

Γονιδιωματική
Συγκριτική γονιδιωματική[4]

Τμήμα Γεωπονίας, Ιχθυολογίας
και Υδάτινου Περιβάλλοντος

Μεζίτη Αλεξάνδρα

Μέγεθος και οργάνωση γονιδιωμάτων

Μελετάμε τα γονιδιώματα για να καταλάβουμε πως λειτουργεί ο οργανισμός μας αλλά και ποιες είναι οι σχέσεις μας με τους άλλους οργανισμούς.

Που επικεντρωνόμαστε;;;

- Αλληλουχίες γονιδιωμάτων
- Πρότυπα έκφρασης RNA και πρωτεϊνών
- Χωρική οργάνωση μακρομορίων, των συμπλόκων τους, των οργανιδίων, ολόκληρων κυττάρων και των ιστών.
- Λειτουργικά δίκτυα και εσωτερική δομή και λογική των συστημάτων προσαρμογής.

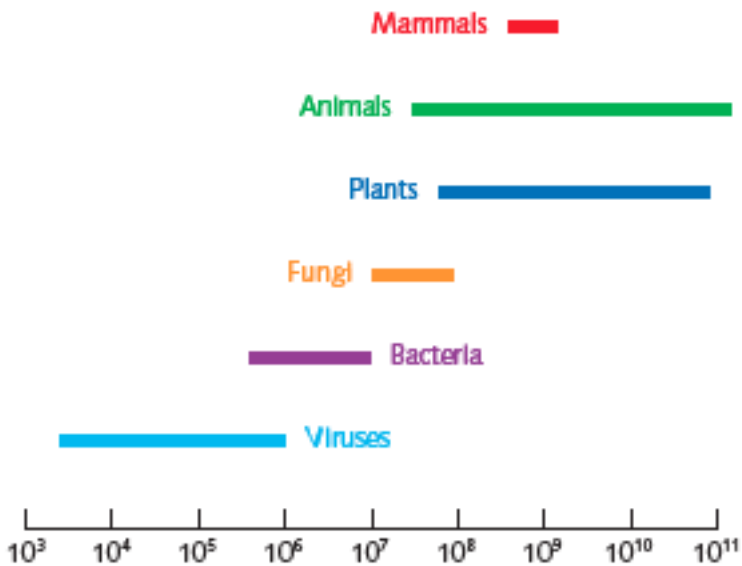
Η μελέτη της 'Ιστορίας' είναι επίσης σημαντική

- S.E. Luria «Για να βρούμε τα κοινά χαρακτηριστικά της ζωής, δεν χρειάζεται να μελετάμε όλα τα χαρακτηριστικά των οργανισμών αλλά αρκεί να προσπαθήσουμε να συγκρίνουμε κοινά χαρακτηριστικά σε οργανισμούς πολυ διαφορετικούς».
- **Ποιες ομοιότητες ανάμεσα στην *E. coli* και των άνθρωπο αποδεικνύουν την κοινή τους προέλευση;**
- Σύγκριση παρόμοιων οργανισμών και αναγνώριση των διαφορών.
- **Που οφείλονται οι φαινοτυπικές κ λειτουργικές διαφορές ανάμεσα στον άνθρωπο και τον χιμπαντζή, που παρουσιάζουν >96% ομοιότητα στα γονιδιώματά τους;**

Μέγεθος γονιδιωμάτων

Η βασική ένσταση στη θεωρία του Δαρβίνου ήταν ότι δεν αναγνώριζε κάποια ξεχωριστή θέση για τους ανθρώπους σε σχέση με τα άλλα ζώα.

Η γονιδιωματική το κάνει αυτό για άλλη μία φορά.



Introduction to Genomics, Lesk A.M.

Οι περισσότεροι αναμένουν τη διαφορετικότητα του ανθρώπου να αντικατοπτρίζεται στο γονιδίωμα.

Αυτό συμβαίνει, αλλά όχι με κάποιο προφανή τρόπο.

****Το ανθρώπινο γονιδίωμα δεν είναι ξεχωριστό****

Συσχέτιση μεγέθους γονιδιώματος με την περιπλοκότητα του οργανισμού

!!!δεν ισχύει πάντα!!!

Μέγεθος γονιδιωμάτων

Organism	Number of base pairs	Number of genes	Comment
→ ϕX-174	5 386	10	Virus infecting <i>E. coli</i>
Influenza A	13 590	10	Strain A/goose/guandong/1/96(H5N1)
→ Human mitochondrion	16 569	37	Subcellular organelle
Epstein–Barr virus (EBV)	172 282	80	Cause of mononucleosis
→ <i>Nanoarchaeum equitans</i>	490 885	552	Archaeon, smallest known genome of a cellular organism
<i>Mycoplasma pneumoniae</i>	816 394	680	Cause of cyclic pneumonia epidemics
→ <i>Rickettsia prowazekii</i>	1 111 523	834	Bacterium, cause of epidemic typhus
<i>Thermoplasma acidophilum</i>	1 564 905	1 509	Archaeal prokaryote, lacks cell wall
<i>Helicobacter pylori</i>	1 667 867	1 589	Chief cause of stomach ulcers
→ <i>Synechocystis</i>	3 573 470	4 003	Cyanobacterium, 'blue-green alga'
<i>Vibrio cholera</i>	4 033 460	3 890	Cause of cholera
<i>Bacillus subtilis</i>	4 214 814	4 779	Popular in molecular biology
→ <i>Escherichia coli</i>	4 639 221	4 485	Molecular biologists' all-time favourite
→ <i>Saccharomyces cerevisiae</i>	12 495 682	5 770	Yeast, first eukaryotic genome sequenced
<i>Caenorhabditis elegans</i>	100 258 171	19 099	'The worm'
→ Human	3.3×10^9	23 000	
Wheat	16×10^9	30 000	
→ Salamander	10^{11}	?	
<i>Psilotum nudum</i>	2.5×10^{11}	?	Whisk fern, a simple plant
<i>Amoeba dubia</i>	6.7×10^{11}	?	Protozoan

Μέγεθος γονιδιωμάτων

Junk DNA;;;

- Η έλλειψη γνώσης της λειτουργίας δε σημαίνει και ότι δεν υπάρχει
- Νέα είδη RNA που εμπλέκονται κυρίως σε διαδικασίες ελέγχου.
- Χώρος ανάμεσα στα γονίδια → ρυθμοί ανασυνδυασμού → ρυθμός εξέλιξης → χαρακτηριστικό του είδους → πιέσεις επιλογής

Αριθμός γονιδίων;;;

- Συσχέτιση ανάμεσα στην πολυπλοκότητα του οργανισμού και τον αριθμό των γονιδίων
- Το εναλλακτικό κόψιμο (alternative splicing) δεν επιτέπει τον εύκολο καθορισμό του αριθμού των γονιδίων → πολλές πρωτεΐνες από ένα γονίδιο (π.χ. ανθρώπινο ανοσοποιητικό, ~100 εξώνια).

Species	Genome size (Mb)	Coding (%)	Approximate number of genes	Estimated gene density (kb/gene)
<i>E. coli</i>	4.64	88	4 485	1.03
Yeast	12.5	70	6 000	2.1
Puffer fish	365	15	23 000	10
<i>A. thaliana</i>	115	29	23 000	6
Human	3289	1.3	23 000	143

Οργάνωση Προκαρυωτικού γονιδιώματος

Κυκλικό δίκλωνο μόριο DNA μήκους 0.6-10 Mbp.

Συμπυκνώμενο με πρωτεΐνες που μοιάζουν με ιστόνες στο νουκλεοειδές.

Υπαρξη πλασμιδίων ~ 1Kbp

Αλλά...

→ Υπάρχουν και γραμμικά DNA (*Borrelia burgorferi*) αλλά και πλασμίδια.

→ Περισσότερα από ένα χρωμοσώματα (*Vibrio cholerae*).

→ Στοιχεία εισαγωγής (insertion elements) αντίστοιχα με τα μεταθετά στοιχεία στους Ευκαρυώτες.

Το μοντέλο της *Escherichia coli*

→ 4.6 Mbp, ~ 4.500 γονίδια

→ Γονίδια και στις δύο αλυσίδες

→ 87,8% → πρωτεΐνες

→ 0,8% δομικά RNA

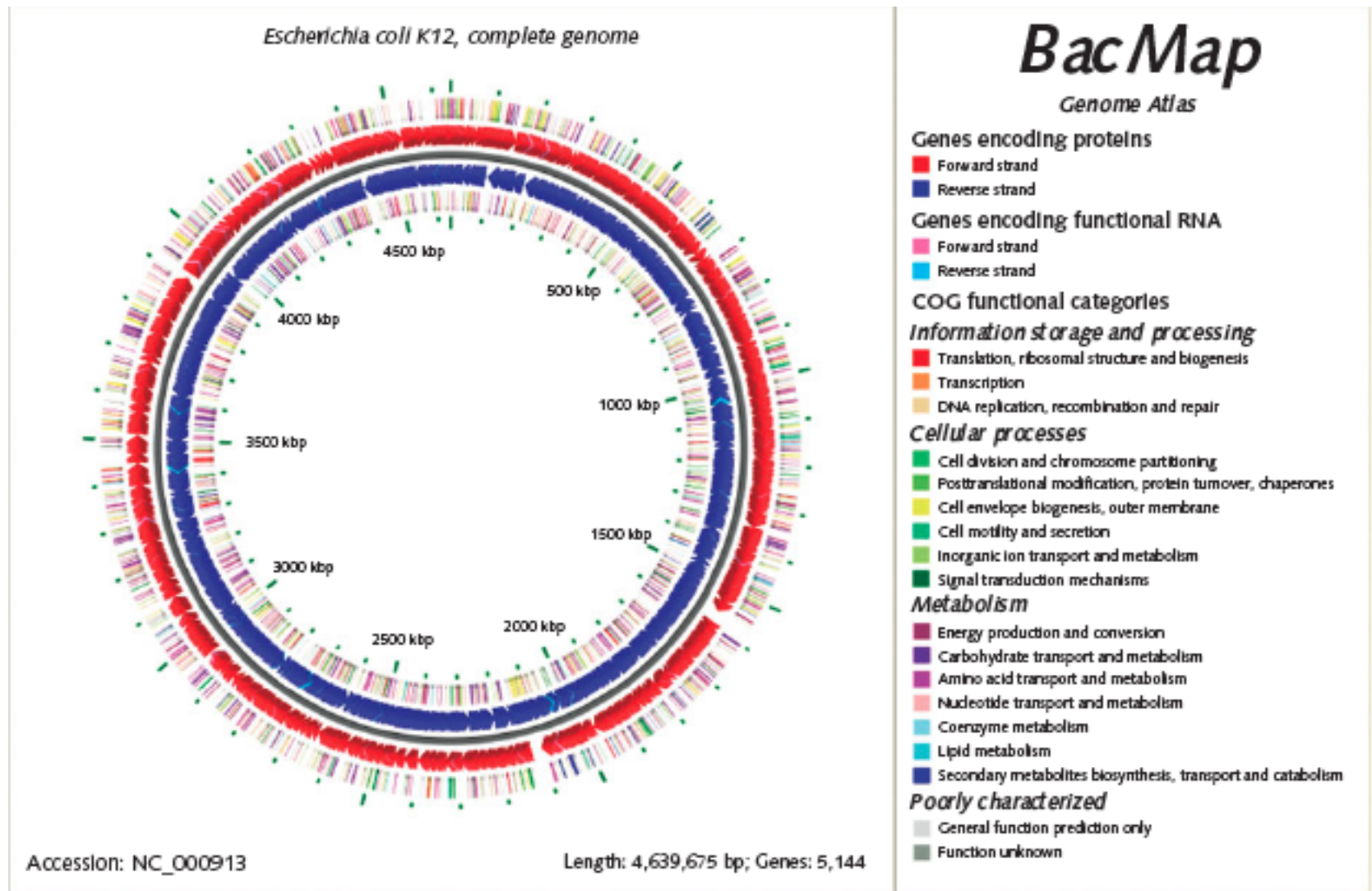
Μεταφορά DNA

→ Μετασηματισμός (μεταφορά 'γυμνου' DNA)

→ Σύντηξη → μεταφορά μέρους ή όλου του DNA από ένα κύτταρο σε ένα άλλο

→ Μεταγωγή (transjunction) → μεταφορά DNA μέσω ιών

Οργάνωση Προκαρυωτικού γονιδιώματος



Click tick marks to expand the view.

Valid XHTML 1.0; Valid CSS.

Displayed PNG file size: 186 kb.

Centered on base 1; Zoom = 1.

Οργάνωση Προκαρυωτικού γονιδιώματος



BacMap: An Interactive Atlas for Exploring Bacterial Genomes

[Home](#) | [About](#) | [Help](#)

[Back](#)

BacMap BLAST Search

Sequence in FASTA format:

Species: *Escherichia coli*, Accession: NC_008253

Program:

Database:

BLAST Options

Mask: Low Complexity Lowercase

Expect:

Matrix:

Return: hits maximum

Δυνατότητα για BLAST με τον οργανισμό που θέλουμε

Στατιστικά δεδομένα

→ Μήκος πρωτεϊνών

→ Σύνθεση αμινοξέων

→ GC content σε διαφορετικές περιοχές

Μεταφόρτωση δεδομένων από τον μικροοργανισμό που μας ενδιαφέρει

→ Πρωτεΐνες

→ Γονίδια

→ Λειτουργικά RNAs

→ Genbank

Definition: *Escherichia coli* 536, complete genome.

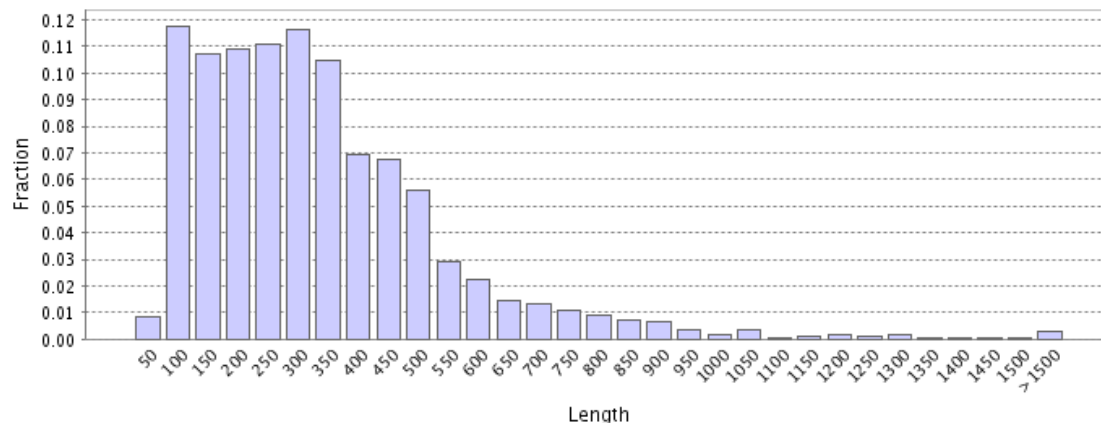
Accession: NC_008253

Length: 4,938,920 bp

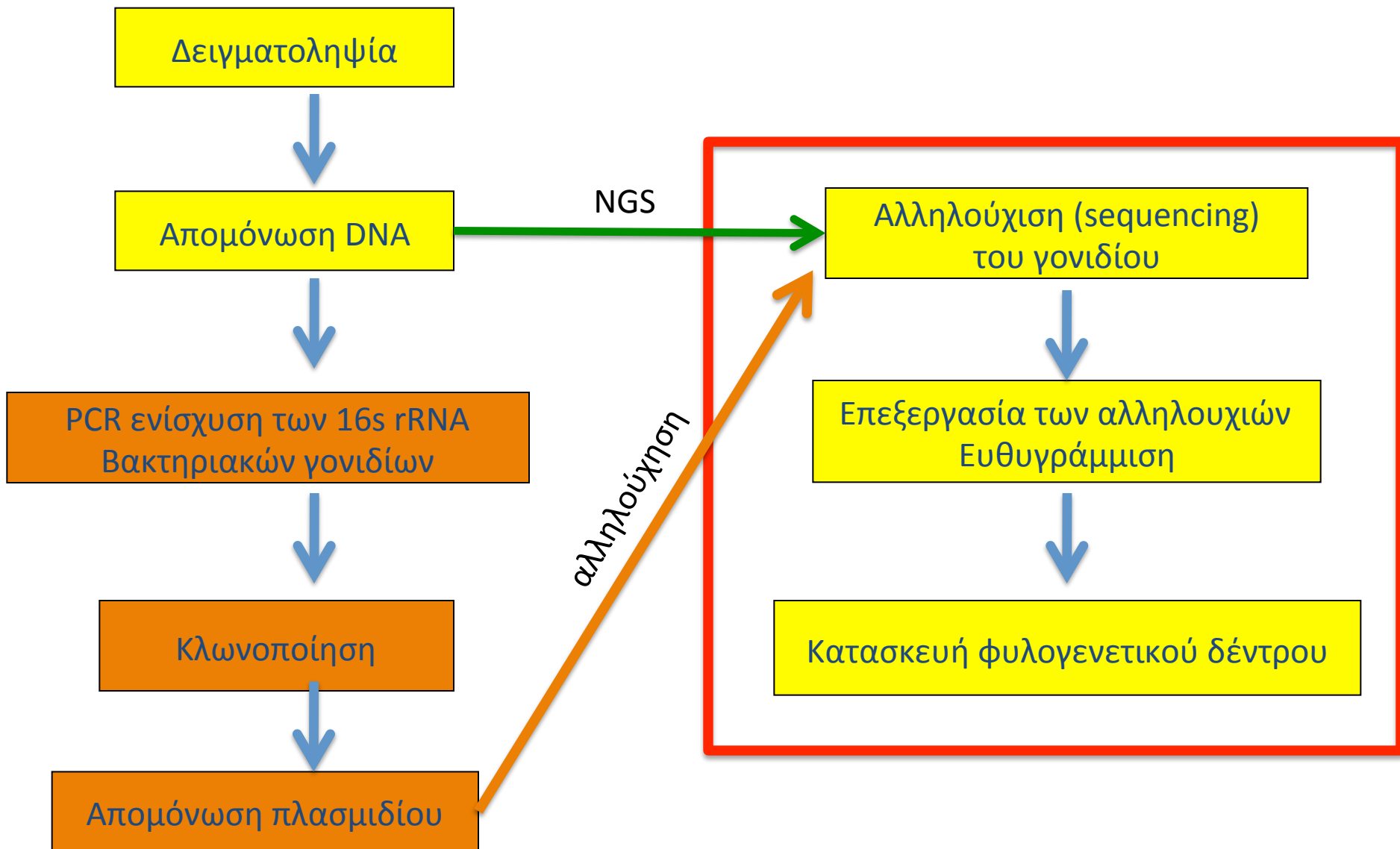
[View map for NC_008253](#)

[View ALL maps](#)

Protein lengths for NC_008253



Μεθοδολογία



Επεξεργασία αλληλουχιών από illumina

- Ανάλυση 16S rRNA γονιδίων (Αλληλούχηση με συγκεκριμένους εκκινητές)
- Διαφορετικά λογισμικά
 - mothur, <https://www.mothur.org/>
 - Qiime, <http://qiime.org/>
- Στόχος
 - Μείωση των λαθών της αλληλούχησης και της PCR.
 - Ένωση των reads (paired) σε contigs.
 - Ευθυγράμμιση (alignment) των αλληλουχιών χρησιμοποιώντας την επιθυμητή βάση δεδομένων (Προκαρυώτες, Ευκαρυώτες).
 - Απομάκρυνση χιμαιρικών contigs.
 - Ταξινόμηση χρησιμοποιώντας την επιθυμητή βάση δεδομένων (Προκαρυώτες, Ευκαρυώτες).
 - Ομαδοποίηση αλληλουχιών σε λειτουργικές ταξινομικές ομάδες (Operational Taxonomic Units, OTUs), σε επιθυμητό κατώφλι ομοιότητας (π.χ. 97% για επίπεδο είδους στους Προκαρυώτες).
 - Χρησιμοποίηση OTUs για τις επόμενες αναλύσεις → Δείκτες ποικιλότητας, υπολογισμός κάλυψης, ομοιότητες ανάμεσα στα δείγματα, φυλογενετική ανάλυση, στατιστικές αναλύσεις, απεικόνιση αποτελεσμάτων).

Επεξεργασία αλληλουχιών από illumina

- Μεταφόρτωση mothur και των άλλων λογισμικών που χρειάζονται για την ανάλυση
https://www.mothur.org/wiki/Download_mothur

- Μείωση των λαθών της αλληλούχησης και της PCR-δημιουργία contigs.

```
mothur > make.contigs(file=stability.files, processors=8)
```

```
F3D0      F3D0_S188_L001_R1_001.fastq      F3D0_S188_L001_R2_001.fastq
F3D141    F3D141_S207_L001_R1_001.fastq    F3D141_S207_L001_R2_001.fastq
F3D142    F3D142_S208_L001_R1_001.fastq    F3D142_S208_L001_R2_001.fastq
```

```
mothur > screen.seqs(fasta=stability.trim.contigs.fasta, group=stability.contigs.groups,
maxambig=0, maxlength=275)
```

- Ευθυγράμμιση (alignment) των αλληλουχιών

```
mothur > align.seqs(fasta=stability.trim.contigs.good.unique.fasta, reference=silva.v4.fasta)
```

- Απομάκρυνση χιμαιρικών αλληλουχιών

```
mothur > chimera.vsearch
```

```
(fasta=stability.trim.contigs.good.unique.good.filter.unique.precluster.fasta,
count=stability.trim.contigs.good.unique.good.filter.unique.precluster.count_table,
dereplicate=t)
```

Επεξεργασία αλληλουχιών από illumina

- Ταξινόμηση αλληλουχιών

```
mothur > classify.seqs
```

```
(fasta=stability.trim.contigs.good.unique.good.filter.unique.precluster.pick.fasta,  
count=stability.trim.contigs.good.unique.good.filter.unique.precluster.denovo.vsearch.pick.c  
ount_table,  
reference=trainset9_032012.pds.fasta, taxonomy=trainset9_032012.pds.tax, cutoff=80)
```

- Ομαδοποίηση αλληλουχιών σε OTUs

```
mothur > dist.seqs #Υπολογισμός αποστάσεων των αλληλουχιών απο το alignment#
```

```
(fasta=stability.trim.contigs.good.unique.good.filter.unique.precluster.pick.pick.pick.fasta,  
cutoff=0.03)
```

```
mothur > cluster #ομαδοποίηση #
```

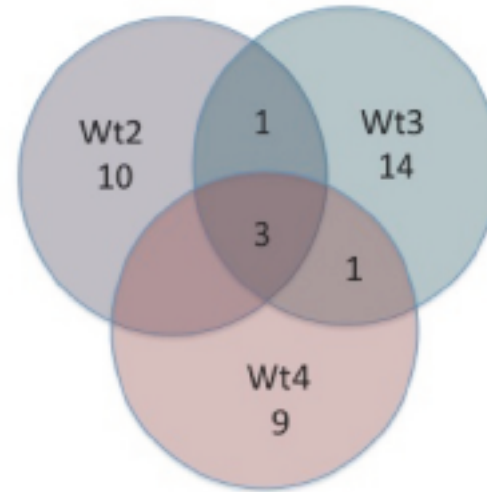
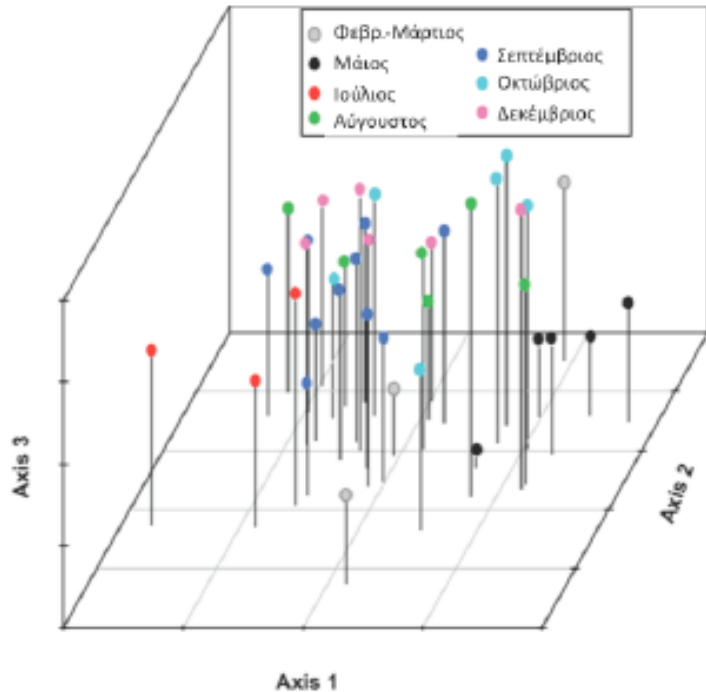
```
(column=stability.trim.contigs.good.unique.good.filter.unique.precluster.pick.pick.pick.dist,  
count=stability.trim.contigs.good.unique.good.filter.unique.precluster.denovo.vsearch.pick.p  
ick.pick.count_table)
```

```
mothur > make.shared
```

```
(list=stability.trim.contigs.good.unique.good.filter.unique.precluster.pick.pick.pick.opti_mcc.  
unique_list.list,  
count=stability.trim.contigs.good.unique.good.filter.unique.precluster.denovo.vsearch.pick.p  
ick.pick.count_table, label=0.03)
```

Επεξεργασία αλληλουχιών από illumina

- Χρησιμοποίηση OTUs για τις επόμενες αναλύσεις



Διάγραμμα Venn για την εύρεση Κοινών φυλότυπων μεταξύ δειγμάτων

Non-metric multidimensional scaling (NMDS) σε δείγματα από το έντερο καραβίδας-ανάλυση βακτηριακών κοινοτήτων