

Lab05 Analiza mikrobiomu

Dane wejściowe

1. Wejdź na usegalaxy.com
2. Znajdź pliki znajdujące się w Shared data → Histories → lab05_microbiome
 - 2.1.dane referencyjne:
 - 2.1.1.silva.v4.fasta
 - 2.1.2.trainset9.032012.pds.fasta
 - 2.1.3.trainset9.032012.pds.tax
 - 2.1.4.mouse.dpw.metadata
 - 2.2.kolekcję plików fastq: mouse gut microbiome

Analiza jakości

Redukcja błędów sekwencjonowania i PCR

1. Utwórz kontigi z nakładających się odczytów. Znajdź program **Mothur Suite** -> **Make.contig**
 - 1.1.Select a way to provide forward and reverse fastq files ?: Multiple pairs - Combo mode (list:paired collection)
 - 1.2.Fastq pairs (collection) : mouse gut microbiome
2. Przeprowadź analizę jakości na pliku **trim.contigs.fasta** z poprzedniego kroku za pomocą **Summary.seqs**
3. Przeprowadź filtrowanie za pomocą **Screen.seqs**
 - 3.1. Fasta: trim.contigs.fasta z poprzedniego kroku
 - 3.2.Group: plik group utworzony przez Make.contigs
 - 3.3.Maxlength: 275
 - 3.4.Maxambig:0
4. Usuń duplikaty za pomocą **Mothur Suite** -> **Unique.seqs**
 - 4.1.fasta: good.fasta (utworzony przez Screen.seqs)
 - 4.2.Output format: Name file
5. Ile znaleziono sekwencji unikalnych, a ile duplikatów?
6. Zlicz wystąpienie każdej sekwencji w próbkach za pomocą **Count.seq**
 - 6.1.name: name utworzony przez Unique.seq
 - 6.2.Use a group file: Yes
 - 6.3.Group: plik utworzony przez Screen.seqs

Mapowanie sekwencji

1. Przeprowadź mapowanie odczytów do sekwencji referencyjnej za pomocą **Align.seqs**
 - 1.1.fasta: plik fasta utworzony przez Unique.seqs
 - 1.2.Reference: silva.v4.fasta
2. Przeprowadź analizę jakości dopasowania za pomocą **Summary.seqs**
 - 2.1.fasta: plik z fasta z poprzedniego kroku
 - 2.2.Count: plik count_table utworzony przez Count.seqs
 - 2.3.Output logfile: yes
3. Zapisz statystyki do protokołu
4. Odfiltruj kolumny zawierające tylko niedopasowania za pomocą **Filter.seq**
 - 4.1.fasta: good.fasta utworzona przez Screen.seq
 - 4.2.Vertical: yes
 - 4.3.trump: ``
 - 4.4.output logfile: yes
5. Usuń duplikaty sekwencji za pomocą **Unique.seqs**
 - 5.1.fasta: filtered fasta z Filter.seq
 - 5.2.name file or count table: count table z Count.seq

Klastrowanie sekwencji

1. Połącz sekwencje o wysokim podobieństwie za pomocą **Pre.cluster**
 - 1.1.fasta: plik fasta utworzony przez Unique.seqs
 - 1.2.name file or count table: count table utworzony przez Unique.seqs
 - 1.3.Diffs: 2
2. Ile sekwencji unikalnych pozostało?

Klasyfikacja sekwencji do jednostki taksonomicznej

1. Zaklasyfikuj sekwencję do jednostki taksonomicznej za pomocą **Classify.seqs**
 - 1.1.fasta: fasta utworzona przez Pre.cluster
 - 1.2.Reference: trainset9032012.pds.fasta
 - 1.3.taxonomy: trainset9032012.pds.tax
 - 1.4.Count: count table utworzona przez Pre.cluster

Klastrowanie sekwencji w OTU

1. Utwórz OTU za pomocą **Cluster.split**
 - 1.1.Split by: Classification using fasta
 - 1.2.Fasta: fasta z Pre.cluster
 - 1.3.Count: Count table z Pre.cluster
 - 1.4.Taxonomy z Classify.seq
 - 1.5.Taxlevel: 4
 - 1.6.Cutoff: 0.03
2. Sprawdź ile sekwencji w każdym OTU należy do danej grupy **Make.shared**
 - 2.1.Select input type: OTU list
 - 2.2.list: plik utworzony przez Cluster.split
 - 2.3.count: Count table z Pre.cluster
 - 2.4.label: 0.03
3. Zaklasyfikuj OTU do jednostki taksonomicznej za pomocą **Classify.otu**
 - 3.1.list: plik utworzony przez Cluster.split
 - 3.2.count: Count table z Pre.cluster
 - 3.3.taxonomy: taxonomy utworzone przez Classify.seqs
 - 3.4.label: 0.03
4. Które próbki zawierały OTU zaklasyfikowane jako Staphylococcus?

Porównanie mikrobiomów z poszczególnych prób

1. Utwórz diagram Venn'a za pomocą **Venn**
 - 1.1. OTU shared: plik utworzony przez Make.shared
 - 1.2. groups: F3D0, F3D1, F3D2, F3D3
2. Utwórz dendrogram obrazujący zależności pomiędzy próbami
3. Zastosuj **Dist.shared** w celu obliczenia dystansów
 - 3.1. shared: shared utworzony przez Make.shared
 - 3.2. calc: thetayc.jclass
4. Wizualizuj zależności za pomocą **Tree.shared**
 - 4.1. Select input format: Phylip Distance Matrix
 - 4.2. phylip: plik dist z poprzedniego kroku
5. Pobierz plik utworzony za pomocą Tree.shared i zwizualizuj za pomocą <https://icytree.org/>
6. Wizualizacja mikrobiomu
 - 6.1. Wybierz z listy program **Make.biom**
 - 6.1.1. Shared: plik shared z Make.share
 - 6.1.2. constaxonomy: plik taxonomy utworzony przez Classify.otu
 - 6.1.3. metadata: mouse.dpw.metadata
 - 6.2. Pobierz plik i zwizualizuj za pomocą programu **Phinch** <http://phinch.org/>
[Przed załadowaniem pliku zmień rozszerzenie na .biom]