Some sample questions

In the case of multiple choice, only one answer is correct.

1. Microarray: What is a probe and what is a sample?
    - The probe is the labeled RNA or DNA and the sample is attached to the microarray surface.
    - The probe is the is attached to the microarray surface and the sample is the labeled RNA or DNA.
    - Both probe and sample are actually the same namely the molecules attached to the microarray surface.

2. Microarray: What is the purpose of the mismatch probes on an Affymetrix microarray array?
    - They are there to measure also a good signal in the case of single nucleotide polymorphisms.
    - They represent sequences, for which the target sequence showed ambiguities.
    - They are supposed to measure unspecific hybridization.

3. Microarray: Why do you have to normalize microarray data to compare two conditions? (Hint: think about the nature of the signal)

4. General: What is a MA-plot?
    - The MA-plot is a plot, where the intensities of one experiment is plotted against the other.
    - The MA-plot is a plot, where the sum of the intensities of two experiments is plotted against the difference of the two intensities.
    - The MA-plot is a plot, where the log-intensities of one experiment is plotted against the untransformed intensities.
    - The MA-plot is a plot, where the sum of the log-intensities of two experiments is plotted against the difference of the log-intensities.

5. Microarray: Variance Stabilization is a transformation, which achieves:
    - That the variance of the data is one.
    - That the variance along the total range of the data is approximately constant.
    - That the law of error propagation is not valid anymore.

6. Microarray: What is a probeset?

7. Microarray: What is probeset summarization and why do we need it?

8. RNA-seq: Why do you have to remove rRNA and tRNA before performing RNA-seq?
    - Because otherwise one sequences just these, since these two classes make up to 90% of the total RNA.
    - Because in particular tRNA binds to the mRNA and renders sequencing difficulat.
    - Because some rRNAs and tRNAs have the same sequence as the sequencing adaptors
9. RNA-seq: You have obtained a fastq file with RNA-seq reads, what do you have to do to get gene expression estimates from these reads?

10. RNA-seq: Why is mapping of RNA-seq reads more difficult than mapping re-sequencing reads or ChIP-seq reads?
    - RNA is so unstable, such that the sequence is of bad quality
    - mRNAs are spliced together from exons, which are separated by introns, therefore there are some reads which span two exons, a sequence that is not present in the genome.
    - The RNA polymerase makes some many mistakes that there are a lot of errors in the RNA
    - The dynamic range of RNA-seq reads is much higher than in re-sequencing reads, such that mapping gets difficult, because on hits certain regions over and over again

11. RNA-seq: For what is RPKM an abbreviation and for which two biases in the tag counts does it account fort

12. General: How is the variance of a poisson distribution defined?
    - It is N p, where N equals the sample size and p the probability of success, i.e. the variance is equal to the mean.

- It is N^2 p where N equals the sample size and p the probability of success, i.e. the variance is almost the mean squared.
- The poisson distribution has no variance.

13. RNA-seq: What is the main concept used in *de novo* transcriptome assembly?
    - Suffix trees
    - Suffix arrays
    - Gapped kmer seeds
    - De Bruijn graphs

14. ChIP-seq: name the 9 steps required for a ChIP-seq experiment (until you have the reads)!

15. Multiple-testing: Suppose you perform 2 statistical test, where you would declare the null-hypothesis as significantly rejected when the p-value is smaller than 0.05. What is the probability that you have wrongly rejected at least one null-hypothesis? (just a number)

16. ChIP-seq: Why is there strand-dependent bimodality in ChIP-seq peaks?

17. General: What are the two main reason why one performs Principle Component Analysis

18. DNA binding motifs: What is a PWM?
    - It is a description of the amino acids that bind to DNA.
    - It is an abbreviation for protein water mass
    - It is an abbreviation for position-specific weight matrix

19. DNA binding motifs: why is the most frequent k-mer in a set of peak sequences not the best candidate for a binding sequence