Exercise 2: Questions and hints

Mars Soil Sample Analysis: If you obtained a soil sample from Mars for identifying microorganisms and assembling their genomes, which sequencing technology would be optimal? Consider factors like the detection of novel organisms and the precision required for genome assembly. Discuss your choice, focusing on read length, accuracy, and cost implications.

- **Chosen Platform:** Oxford Nanopore Technologies (ONT) or Pacific Biosciences (PacBio)
- **Rationale:**
    - **Long Reads for Genome Assembly:** These technologies can produce much longer reads, crucial for complex or novel microbial genomes assembly. Long reads overcome challenges like repetitive regions, aiding effective de novo assembly.
    - **Detecting Novel Organisms:** More likely to capture and accurately assemble genomes of potentially novel organisms in Martian soil.
- **Platform Differences:**
    - **Oxford Nanopore Technologies:** Known for portability and ultra-long reads. Its real-time sequencing is advantageous for adaptive sequencing strategies.
    - **Pacific Biosciences:** Offers high accuracy with HiFi reads, preferred for the highest accuracy in genome assembly.
    - **Cost Considerations:** Higher cost per base than Illumina is justified by their capability to assemble complex or unknown genomes.

Gut Microbiome Study: In researching the impact of dietary changes on the gut microbiome, what type of sample would you collect, and which sequencing technology would be most suitable? Provide your rationale for this choice.

- **Chosen Platform:** Illumina Sequencing
- **Rationale:**
    - **Suitability for 16S rRNA Sequencing:** Highly effective for 16S rRNA gene sequencing, commonly used to study bacterial communities in the gut.
    - **Cost-Effectiveness:** Offers a balance between cost, read length, and throughput, making it affordable for large-scale microbiome studies.
    - **Short Reads and High Throughput:** Suitable for both 16S rRNA and shotgun metagenomics approaches, with the ability to process multiple samples simultaneously for reduced costs.