Charlee Cobb - Transcriptomics, Exercise 6

Charlee Cobb

2023-02-27

Give three different applications of Next-generation sequencing technologies. Genome annotation, comparative genomics, and SNP variation studies.

Why is it important to know the structure of the gene? This reveals history of gene distribution and can explain phenotypic traits.

Why are important parts of the genome less likely to contain changes? Because mutations in a genome that harm an individual are less likely to be passed down, minimizing the prevalence of changes in the genome. So the genetic code stays robust because of natural selection.

What can we learn by comparing the human genome to the mouse genome? There are a lot of similarities between the human and mouse genome, and studying the differences and similarities lead to more respect and understanding of our evolution.

What are SNPs and how can they be used to improve our health? SNPs are single nucleotide polymorphisms which are changes in the genetic code at a nucleotide level. So an A at a certain location in DNA is G in the same location in different populations.

What do we hope to learn by sequencing 1 million individuals? The goal in sequencing 1 million individuals is to increase diversity in the data, which broadens our data set for genetic analysis. This adds more representation and diversity studies and diagnosis.

Do we have more human cells or bacteria cells on our bodies? We have more bacteria cells than human cells on our bodies, about 10 times more. So metagenomics analysis is increasingly important.

How did Next-generation sequencing technology change the way we do biology?*** It increased the throughput of data and maintained high accuracy in sequencing. This made it cheaper and safer to sequence DNA, RNA, and Proteins. Next-generation sequencing mitigate many boundaries and limitation of biological experiements.

Explain how Dr. Sanger used the properties of DNA replication to sequence DNA. What is the difference between the dNTPs and ddNTPs? Sanger utilized the fact that DNA matches with complimentary strands to create a sequencing technique that detects nucleotides on the hybridizing strand. As the nucleotides are bonding, you make multiple fragments and add a "stop" nucleotide randomly in the sequence to terminate the building of the DNA strand. Then with gels, we can align the strands and detect the flourcesnt "stop nucleotides in order, revealing the sequence of DNA.

dNTPs are used to continue elongation of the hybridizing strand. ddNTPs are the "stop" nucleotide that terminate the elongation process.

How was the sequencing automated? Using robotics elements. The four nucleotides where each given a color, and automated machines were able to detect the color using a laser.

At which step of the sequencing workflow do the Next-generation sequencing technologies capitalize to increase the throughput? They increased the number of capillaries used, as machines could handle up to 384 capillaries at a time.

What is bridge amplification? Bridge amplification is used by illumina sequencing products. It fragments the DNA, then uses a polyamerase to build a complimentary strand of DNA. The DNA fragments

are held down on a plate by oligonucleotides adapters. After the first strand is made and attached to the adapters, the original strand is washed away, and remade by the new strand when it bridges to another adapter. Everytime a nucleotide is added, it gives off a flourescent light, signifying which nucleotide it is.

What does the number of clusters and number of cycles mean in terms of Illumina? Cluster is the area in which the amplification reaction is happening on the plate. Cycles reference read length, so how many nucleotides were attached in sequencing.

Why may it be beneficial to sequence both ends of the fragment? It's helpful to sequence both ends of a fragment because we can construct better alignment with two sides of the DNA sample.