάτι τέτοιο, τα καταβολικά γονίδια του *ipaH* που εμφανίζονται σε πλασμίδια, βρίσκονται σε πλασμίδια επειδή για τα συγκεκριμένα στελέχη/βακτήρια δεν έχει «ωριμάσει» εξίσου το καταβολικό μονοπάτι, δεν έχει ασκηθεί παρόμοια εξελικτική πίεση. Αντιθέτως, εφόσον παρατηρούνται τα γονίδια “ακυλαμιδάσης” στα δυο στελέχη και δυο δείγματα που μελετήθηκαν, θα μπορούσε να ισχύει το αντίστροφο, ότι τα πλασμιδιακά γονίδια του *ipaH* προέκυψαν από τα γονίδια της “ακυλαμιδάσης” μέσω εξελικτικής πίεσης.

Τα μεταθετά στοιχεία δεν φαίνεται να έχουν ιδιαίτερα σημαντικό ρόλο στην διασπορά του μεταβολικού μονοπατιού καθώς δεν εντοπίζονται κοντά στα γνωστά καταβολικά γονίδια. Το γεγονός αυτό δεν είναι εντελώς αναπάντεχο καθώς το γένος *Arthrobacter* (πολύ συγγενικό του *Paenarthrobacter* (Busse et al. 2016)) φαίνεται ότι δεν χρησιμοποιεί τόσο πολύ την οριζόντια γονιδιακή μεταφορά ως μηχανισμό για διασπορά του γενετικού του υλικού και την απόκτηση νέων καταβολικών δυνατοτήτων (Mihasan et al. 2013 · Yao et al. 2015). Είναι πιθανότερο ότι μετατρέπουν-εξελίσσουν προϋπάρχοντα γονίδια τους για να προσαρμοστούν στις μεταβαλλόμενες συνθήκες του περιβάλλοντος.

Η εμφάνιση του *Microbacterium* στο δείγμα T1.6/T1.8 αποτελεί προϋπόθεση για την θέσπιση θεωριών βακτηριακής κοινοπραξίας. Μάλιστα, στέλεχος του *Microbacterium* έχει ταυτοποιηθεί σε προηγούμενη μελέτη ότι μπορεί να αποικοδομήσει το iprodione (*Cao et al. · 2018 Yang et al. 2017*). Έτσι, προέκυψαν τρεις θεωρίες αναφορικά με το *Microbacterium*:

Αρχικά, είναι πιθανό ότι τα πλασμίδια του δείγματος T1.6/T1.8 ανήκουν σε ένα από τα δυο βακτήρια ή είναι μοιρασμένα μεταξύ αυτών, κάτι το οποίο μπορεί να σημαίνει ότι κανένα από τα δυο δεν καταβολίζει εξ'ολοκλήρου το iprodione, αλλά μόνο ένα ή δυο από τα ενδιάμεσα προϊόντα του μονοπατιού καταβολισμού του.

Κατά συνέχεια, είναι πιθανό να μπορούν να καταβολίσουν και τα δυο βακτήρια το iprodione. Το *Paenarthrobacter/Arthrobacter* θα το επιτύγχανε αυτό με αξιοποίηση των γονιδίων της “ακυλαμιδάσης” και του *duaH* του βασικού χρωμοσώματος, και του *ddaH* του πλασμιδίου 1. Έτσι, το *Microbacterium* μπορεί να αξιοποιεί το γονίδιο του *ipaH* του πλασμιδίου 2 και να φέρει στο βασικό του χρωμόσωμα τα υπόλοιπα απαραίτητα γονίδια.

Τέλος, είναι πιθανό ο καταβολισμός να πραγματοποιείται εξ'ολοκλήρου από το *Paenarthrobacter/Arthrobacter* και το *Microbacterium* να αποτελεί παρασιτικό οργανισμό ή να προσφέρει στην κοινοπραξία με άλλο τρόπο. Σε αυτή την περίπτωση το πιθανότερο είναι ότι τα πλασμίδια ανήκουν στο *Paenarthrobacter/Arthrobacter*.

5. Συμπεράσματα-μελλοντική έρευνα

Ολοκληρώνοντας, τα ανοιχτά πλαίσια ανάγνωσης που αναμένονταν να βρεθούν για τον καταβολισμό του Iprodione, εμφανίστηκαν στα υπό μελέτη δείγματα, ενισχύοντας την υπόθεση για το προτεινόμενο καταβολικό μονοπάτι. Από την ανάλυση synteny και την ανάλυση του πανγονιδιώματος μπορούν να γίνουν υποθέσεις αναφορικά με την εξελικτική θέση και την προέλευση του καταβολικού μονοπατιού στο γένος *Paenarthrobacter*.

Συνολικά όμως, για την παραγωγή αδιάσειστων στοιχείων είναι απαραίτητο να πραγματοποιηθεί περαιτέρω έρευνα. Οι βασικοί πυλώνες αυτής μπορούν να αποτελούν: την εύρεση και ανάλυση γενετικών στοιχείων και ρυθμιστικών αλληλουχιών που επηρεάζουν τα υπό μελέτη γονίδια, ειδικά στους γενετικούς τόπους όπου αυτά εντοπίζονται, όπως επίσης και την ενδελεχή πανγονιδιωματική ανάλυση, συμπεριλαμβάνοντας πιθανώς και τα γένη *Arthrobacter*, *Pseudarthrobacter* και *Paenarthrobacter*.

Μεταγενέστερα, με αξιοποίηση συνθετικής βιολογίας, ει δυνατόν, μπορεί να εκφραστούν επιλεκτικά τα γονίδια σε οργανισμούς-μοντέλα και να μελετηθούν ξεχωριστά από τον οργανισμό στον οποίο πρωτοεμφανίστηκαν.

6. Βιβλιογραφία

- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997 Sep 1;25(17):3389-402. doi: 10.1093/nar/25.17.3389. PMID: 9254694; PMCID: PMC146917.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) "Basic local alignment search tool." *J. Mol. Biol.* 215:403-410.
- Blystone, C.R., Lambright, C.S., Furr, J., Wilson, V.S., Gray Jr., L.E., 2007. Iprodione delays male rat pubertal development, reduces serum testosterone levels, and decreases ex vivo testicular testosterone production. *Toxicol. Lett.* 174, 74e81.
- Britannica, The Editors of Encyclopaedia. "Hydrolase". *Encyclopedia Britannica*, Invalid Date, <https://www.britannica.com/science/hydrolase>. Accessed 27 September 2021.
- Burns, R.G., 1986. Extracellular enzyme-substrate interactions in soil. In: Slatter, J.A., Whitenburry, R., Wimpenny, J.W.T. (Eds.), *Microbes in Their Natural Environment*, 34th Symposium of the Society of General Microbiology. Cambridge University Press, Cambridge, pp. 249E298.
- Busse, Hans-Jürgen (1 January 2016). "Review of the taxonomy of the genus *Arthrobacter*, emendation of the genus *Arthrobacter sensu lato*, proposal to reclassify selected species of the genus *Arthrobacter* in the novel genera *Glutamicibacter gen. nov.*, *Paeniglutamicibacter gen. nov.*, *Pseudoglutamicibacter gen. nov.*, *Paenarthrobacter gen. nov.* and *Pseudarthrobacter gen. nov.*, and emended description of *Arthrobacter roseus*". *International Journal of Systematic and Evolutionary Microbiology*. 66 (1): 9–37.
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. *BMC Bioinformatics.* 2009 Dec 15;10:421. doi: 10.1186/1471-2105-10-421. PMID: 20003500; PMCID: PMC2803857.
- Campos M, Karas PS, Perruchon C, Papadopoulou ES, Christou V, Menkissoglou-Spiroudi U, Diez MC, Karpouzas DG. Novel insights into the metabolic pathway of iprodione by soil bacteria. *Environ Sci Pollut Res Int.* 2017 Jan;24(1):152-163. doi: 10.1007/s11356-016-7682-1. Epub 2016 Oct 5. PMID: 27704380.
- Campos, M, Karas, P S, Perruchon, C, Papadopoulou, E S, Christou, V, Menkissoglou-Spiroudi, U, Diez, M C, Karpouzas, D G (2017). Novel insights into the metabolic pathway of iprodione by soil bacteria. *Environ Sci Pollut Res Int* 24, 152- 163, 10.1007/s11356-016- 7682-1
- Campos, M, Perruchon, C, Vasilieiadis, S, Menkissoglu-Spiroudi, U,

- Karpouzas, D G, Diez, M C (2015). Isolation and characterization of bacteria from acidic pristine soil environment able to transform iprodione and 3,5-dichloraniline. Int Biodeterior Biodegrad 104, 201-211, 10.1016/j.ibiod.2015.06.009*
- *Campos, M., C. Perruchon, S. Vasilieiadis, U. Menkissoglu-Spiroudi, D.G. Karpouzas, M.C. Diez, Isolation and characterization of bacteria from acidic pristine soil environment able to transform iprodione and 3,5-dichloraniline, International Biodeterioration & Biodegradation, Volume 104, 2015, Pages 201-211, ISSN 0964-8305, <https://doi.org/10.1016/j.ibiod.2015.06.009>.*
 - *Canonical, n.d.*
 - *Cao L, Shi W, Shu R, Pang J, Liu Y, Zhang X, Lei Y. Isolation and characterization of a bacterium able to degrade high concentrations of iprodione. Can J Microbiol. 2018 Jan;64(1):49-56. doi: 10.1139/cjm-2017-0185. Epub 2017 Dec 8. PMID: 29219613.*
 - *Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Res. 2016 Jan 4;44(D1):D67-72. doi: 10.1093/nar/gkv1276. Epub 2015 Nov 20. PMID: 26590407; PMCID: PMC4702903.*
 - *Conn & Dimmick 1947*
 - *Copping, L.G., Duke, S.O., 2007. Natural products that have been used commercially as crop protection agents. Pest Manag. Sci. 554, 524e554.*
 - *Durbin, Richard; Sean R. Eddy; Anders Krogh; Graeme Mitchison (1998). Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids. Cambridge University Press. ISBN 0-521-62971-3*
 - *Eddy, SR (1998). "Profile hidden Markov models". Bioinformatics. 14 (9): 755–63. doi:10.1093/bioinformatics/14.9.755. PMID 9918945.*
 - *Eschbach M, Möbitz H, Rompf A, Jahn D. Members of the genus Arthrobacter grow anaerobically using nitrate ammonification and fermentative processes: anaerobic adaptation of aerobic bacteria abundant in soil. FEMS Microbiol Lett. 2003 Jun 27;223(2):227-30. doi: 10.1016/S0378-1097(03)00383-5. PMID: 12829291*
 - *FAO, 2004 FAO Specifications and Evaluations for Agricultural Pesticides: Iprodione.*
 - *Fetzner, S., 1998. Bacterial dehalogenation. Appl. Microbiol. Biotechnol. 50, 633e657.*
 - *Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. 2011 Jul;39(Web Server issue):W29-37. doi: 10.1093/nar/gkr367. Epub 2011 May 18. PMID: 21593126; PMCID: PMC3125773.*
 - *Genome Annotation Josep F. Abril, Sergi Castellano, in Encyclopedia of Bioinformatics and Computational Biology, 2019.*
 - *Genome annotation: from sequence to biology Lincoln Stein Nature Reviews Genetics volume 2, pages 493–503 (01 July 2001).*
 - *Grabke, A., Fernandez-Ortuno, D., Amiri, A., Li, X., Peres, N.A., Smith, P., Schnabel, G., ~ 2014. Characterization of iprodione resistance in Botrytis cinerea from strawberry and blackberry. Phytopathology 104, 396e402.*

- Gribble, G., 1992. Naturally occurring organohalogen compounds - a survey. *J. Nat. Prod.* 55, 1353e1395.
- Hani, E. K., & Chan, V. L. (1995). Expression and characterization of *Campylobacter jejuni* benzoylglycine amidohydrolase (Hippuricase) gene in *Escherichia coli*. *Journal of bacteriology*, 177(9), 2396–2402. <https://doi.org/10.1128/jb.177.9.2396-2402.1995>.
- <https://pubchem.ncbi.nlm.nih.gov/compound/Iprodione>
- Katsoula A, Vasileiadis S, Sapountzi M, Karpouzas DG. The response of soil and phyllosphere microbial communities to repeated application of the fungicide iprodione: accelerated biodegradation or toxicity? *FEMS Microbiol Ecol.* 2020 Jun 1;96(6):fiae056. doi: 10.1093/femsec/fiae056. PMID: 32221586.
- Kitagawa, W., Takami, S., Miyauchi, K., Eiji, M., Kamagata, Y., Tiedje, J.M., Fukuda, M., 2002. Novel 2,4-dichlorophenoxyacetic acid degradation genes from oligotrophic *Bradyrhizobium* sp. strain HW13 isolated from a pristine environment. *J. Bacteriol.* 184, 509e518.
- Lionel Guy, Jens Roat Kultima, and Siv G. E. Andersson: *genoPlotR: comparative gene and genome visualization in R.* *Bioinformatics* 2010;26(18):2334-2335
- Martin, C., Vega, D., Bastide, J., Davet, P., 1990. Enhanced degradation of iprodione in soil after repeated treatments for controlling *Sclerotinia minor*. *Plant Soil* 127, 104e142.
- Mihasan M, Brandsch R. pAO1 of *Arthrobacter nicotinovorans* and the spread of catabolic traits by horizontal gene transfer in gram-positive soil bacteria. *J Mol Evol.* 2013 Aug;77(1-2):22-30. doi: 10.1007/s00239-013-9576-x. PMID: 23884627.
- Miñambres, G.G., Conles, M.Y., Lucini, E.I. et al. Application of thymol and iprodione to control garlic white rot (*Sclerotium cepivorum*) and its effect on soil microbial communities. *World J Microbiol Biotechnol* 26, 161 (2010). <https://doi.org/10.1007/s11274-009-0155-7>, 2010.
- Mistry, J., S. Chuguransky, L. Williams, M. Qureshi, G.A. Salazar, E.L.L. Sonnhammer, S.C.E. Tosatto, L. Paladin, S. Raj, L.J. Richardson, R.D. Finn, A. Bateman: *Pfam: The protein families database in 2021.*
- Mitchel, J.A., Cain, R.B., 1996. Rapid onset of the accelerated biodegradation of dicarboxamide fungicides in a UK soil with a long history of agrochemical exclusion. *Pesticide Sci.* 48, 1e11.
- Mount, D. W. (2004). *Bioinformatics: Sequence and Genome Analysis* (2nd ed.). Cold Spring Harbor Press. ISBN 978-0-87969-712-9.
- Nautiyal, C., Srivastava, S., Chauhan, P., 2008. Rhizosphere colonization;: molecular determinants from plant-microbe coexistence perspective. In: *molecular mechanisms of plant and microbe coexistence.* In: *Soil Biology Series.* SpringerVerlag, Berlin, pp. 99e123. Profiler, P.B.T., 2013. Developed by Environmental Science Center for the Office of Pollution Prevention and Toxics. U.S. Environmental Protection Agency. Online at. <http://www.pbtprofiler.net>.

- *NCBI Resource Coordinators. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res.* 2016;44(D1):D7-D19. doi:10.1093/nar/gkv1290.
- *Neil D. Rawlings, Alan J. Barrett, Evolutionary families of metallopeptidases, Methods in Enzymology, Academic Press, Volume 248, 1995, Pages 183-228, ISSN 0076-6879, ISBN 9780121821494, https://doi.org/10.1016/0076-6879(95)48015-3.*
- *Okonechnikov, K., Olga Golosova, Mikhail Fursov, the UGENE team Bioinformatics 2012: Unipro UGENE: a unified bioinformatics toolkit.*
- *Orla-Jensen S. (1919). The Lactic Acid Bacteria. Copenhagen: Host & Sons*
- *Pesticides: a review article.al-Saleh IA Journal of Environmental Pathology, Toxicology and Oncology : Official Organ of the International Society for Environmental Toxicology and Cancer, 01 Jan 1994, 13(3):151-161*
- *(«International Code of Conduct on the Distribution and Use of Pesticides» (PDF). Food and Agriculture Organization of the United Nations. 2002.*
- *Pevsner, Jonathan (2009). Bioinformatics and functional genomics (2nd ed.). Hoboken, N.J: Wiley-Blackwell. ISBN 9780470085851.*
- *RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL <http://www.rstudio.com>.*
- *Sanchez, M., Vasquez, M., Gonzalez, B., 2004. A previously unexposed forest soil microbial community degrades high levels of the pollutant 2, 4, 6-trichlorophenol. Appl. Environ. Microbiol. 70, 7567e7570.*
- *Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014 Jul 15;30(14):2068-9*
- *Shin S, Yun YS, Koo HM, Kim YS, Choi KY, Oh BH. Characterization of a novel Ser-cisSer-Lys catalytic triad in comparison with the classical Ser-His-Asp triad. J Biol Chem. 2003 Jul 4;278(27):24937-43. doi: 10.1074/jbc.M302156200. Epub 2003 Apr 23. PMID: 12711609.*
- *Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M. ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res. 2006 Jan 1;34(Database issue):D32-6. doi: 10.1093/nar/gkj014. PMID: 16381877; PMCID: PMC1347377.*
- *Slade, E., Fullerton, R., Stewart, A., Young, H., 1992. Degradation of the dicarboximide fungicides iprodione, vinclozolin and procymidone in Patumahoe clay loam soil, New Zealand. Pesticide Sci. 35, 95e100.*
- *Smith, T F, Waterman, M S (1981). Identification of common molecular subsequences. J Mol Biol 147, 195-197, https://doi.org/10.1016/0022 - 2836(81)90087-5*
- *Snipen, L., Liland, K.H. microman: an R-package for microbial pan-genomics. BMC Bioinformatics 16, 79 (2015). https://doi.org/10.1186/s12859-015-0517-0.*
- *States DJ, Gish W. Combined use of sequence similarity and codon bias for coding region identification. J Comput Biol. 1994 Spring;1(1):39-50. doi: 10.1089/cmb.1994.1.39. PMID: 8790452.*
- *The UniProt Consortium, UniProt: the universal protein knowledgebase in*

- 2021, *Nucleic Acids Research*, Volume 49, Issue D1, 8 January 2021, Pages D480–D489, <https://doi.org/10.1093/nar/gkaa1100>.
- U.S. Environmental Protection Agency (USEPA), 1998. *Registration Eligibility Decision (RED) iprodione. Prevention, pesticides and toxic substances*. Online at. <http://www.epa.gov/oppsrrd1/REDs/2335.pd>.
 - Valiña AL, Mazumder-Shivakumar D, Bruice TC (December 2004). "Probing the Ser-Ser-Lys catalytic triad mechanism of peptide amidase: computational studies of the ground state, transition state, and intermediate". *Biochemistry*. **43** (50): 15657–72. doi:10.1021/bi049025r.
 - Walker, A., 1987a. *Further observations on the enhanced degradation of iprodione and vinclozolin in soil*. *Pesticide Sci.* **21**, 219e231.
 - Wang, Y; Tang, H; Debarry, JD; Tan, X; Li, J; Wang, X; Lee, TH; Jin, H; Marler, B; Guo, H; Kissinger, JC; Paterson, AH (April 2012). "MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity". *Nucleic Acids Research*. **40** (7): e49. doi:10.1093/nar/gkr1293. PMC 3326336. PMID 22217600.
 - Wick R.R., Schultz M.B., Zobel J. & Holt K.E. (2015). *Bandage: interactive visualisation of de novo genome assemblies*. *Bioinformatics*, **31**(20), 3350-3352.
 - Yang Z, Jiang W, Wang X, Cheng T, Zhang D, Wang H, Qiu J, Cao L, Wang X, Hong Q. *An Amidase Gene, ipaH, Is Responsible for the Initial Step in the Iprodione Degradation Pathway of Paenarthrobacter sp. Strain YJN-5*. *Appl Environ Microbiol.* 2018 Sep 17;84(19):e01150-18. doi: 10.1128/AEM.01150-18. PMID: 30054359; PMCID: PMC6147002.
 - Yang, Z., Wu, G., Jin, W., Cao, L., Yan, X., & Hong, Q. (2017). *Isolation, identification and characterization of an iprodione-degrading strain YJN-G*. *Chin. J. Appl. Environ. Biol.*, **1**, 164-168.
 - Yao Y, Tang H, Su F, Xu P. *Comparative genome analysis reveals the molecular basis of nicotine degradation and survival capacities of Arthrobacter*. *Sci Rep.* 2015 Feb 27;5:8642. doi: 10.1038/srep08642. PMID: 25721465; PMCID: PMC4342571.
 - Zhang M, Ren Y, Jiang W, Wu C, Zhou Y, Wang H, Ke Z, Gao Q, Liu X, Qiu J, Hong Q. *Comparative genomic analysis of iprodione-degrading Paenarthrobacter strains reveals the iprodione catabolic molecular mechanism in Paenarthrobacter sp. strain YJN-5*. *Environ Microbiol.* 2021 Feb;23(2):1079-1095. doi: 10.1111/1462-2920.15308. Epub 2020 Nov 17. PMID: 33169936.

7. Παραρτήματα

Παράρτημα Ι

Αναλυμένα αποτελέσματα για τις στοιχίσεις IpaH στα βασικά χρωμοσώματα του δείγματος T1.6 (*Paenarthrobacter* και *Microbacterium*)

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
AHEJKLAN_02680 Acylamidase		135	135	97%	6e-36	27.66%	482	Query_68035
AHEJKLAN_03257 Glutamyl-HRNA(Gln) amidotransferase subunit A		114	114	96%	3e-28	27.44%	517	Query_68604
AHEJKLAN_01304 Acylamidase		113	113	95%	4e-28	29.55%	478	Query_66686
AHEJKLAN_01855 Acylamidase		75.9	75.9	46%	1e-15	33.04%	467	Query_67227
AHEJKLAN_01141 Glutamyl-HRNA(Gln) amidotransferase subunit A		50.4	50.4	26%	2e-07	30.40%	572	Query_66524

Αποτελέσματα στοίχισης της αλληλουχίας της IpaH στα ανοιχτά πλαίσια ανάγνωσης του βασικού χρωμοσώματος του *Raenarthrobacter* του δείγματος T1.6.

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
IGGBDPDB_02218 Glutamyl-HRNA(Gln) amidotransferase subunit A		107	107	92%	7e-26	28.12%	505	Query_70877
IGGBDPDB_00978 Glutamyl-HRNA(Gln) amidotransferase subunit A_chloroplastic/mitochondrial		49.7	49.7	41%	2e-07	29.06%	571	Query_69662

Αποτελέσματα στοίχισης της αλληλουχίας της IpaH στα ανοιχτά πλαίσια ανάγνωσης του βασικού χρωμοσώματος του *Microbacterium* του δείγματος T1.6.

Παράρτημα II

Αναλυμένα αποτελέσματα για τις στοίχισεις DdaH στα ανοιχτά πλαίσια ανάγνωσης των βασικού χρωμοσώματος του *Raenarthrobacter* του δείγματος T1.6

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
AHEJKLAN_01401 hypothetical protein		157	157	91%	4e-47	38.37%	315	Query_50721
AHEJKLAN_01508 hypothetical protein		40.8	40.8	31%	6e-05	30.34%	492	Query_50828

Αποτελέσματα στοίχισης της αλληλουχίας της DdaH στα ανοιχτά πλαίσια ανάγνωσης του βασικού χρωμοσώματος του *Raenarthrobacter* του δείγματος T1.6.

Παράρτημα III

Αναλυμένα αποτελέσματα για τις στοίχισεις DuaH στα ανοιχτά πλαίσια ανάγνωσης των βασικού χρωμοσώματος του *Microbacterium* του δείγματος T1.6

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
IGGBDPDB_03413 Hippurate hydrolase		343	343	93%	9e-117	51.27%	404	Query_66838
IGGBDPDB_01610 Carboxypeptidase G2		35.8	35.8	21%	0.004	28.41%	383	Query_65062

Αποτελέσματα στοίχισης της αλληλουχίας της DuaH στα ανοιχτά πλαίσια ανάγνωσης του

βασικού χρωμοσώματος του *Microbacterium* του δείγματος T1.6.

Παράρτημα IV

Αποτελέσματα στοιχίσεων BLAST για το δείγμα T1.8

IraH: Δεν βρέθηκαν ομόλογες αλληλουχίες στα ανοιχτά πλαίσια ανάγνωσης των πλασμιδίων 1 και 3. Στο πλασμίδιο 2 βρέθηκε η καλύτερη στοίχιση:

HONKHINH_00051 Acylamidase

Sequence ID: Query_25526 Length: 469 Number of Matches: 1

Range 1: 1 to 469 [Graphics](#)

[▼ Next Match](#) [▲ Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
921 bits(2381)	0.0	Compositional matrix adjust.	456/469(97%)	458/469(97%)	0/469(0%)
Query 1		MSDQLWSKSATELAGLVRKVVSADEVVEAHLQRIEDVNPENAIHAVVLADDARQSARIA			60
Sbjct 1		MSDQLWSKSATELAGLVRKVVSADEVVEAHLQRIEDVNPENAIHAVVLADDARQSARIA			60
Query 61		DEKTRTEPDQLGRLHGVPITLKVNI DLVGSATSDAVPAFKDFYPPMNTPLVDRLLGEGAI			120
Sbjct 61		DEKTRTEPDQLGRLHGVPITLKVNI DLVGSATSDAVPAFKDFYPPMNTPLVDRLLGEGAI			120
Query 121		VVGRTNMSDMGMRMTTDDSTLHGLTRNPWHPGRTAGGSSGGEGAALASGMSALGVGNLGV			180
Sbjct 121		VVGRTNMS+MGMR TTDSTLHGLTRNPWHPGRTAGGSSGGEGAALASGMSALGVGNL G			180
Query 181		SLRNPAHCCGISTLKPTPGRIPWADSI LPPDGPLSFQMLLVHGPMARQVADVRLGMEIMS			240
Sbjct 181		SLRNPAHCCGISTLKPTPGRIPWADSI LPPDGPLSFQMM VHGPMARQVADVRLGMEIMS			240
Query 241		GAHPRDPYSVDVPLYRADSNQPFKVAVMAAPPGAPTDPVSSVVRKAGEAVAAAAGYEVDE			300
Sbjct 241		GAHPRDPYSVDVPLYRADSNQPFKVAVMAAPPGAPTDPVSSVVRKAGEAVAAAAGYEVDE			300
Query 301		IDAPEYLETRQIWLDFLMTEVNVLQDH ITEAMGPAGRQFLRGFLDHAGPLDLPGYIDL FV			360
Sbjct 301		IDAPEYLETRQIWLDFLMTEVNVLQD LITEAMGPAGRQFLSDLFDHAGPLDLPGYIDL FV			360
Query 361		RRRRMSRVWNDFQEYYPVIIAPT WLDVAFEHWDWLKSVADTVDNCGPVVPANV LGLPAAV			420
Sbjct 361		RRRRMSRVWNDFQEYYPVIIAPT WLDVAFEHWDWLKSVADTVDM LGPIQANV LGLPAAV			420
Query 421		TFGGLAKGMPVGVQCIAAAFHDDQAMQVAEVI EAAVGATRPVSPFINAG			469
Sbjct 421		TFGGLAKGMPVGVQCIAAAFHDDQAMQVAEVI EAAVGATRPVSPFINAG			469

Στοίχιση και στατιστικά του καλύτερου αποτελέσματος της αναζήτησης BLAST βάσει bit score της πρωτεϊνικής αλληλουχίας του IraH έναντι της μεταφρασμένης αλληλουχίας

α

ν

Στα χρωμοσώματα των *Paenarthrobacter* και *Microbacterium* βρέθηκαν στοιχίσεις με μικρότερη ομολογία:

λ

τ

ο

ύ

π

λ

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
HCJADMNC_00921 Acylamidase		135	135	97%	6e-36	27.66%	482	Query_37292
HCJADMNC_01497 Glutamyl-tRNA(Gln).amidotransferase subunit_A		114	114	96%	3e-28	27.44%	517	Query_37860
HCJADMNC_04190 Acylamidase		113	113	95%	4e-28	29.55%	478	Query_40516
HCJADMNC_00096 Acylamidase		75.9	75.9	46%	1e-15	33.04%	467	Query_36484
HCJADMNC_04027 Glutamyl-tRNA(Gln).amidotransferase subunit_A		50.4	50.4	26%	2e-07	30.40%	572	Query_40354

Αποτελέσματα στοίχισης της αλληλουχίας της IpaH στα ανοιχτά πλαίσια ανάγνωσης του βασικού χρωμοσώματος του *Paenarthrobacter* του δείγματος T1.8.

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
KIOKEBGJ_00587 Glutamyl-tRNA(Gln).amidotransferase subunit_A		107	107	92%	7e-26	28.12%	505	Query_45536
KIOKEBGJ_01872 Glutamyl-tRNA(Gln).amidotransferase subunit_A_chloroplastic/mitochondrial		49.7	49.7	41%	2e-07	29.06%	571	Query_46801

Αποτελέσματα στοίχισης της αλληλουχίας της IpaH στα ανοιχτά πλαίσια ανάγνωσης του βασικού χρωμοσώματος του *Microbacterium* του δείγματος T1.8.

DdaH: Δεν βρέθηκαν ομόλογες αλληλουχίες στα ανοιχτά πλαίσια ανάγνωσης των πλασμιδίων 2 και 3, όπως επίσης και στο βασικό χρωμόσωμα του *Microbacterium*. Στο πλασμίδιο 1 βρέθηκε η καλύτερη στοίχιση:

EHCEJHHC_00053 Peptidoglycan deacetylase

Sequence ID: **Query_62817** Length: **279** Number of Matches: **1**

Range 1: 1 to 279 [Graphics](#)

[▼ Next Match](#) [▲ Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
571 bits(1472)	0.0	Compositional matrix adjust.	276/279(99%)	278/279(99%)	0/279(0%)
Query 1		MATICLTFDFDAISAWAGTFGLTTPQYVSRGEYGANVATPRILDLLEREGVKSTWYIPGM			60
Sbjct 1		MATICLTFDFDAISAWAGTFGLTTPQYVSRGEYGANVATPRILDLLEREGVKSTWYIPGM			60
Query 61		DVDTYPEVCRRIRDAGHEIGHHGYAHEGPTALSEQAERDVLERGLDALDRVLGVPYGYR			120
Sbjct 61		DVDTYPEVCRRIRDAGHEIGHHGYAHEGPTALSEQAERDVLERGLDALDRVLGVPYGYR			120
Query 121		SPAFDLSPNSTRLLHEYGFAYDSSMMANDFELYKCRGTGDI IHTDKAVEFGPTLDLVEVPV			180
Sbjct 121		SPAFDLSPNSTRLLHEYGFAYDSSMMANDFELYKCRGTGDI IHTDKAVEFGPTLDLVEVPV			180
Query 181		TWTLDDFPFTEFVLAPGMLVPGSTDVEAMAQRWMADLDFMIEEYPTGVFTQTFHPQSIGR			240
Sbjct 181		TWTLADFPFTEFVLAPGMLVPGSTDVEAMAQRWMADLDFMIEEHPPTGVFTQTFHPQSIGR			240
Query 241		GSRLRILERI IIRRAKDHGAEFSTVNDAVVKWGARGQTLR	279		
Sbjct 241		GSRLRILERI IIRRAK+HGAEFSTVNDAVVKWGARGQTLR	279		

Στοίχιση και στατιστικά του καλύτερου αποτελέσματος της αναζήτησης BLAST βάσει bit score της πρωτεϊνικής αλληλουχίας του DdaH έναντι της μεταφρασμένης αλληλουχίας

α
ν
ο
ι
χ

Στο χρωμόσωμα του *Paenarthrobacter* βρέθηκε στοιχίση με μικρότερη ομολογία:

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
HCJADMNC_04287 hypothetical protein		157	157	91%	4e-47	38.37%	315	Query_40919
HCJADMNC_04394 hypothetical protein		40.8	40.8	31%	6e-05	30.34%	492	Query_41026

Αποτελέσματα στοίχισης της αλληλουχίας της DdaH στα ανοιχτά πλαίσια ανάγνωσης του βασικού χρωμοσώματος του *Paenarthrobacter* του δείγματος T1.8.

DuaH: Δεν βρέθηκαν ομόλογες αλληλουχίες στα ανοιχτά πλαίσια ανάγνωσης των πλασμιδίων 1, 2 και 3, όμως εμφανίστηκαν στο βασικό χρωμόσωμα του *Microbacterium* και του *Paenarthrobacter*. Στο βασικό χρωμόσωμα του *Paenarthrobacter* βρέθηκε η καλύτερη στοίχιση:

HCJADMNC_00924 Hippurate hydrolase

Sequence ID: Query_70553 Length: 409 Number of Matches: 1

Range 1: 1 to 409 [Graphics](#)

[Next Match](#) [Previous Match](#)

Score	Expect	Method	Identities	Positives	Gaps
728 bits(1878)	0.0	Compositional matrix adjust.	361/409(88%)	381/409(93%)	1/409(0%)
Query 1		MTQTATVQLDTEQTTALHELYRRLHAHPELSMQEHATAALIEEQDLRLGADHFRCGRGTG			60
Sbjct 1		MT T T L+LD++QTT LH LYR LHAHPELSMQE ATA LIEEQLD LG +HFRCG TG			60
Query 61		VVGILRNGEGPVVAFRADTDGLPIEEATGLDYASTDTGML-DGSEVPVMHGCGHDTHVAS			119
Sbjct 61		VVGILRNG GPVVAFRAD+DGLPIEEATGL+YASTDTG L DG+ VPVMHGCGHDTHVAS			120
Query 120		LLGAAEILAGNQDAWSGTLVLI FQPGEETAAGALAMLEDGLWDRAPREVVFGQHVMPRP			179
Sbjct 121		LLGAAE+LAGN +AWSGTLVLI FQPGEETAAGALAML+DGLWD+APRPEVVFGQHVMPRP			180
Query 180		VGTVSISSGPVAAMADSLRVTVHGQQSHGSQPQDSIDPIVMAAHMVTRLQTVSRELDPR			239
Sbjct 181		VGTVDISSGPVAAMADSLRVTVHGQQSHGSQPQDSIDPIVMAAHMVTRLQGVISRELDPR			240
Query 240		KSAVVTVGTTFHAGLKENIIPASAEFTLNIRTFDEEVRGQVLTAVRRIIEAEAAASGAPGP			299
Sbjct 241		KSAVVTVGTTFHAGLKENIIPASAEFTLNIRTFDE+VR QVL AVRRII AEAAASGAP P			300
Query 300		KIEELYRFPQCYNPDPAVPAVLEALRDALGAESVEVTPPMMGSEDFGHLGTAIGVPSVFW			359
Sbjct 301		IEE+YRFPQC+NDPDAVP+V+EALR ALG E+VE+TPPMMGSEDFGHLGTAIGVPSVFW			360
Query 360		MFGGTPGAVFDGPGVPVNHSPPFAPELEGLAGGISA AVAVLLSRLGN		408	
Sbjct 361		MFGGTP A +GPGVPVNHSPPFAPELEGLAGG+SA AVAVLLSRLGN		409	

Στοιχίση και στατιστικά του καλύτερου αποτελέσματος της αναζήτησης BLAST βάσει bit score της πρωτεϊνικής αλληλουχίας του *DuaH* έναντι της μεταφρασμένης αλληλουχίας ανοιχτού πλαισίου ανάγνωσης του βακτηριακού χρωμοσώματος του *Paenarthrobacter* του δείγματος T1.8.

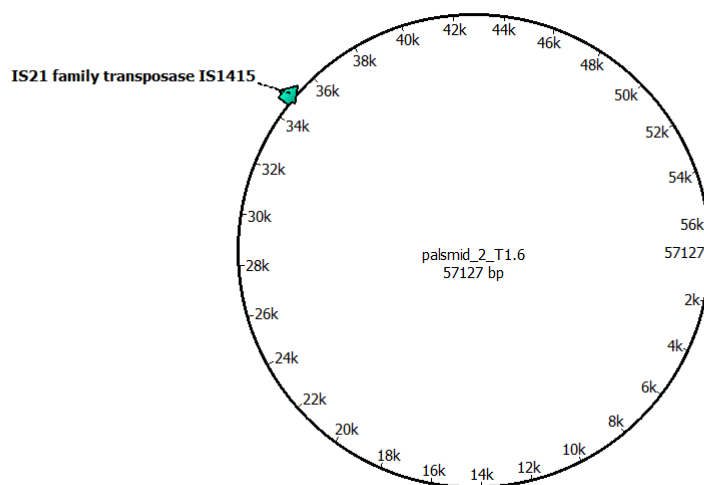
Στο χρωμόσωμα του *Microbacterium* βρέθηκε στοιχίση με μικρότερη ομολογία:

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
KIOKEBGJ_02998 Hippurate hydrolase		343	343	93%	9e-117	51.27%	404	Query_27260
KIOKEBGJ_01240 Carboxypeptidase G2		35.8	35.8	21%	0.004	28.41%	383	Query_25519

Αποτελέσματα στοιχίσης της αλληλουχίας της DuaH στα ανοιχτά πλαίσια ανάγνωσης του βασικού χρωμοσώματος του *Microbacterium* του δείγματος T1.8.

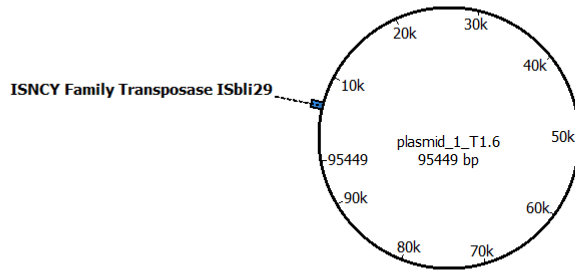
Παράρτημα V

Κατανομή γονιδίων τρανσποζάσης στο πλασμίδιο 2 του δείγματος T1.6



Παράρτημα VI

Κατανομή γονιδίων τρανσποζάσης στο πλασμίδιο 1 του δείγματος T1.6



Παράρτημα VII

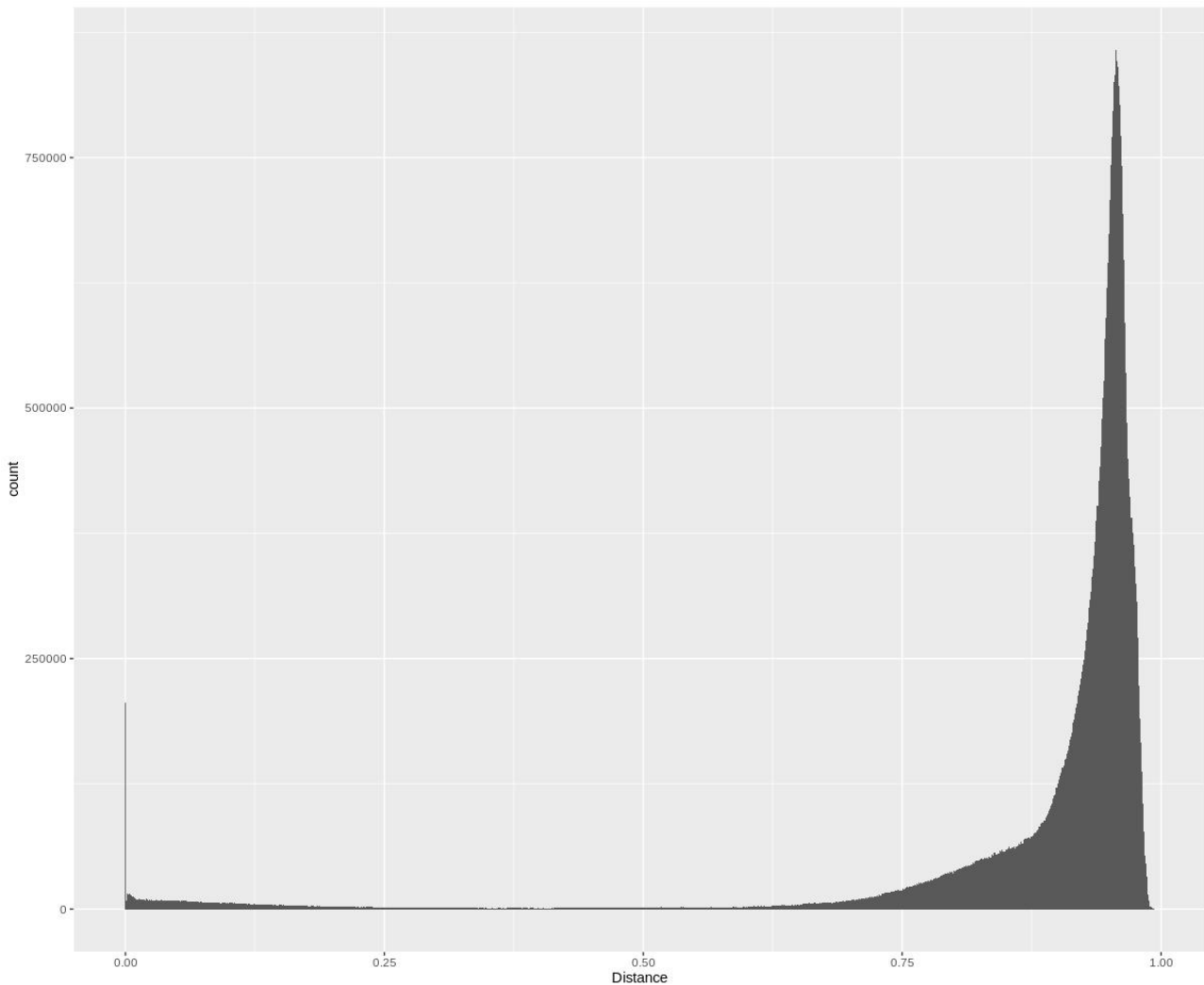
Τμήματα DNA που αναλύθηκαν, τα οποία περιέχουν τα υπό μελέτη ανοιχτά πλαίσια ανάγνωσης και τμήματα DNA που χρησιμοποιήθηκαν για τον έλεγχο synteny.

	IpaH	DdaH	DuaH
T1.6	<i>Plasmid 2</i>	<i>Plasmid 1</i>	<i>chromosome</i>
<i>Start.</i>	23.044	7.899	2.771.481
<i>End.</i>	32.937	17.518	2.789.226
Marco	<i>chromosome</i>	<i>chromosome</i>	<i>chromosome</i>
<i>Start.</i>	809.769	4.373.500	809.769
<i>End.</i>	825.652	4.385.057	825.652
YJN-D	<i>Plasmid 3</i>	<i>Plasmid 2</i>	<i>chromosome</i>
<i>Start.</i>	2.515	6.242	1.695.000
<i>End.</i>	14.000	16.038	1.710.000
YJN-5	<i>Plasmid 1</i>	<i>Plasmid 3</i>	<i>Plasmid 2</i>
<i>Start.</i>	268.150	44.200	26.000
<i>End.</i>	278.369	55.000	41.000

Πίνακας με τα τμήματα DNA που αναλύθηκαν, τα οποία περιέχουν τα υπό μελέτη ανοιχτά πλαίσια ανάγνωσης και με τα τμήματα που χρησιμοποιήθηκαν για τον έλεγχο synteny για τα δείγματα συγκριτικά με τα βακτήρια *Paenarthrobacter YJN-D* και *Paenarthrobacter YJN-5*.

Παράρτημα VIII

Συμπληρωματικό υλικό της ανάλυσης του πανγονιδιώματος



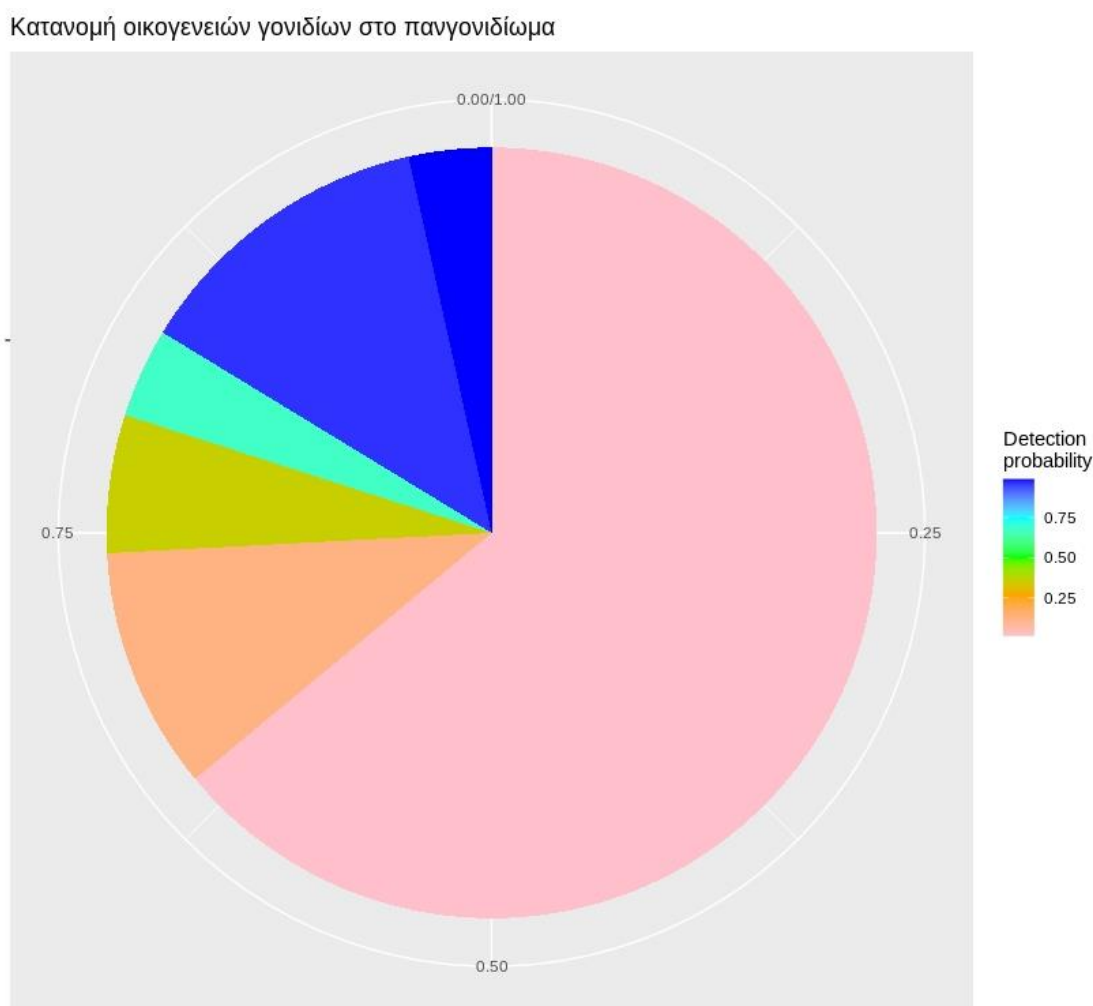
Εικόνα VIIIa: ιστόγραμμα της “απόστασης” των πρωτεϊνών μετά τον έλεγχο ομολογίας, για όλες τις πρωτεΐνες των 35 υπό μελέτη γονιδιωμάτων. Στον κάθετο άξονα αναγράφεται ο αριθμός των πρωτεϊνών, ενώ στον οριζόντιο άξονα αναγράφεται η τιμή της “απόστασης”.

Στην Εικόνα VIIIa του παραρτήματος παρατηρείται ότι μια στήλη εμφανίζεται στο 0, αυτή η στήλη αντιπροσωπεύει τις στοιχίσεις των πρωτεϊνών με τον εαυτό τους και πιθανώς στοιχίσεις μεταξύ εξαιρετικά συντηρημένων πρωτεϊνών κατά μήκος των γονιδιωμάτων. Η απόσταση 0.0 σημαίνει πως οι δυο αλληλουχίες που στοιχήθηκαν είναι όμοιες. Επίσης παρατηρείται μεγάλη συσσώρευση κοντά στην τιμή 1.0 η οποία αντιπροσωπεύει ανομοιότητα μεταξύ πρωτεϊνών. Η περιοχή μεταξύ των τιμών 0.01 και 0.25 αναπαριστά το κεντρικό γονιδίωμα· σύνολα πρωτεϊνών τα οποία είναι συντηρημένα και απαντούν γενικώς αναλλοίωτα στο σύνολο των ειδών του γένους. Η περιοχή των τιμών μεγαλύτερες του 0.75 αναπαριστούν κυρίως ειδοειδικά και στελεχοειδικά γονίδια.

Επιπλέον, έγινε μια πρόβλεψη για το μέγεθος του πανγονιδιώματος, σύμφωνα με την οποία

το κλειστό πανγονιδίωμα θα αποτελείται από 17.732 συστάδες πρωτεϊνών. Παρόλαυτα, η αξιοπιστία της πρόβλεψης δεν είναι μεγάλη. Η πρόβλεψη έγινε με χρήση μια συνάρτησης εντός του πακέτου microman, της heaps.est(). Η συνάρτηση αυτή προσεγγίζει το μέγεθος του πανγονιδιώματος όταν και εάν αυτό ολοκληρωθεί, δηλαδή όταν δεν θα επηρεάζουν νεοπροστιθέμενα γονιδιώματα πλέον το πανγονιδίωμα.

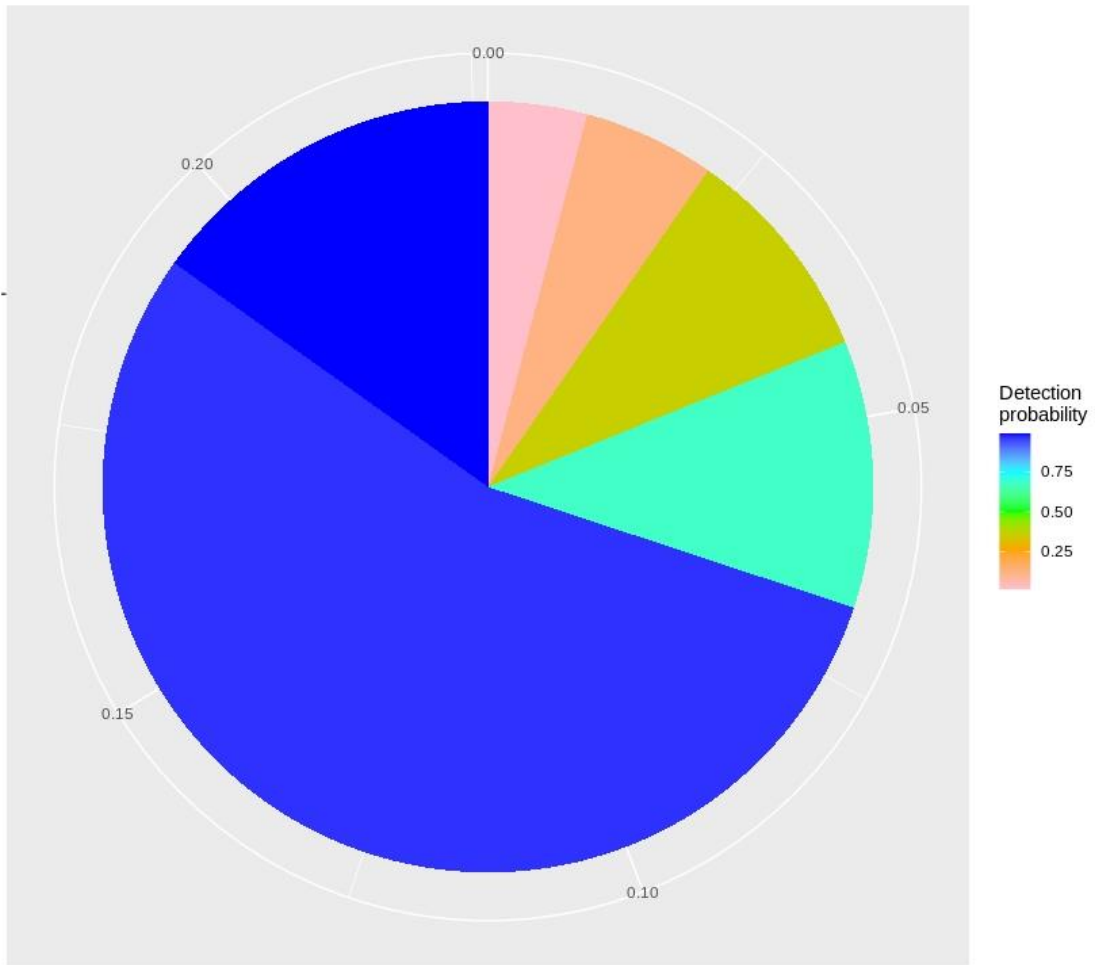
Εναλλακτικά, έγινε προσαρμογή των δεδομένων σε διωνυμικά μοντέλα μείγματος. Το μοντέλο το οποίο βέλτιστα περιγράφει τα δεδομένα αποτελείται από 6 μέρη, το οποίο σημαίνει ότι είναι δυνατόν να χωριστούν οι συστάδες γονιδίων σε 6 κατηγορίες(οικογένειες) ανάλογα με την πιθανότητα να βρίσκονται στα γονιδιώματα.



Εικόνα VIIIb: Σχηματική χρωματική απεικόνιση των οικογενειών συστάδων γονιδίων ανάλογα με την πιθανότητα τους να εμφανιστούν στο πανγονιδίωμα. Τα χρώματα αναπαριστούν την πιθανότητα να εμφανιστεί η οικογένεια γονιδίων στο πανγονιδίωμα.

Οι έξι οικογένειες αναπαρίστανται με χρώματα: με σκούρο μπλε αναπαριστάται η οικογένεια που είναι πιθανό ότι θα βρεθεί σε όλα τα γονιδιώματα, με ροζ αναπαριστάται η οικογένεια που είναι πιθανό ότι θα βρεθεί σε ένα γονιδίωμα. Τα ενδιάμεσα χρώματα αποτελούν φάσμα μεταξύ των δύο αυτών καταστάσεων.

Κατανομή οικογενειών γονιδίων σε τυχαίο γονιδίωμα

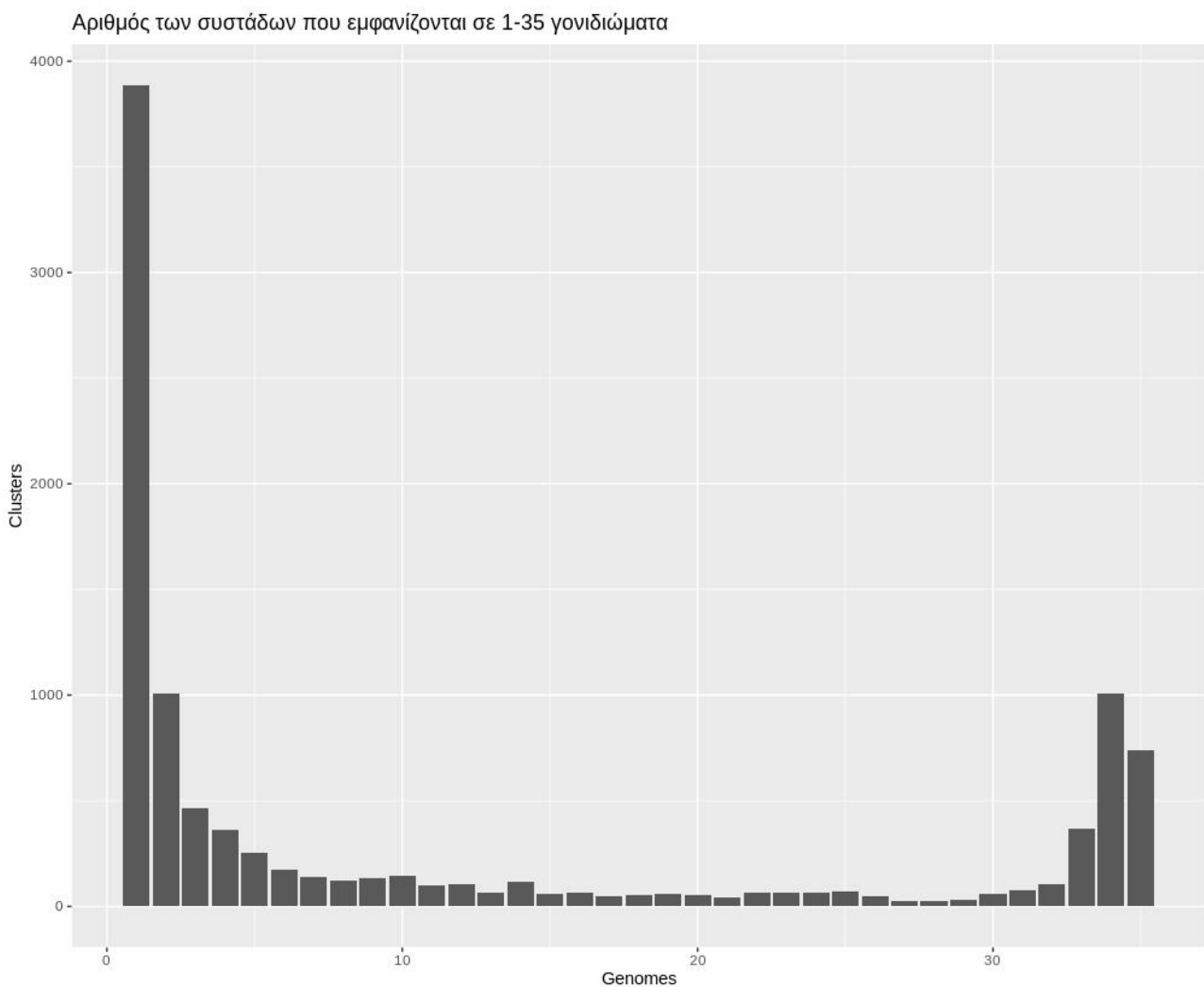


Εικόνα VIIIc: Σχηματική χρωματική απεικόνιση των οικογενειών συστάδων γονιδίων ανάλογα με την πιθανότητα τους να εμφανιστούν σε ένα τυχαίο γονιδίωμα. Τα χρώματα αναπαριστούν την πιθανότητα να εμφανιστεί η οικογένεια γονιδίων στο γονιδίωμα.

Οι έξι οικογένειες αναπαρίστανται με χρώματα: με σκούρο μπλε αναπαριστάται η οικογένεια γονιδίων που είναι πιθανό ότι θα βρεθεί σε όλα τα γονιδιώματα, με ροζ αναπαριστάται η οικογένεια που είναι πιθανό ότι θα βρεθεί σε ένα γονιδίωμα. Τα ενδιάμεσα χρώματα αποτελούν φάσμα μεταξύ των δύο αυτών καταστάσεων.

Μεταξύ των δυο γραφημάτων παρατηρείται σημαντική διαφορά στην κατανομή των “ροζ” γονιδίων. Η οικογένεια γονιδίων αυτή αποτελεί τα στελεχοειδικά και ειδοειδικά γονίδια. Σε ένα τυχαίο γονιδίωμα είναι λιγότερο πιθανό να υπάρχουν πολλά ειδικά γονίδια συγκριτικά με το σύνολο των 35 γονιδιωμάτων του πανγονιδιώματος. Αντιστρόφως, η οικογένεια των “μπλε” γονιδίων είναι πολύ πιο επιφανής σε ένα τυχαίο γονιδίωμα απ’ότι στο σύνολο των 35 γονιδιωμάτων.

Για τις αλληλουχίες που αντιστοιχούν στις τιμές 0.0 έως 0.8 του γραφήματος της Εικόνας VIIa έγινε ομαδοποίηση σε συστάδες. Οι συστάδες αυτές τοποθετήθηκαν σε γράφημα έναντι της ύπαρξής τους στα γονιδιώματα, όπως φαίνεται στην Εικόνα VIIId.



Εικόνα VIIId: Ραβδόγραμμα του αριθμού των συστάδων που εμφανίζονται σε 1 έως 35 γονιδιώματα. Στον κάθετο άξονα αναγράφεται το πλήθος των συστάδων, ενώ στον οριζόντιο άξονα αναγράφεται ο αριθμός των γονιδιωμάτων στα οποία οι συστάδες εμφανίζονται.