UNIT-III

Requirements for animal cell culture lab:

Any laboratory, in which tissue culture techniques are performed, regardless of the specific purpose, must contain a number of basic facilities. These usually include the following:

• A general washing area

• A media preparation, sterilization, and storage area

• An aseptic transfer area

• Environmentally controlled incubators or culture rooms

• An observation/data collection area

. Equipments

Washing Area

The washing area should contain large sinks, some lead-lined to resist acids and alkalis, draining boards, and racks, and have access to demineralized water, distilled water, and double distilled water. Space for drying ovens or racks, automated dishwashers, acid baths, pipette washers and driers, and storage cabinets should also be available in the washing area

Media Preparation Area

The media preparation area should have ample storage space for the chemicals, culture vessels and closures, and glassware required for media preparation and dispensing. Bench space for hot plates/stirrers, pH meters, balances, water baths, and media dispensing equipment should be available. Other necessary equipment may include air and vacuum sources, distilled and double-distilled water, Bunsen burners with a gas source, refrigerators and freezers for storing stock solutions and chemicals, a microwave or a convection oven, and an autoclave or domestic pressure cooker for sterilizing media, glassware, and instruments.

Transfer Area

Under very clean and dry conditions, tissue culture techniques can be successfully performed on an open laboratory bench. However, it is advisable that a laminar flow hood or sterile

transfer room be utilized for making transfers. Within the transfer area there should be a source of electricity, gas, compressed air, and vacuum. The most desirable arrangement is a small dust-free room equipped with an overhead ultraviolet light and a positivepressure ventilation unit. The ventilation should be equipped with a high-efficiency particulate air (HEPA) filter. A 0.3-µm HEPA filter of 99.97-99.99% efficiency works well.

Culture Room

All types of tissue cultures should be incubated under conditions of well-controlled temperature, humidity, air circulation, and light quality and duration. These environmental factors may influence the growth and differentiation process directly during culture or indirectly by affecting their response in subsequent generations. Protoplast cultures, low-density cell suspension cultures, and anther cultures are particularly sensitive to environmental cultural condition.

Basic Equipment

- Cell culture hood (i.e., laminar-flow hood or biosafety cabinet)
- Incubator (humid CO2 incubator recommended)
- Water bath
- Centrifuge
- Refrigerator and freezer (–20°C)
- Cell counter (e.g., Countess® Automated Cell Counter or hemacytometer)
- Inverted microscope
- Liquid nitrogen (N2) freezer or cryostorage container
- Sterilizer (i.e., autoclave)

Culture media:

One of the most important factors in animal cell culture is the medium composition. In vitro growth and maintenance of animal cells require appropriate nutritional, hormonal, and stromal factors that

resemble their milieu in vivo as closely as possible. Important environmental factors are the medium in which the cells are surrounded, the substratum upon which the cells grow, temperature, oxygen and carbon dioxide concentration, pH, and osmolality. In addition, the cell requires chemical substances that cannot be synthesized by the cells themselves. Any successful medium is composed of isotonic, low-molecular-weight compounds known as basal medium and provides inorganic salts, an energy source, amino acids, and various supplements.

Basic components in culture media:

The 10 basic components that make up most of the animal cell culture media are as follows: inorganic salts (Ca^{2+}, Mg^{2+}, Na^{+}, K^{+}), nitrogen source (amino acids), energy sources (glucose, fructose), vitamins, fat and fat soluble component (fatty acids, cholesterols), nucleic acid precursors, growth factors and hormones, antibiotics, pH and buffering systems, and oxygen and carbon dioxide concentrations.

Complete formulation of media that supports growth and maintenance of a mammalian cell culture is very complex. For this reason, the first culture medium used for cell culture was based on biological fluids such as plasma, lymph serum, and embryonic extracts. The nutritional requirements of cells can vary at different stages of the culture cycle. Different cell types have highly specific requirements, and the most suitable medium for each cell type must be determined experimentally. Media may be classified into two categories: (1) natural media and (2) artificial media.

Natural media:

Natural media consist of naturally occurring biological fluids sufficient for the growth and proliferation of animals cells and tissues. This media useful for promoting cell growth are of the following three types:

1.Coagulant or clots: Plasma separated from heparinized blood from chickens or other animals is commercially available in the form of liquid plasma.

2.Biological fluids: This includes body fluids such as plasma, serum lymph, amniotic fluid, pleural fluid, insect hemolymph, and fetal calf serum. These fluids are used as cell culture media after testing for toxicity and sterility.

3.Tissue extract: Extracts of liver, spleen, bone marrow, and leucocytes are used as cell culture media. Chicken embryo extract is the most common tissue extract used in some culture media.

Artificial media

The media contains partly or fully defined components that are prepared artificially by adding several nutrients (organic and inorganic). It contains a balanced salt solution with specific pH and osmotic pressure designed for immediate survival of cells. Artificial media supplemented with serum or with suitable formulations of organic compounds supports prolonged survival of the cell culture.The artificial media may be grouped into the following four classes: serum-containing media, serum-free media, chemically defined media, and protein-free media.

Cell lines:

A cell line is a permanently established cell culture which will proliferate forever if a suitable fresh medium is provided continuously, whereas cell strains have been adapted to culture but, unlike cell lines, have a finite division potential. A cell strain is obtained either from a primary culture or a cell line.

Types of Cell Lines:

Finite Cell Lines :

The cells in culture divide only a limited number of times, before their growth rate declines and they eventually die. The cell lines with limited culture life spans are referred to as finite cell lines. The cells normally divide 20 to 100 times (i.e. is 20-100 population doublings) before extinction. The actual number of doublings depends on the species, cell lineage differences, culture conditions etc. The human cells generally divide 50-100 times, while murine cells divide 30-50 times before dying.

Continuous Cell Lines :

A few cells in culture may acquire a different morphology and get altered. Such cells are capable of growing faster resulting in an independent culture. The progeny derived from these altered cells has unlimited life (unlike the cell strains from which they originated). They are designated as continuous cell lines.

The continuous cell lines are transformed, immortal and tumorigenic. The transformed cells for continuous cell lines may be obtained from normal primary cell cultures (or cells strains) by treating them with chemical carcinogens or by infecting with oncogenic viruses.

Immobilization of an Enzyme

An immobilized enzyme is an enzyme attached to an inert, insoluble material—such as calcium alginate (produced by reacting a mixture of sodium alginate solution and enzyme solution with calcium chloride). There are various ways by which one can immobilize an enzyme:

Affinity-tag binding: Enzymes may be immobilized to a surface, e.g. in a porous material, using non-covalent or covalent Protein tags. This technology has been established for protein purification purposes. This technique is the generally applicable, and can be performed without prior enzyme purification with a pure preparation as the result. Porous glass and derivatives thereof are used, where the porous surface can be adapted in terms of hydrophobicity to suit the enzyme in question.

Adsorption on glass, alginate beads or matrix: Enzyme is attached to the outside of an inert material. In general, this method is the slowest among those listed here. As adsorption is not a chemical reaction, the active site of the immobilized enzyme may be blocked by the matrix or bead, greatly reducing the activity of the enzyme.

Entrapment: The enzyme is trapped in insoluble beads or microspheres, such as calcium alginate beads. However, these insoluble substances hinder the arrival of the substrate, and the exit of products.

Cross-linkage: Enzyme molecules are covalently bonded to each other to create a matrix consisting of almost only enzyme. The reaction ensures that the binding site does not cover the enzyme's active site, the activity of the enzyme is only affected by immobility. However, the inflexibility of the covalent bonds precludes the self-healing properties exhibited by chemoadsorbed self-assembled monolayers. Use of a spacer molecule like poly(ethylene glycol) helps reduce the steric hindrance by the substrate in this case.

Covalent bond: The enzyme is bound covalentely to an insoluble support (such as silica gel or macroporous polymer beads with epoxide groups). This approach provides the strongest enzyme/support interaction, and so the lowest protein leakage during catalysis.

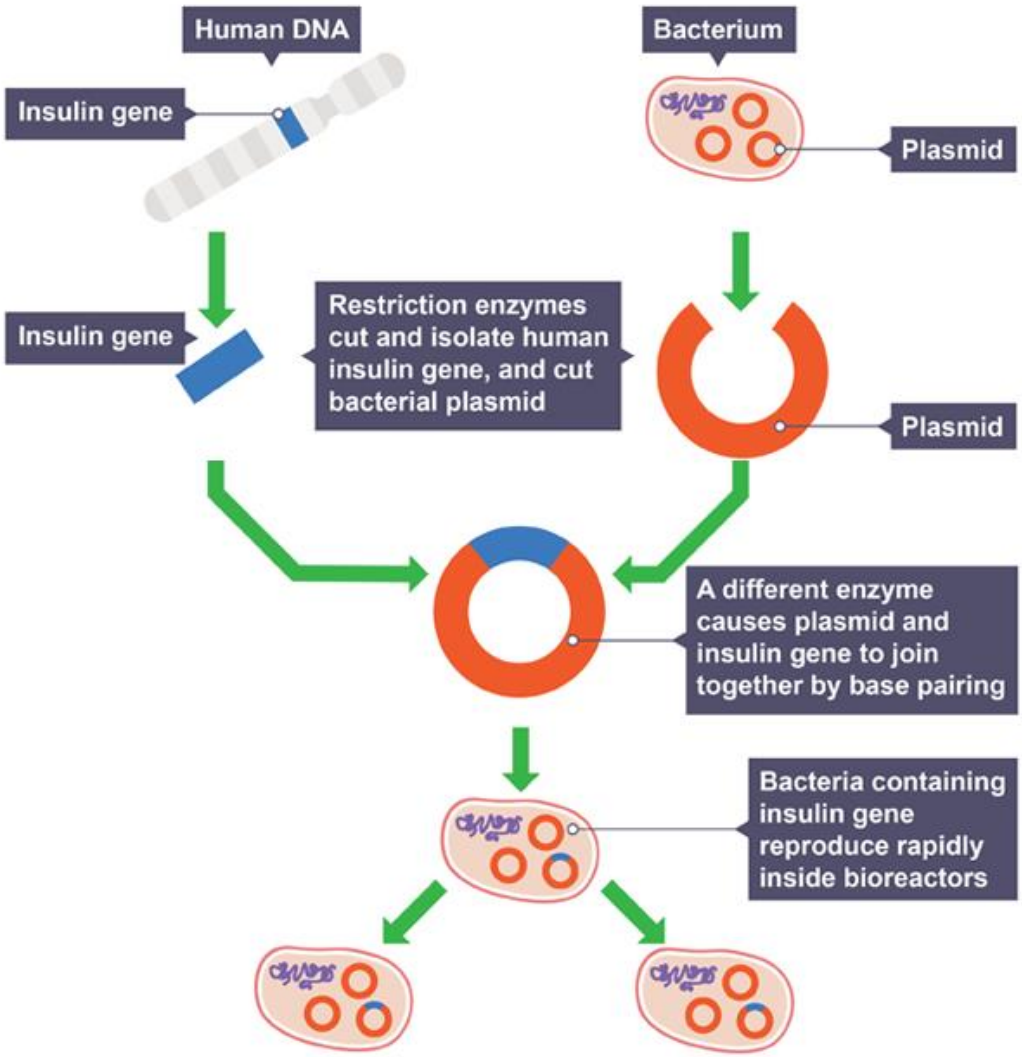
Recombinant insulin production

Recombinant DNA is a technology scientists developed that made it possible to insert a human gene into the genetic material of a common bacterium. This "recombinant" micro-organism could now produce the protein encoded by the human gene. Scientists build the human insulin gene in the laboratory.

Insulin production

Before genetic engineering, insulin was obtained from pigs and cattle.

Due to an increase in the number of diabetics, more insulin is required than ever before.



Process

1. The human insulin gene is removed using a restriction enzyme.
2. A bacterial plasmid is cut open using the same restriction enzyme.
3. Restriction enzymes leave 'sticky ends', where one of the two DNA strands is longer than the other.
4. Using the same restriction enzyme to cut both the human DNA and bacterial plasmid results in complementary sticky ends that join by base pairing.
5. A different enzyme is used to join the insulin gene and the bacterial plasmid.
6. The bacterial plasmid containing the insulin gene is placed into a bacterial cell.
7. The bacterial cell is placed in a fermenter to allow reproduction under perfect conditions (warmth, moisture and oxygen).
8. Downstreaming occurs – this is when insulin is extracted, purified and packaged.
9. The pure insulin produced can be used to treat diabetes.

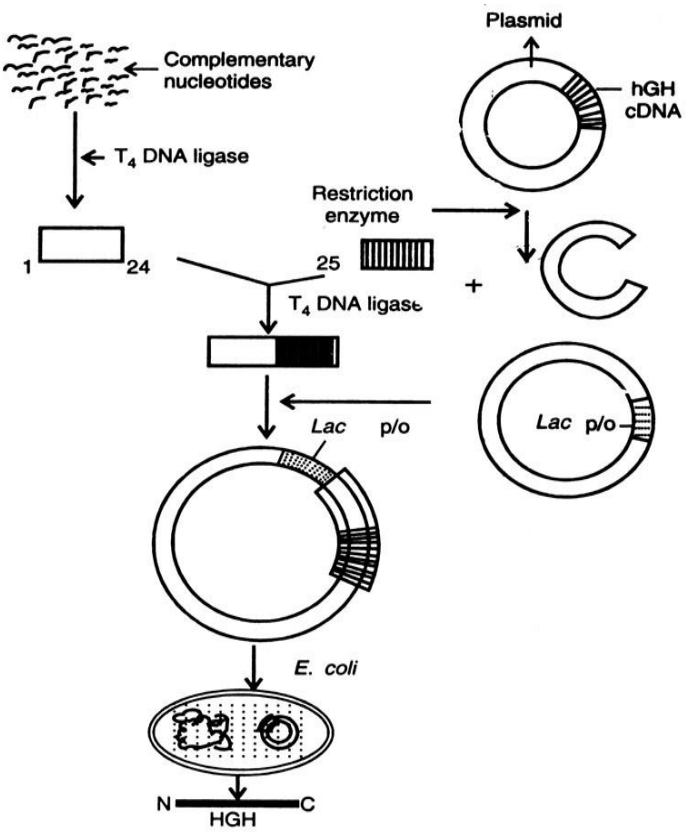
Advantages of genetically engineered insulin:

- Not limited by the slaughter of animals.
- Large quantities can be made quickly.
- No risk of transferring infections.
- More effective at treating diabetes as animal insulin is different to human insulin.
- No ethical issues concerning the use of animals.

Production of human growth hormone (hGH) by recombinant DNA technology

- The gene for human growth hormone (hGH) is isolated from human pituitary gland.

- Insertion of whole hGH gene into plasmid vector and cloning into *E.coli* results into production of biologically inactive hormone because bacteria can translate the region of gene that are not translated in human thereby producing a prehormone containing an extra 26 aminoacids which might be difficult to remove.
- Hence the segment of gene that codes for the first 24 aminoacids of hormone is constructed chemically from blocks of nucleotide.



Step I: Chemical synthesis of gene for first 24 aminoacids:

- From the known aminoacids sequence of hGH, gene for first 24 aminoacids are constructed chemically. These genes are constructed in three small fragments and then they are joined by T4 DNA ligase to get whole gene for first 24 aminoacids.

Step II: Isolation of mRNA for hGH

- In this step mRNA for hGH is isolated from human pituitary gland tissue.

Step III: Reverse transcription

- Using reverse transcriptase enzyme complimentary DNA (cDNA) is synthesized from mRNA.
- The cDNA obtained by reverse transcription process, is the gene for hGH.
- The full gene is cut with restriction endonuclease enzyme to remove first 24 gene.

Step IV: Joining of synthetic gene and cDNA

- In this step synthetic gene (gene for first 24 aminoacids) and cDNA are joined in order to obtain full gene with its own initiation codon (AUG). T4 DNA ligase join these genes.

Step V: selection of suitable vector and recombination:

- Expression vector phGH407 derived from plasmid vector PBR322 is used as carrier vector.
- HGH gene is ligated into a restriction site just downstream of Lac; promotor/operator region of the expression vector.

Step VI: selection and recombination into suitable host cell

- *E. coli* is used as suitable host cell.
- The recombinant expression vector is then transformed into *E.coli.*
- The recombinant *E. coli* then starts producing hGH.
- The recombinant *E. coli* are isolated from the culture and mass production by fermentation technology to obtain hGH.

Use of recombinant human growth hormone (hGH)

1. Treatment of children suffering from growth deficiency
2. Treat the patient with Tumer's syndrome and chronic renal insufficiency
3. To treat patient with renal carcinoma
4. Bovine somatotropin hormone is used to increases milk production in lactating cows and also to increase body mass of cattles

Questions:

5 marks:

1.Give a brief account on culture media

2. What are the requirements for animal cell culture lab

3.Comment on cell lines

4.Write about enzyme immobilization

10 marks:

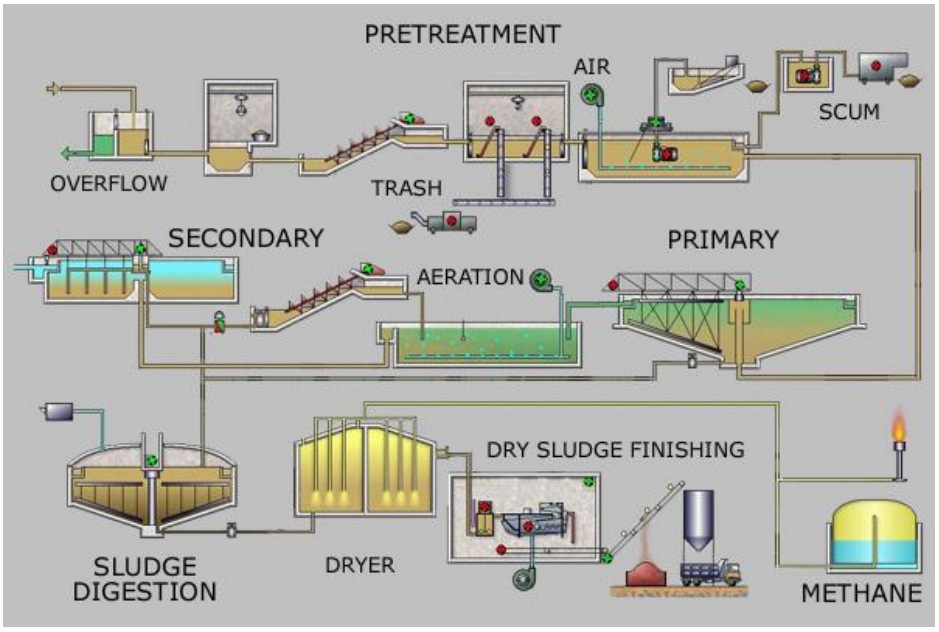
1, Explain the production of recombinant insulin

2. Describe the production of somatotropin by rDNA technology

UNIT-IV

Sewage treatment

Sewage treatment is the process of removing contaminants from municipal wastewater, containing mainly household sewage plus some industrial wastewater. Physical, chemical, and biological processes are used to remove contaminants and produce treated wastewater (or treated effluent) that is safe enough for release into the environment. A by-product of sewage treatment is a semi-solid waste or slurry, called sewage sludge. The sludge has to undergo further treatment before being suitable for disposal or application to land.



Primary treatment

In the primary sedimentation stage, sewage flows through large tanks, commonly called "pre-settling basins", "primary sedimentation tanks" or "primary clarifiers". The tanks are used to settle sludge while grease and oils rise to the surface and are skimmed off. Primary settling tanks are usually equipped with mechanically driven scrapers that continually drive the collected sludge towards a hopper in the base of the tank where it is pumped to sludge treatment facilities.Grease and oil from the floating material can sometimes be recovered for saponification (soap making).

Secondary treatment

Secondary treatment is designed to substantially degrade the biological content of the sewage which are derived from human waste, food waste, soaps and detergent. The majority of municipal plants treat the settled sewage liquor using aerobic biological processes. To be effective, the biota require both oxygen and food to live. The bacteria and protozoa consume biodegradable soluble organic contaminants (e.g. sugars, fats, organic short-chain carbon molecules) and bind much of the less soluble fractions into floc.

Secondary treatment systems are classified as fixed-film or suspended-growth systems.

Fixed-film or attached growth systems include trickling filters, constructed wetlands, bio-towers, and rotating biological contactors, where the biomass grows on media and the sewage passes over its surface. The fixed-film principle has further developed into moving bed biofilm reactors (MBBR) and Integrated Fixed-Film Activated Sludge (IFAS) processes. An MBBR system typically requires a smaller footprint than suspended-growth systems.

Suspended-growth systems include activated sludge, where the biomass is mixed with the sewage and can be operated in a smaller space than trickling filters that treat the same amount of water. However, fixed-film systems are more able to cope with drastic changes in the amount of biological material and can provide higher removal rates for organic material and suspended solids than suspended growth systems.

Some secondary treatment methods include a secondary clarifier to settle out and separate biological floc or filter material grown in the secondary treatment bioreactor.

Tertiary treatment

The purpose of tertiary treatment is to provide a final treatment stage to further improve the effluent quality before it is discharged to the receiving environment (sea, river, lake, wet lands, ground, etc.). More than one tertiary treatment process may be used at any treatment plant. If disinfection is practised, it is always the final process. It is also called "effluent polishing".

Biosensor

A biosensor is an analytical device, used for the detection of a chemical substance, that combines a biological component with a physicochemical detector. The sensitive biological element, e.g. tissue, microorganisms, organelles, cell receptors, enzymes, antibodies, nucleic acids, etc., is a biologically derived material or biomimetic component that interacts with, binds with, or recognizes the analyte under study. The biologically sensitive elements can also be created by biological engineering. The transducer or the detector element, which transforms one signal into another one, works in a physicochemical way: optical, piezoelectric, electrochemical, electrochemiluminescence etc., resulting from the interaction of the analyte with the biological element, to easily measure and quantify. The biosensor reader device connects with the associated electronics or signal processors that are primarily responsible for the display of the results in a user-friendly way.

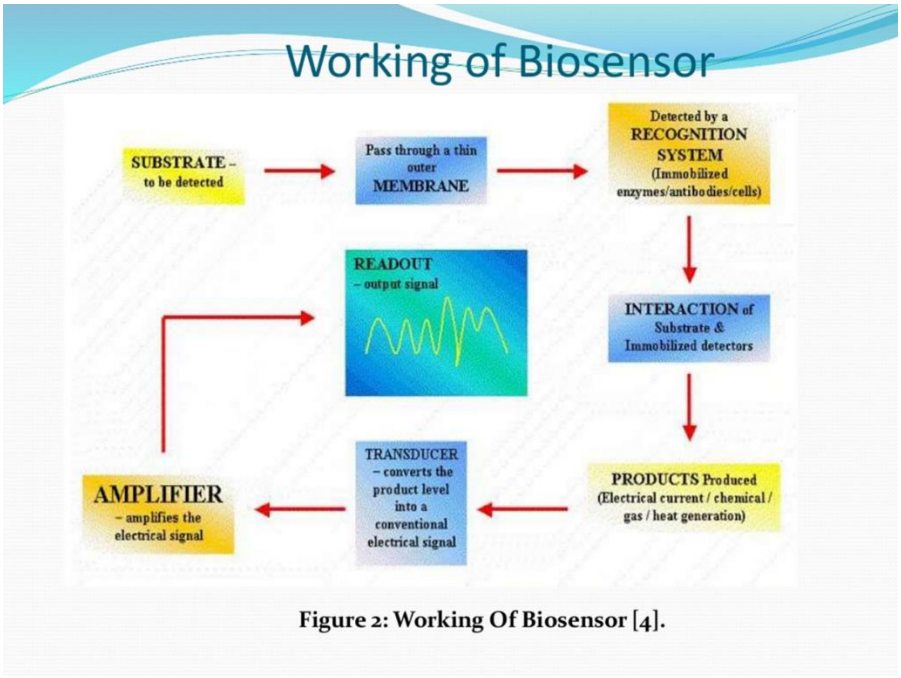


Figure 2: Working Of Biosensor [4].

Every biosensor comprises:

A biological component that acts as the sensor

An electronic component that detects and transmits the signal

Biosensor elements

A variety of substances may be used as the bioelement in a biosensor. Examples of these include:

Nucleic acids

Proteins including enzymes and antibodies. Antibody-based biosensors are also called immunosensors.

Plant proteins or lectins

Complex materials like tissue slices, microorganisms and organelles

The signal generated when the sensor interacts with the analyte may be electrical, optical or thermal. It is then converted by means of a suitable transducer into a measurable electrical parameter – usually a current or voltage.

Biosensors can be classified by their biotransducer type. The most common types of biotransducers used in biosensors are: electrochemical biosensors, optical biosensors, electronic biosensors, piezoelectric biosensors, gravimetric biosensors, pyroelectric biosensors, magnetic biosensors, Electrochemical

Applications

Biosensor probes are becoming increasingly sophisticated, mainly owing to a combination of advances in two technological fields: microelectronics and biotechnology. Biosensors are highly valuable devices for measuring a wide spectrum of analytes including organic compounds, gases, ions and bacteria.

Bioremediation

Bioremediation is a process used to treat contaminated media, including water, soil and subsurface material, by altering environmental conditions to stimulate growth of microorganisms and degrade

the target pollutants. In many cases, bioremediation is less expensive and more sustainable than other remediation alternatives. Biological treatment is a similar approach used to treat wastes including wastewater, industrial waste and solid waste.

Most bioremediation processes involve oxidation-reduction reactions where either an electron acceptor (commonly oxygen) is added to stimulate oxidation of a reduced pollutant (e.g. hydrocarbons) or an electron donor (commonly an organic substrate) is added to reduce oxidized pollutants (nitrate, perchlorate, oxidized metals, chlorinated solvents, explosives and propellants).In both these approaches, additional nutrients, vitamins, minerals, and pH buffers may be added to optimize conditions for the microorganisms. In some cases, specialized microbial cultures are added (bioaugmentation) to further enhance biodegradation. Some examples of bioremediation related technologies are phytoremediation, mycoremediation, bioventing, bioleaching, landfarming, bioreactor, composting, bioaugmentation, rhizofiltration, and biostimulation.

Aerobic

Aerobic bioremediation is the most common form of oxidative bioremediation process where oxygen is provided as the electron acceptor for oxidation of petroleum, polyaromatic hydrocarbons (PAHs), phenols, and other reduced pollutants. Oxygen is generally the preferred electron acceptor because of the higher energy yield and because oxygen is required for some enzyme systems to initiate the degradation process.

Common approaches for providing oxygen above the water table include landfarming, composting and bioventing. During landfarming, contaminated soils, sediments, or sludges are incorporated into the soil surface and periodically turned over (tilled) using conventional agricultural equipment to aerate the mixture. Composting accelerates pollutant biodegradation by mixing the waste to be treated with a bulking agent, forming into piles, and periodically mixed to increase oxygen transfer. Bioventing is a process that increases the oxygen or air flow into the unsaturated zone of the soil which increases the rate of natural in situ degradation of the targeted hydrocarbon contaminant.

Anaerobic

Anaerobic bioremediation can be employed to treat a broad range of oxidized contaminants including chlorinated ethenes (PCE, TCE, DCE, VC), chlorinated ethanes (TCA, DCA), chloromethanes (CT, CF), chlorinated cyclic hydrocarbons, various energetics (e.g., perchlorate,[6] RDX, TNT), and nitrate.

Biogas

Biogas is the mixture of gases produced by the breakdown of organic matter in the absence of oxygen (anaerobically), primarily consisting of methane and carbon dioxide. Biogas can be produced from raw materials such as agricultural waste, manure, municipal waste, plant material, sewage, green waste or food waste. Biogas is a renewable energy source. In India, it is also known as "Gobar Gas".

Production

Biogas is produced by microorganisms, such as methanogens and sulfate-reducing bacteria, performing anaerobic respiration. Biogas can refer to gas produced naturally or industrially.

Natural

In soil, methane is produced in anaerobic zones environments by methanogens, but is mostly consumed in aerobic zones by methanotrophs. Methane emissions result when the balance favors methanogens. Wetland soils are the main natural source of methane. Other sources include oceans, forest soils, termites, and wild ruminants.

Industrial

The purpose of industrial biogas production is the collection of biomethane, usually for fuel. Industrial biogas is produced either;As landfill gas (LFG), which is produced by the breakdown of biodegradable waste inside a landfill due to chemical reactions and microbes, or As digested gas, produced inside an anaerobic digester.

Biogas plants

A biogas plant is the name often given to an anaerobic digester that treats farm wastes or energy crops. It can be produced using anaerobic digesters (air-tight tanks with different configurations). These plants can be fed with energy crops such as maize silage or biodegradable wastes including

sewage sludge and food waste. During the process, the micro-organisms transform biomass waste into biogas (mainly methane and carbon dioxide) and digestate. Higher quantities of biogas can be produced when the wastewater is co-digested with other residuals from the dairy industry, sugar industry, or brewery industry. For example, while mixing 90% of wastewater from beer factory with 10% cow whey, the production of biogas was increased by 2.5 times compared to the biogas produced by wastewater from the brewery only.

Transgenic animals

Transgenic animals are animals that have been deliberately bred for research and that contain elements of two different species - they are creatures that blur the barrier between species.

These animals are often deliberately created with genetic defects, and these defects may well cause the animal to have a bad quality of life. A mouse has been created, for example, that has been genetically modified to develop cancer.

Ethical issues of transgenic animals

Transgenic animals raise several particular moral issues (quite apart from any damage they might do to the environment):

Are animals that combine species an unethical alteration of the natural order of the universe?

Is it unethical to modify an animal's genetic make-up for a specific purpose, without knowing in advance if there will be any side-effects that will cause suffering to the animal?

Does 'creating' animals by genetic engineering amount to treat the animals entirely as commodities?

Is it unethical to create 'diseased' animals that are very likely to suffer?

Suffering may last for a long time in these animals as researchers want to conduct long-term investigations into the development of diseases

Religious views of transgenic animals

Against transgenic animals:

God laid down the structure of creation and any tampering with it is sinful.

Manipulating DNA is manipulating 'life itself' - and this is tampering with something that God did not intend humanity to meddle with.

In favour of transgenic animals:

As human beings have been given 'dominion' over the animals, they are entitled to tamper with them.

Palaeontology shows that the structure of creation has changed over time as some species became extinct and new ones came into being. They say that this shows that there is nothing fixed about the structure of creation.

Transgenic animals and religious food laws

Transgenic animals pose problems for religions that restrict the foods that their believers can eat, since they may produce animals that appear to be one species, but contain some elements of a forbidden species.

Creation of Transgenic Animals

The term transgenic animal refers to an animal in which there has been a deliberate modification of the genome - the material responsible for inherited characteristics - in contrast to spontaneous mutation). Foreign DNA is introduced into the animal, using recombinant DNA technology, and then must be transmitted through the germ line so that every cell, including germ cells, of the animal contain the same modified genetic material.

Methods of creation of transgenic animals:

For practical reasons, i.e., their small size and low cost of housing in comparison to that for larger vertebrates, their short generation time, and their fairly well defined genetics, mice have become the main species used in the field of transgenics.

The three principal methods used for the creation of transgenic animals are DNA microinjection, embryonic stem cell-mediated gene transfer and retrovirus-mediated gene transfer.

a) DNA microinjection.

This method involves the direct microinjection of a chosen gene construct (a single gene or a combination of genes) from another member of the same species or from a different species, into the pronucleus of a fertilized ovum. It is one of the first methods that proved to be effective in mammals The introduced DNA may lead to the over- or under-expression of certain genes or to the expression of genes entirely new to the animal species. The insertion of DNA is, however, a random process, and there is a high probability that the introduced gene will not insert itself into a site on the host DNA that will permit its expression. The manipulated fertilized ovum is transferred into the oviduct of a recipient female, or foster mother that has been induced to act as a recipient by mating with a vasectomized male.

A major advantage of this method is its applicability to a wide variety of species.

b) Embryonic stem cell-mediated gene transfer.

This method involves prior insertion of the desired DNA sequence by homologous recombination into an in vitro culture of embryonic stem (ES) cells. Stem cells are undifferentiated cells that have the potential to differentiate into any type of cell (somatic and germ cells) and therefore to give rise to a complete organism. These cells are then incorporated into an embryo at the blastocyst stage of development. The result is a chimeric animal. ES cell-mediated gene transfer is the method of choice for gene inactivation, the so-called knock-out method.

This technique is of particular importance for the study of the genetic control of developmental processes. This technique works particularly well in mice. It has the advantage of allowing precise targeting of defined mutations in the gene via homologous recombination.

c) Retrovirus-mediated gene transfer.

To increase the probability of expression, gene transfer is mediated by means of a carrier or vector, generally a virus or a plasmid. Retroviruses are commonly used as vectors to transfer genetic material into the cell, taking advantage of their ability to infect host cells in this way. Offspring derived from this method are chimeric, i.e., not all cells carry the retrovirus. Transmission of the transgene is possible only if the retrovirus integrates into some of the germ cells.

For any of these techniques the success rate in terms of live birth of animals containing the transgene is extremely low. Providing that the genetic manipulation does not lead to abortion, the

result is a first generation (F1) of animals that need to be tested for the expression of the transgene. Depending on the technique used, the F1 generation may result in chimeras. When the transgene has integrated into the germ cells, the so-called germ line chimeras are then inbred for 10 to 20 generations until homozygous transgenic animals are obtained and the transgene is present in every cell. At this stage embryos carrying the transgene can be frozen and stored for subsequent implantation.

A representative, but non-inclusive, list of purposes for which transgenic animals have been used indicates the wide ranging application of this biotechnology:

In medical research, transgenic animals are used to identify the functions of specific factors in complex homeostatic systems through over- or under-expression of a modified gene (the inserted transgene);

In toxicology: as responsive test animals (detection of toxicants);

In molecular biology, the analysis of the regulation of gene expression makes use of the evaluation of a specific genetic change at the level of the whole animal;

In the pharmaceutical industry, targeted production of pharmaceutical proteins, drug production and product efficacy testing;

In biotechnology: as producers of specific proteins;

Genetically engineered hormones to increase milk yield, meat production; genetic engineering of livestock and in aquaculture affecting modification of animal physiology and/or anatomy; cloning procedures to reproduce specific blood lines; and

Developing animals specially created for use in xenografting.

Questions

5 marks

1.Write a note on biosensor

2. Give a brief account on bioremediation

3. Comment on biogas

4. Write the ethical issues of transgenic animals

10 marks

1.Explain the process 0f sewage treatment

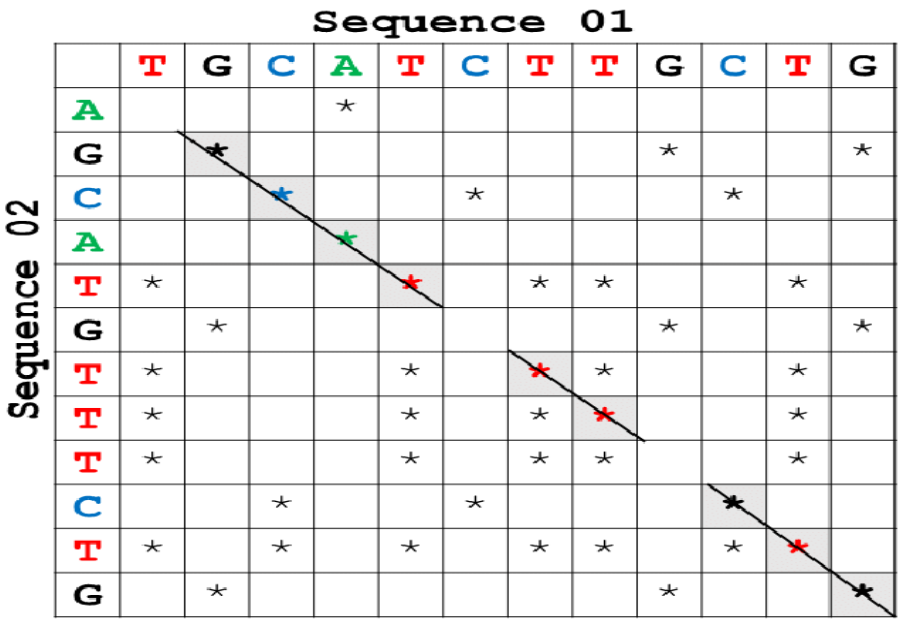
2. Describe the methods of creation of transgenic animals

UNIT-V

Dot matrix analysis

• A dot matrix analysis is a method for comparing two sequences to look for possible alignment (Gibbs and McIntyre 1970)

• One sequence (A) is listed across the top of the matrix and the other (B) is listed down the left side



• Starting from the first character in B, one moves across the page keeping in the first row and placing a dot in many column where the character in A is the same

• The process is continued until all possible comparisons between A and B are made

• Any region of similarity is revealed by a diagonal row of dots

• Isolated dots not on diagonal represent random matches Dot matrix analysis

• Detection of matching regions can be improved by filtering out random matches and this can be achieved by using a sliding window

• It means that instead of comparing a single sequence position more positions is compared at the same time and dot is printed only if a certain minimal number of matches occur

• Dot matrix analysis can also be used to find direct and inverted repeats within the sequences Sequence comparison with dot matrices

• Basic Method: For two sequences of lengths M and N, lay out an M by N grid (matrix) with one sequence across the top and one sequence down the left side. For each position in the grid, compare the sequence elements at the top (column) and to the left (row). If and only if they are the same, place a dot at that position.

Interpretation of dot matrices

• Regions of similarity appear as diagonal runs of dots

• Reverse diagonals (perpendicular to diagonal) indicate inversions

• Reverse diagonals crossing diagonals (Xs) indicate palindromes

Uses for dot matrices

• Can use dot matrices to align two proteins or two nucleic acid sequences

• Can use to find amino acid repeats within a protein by comparing a protein sequence to itself. Repeats appear as a set of diagonal runs stacked vertically

Can use to find self base-pairing of an RNA (e.g., tRNA) by comparing a sequence to itself complemented and reversed

• Excellent approach for finding sequence transpositions

Difference between global and local sequence alignment

Global Sequence Alignment	Local Sequence Alignment
In global alignment, an attempt is made to align the entire sequence (end to end alignment)	Finds local regions with the highest level of similarity between the two sequences.
A global alignment contains all letters from both the query and target sequences	A local alignment aligns a substring of the query sequence to a substring of the target sequence.
If two sequences have approximately the same length and are quite similar, they are suitable for global alignment.	Any two sequences can be locally aligned as local alignment finds stretches of sequences with high level of matches without considering the alignment of rest of the sequence regions.
Suitable for aligning two closely related sequences.	Suitable for aligning more divergent sequences or distantly related sequences.
Global alignments are usually done for comparing homologous genes like comparing two genes with same function (in human vs. mouse) or comparing two proteins with similar function.	Used for finding out conserved patterns in DNA sequences or conserved domains or motifs in two proteins.
A general global alignment technique is the Needleman–Wunsch algorithm.	A general local alignment method is Smith–Waterman algorithm.
Examples of Global alignment tools: • > EMBOSS Needle • > Needleman-Wunsch Global Align Nucleotide Sequences (Specialized BLAST)	**Examples of Local alignment tools:** • > BLAST • > EMBOSS Water • > LALIGN

Multiple sequence alignment

A multiple sequence alignment (MSA) is a sequence alignment of three or more biological sequences, generally protein, DNA, or RNA. In many cases, the input set of query sequences are assumed to have an evolutionary relationship by which they share a linkage and are descended from a common ancestor. From the resulting MSA, sequence homology can be inferred and phylogenetic analysis can be conducted to assess the sequences' shared evolutionary origins. Alignment methods

Progressive alignment construction

The most widely used approach to multiple sequence alignments uses a heuristic search known as progressive technique (also known as the hierarchical or tree method) developed by Da-Fei Feng and Doolittle in 1987. Progressive alignment builds up a final MSA by combining pairwise alignments beginning with the most similar pair and progressing to the most distantly related. All progressive alignment methods require two stages: a first stage in which the relationships between the sequences are represented as a tree, called a guide tree, and a second step in which the MSA is built by adding the sequences sequentially to the growing MSA according to the guide tree. The initial guide tree is determined by an efficient clustering method such as neighbor-joining or UPGMA, and may use distances based on the number of identical two-letter sub-sequences (as in FASTA rather than a dynamic programming alignment).

Progressive alignments are not guaranteed to be globally optimal. The primary problem is that when errors are made at any stage in growing the MSA, these errors are then propagated through to the final result. Performance is also particularly bad when all of the sequences in the set are rather distantly related. Most modern progressive methods modify their scoring function with a secondary weighting function that assigns scaling factors to individual members of the query set in a nonlinear fashion based on their phylogenetic distance from their nearest neighbors. This corrects for non-random selection of the sequences given to the alignment program.

Iterative methods

A set of methods to produce MSAs while reducing the errors inherent in progressive methods are classified as "iterative" because they work similarly to progressive methods but repeatedly realign the initial sequences as well as adding new sequences to the growing MSA. One reason progressive methods are so strongly dependent on a high-quality initial alignment is the fact that these alignments are always incorporated into the final result — that is, once a sequence has been aligned into the MSA, its alignment is not considered further. This approximation improves efficiency at the cost of accuracy. By contrast, iterative methods can return to previously calculated pairwise alignments or sub-MSAs incorporating subsets of the query sequence as a means of optimizing a general objective function such as finding a high-quality alignment score.

Consensus methods

Consensus methods attempt to find the optimal multiple sequence alignment given multiple different alignments of the same set of sequences. There are two commonly used consensus methods, M-COFFEE and MergeAlign. M-COFFEE uses multiple sequence alignments generated by seven different methods to generate consensus alignments. MergeAlign is capable of generating consensus alignments from any number of input alignments generated using different models of sequence evolution or different methods of multiple sequence alignment. The default option for MergeAlign is to infer a consensus alignment using alignments generated using 91 different models of protein sequence evolution.

Nucleotide sequences databases

There are three chief databases that store and make available raw nucleic acid sequences to the public and researchers alike: GenBank, EMBL, DDBJ.

They are referred to as the primary nucleotide sequence databases since they are the repository of all nucleic acid sequences.

GenBank is physically located in the USA and is accessible through the NCBI portal over the intern.

EMBL (European Molecular Biology Laboratory) is in UK and DDJB (DNA databank of Japan) is in Japan.

All three accept nucleotide sequence submissions and then exchange new and updated data on a daily basis to achieve optimal synchronization between them.

These three databases are primary databases, as they house original sequence data.

They collaborate with Sequence Read Archive (SRA), which archives raw reads from high-throughput sequencing instruments.

a. GenBank

The GenBank sequence database is open access, annotated collection of all publicly available nucleotide sequences and their protein translations. This database is produced and maintained by the National Center for Biotechnology Information (NCBI) as part of the International Nucleotide Sequence Database Collaboration (INSDC). receive sequences produced in laboratories

throughout the world from more than 100,000 distinct organisms. GenBank has become an important database for research in biological fields and has grown in recent years at an exponential rate by doubling roughly every 18 months.

b. EMBL (European Molecular Biology Laboratory)

The European Molecular Biology Laboratory (EMBL) Nucleotide Sequence Database is a comprehensive collection of primary nucleotide sequences maintained at the European Bioinformatics Institute (EBI). Data are received from genome sequencing centers, individual scientists and patent offices.

c. DDBJ (DNA databank of Japan)

It is located at the National Institute of Genetics (NIG) in the Shizuoka prefecture of Japan. It is the only nucleotide sequence data bank in Asia. Although DDBJ mainly receives its data from Japanese researchers, it can accept data from contributors from any other country.

Protein Databases- Types and Importance

- As biology has increasingly turned into a data-rich science, the need for storing and communicating large datasets has grown tremendously.
- The obvious examples are the nucleotide sequences, the protein sequences, and the 3D structural data produced by X-ray crystallography and macromolecular NMR.
- The biological information of proteins is available as sequences and structures. Sequences are represented in a single dimension whereas the structure contains the three-dimensional data of sequences.
- A biological database is a collection of data that is organized so that its contents can easily be accessed, managed, and updated.
- A protein database is one or more datasets about proteins, which could include a protein's amino acid sequence, conformation, structure, and features such as active sites.
- Protein databases are compiled by the translation of DNA sequences from different gene databases and include structural information. They are an important resource because proteins mediate most biological functions.

There is a number of primary protein sequence databases and each requires some specific consideration.

a. Protein Information Resource (PIR) – Protein Sequence Database (PIR-PSD):

- The PIR-PSD is a collaborative endeavor between the PIR, the MIPS (Munich Information Centre for Protein Sequences, Germany) and the JIPID (Japan International Protein Information Database, Japan).
- The PIR-PSD is now a comprehensive, non-redundant, expertly annotated, object-relational DBMS.
- A unique characteristic of the PIR-PSD is its classification of protein sequences based on the superfamily concept.
- The sequence in PIR-PSD is also classified based on homology domain and sequence motifs.
- Homology domains may correspond to evolutionary building blocks, while sequence motifs represent functional sites or conserved regions.
- The classification approach allows a more complete understanding of sequence function-structure relationship.

b. SWISS-PROT

- The other well known and extensively used protein database is SWISS-PROT. Like the PIR-PSD, this curated proteins sequence database also provides a high level of annotation.
- The data in each entry can be considered separately as core data and annotation.
- The core data consists of the sequences entered in common single letter amino acid code, and the related references and bibliography. The taxonomy of the organism from which the sequence was obtained also forms part of this core information.
- The annotation contains information on the function or functions of the protein, post-translational modification such as phosphorylation, acetylation, etc., functional and structural domains and sites, such as calcium binding regions, ATP-binding sites, zinc

fingers, etc., known secondary structural features as for examples alpha helix, beta sheet, etc., the quaternary structure of the protein, similarities to other protein if any, and diseases that may arise due to different authors publishing different sequences for the same protein, or due to mutations in different strains of an described as part of the annotation.

TrEMBL (for Translated EMBL) is a computer-annotated protein sequence database that is released as a supplement to SWISS-PROT. It contains the translation of all coding sequences present in the EMBL Nucleotide database, which have not been fully annotated. Thus it may contain the sequence of proteins that are never expressed and never actually identified in the organisms.

c. Protein Databank (PDB):

- PDB is a primary protein structure database. It is a crystallographic database for the three-dimensional structure of large biological molecules, such as proteins.
- In spite of the name, PDB archive the three-dimensional structures of not only proteins but also all biologically important molecules, such as nucleic acid fragments, RNA molecules, large peptides such as antibiotic gramicidin and complexes of protein and nucleic acids.
- The database holds data derived from mainly three sources: Structure determined by X-ray crystallography, NMR experiments, and molecular modeling'

Applications of Bioinformatics

Genomics

The study of genes and their expression is called as Genomics. This field generates a vast amount of data from gene sequences, their interrelation and functions. To manage this vast enormous data, bioinformatics plays a very important role. With the complete genome sequences for an increasing number of organisms, bioinformatics is beginning to provide both conceptual bases and practical methods for detecting systemic functional behaviours of the cell and the organism . Bioinformatics plays a vital role in the areas of structural genomics, functional genomics and nutritional genomics.

Proteomics

The study of protein structure, function, and interactions produced by a particular cell, tissue, or organism is called as proteomics. It deals with techniques of genetics, biochemistry and molecular biology. The vast data can be managed and access easily by using bioinformatics tools, software and databases. Till now, many algorithms in the field of proteomics viz. image analysis of 2D gels, peptide mass fingerprinting and peptide fragmentation fingerprinting have been developed .

Transcriptomics

The study of sets of all messenger RNA molecules in the cell is called as transcriptomics. Analysis of vast data is done by numerous software packages. In this way, bioinformatics is used for transcriptome analysis where mRNA expression levels can be determined . RNA sequencing (RNAseq) also has been included under transcriptomics. It is carried out using next generation sequencing to determine the presence and quantity of RNA in a sample at a given time. It is used to analyze the continuously changing cellular transcriptome.

Cheminformatics

Cheminformatics (chemical informatics) focuses on storing, indexing, searching, retrieving, and applying information about chemical compounds. It involves organization of chemical data in a logical form to facilitate the retrieval of chemical properties, structures and their relationships. Using bioinformatics, it is possible through computer algorithm to identify and structurally modify a natural product, to design a compound with the desired properties and to assess its therapeutic effects, theoretically. Cheminformatics analysis includes analyses such as similarity searching, clustering, QSAR modeling, virtual screening, etc. .

Drug Discovery

Bioinformatics is playing an increasingly important role in nearly all aspects of drug discovery, drug assessment and drug development. The increasing pressure to generate more and more drugs in a short period of time with low risk has resulted in remarkable interest in bioinformatics. In fact, now there is an existence of new separate field known as computer-aided drug design (CADD). Bioinformatics provides a huge support to overcome the cost and time context in various ways. It provides wide range of drug-related databases and softwares which can be used for various purposes related to drug designing and development process .

Evolutionary Studies/Phylogenetics

The study of evolutionary relationship among individuals or group of organisms is defined as phylogenetics. Taxonomists find the evolutionary relationship using various anatomical methods that takes too much time. Using Bioinformatics, phylogenetic trees are constructed based on the sequence alignment using various methods. Various algorithmic methods are developed for the construction of phylogenetic tree that are used depending on the various evolutionary lineages .

Crop Improvement

Sustainable agricultural production is an urgent issue in response to global climate change and population increase. Innovations in omics based research improve the plant based research. The integrated 'omics' strategies clarify the molecular system of the plant which are used to improve the plant productivity. Genomics strategy, especially comparative genomics helps in understanding the genes and their functions, and also the biological properties of each species. Bioinformatics databases are also used in designing new techniques and experiments for increased plant production .

Veterinary Science

Food production from livestock can meet demand of human population for food. For better bio-economy, there is a need of efficient animal production and reproduction. This is achieved with better understanding of livestock species. Current and new methods in livestock species using data from experimental or field studies with bioinformatics are helping in understanding the systems genetics of complex traits and provide biologically meaningful and accurate predictions. Finally, almost all of the next generations-omics tools and methods that are used in other fields of biological sciences, can also be used in veterinary sciences .

Forensic Science

Forensic science includes the study regarding identification and relatedness of individuals. It is inherently interdisciplinary with bioinformatics as both are dependent on computer science and statistics. This field is based on the molecular data and many databases are being developed to store the DNA profiles of known offenders. This field is being pushed due to technological and

statistical advances in microarray, Bayesian networks, machine learning algorithms, TFT biosensors and others. This provides the effective way of evidence organization and inference .

Biodefense

Biodefense includes measures to restore biosecurity to a group of organisms who are subjected to biological threats or infectious diseases (in context of bio-war or bioterrorism).

Waste Cleanup

Today, the major concern all over the Globe is environmental pollutants. The main concern of the environmentalists is waste generated from the industries. Genomic and bioinformatics data provide a wealth of information that would be greatly enhanced by structural characterisation of some protein. Bioinformatics provides data of microbial genomics, proteomics, systems biology, computational biology, and bioinformatics tools for understanding of the mechanisms of biodegradative pathways .

Climate Change Studies

Another Global concern is the Climate change because of loss of sea ice, accelerated sea level rise and longer and more intense heat waves. To solve this issue, bioinformatics may help by way of sequencing microbial genome which can reduce levels of carbon dioxide and other greenhouse gases. This plays an important role in stabilizing the global climate change.

Bioenergy/Biofuels

Biofuels offer great promise in contributing to the growing global demand for alternative sources of renewable energy. Bioinformatics is important in understanding and analysis of biofuel producing pathways. Recent progress in algal genomics, in conjunction with other “omics” approaches, has accelerated the ability to identify metabolic pathways and genes that are potential targets in the development of genetically engineered micro-algal strains with optimum lipid content .

Questions:

5 marks

1. Give a brief account on Dot matrix analysis

2. Differentiate local and global sequence alignment
3. Write notes on EMBL, DDBJ and Genbank
4. Write a note on multiple sequence alignment

10 marks

1. Describe about protein sequence databases
2. Give a detailed account on applications of bioinformatics