Danmarks Tekniske Universitet / Technical University of Denmark

Skriftlig prøve / Written examination: dato/date: 20/12-2018

Side 1 af 9 sider / Page 1 of 9 pages

Kursus navn / Course titel: Metagenomics and Microbiome Analysis

Kursus nummer/ Course number: 36636

Hjælpemidler / Aids allowed: NO

Varighed / Exam duration: 2 hours

Vægtning/ Weighting: 40%

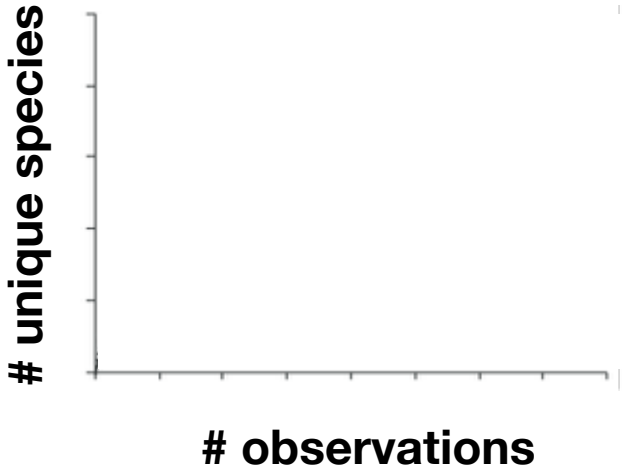
Remark: The last 2 pages marked as notes can be used if extra space is necessary

A. Metagenomics basics (25%)

1. Explain what it means to do metagenomics?

2. What is the main difference between 16S and shotgun metagenomics?

3. Draw the rarefaction curves of two different samples in the plot below. Sample 1 (S1) has lower richness compared to sample 2 (S2).



4. Explain in your own words what alpha and beta diversity means.

5. How would you design an intervention study where you want to study the effect of coffee consumption on the human microbiome?

B. Laboratory questions (35%)

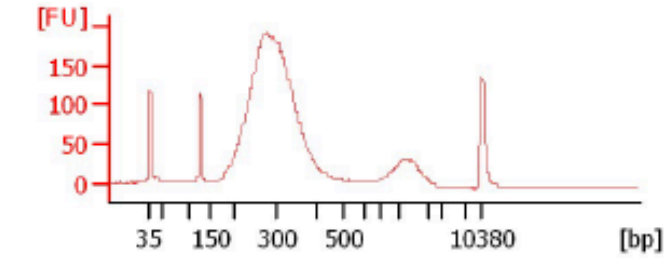
6. State four important considerations to make when collecting and handling metagenomics samples:

7. Homogenisation of the sample is important. Briefly explain why:

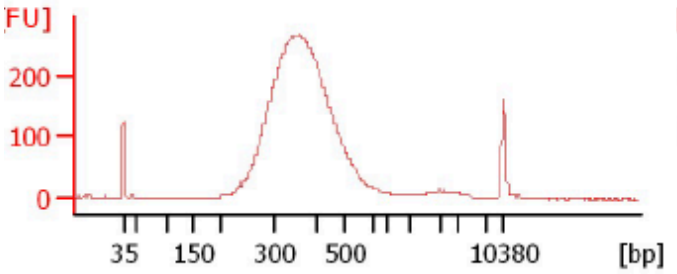
8. What is an adapter and what elements does it consists of?

9. When you evaluate your final libraries on the bioanalyser electrograms which parameters will you look at? Please indicate and explain based on electrogram A below (next page):

A



B



10. Which of the two above (A or B) would you send for sequencing and why?

11. What is the approximate length of the library above? Is that the insert size?

12. Explain why we make double beads purification?

C. NGS analysis basics and quantitative metagenomics (40%)

13. Explain what the 4 lines are in the read below:

```
@ILLUMINA-C90280_0030_FC:5:1:2675:1090#NNNNNN/1
ATTCCCGGCCTTTTTCCAGGCCTGCCTGCTCGAGC
+
BAAAGECEE<EEDFEDF3DBDBB=A+==>9>>88?
```

14. What is the purpose of *de novo* assembly of Illumina data when doing metagenomics?

15. How do you create a non-redundant gene catalogue from raw sequencing reads? You can draw the workflow

16. What is the problems with using databases (e.g. aligning to them with blast) in determining which species are present in your metagenomics samples?

17. How does metagenomics binning work? Base your explanation on canopy clustering.

18. What is the difference between rarefying and normalizing sample counts? Which one is the best to use?

19. Explain why we use compositional analysis and what makes it better compared to rarefying, and what makes it better compared to normalization

Notes: