Some of the techniques used in the human genome project

Science, Genetics

Introduction

DNA is the key to understanding all life forms. Many techniques have been discovered throughout the years in the race to be the first research team to code the human genome. The human genome project (HGP) began in 1990. It was an international project which lasted thirteen years costing $3. 8 billion dollars from the US government Bodies using Sanger sequencing. Celera Genomics a private company, launched a project using shotgun sequencing to complete the human genome. The HGP led to the discovery of many DNA sequencing techniques as well as a boom in technological advances data analysis and storage. With the aim to gain greater knowledge on our health and molecular composition as well as aiding in medical diagnostics the HGP has countless applications.

Sanger Sequencing and the Plus and Minus Method

In 1977 Fredrick Sanger developed a DNA sequencing method called “ Sanger Sequencing “ or “ Chain termination method”. The single stranded DNA (up to 900bp) sample is separated into four reaction tubes. Each tube contains the normal deoxynucleosidetriphosphatesd, DNA polymerase, and one of the four dideoxynucleotides (ddNTP’s) which lack a three prime OH group at a specific concentration. ddNTP’s can be radiolabelled to aid visualisation. Once the ddNTP is added the sequence the reaction stops due to the absence of OH group as the DNA polymerase cannot add another nucleotide. The DNA fragments are then separated by their mass in a gel. The DNA bands can be visualised using auto radiography or UV light method and the sequence is determined.

At the time, Sanger sequencing was in rivalry with the “ plus and minus” DNA sequencing technique. The phi X 174 (or ΦX174) bacteriophage genome was sequenced using the plus and minus method 1977 then re-sequenced using the chain termination method. ΦX174 a gram negative bacterium found in the repertory tract of humans. It causes repository tract infections as well as B type strain causing Meningitis.

Sanger sequencing overcame the problems associated with the plus and minus method as it is quicker with longer read lengths. To this day Sanger sequencing is still used in many laboratories as it has proven to be very specific. Andersen-Tawil syndrome a multisystemic autosomal dominant disease caused by a heterozygous mutation in the KCNJ2 gene, is diagnosed using Sanger sequencing. Symptoms include muscle weakness and abnormalities to the facial and skeletal features.

Polymerase Chain Reaction

The Polymerase chain reaction (PCR) was first discovered by Kary Mullins in 1985 had an immense impact in the unravelling of the human genome. This enzymatic assay amplifies DNA (approx. 0. 1-10Kb) in an exponential fashion generating many copies of the DNA in a relatively short generation time. A PCR thermal cycler was developed in 1988 which reinforced PCR as one of the major discoveries of its time. Developed by Mullins, the thermal cycler was an automated system developed to perform the specific temperature changes required for the PCR reaction. The amplified DNA products are them analysed using polyacrylamide gel electrophoresis or Agarose gel

electrophoresis methods. PCR is deemed a qualitative method as it can detect the presence or absence of a particular DNA sequence.

Quantitative PCR (Q-PCR) is used to quantify nucleic acids. Q-PCR can be dye based where fluorescent labeling allows the quantification of the amplified DNA. During each cycle, fluorescence is measured in real time. In probe based Q-PCR many targets can be analysed at the same time but probes must be specific. Q-PCR uses a standard curve and statistical methods to quantify the amplified DNA Q-PCR was used in the detection of Coxiella burnetti, which causes of Q fever that infects mammals including humans and cattle.

In Reverse transcription PCR (RT-PCR) the primary sequence is RNA and is amplified using complementary DNA. RT-PCR can be analysed using quantitative PCR methods in real time. RT-PCR was used in diagnosing the Influenza A virus, which causes annual epidemic disease resulting in approximately 20, 000 to 36, 000 deaths per year in the United States.

PCR is not without its limitations, contamination to leads ambiguous results. Primers need to be highly specific, leading to previous knowledge of the DNA sequence. The amount of free nucleotides present dictate and control the reaction. G-C rich regions can make secondary hairpin loop structures hindering DNA polymerase. The pH must be kept constant and the fidelity of the DNA polymerase are all factors that need to be addressed. PCR methods were used in the detection of C. puneumoniae in the cerebral spinal fluids in multiple sclerosis patients.

Conclusion

Although these techniques mentioned are only the tip of the iceberg when it comes to the techniques used in the HGP, they were very beneficial not only in the generation of data but also for medical diagnostics. Research is ever changing, and it is exciting to see what techniques are yet to come.