

Welcome back

Today's schedule

TIME	Tuesday 20th of March 2018
9:00-9:15	Welcome back and intro
9:15-9:30	<i>E. coli</i> diagnostics and the CGE databases
9:30-9:45	
9:45-10:00	
10:00-10:15	Exercise 2
10:15-10:30	Computer work w. VTEC (ResFinder, SerotypeFinder, VirulenceFinder, myDBFinder)
10:30-10:45	
10:45-11:00	
10:45-11:00	Coffee
11:00-11:15	Exercise 2 continued
11:15-11:30	
11:30-11:45	
11:45-12:00	Wrap-up of computer work
12:00-12:15	Lunch
12:15-12:30	
12:30-12:45	
12:45-13:00	

13:00-13:15	Whole genome based phylogeny (Johanne Ahrenfeldt, DTU)
13:15-13:30	
13:30-13:45	
13:45-14:00	Introduction to Exercise 3
14:00-14:15	Exercise 3 Computer work - phylogeny (NDtree and CSIPhylogeny)
14:15-14:30	
14:30-14:45	Coffee
14:45-15:00	Exercise 3 continued
15:00-15:15	Wrap-up of computer work
15:15-15:30	
15:30-15:45	
15:45-16:00	Alternative tools for phylogeny
15:45-16:00	Identification of unique core sequences
16:00-16:15	Exercise 4 Computer work - RUCS
16:15-16:30	
16:30-16:45	
16:45-17:00	
	Wrap-up of computer work
	Workshop dinner in the evening

Yesterday's multiple choice questions

6. In a file containing raw sequence reads, one of the bases is associated with the quality score "N". What is the probability that this base is incorrect?

A: 0.0000016 %

B: 0.015 %

C: 45 %

D: 78 %

E: 0.00316 %

Answer:

N corresponds to the Quality score 45 according to table slide 38 in pres. on Sequencing techniques

Error probability, $P = 10^{-Q/10} = 10^{-4,5} = 0,0000316 = 0,00316\%$

10. A draft genome consists of seven contigs. Their lengths are listed below. What is the N50 value of the draft genome?

Contig1: 2500
Contig2: 150.000
Contig3: 100.000
Contig4: 90.000
Contig5: 25.000
Contig6: 500
Contig7: 130.000

- A: 498.000
- B: 249.000
- C: 130.000
- D: 100.000
- E: 150.000

Answer:

First, calculate half of the assembly length =

$$(2500 + 150.000 + 100.000 + 90.000 + 25.000 + 500 + 130.000) / 2 = 249.000$$

Sorting the contigs according to length and starting from the longest contig, what is the length of the last contig added to the set of contigs that together cover at least half of the assembly length?

$$150.000 + 130.000 > 249.000. \text{ N50} = 130.000$$