**Question 1**

2nd and 3rd generation sequencing techniques have very different performance characteristics.

1. Please explain in a concise paragraph what the parameters: “read-length”, “coverage”, “throughput” and “error rate” describe.
2. Name one example each of a 2nd and a 3rd generation sequencing technique and give a rough value of their read-length and error rate.
3. Use your knowledge of the molecular process underlying these two sequencing techniques to explain the big differences in read-length, error rate and error type between these two techniques.

**Question 2**

Many of the biological DNA sequences investigated by sequencing (e.g. entire genes or even entire genomes) are much longer than the reads that can be obtained by any of the available sequencing technologies. Also, the error rates of the raw sequencing reads obtained by 2nd and 3rd generation sequencing technologies are very high relative to the frequency of real genetic variations that are investigated in most research projects. The questions below are about how short and error-prone primary sequencing data can be assembled into long and reliable sequences. In answering the questions please feel free to use diagrams or drawings to illustrate your point.

1. How does the shotgun approach use sequencing data of short, random DNA fragments to reconstruct the continuous sequence of the much longer unfragmented original DNA molecules. And, what determines the shortest read length that can, theoretically, be used to determine a long DNA sequence (e.g. a human genome) with the shotgun approach.
2. Explain how raw sequencing data with relatively high error rates can be assembled into final DNA sequences with very low error rates. In particular, explain how one distinguishes between a sequencing error (e.g. base call error) and an actual change in the DNA sequence (e.g. single nucleotide polymorphism).

**Question 3**

You are working as the sequencing specialist in a medical team that is treating an individual with a severe genetic disease. You are suspecting that the cause of the disease is a large scale chromosomal rearrangement.

1. Why may complete genome-resequencing with short-read, sequencing-by-synthesis technology not be the ideal technology to investigate a chromosomal rearrangement?
2. What other technology may provide a faster and more reliable way to detect and localize a chromosomal rearrangement?

**Question 4**

In the lab, you have just “cloned” a PCR product that is about 4000 bp long. That is to say you ligated this PCR product into a plasmid and have purified many copies of this plasmid. You now want to determine the sequence of the cloned PCR product.

1. What sequencing technique will you likely want to use and why?
2. Describe how you will design the sequencing primer?
3. Describe how you will continue with your work once you received your first sequencing results?

**Question 5**

The massspectrometers used for peptide-fragment-based protein massspectrometry use two mass-filter stages.

1. What type of ions are separated in these two mass-filter stages?
2. What takes place between these two mass-filter stages?
3. Why is it necessary to use two mass-filter stages in order to identify a peptide?

**Question 6**

Peptide-fragment-based protein massspectrometry can be used to analyze the presence of proteins in complex samples (e.g. extracts from cells or tissues).

1. Please describe what is the raw data that is recorded in a typical peptide-fragment-based protein massspectrometry experiment.
2. How will you use this this raw data to determine, which proteins were present in the sample? Explain the importance of proteotypic peptides in this context.

**Question 7**

A colleague of yours from a marine biology laboratory has discovered a new worm species that lives at the bottom of the Pacific Ocean. Remembering what you had told her about studying protein expression in human tissues via protein MS, the first thing she now wants to do is to analyze the proteome of this new organism with peptide-fragment-based protein massspectrometry.

1. Would the approach you are using to study the presence of proteins in human tissue samples work for the samples from the worm? Explain why.
2. What should your colleague analyze first before starting to analyze the proteome.