

REGULATION OF GENE EXPRESSION & LAC OPERON

1. Regulation of Gene Expression

Gene Regulation is the biological process in which genes are regulated to control various cellular metabolic activities. The process by which protein synthesis is initiated or stopped under genetic control is called the regulation of gene expression.

In eukaryotes, gene expression can be regulated at several levels during protein synthesis:

- **(i) Transcriptional level:** Formation of primary transcript.
- **(ii) Processing level:** Processing of primary transcript (mRNA splicing).
- **(iii) Transport level:** Transport of mRNA from nucleus to cytoplasm.
- **(iv) Translational level:** Translation of mRNA into protein.

2. The Lac Operon Concept

Jacob and Monod (1961) proposed the operon concept to explain gene regulation. They studied the hydrolysis of lactose in *Escherichia coli* (*E. coli*).

According to this concept, a group of structural genes is regulated as a functional unit called an "**operon**". The operon that regulates the hydrolysis of the disaccharide lactose is known as the **Lac Operon**.

Components of the Lac Operon:

1. **Structural Genes:** Genes that code for specific functional enzymes involved in lactose degradation (**Z, Y, and A**).
2. **Operator Gene:** Controls the synthesis of mRNA over structural genes. It is switched *OFF* when bound by the repressor protein and switched *ON* when the repressor is removed by an inducer.
3. **Promoter Gene:** DNA region where RNA polymerase binds. In the presence of RNA polymerase, it passes signals to structural genes for transcription when operator is open.
4. **Regulatory Gene (i-gene):** Codes for a protein substance called the **repressor protein**.
 - **Positive Regulation:** When an inducer binds to the repressor, it undergoes a conformational change preventing it from binding to the operator. RNA polymerase binds and transcription starts.
 - **Negative Regulation:** When repressor binds to the operator gene, RNA polymerase is blocked and transcription stops.

Inducer: Lactose acts as the natural inducer in *E. coli*.

Structural Genes of Lac Operon & Their Functions

In *E. coli*, three structural genes are coordinated together for lactose metabolism:

Structural Gene	Enzyme Coded	Biological Function
lac Z	β-Galactosidase	Hydrolyzes disaccharide lactose into glucose and galactose.
lac Y	Permease	Increases membrane permeability to allow lactose entry into the cell.
lac A	Transacetylase	Transfers an acetyl group to β-galactosides for metabolic processing.

All three structural genes are strictly required for the complete catabolic breakdown of lactose in bacterial cells.

3. Human Genome Project (HGP)

With the advancements in genetic engineering techniques, DNA molecules can now be isolated, cloned, and sequenced. Using these modern techniques, the precise sequence of nucleotide bases in human DNA can be determined.

In **1990**, an ambitious international scientific research mega-project was launched to determine the complete DNA sequence of the human genome, titled the **Human Genome Project (HGP)**.

Scope, Cost & Bioinformatics:

- The human genome consists of approximately **3×10^9 base pairs** (3 billion bp).
- If estimated at US \$3 per base pair, the entire project involved an estimated expenditure of **9 billion US Dollars**.
- Managing and analyzing this vast quantity of genomic data was made possible by establishing a new branch of biology called **Bioinformatics** (storing, managing, and analyzing biological data using computational tools).

Goals of Human Genome Project:

- To determine the sequence of all 3 billion chemical base pairs in human DNA.
- To identify and map all approximately **20,000 to 25,000 genes** present in human genome.
- To store all accumulated sequence information in public databases.

Goals of HGP (Continued):

- To improve computational algorithms and tools for genomic data analysis.
- To identify specific genes responsible for causing genetic disorders.
- To transfer related genomic technologies to other sectors, including industrial R&D.
- To address the **ELSI** (Ethical, Legal, and Social Issues) that may arise from genomic research.

Project History: HGP was officially launched in **October 1990** jointly by the US Department of Energy (DOE) and the National Institutes of Health (NIH). The core human genome sequencing project was successfully completed in **2003**.

4. Methodologies of HGP

Two main approach strategies were utilized for genome sequencing:

1. **Expressed Sequence Tags (ESTs):** Focused on identifying only those genes that are expressed as RNA transcripts.
2. **Sequence Annotation:** Blind sequencing of the entire genome (both coding and non-coding sequences) and later assigning functional regions.

Sequencing Protocol & Cloning Vectors:

- Total cellular DNA was extracted and cleaved into smaller random fragments.
- Fragments were cloned into specialized host organisms using artificial vector systems:
 - **BAC:** Bacterial Artificial Chromosomes (for bacterial hosts)
 - **YAC:** Yeast Artificial Chromosomes (for yeast hosts)
- Presently, **Polymerase Chain Reaction (PCR)** is used for rapid cloning of DNA fragments.
- DNA fragments were sequenced using automated DNA sequencers based on the principle developed by Nobel Laureate **Frederick Sanger**.
- Sequences were arranged based on overlapping regions using supercomputer programs.
- All 22 autosomes and sex chromosomes (X and Y) have been sequenced. Chromosome 1 sequence was completed in **May 2006**.

5. Salient Features of Human Genome

- The human genome contains **3,164.7 million base pairs** (3.16 billion bp).
- The total number of genes in the human genome is estimated at **30,000**.
- The average gene consists of **3,000 base pairs**. The largest known gene is **Dystrophin** (contains 2.4 million / 24 lakh base pairs).
- Single Nucleotide Polymorphisms (**SNPs**) occur at **1.4 million locations**, assisting in locating disease-associated chromosomal regions.
- Functions of over **50%** of discovered genes remain unknown.
- **Chromosome 1** contains the highest number of genes (**2,968 genes**), while the **Y chromosome** contains the fewest (**231 genes**).
- **Less than 2%** of the human genome codes for functional proteins.
- **99.9%** of nucleotide bases are identical in all humans; only **0.1%** (approx. 3.2 million bp) accounts for individual variations.
- Repetitive nucleotide sequences constitute a major portion of the human genome.

6. Applications of HGP

- Provides valuable insights to diagnose, treat, and prevent inherited genetic diseases.
- Establishes unique genetic identity profiles for individuals.
- Enables targeted gene therapy treatments for complex hereditary disorders.
- Enhances fundamental understanding of human genetics, development, and evolution.

7. DNA Fingerprinting

DNA Fingerprinting is a molecular technique used to identify individuals based on their unique DNA profiles. It was developed in 1984 by British geneticist **Sir Alec Jeffreys** (introduced in India in 1989 by Dr. Lalji Singh).

Human DNA contains non-coding region sequences of 10–100 base pairs repeated in tandem, called **Short Tandem Repeats (STRs)** or repetitive DNA, which are inherited directly from parents and provide 0.1% individual uniqueness.

Satellite DNA & VNTR Probes

Repetitive DNA is separated from bulk genomic DNA as satellite peaks during density gradient centrifugation. It is categorized based on sequence length and unit repeat frequency:

Parameter	Mini-Satellite (VNTR)	Micro-Satellite (STR)
Sequence Length	15 to 50 base pairs	Less than 10 base pairs
Tandem Repeats	10 to 100 repeats	Up to 1,000 repeats

Sir Alec Jeffreys utilized radio-labeled mini-satellite DNA as a probe showing high degrees of polymorphism, termed **Variable Number of Tandem Repeats (VNTRs)**, forming the foundation of DNA fingerprinting.

Biological Samples: Blood stains, hair follicles, skin, bone, saliva, semen, and urine (since genomic DNA is identical in all somatic cells).

8. Steps in DNA Fingerprinting Technique

- 1. DNA Isolation:** Extraction and purification of genomic DNA from biological samples.
- 2. Enzymatic Digestion:** Cleavage of DNA into smaller fragments using restriction endonucleases.
- 3. Electrophoresis:** Separation of DNA fragments based on molecular size using Agarose Gel Electrophoresis.
- 4. Southern Blotting:** Transfer of single-stranded DNA fragments onto a nylon/nitrocellulose membrane (fixed at 80°C for 3 hours).
- 5. Hybridization:** Probing the membrane with radiolabeled VNTR probes to bind complementary sequence repeats.
- 6. Autoradiography:** Exposure to X-ray film to reveal distinct dark bands corresponding to individual-specific DNA band patterns.

9. Applications of DNA Fingerprinting

- **Forensics:** Solving crimes such as murder, assault, and sexual assault by matching crime scene DNA.
- **Paternity Disputes:** Settling parentage and inheritance claims in legal cases.
- **Medical Diagnostics:** Detecting genetic disorders like Sickle Cell Anemia, Hemophilia, and Thalassemia.
- **Agriculture & Breeding:** Verifying plant genetic traits and selective livestock breeding.