Genomics

It refer study of mapping, sequencing and analysis of the genome. It helps to understand the structure, function ,co-ordinate regulation of all genes in an organisms.

Genes carry information for making all the proteins required by the organisms. These protein determine how the organisms look, how well its body metabolizes food or fight infection and even how it behaves.

DNA is made up of four similar chemicals A,T,Cand G.that are repeated millions or billion of times , for example human genome contain 3 million base pairs.

This sequence brings out

Life is diversity

Differentiation of one species from others

Evolution of these organisms

Local variation within species

Genome projects

It refer mapping and sequencing of DNA of organisms is called as genome projects.

The following organisms are already sequenced , it includes ***Haemophilus influenza, Mycoplasma genitalium, Saccharomyces cerevisiae, Methanococcus jannschii, Pseudomonas aeroginosa, Drosophila melanogaster***

The sequencing projects are useful

To identify all genes located in the DNA of organisms

To determine the nucleotide sequences of genes and intergenic regions in all chromosomes.

To identify the genes responsible for pathogenicity, immunity and abnormalities in organisms.

To store the sequence data for future use

To transfer sequence data to researcher.

To identify the co-ordinated expression and regulation of all genes in an organisms

To understand the genetic basis of life of different organisms

To determine single base disorder in individuals of a species

Human genome project(HGP)

It refer a project to map and sequence the 3 million nucleotides contained in the human genome and to identify all the genes present in it. The HGP is governed by Human Genome Organization (HUGO) was established in 1990 and completed in 2003.

The genome of any individuals (except for identical twins) is unique, multiple variation of each gene.

There are currently two human genome projects

1, HGP-HUGO

2, Celera Genomics

HGP-HUGO

It was established by Charles DeLisi in 1986, the ultimate goal of this is to understand the human genome, is necessary for the progress of medicine and other health sciences, Atleast 18 Countries have participated in human genome research programs. They are **Australia, Canada, USA, UK, Germany, Korea, Japan, USSR, Sweden, Italy, Korea, France , Brazil, China, Denmark, Denmark, Israel** etc,. Other developing countries are participated through the genome research

Celera Genomics HGP.

It is a united base Biotechnology company that takes commercial potential of human genome. It was a privately funded project co-ordinated by Craig Venter. It was aimed to identify the gene sequence for discovery of new drugs.

Goals of HGP

To determine the sequence of the 3 million chemical base pairs that make up human DNA.

To store the information in databases

To improve the tools for data analysis

To transfer data to private sector

To address the ethical legal and social issues that may arise from projects

Techniques of HGP

Two methods are common for sequencing DNA

1, Maxam – Gilbert Technique

2, Sanger Technique

Genome sequencing

The determination of order of various genes along the length of a chromosome is called as gene sequencing.

Methods of gene sequencing

1, Directed gene sequencing or ordered sequencing

2, Shotgun library

3, Whole genome shot gun sequencing

Directed gene sequencing or ordered sequencing

In this method any one DNA fragment of genome is chosen for DNA sequencing and then the fragment coming next to it as in intact chromosome is subjected to DNA sequencing. Thus the DNA sequencing is proceeded from the first fragment and then next fragment. In this method genes are identified using DNA probes and cloned in **BAC** vector

Shotgun library

In this method lambda phages, cosmids, BAC and YAC are used to construct shotgun libraries and they can accommodate 23-350 kb long DNA,

The genomic DNA is isolated from target organism and treated with restriction enzymes , the DNA fragments are maintained in vectors, based on size the cloned DNA are arranged, then subjected for DNA sequencing using chain termination or automatic DNA sequencing method.

Whole genome shot gun sequencing

In this method, DNA of particular chromosome is isolated from the cell DNA .

The DNA is cut into fragments using exposing it to ultrasonic waves.

The DNA fragments are separated by agarose gel electrophoresis, suitable markers are used to detect the molecular weight of the DNA fragments.

DNA is isolated from each and every band from the gel and purified,

It is cloned in a plasmid, thus all DNA fragments are cloned separately.

The gene cloned plasmids are introduced into E. coli cells separately

Each bacterial clone is isolated and sequenced separately

The DNA is sequenced using automatic DNA sequencing

Gene prediction

The determination of genes, their function and regulation in the genome, it provides a clear knowledge of genome organization. It helps to count the actual number of genes in the genome. The structure of particular gene is the same in all species of organisms, for example many genes in mouse show homology to human genes, homology based gene prediction.

Single nucleotide polymorphism

Variation in the lengths of some DNA between individuals due to single base change is called as single nucleotide polymorphism.

The human genome project has stated that genome sequences of any two unrelated men are 99.9% identical and the remaining 0.1 % shows differences from one another. This difference occurs due to gene mutation and their inheritance over many generation. The single nucleotide position occupied by one base in some individuals but by alternative base in some other individuals causes sequence variation leads to variation in length of DNA fragments.

The single base variation are responsible for many variation in characters among different individuals of a population.

For example variation in hair colour is due to single base change within the gene for pigment production,

But in some cases single base changes express disorders in people for example :Cystic fibrosis.

Functional genomics

It refer study of function of all genes and their roles in regulating metabolic pathways at different stages of development. Mutant genes fail to express their traits (character) in the organisms so that the inactivated genes are isolated and characterized to know their actual function in the genomes, here the gene is mutated and the phenotype host is detected.

Gene --- m RNA ------ Protein -------- Phenotype.

DNA microarray

Ordered arrangements of DNA probes on silicon surface is called as DNA microarray or gene chip.

In DNA chips , 16,000 different DNA are immobilized in an area of 12 cm^2 , Latest development is, 16,000 different DNA in $1cm^2$ area of silicon surface. These microarrays are called high density arrays, using these arrays thousands of diseases or complementary strand can be detected at a time. After washings microarray chip can be reused for further tests.

Steps involved :

1, Genomic DNA is isolated and cut into small pieces using restriction enzymes

2. The selective sequence is amplified using PCR.

3,The PCR product is electrophoresed, denatured and labeled with fluorescent compound.

4, The gene chip is dipped in the solution contains labeled DNA .

5, After washing the gene chip is scanned using laser light

6, The hybridized DNA molecule emit fluorescence light, the color of the light depend on nature of fluorescence compound.

Application

1, It is used to detect the gene expression by analysing m RNA(c DNA)

2, It is used to detect the microbes in the environment.

3, It is used to detect the pathogenic organisms from humans and animals and plants.

4,It is used to analyse the transcriptomes and proteomes.

Draw the diagrams

Gene chips

Gene chip is a small glass plate immobilized with array of oligonucleotides specific for different genes. The oligonucleotides are fixed on a glass plate, this technique was introduced by Steve Fodor.

The gene chip has 64,000 unit per $1cm^{2,}$, the DNA is isolated from cells suspected material and cut with restriction enzymes. The fragments are denatured and are labeled with fluorochrome compound, the hybridized materials emits the light, light is received by the detector, which converts the light rays into a scanning image is fed into computer installed with the processing software, the computer gives details of hybridized DNA,

Application

It is used for profiling diseases such as asthma, rheumatoid arthritis, diabetes, cancer and other diseases in man.

Draw a diagram

Protein microarray

In protein microarray, a small glass plate is immobilized with different kinds protein molecules, protein arrays is much less developed than DNA microarray, because protein is very complex, show wide range of chemical properties and binding interaction are more, further, amplification is not possible with PCR.

Before immobilization the protein molecules, proteins are purified using affinity chromatography.

Protein microarray are classified into three types:

1, Interaction array – in this array the protein molecules are quantitatively estimated when proteins are interacted with other proteins, nucleic acid and other molecules.

2, Functional arrays – native proteins of particular organisms are arrayed on glass slide, allowed to interact with sample of same organisms, used for prediction of protein function and expression of the protein at different time and stages.

3,Affinity capture arrays – protein that have affinity towards specific proteins or targets immobilized in arrays, used for determining concentration of analytes.

PROTEOMICS

It refer , study of characterization of the complete set of proteins produced by the genome of an organisms,

It involves the following main aspects:

Identification and quantification of each and every proteins

Determination of the location of the protein in the cells

Post transcriptional modification of proteins

Protein –protein interaction

Primary, secondary and tertiary structure of proteins

Variation in expression of protein due to environmental factors

Dynamics of protein during cell metabolism
Construction of protein map of the cell
There are three main areas of proteomics

1, Expression proteomics

It refer identification and quantification of all proteins components of cell at different stages of development and under different environmental condition

2, Structural proteomics

It refer the structure of protein , protein complexes and their location in cells.

3, Functional proteomics or interaction proteomics

It refer protein-protein interaction and co-ordinating functioning of the protein that form a complex network in living cells.

METHODS TO STUDY PROTEOMICS

Proteomics study involves cell culture, extraction of proteins from the cell culture, purification of protein, digestion of protein using trypsin, mass spectroscopy to detect their amino acid sequences and identification of proteins.

Standard biochemical methods are used to extract proteins from the cell culture

Primary purification is done with dialysis and desalting

Secondary purification methods includes 2-D gel electrophoresis, gel filteration, affinity chromatography, isoelectric focusing, high performance liquid chromatography, etc.

The proteins are stained with coomassie blue , silver or flurophore for visualizing of the proteins in the gel.

The purified proteins are cut into small pieces using trypsin

Mass spectroscopy is used to determine the primary , secondary structures, amino acid sequences, post transcriptional modifications.

Protein microarray are used to detect the expression protein

Similarly search tools used to identify the proteins by comparing the query sequences with standard databases

Protein structure prediction

It refer discovery of 3-D structure of protein from their primary structure , it represents tertiary structure in the case of monomeric proteins and tertiary and quaternary structures in the case of multimeric proteins

Therefore, protein structure prediction helps to find new drug and so designed to attach on these targets.

Ab initio modelling

Designing 3-D structure of protein of new from its amino acid sequence based on chemical –physical properties, computers are used in this modelling.

Comparative modelling

Detecting the 3-D structure of unknown protein using structure of already known protein as reference structures is called comparative modelling, computers are used to match and align the structures.

It has two types

1,Homology modelling

Prediction of 3-D structure of an unknown protein by matching it with structure of its homologous protein already known.

2, Protein threading

Detecting the structure of an unknown protein by matching its amino acid sequence against already solved structures of proteins in the database.

Identification of homologous

A homologue is an identical protein that serve as a template to find out the structure of unknown protein.

Identification of conserved and variable regions

Comparison of identical structures

The regions of unknown protein that show identity with those of the template are called conserved regions or identical structures

Comparison of unrelated structure.

The regions of unknown protein that show low level similarity with those of template are variable regions or unrelated structure.

It is done with pair wise alignment between template and new proteins

Model generation

Designing a model for the unknown protein from the known structures of the template is called as modelling

Functional proteomics

It refer protein- protein interaction and co-ordinating function of these protein that form complex network in living cells is called as interaction proteomics, used to understand the protein participating in various metabolic pathways and protein- protein interaction

Databases of metabolic pathways

It is very much useful for providing information to fermentation processes and bioprocesses engineering

Eco Cye : This a database of metabolic pathways, transporters, gene regulation in E. coli

Klotho: This is a collection of data of biological compounds and their classification

Meta Cye: data of metabolic pathways and enzymes involved in them various organisms are stored in it.

LIGAND :It contains information about chemical compounds and reaction in living being

Biodegradation Database : data about microbial catabolisms and biotransformation are stored it.

Expression proteomics

It describes change in protein contents at different stages of cell cycle and between different cell types of an organisms. Knowledge of expression proteomics is necessary for identification of some protein markers useful for diagnosis of certain diseases such as cancer.

Expression of proteome can be analysed by protein microarrays.