UNIT 15 REGULATION OF GENE EXPRESSION AND DEVELOPMENT IN EUKARYOTES

Structure

15.1 INTRODUCTION

In Unit 14 you have seen that in prokaryotes gene expression is commonly regulated by an operon consisting of structural genes and adjacent controlling sites. The discovery of operons in prokaryotes initiated search for similar regulatory systems in eukaryotes. But so far no operons have been found in eukaryotes. Neverlheless, there are some commonalities between the gene regulation systems of lower eukaryotes and prokaryotes. It implies that gene regulation in eukaryotes must occur by ways other than by operons. In this unit we shall discuss some salient aspects of regulatory systems in eukaryotes.

The eukaryotic gene regulation can be classified into two categories: short-tent1 and long-term regulation. The short-term regulation is just like the regulation by operons as seen in bacteria and viruses. It involves regulatory events in which gene sets are quickly turned on or off in response to changes in environmental or physiological conditions. On the other hand, the long-term regulation involves those events tbat are required for an organism to develop and differentiate.

Objectives

After studying this unit you should be able to:

- compare the genetic organisation of the eukaryotes and prokaryotes, and relate these features to regulation of gene expression in eukaryotes (Sections 15.2 to 15.5);
- explain short-term gene regulation pointing out at which level it operates (Section 15.3);
- describe the key roles of gene regulation in development and differentiation (Section 15.4); and
- explain the most accepted model for regulation of gene expression in eukaryotes (Section 15.5).

15.2 GENETIC ORGANISATION OF PROKARYOTES VS EUKARYOTES

The prokaryotes and eukaryotes differfrom each other with regard to transcription, translation and spatial organisation of DNA. Given below are some of the differential features of the two, that are relevant to regulation.

1) In an eukaryote, usually one type of the polypeptide chain is translated from a completed mRNA molecule. Thus, polycistronic mRNA of the type seen in prokaryotes is not present in eukaryotes.

2) The eukaryotic DNA is associated with histones forming chromatin, and to numerous nonhistone proteins. Only a small fraction of the DNA is bare. In bacteria, most of their DNA is free, and sometimes proteins are present in the folded chromosome.

3) A significant fraction of the DNA of eukaryotes consists of moderately or highly repetitive nucleotide sequences. Some of these sequences are repeated in tandem copies, but others are not. Bacteria contain camparatively lesser amount of repetitive DNA, most of which is confined to the rRNA and tRNA genes, and a few specific sequences such as the insertion sequences.

4) A large fraction of the eukaryotic DNA is untranslated, as most of the nucleotide sequences do not code for proteins.

5) Some eukaryotic genes are expressed and regulated by certain mechanisms for rearranging some DNA segments in a controlled way and for increasing the number of specific genes when needed.

6) The eukaryotic genes can be split into exons and introns. The introns must be removed during processing of the RNA transcript before translation begins,

7) The mRNA, in eukaryotes is synthesised in their nucleus from where it is transported across the nuclear envelope to the cytoplasm where it is utilised. Owing to the absence of nucleus in bacteria, such a compartmentalisation is not seen.

Now that you have studied the salient differences between the genetic organisation in prokaryotes and eukaryotes, you should note in the following sections how some of these features are incorporated into particular modes of eukaryotic gene regulation.

15.3 THE SHORT-TERM GENE REGULATION

In this section, we shall take up some examples of short term regulation of gene expression in eukaryotes. Short-term regulation as you know involves rapid responses to changes in environmental or physiological conditions. *It operates at the transcriptional level.* As you study the following examples, keep in mind the short-term regulation of gene expression in bacteria and compare the molecularevents involved.

153.1 Quinic Acid Metabolic Genes in *Neurospora crassa*

In *Neurospora,* the genes involved in metabolism of quinic acid (qa) as a source of carbon, constitute a simple and genetically well-characterised system for studying gene regulation in eukaryotes. This system comprises a cluster of five structural genes and two regulatory genes arranged as shown in Figure 15.1.

These five structural genes are *inducible* and their products are synthesised *coordinately* thus showing parallels with some bacterial operons. Two regulatory genes are present immediately adjacent to the structural genes. The regulatory gene *qa-IF* codes for a protein — *activator* and the second one *qu-IS* codes for a *repressor* protein. The activator protein (gene qa-1F) is needed both for the transcription of its own mRNA and for the transcription of all the structural genes except qa-x. The gene qa-1S on the other hand, codes for a'protcin that has a repressor (negative) function. If quinic acid is absent, the repressor protein interacts with qa, the effector molecule, and blocks transcription of qa-1F. Transcription of the qa-x structural gene seems to be controlled mainly by qa-1S and to a much lesser extent by qa-1F. Each of the four structural genes under qa-1F control is transcribed from two to four qa-1F independent promoters. Unlike bacterial operons, this gene cluster gives no evidence for operator-like controlling regions, and so it is not an operon. This regulatory system, however, functions to maintain cellular levels of particular chemicals, enabling the organism to adjust to changing physiological environment.

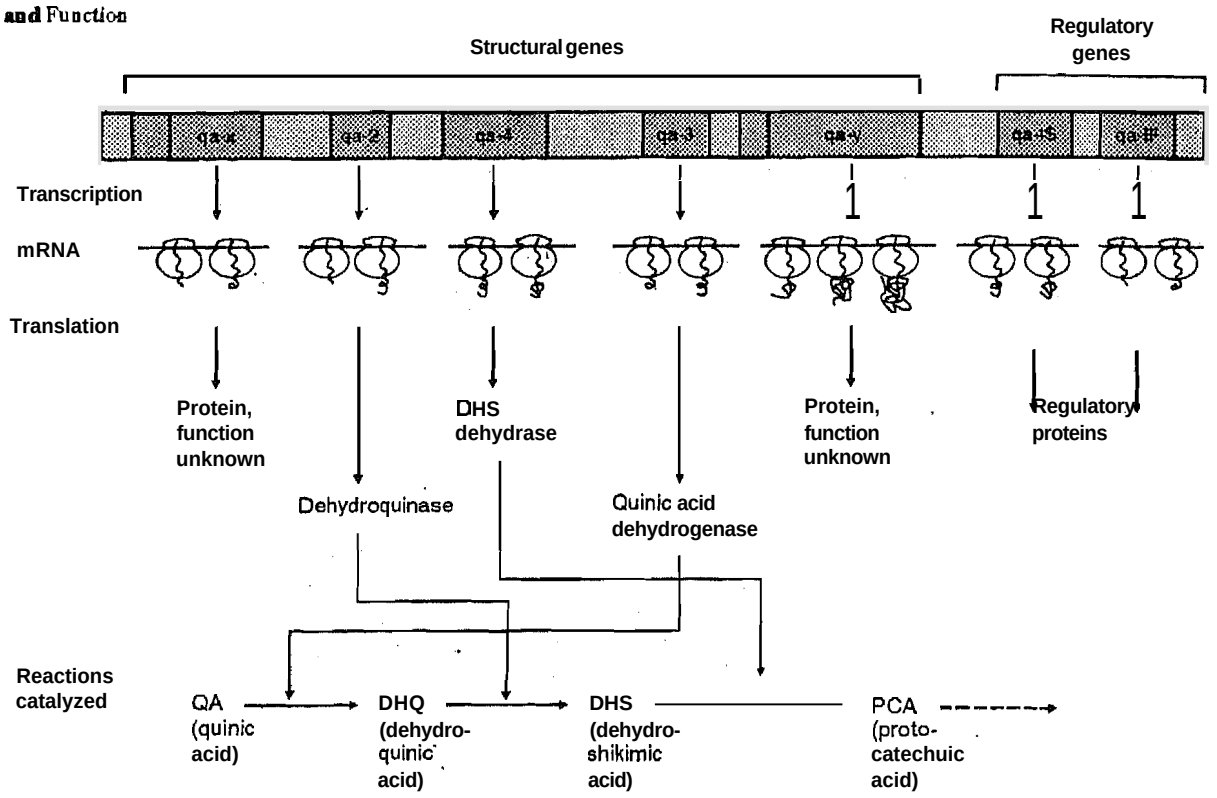
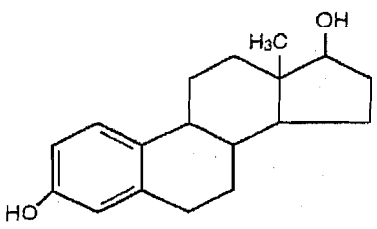
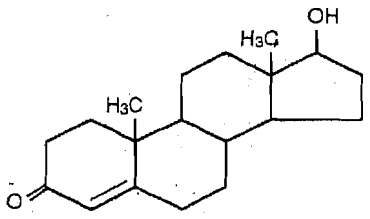


Fig.15.1: Organisation **of quinic acid** (qa) **catabolising** genes in *Neurospora crassa.*



Estrogen



Testosterone

. **Fig. 15.2 : Chemical structures of the steroid** sex hormones estrogen **(female)** and **testosterone** (male). **The steroid hormones** are **relatively** ma. **molecules with molecular weights** around 300. They **have a four-ring structure** and are synthesised from **cholestrol. The various steroid hormones have different** side chains **and different banding patterns within the rings.** These **differences** permit them **to** be **recognised** by **different receptor proteins** that are present in **the** cytoplasm of various target cells.

15.3.2 Hormonal Regulation

Hormones are also known to **regulate** transcription in higher eukaryotcs. They are small molecules or polypeptides that are carried from hormone producing cells to the target cells. The steroid Bormones constitute a prominent class amongst all the hormones. They are made up of small **molecules** synthesised **from** cholestrol. The principal sex hormones are steroids (see Fig. 15.2). Many of the steroid hormones act by turning on the transcriptionof specific sets of genes. It, therefore, means that a hormone that regulates transcription must somehow signal the DNA. The mechanisms by which this occurs is outlined in Figure 15.3. **A** steroid **hormone penetrates** a target cell through diffussion because steroids are hydrophobic (nonpolar) molecules and they pass freely through the cell membrane and the nuclear envelope. The nuclei of target cells contain specific receptor proteins for the steroid hormones. **These** receptor proteins **form** a complex with the hormone. During this process the receptor protein undergoes modifications in three-dimensional shape. This enables the hormone-receptor complex to bind with particular sequences in the DNA and stimulate transcription. The non-target cells do not contain the specific receptor proteins and so are unaffected by a particular **hormone.**

A well studied example of induction of transcription by a **hormone** is the stimulation of synthesis of **ovalbumin** in the chicken oviduct by the steroid sex hormone estrogen (Fig.15.2). When chickens are injected with estrogen, oviduct tissue responds by synthesising ovalbumin mRNA. **This** synthesis continues as long as estrogen is administered.

Once the hormone is withdrawn, the rate of synthesis decreases. Both before the injection of the **hormone** and sixty hours 'after withdrawal, no **ovalbumin** mRNA is detectable. When **estrogen** is given to chickens, only the oviduct synthesises mRNA because other tissues lack the hormone receptors.

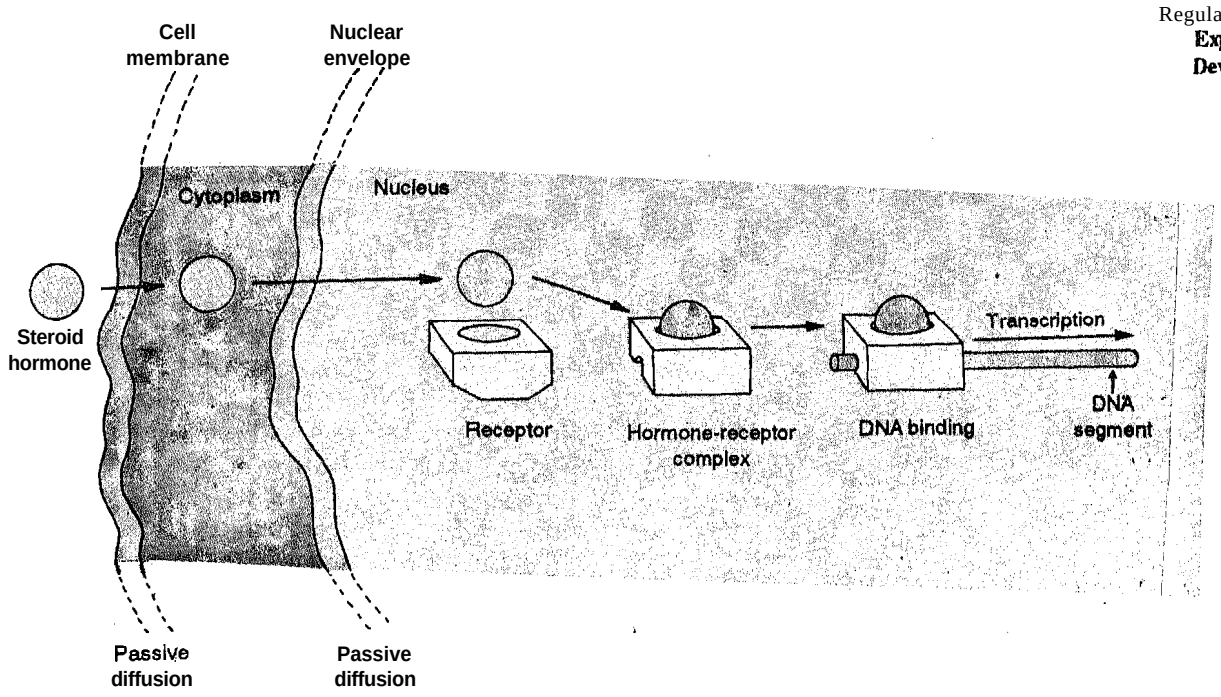


Fig.15.3: **Diagrammatic representation of the action of steroid hormone. The steroid hormone reaches a DNA molecule and** triggers **transcription by binding** with a **receptor in** the **nucleus** to form n **transcriptional activator. The steroid hormone enters** (he **cytoplasm and** nucleus hy **diffusion.**

Besides **animals,** there are several **examples** of hormonal regulation of **gene** expression in plants too. **Plant hormones** are categorised into the following **five main** classes: ethylene gas, abscisic acid, auxins, **cytokinins and** gibberellins (Fig. 15.4). Amongst them, the role of **gibberellins** has been well-studied. These **hormones** have

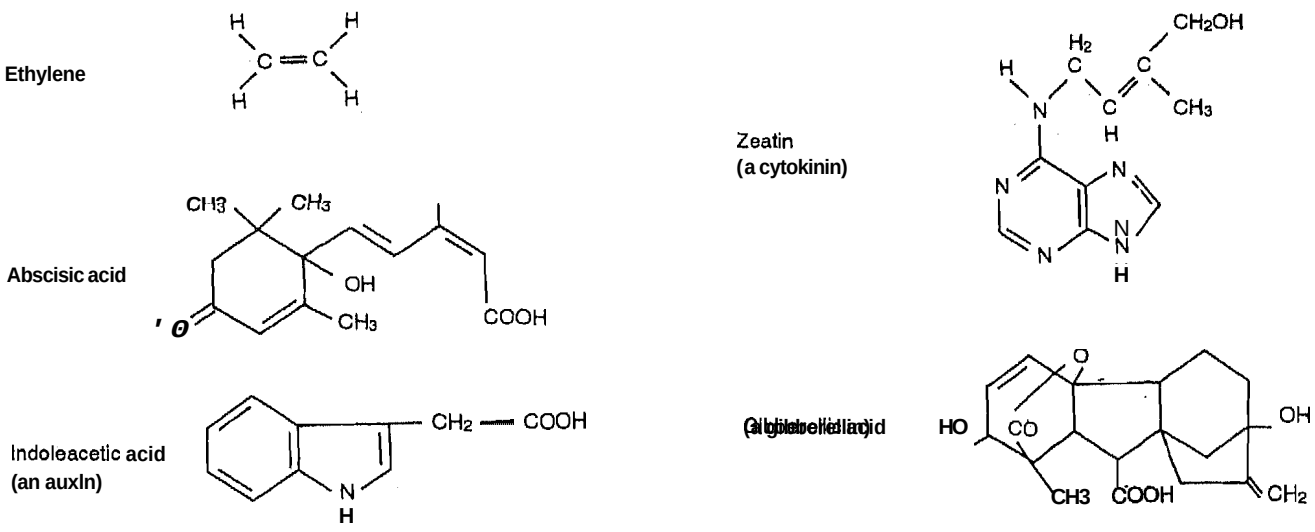


Fig.15.4: Chemical structure of five plant hormones.

actions similar to those of steroid **hormones. Gibberellins** *stimulate transcription* and thereby **increase** the production of specific proteins. These proteins bring out **profound changes** in **plant** cell form during its differentiation. The use of **gibberellins** in making **dwarf plants** grow tall is well known. Apart from **that, gibberellins** are also known to stimulate germination of seeds. It's **role** in **germination** of the wheat seeds is **diagrammatically** shown in Figure 15.5. Have a thorough look at this figure.

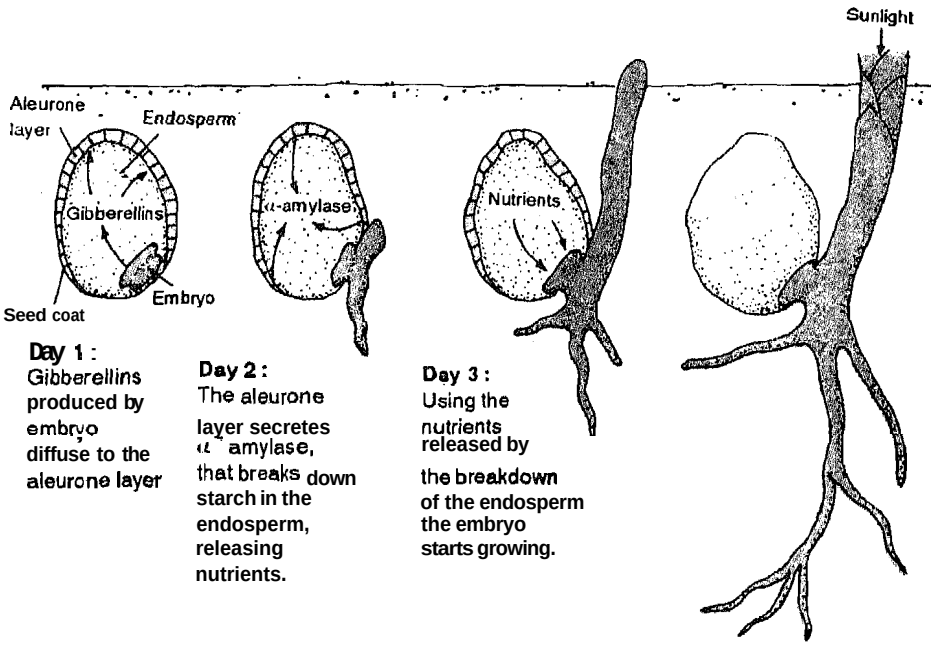


Fig.15.5: The **effect of gibberellins** on **the germination of wheat.**

During ***germination, the*** wheat seed produces **gibberellins** that diffuse to the aleurone layer, which is the outermost layer of the **endosperm.** As a result of the effect of the **gibberellins** on gene expression, and particularly on the genes for the **enzyme ∝-amylase**, the **aleurone** layer synthesises and **secretes ∝-amylase,** which breaks **down the** endosperm, releasing nutrients for the embryo growth. **Eventhough** the action of **gibberellins** is well known, the molecular details of their action are still not clear.

SAQ 1

Which of **the** following is not a true statement?

i) In eukaryotes too **polycistronic** mRNA as seen in prokaryotes is present.

ii) The DNA of eukaryotes is unique as it is associated with various **histone** and **non-histone** proteins.

iii) A large portion of **eukaryotic** DNA is **untranslated.**

iv) The genes of **prokaryotes** are composed of **exons** and **introns.**

v) The site of synthesis of mRNA **in eukaryotes** is the nucleus, from where it is transported to **the** cytoplasm where it is utilised.

vi) The **quinic** acid metabolic **genes** of *Neurospora* being inducible, and their products being synthesised coordinately, show parallels with the bacterial **operons.**

vii) **The** expression **of** quinic acid **metabolising** genes is regulated in **operon** like manner.

viii) The steroid **hormones** act mostly at the post-translational level.

15.4 GENE REGULATION IN DEVELOPMENT AND DIFFERENTIATION

In this section you will study the long-term gene regulation, There are two key terms used in describing the long-term gene regulation. First one is development, It refers to the process of regulated growth resulting from the .interaction of the genome with cytoplasm and **the environment.** It involves a programmed sequence of

phenotypic events that are irreversible. The total phenotypic changes of an organism constitutes its life cycle. The second term is differentiation. This is the most spectacular aspect of development. It involves the formation of different types of cells, tissues and organs from a zygote through the process of specific regulation of gene expression. The process of differentiation results in cells having characteristic structural and functional properties. Therefore, we can say that the processes in differentiation and development are the result of highly programmed patterns of gene activation and gene repression.

A great deal is known to us about the lac operon in bacteria, but in eukaryotes information in this regard is meagre. Equally mysterious is how gene expression is regulated and coordinated during the development of a eukaryotic organism from a fertilised egg or zygote to an adult. In these organisms, the additional complication is that the activities of a multitude of different cells must be coordinated. So, we are dealing not with the control of gene expression in a single cell but with the coordinated regulation of gene activity in a number of different cells, perhaps in different parts of the organism. Despite all these factors, it is not an impossible task. Considerable progress has been made in understanding the regulation of gene expression in the fruit fly *Drosophila melanogaster*.

Drosophila is an ideal system for studying development. It is a relatively simple, segmented organism that develops through distinct larval stages. Since the embryo develops externally, as opposed to the in *utero*, changes can be readily observed. The entire life cycle, from fertilised egg to adult is accomplished in a matter of 2 weeks (see Fig. 15.6).

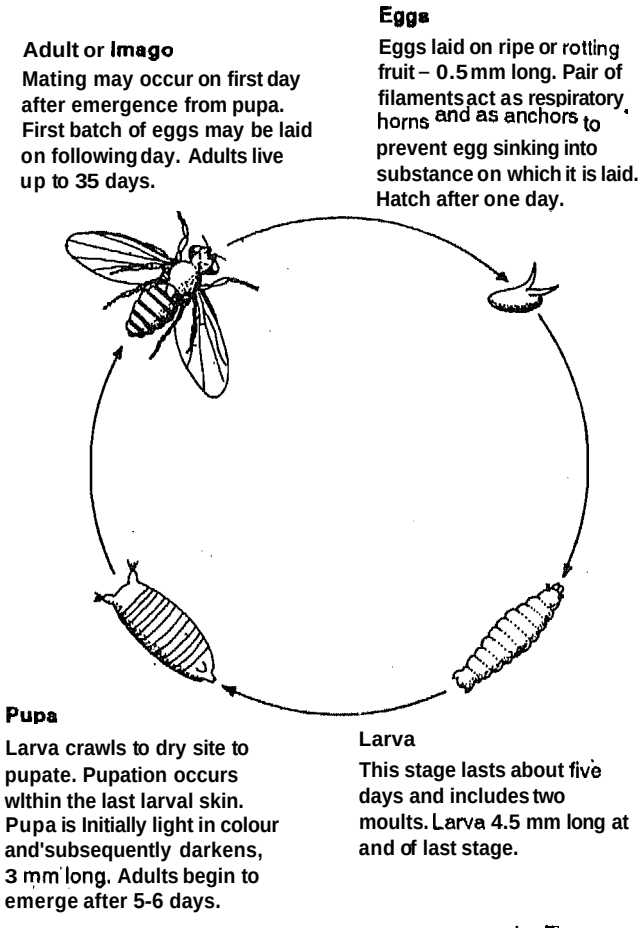


Fig. 15.6: Life cycle of *Drosophila*

The discovery over the years of a number of bizzare developmental mutations in *Drosophila* has been very crucial, as these mutants have turned out to be invaluable tools in understanding differentiation and development. These aberrations which alter the fates of cells during embryogenesis and commit them to new developmental pathways, are known as "homeotic" mutations. For example, the mutation *Antennapedia* causes a pair of normal-looking legs to grow from the head segment, replacing the antenna (see Fig. 15.7a). The *Bithorax* mutations lead to the development of a second pair of partial or complete wings in an inappropriate thoracic segment.

Genes that get altered by these developmental mutations have been studied, and are found to be extremely complex both in terms of structure and of expression. However, one intriguing feature of several of these genes has emerged: ***the homoeobox,*** a short segment of 180 bp coding for a 60 amino acid portion of the gene product which is similar in atleast ten of the relevant ***Drosophila*** genes (Fig. 15.7b). The fact that these genes have a structural relationship, that is, the biological information that each carries is similar in some respect, suggests tbat possibly the gene products are related in terms of function. But the biggest surprise to molecular biologists has been the discovery that homoeoboxes also exist in genes of frog, mouse and man. These genes in vertebrates are totally different from the ***Drosophila*** genes, except for the presence of homoeobox and the fact that the genes in vertebrates are also believed to be involved in development. From this one question arises. Are the genetic instructions for developmental processes universal? Thomas Hunt Morgan and his colleagues showed that the study of a simple organism like ***Drosophila*** could provide information of general relevance to genetics. It is exciting to think that ***Drosophila*** could also provide the key to understanding the complex events that underlie development in man.

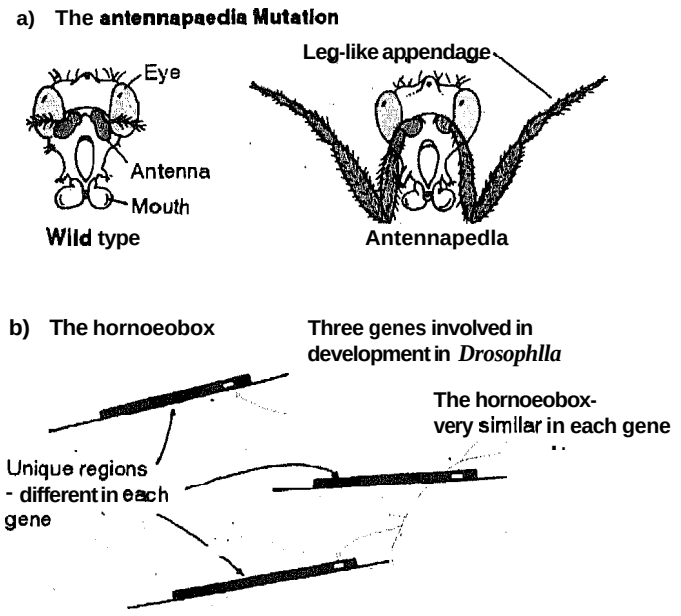
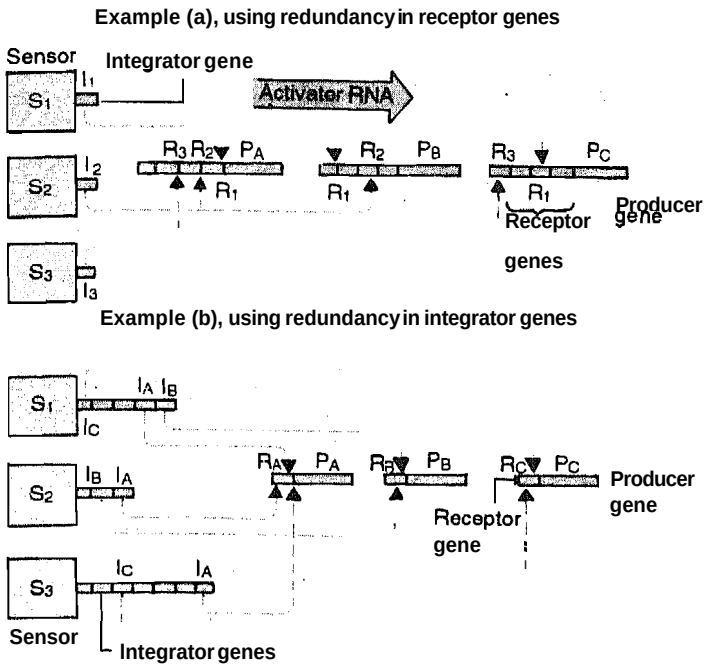


Fig. 15.7 : The development genetics of *Drosophila*. (a) The wild type (left) and *Antennapedia* mutant (right). (b) different genes involved in development contain homoeoboxes.

15.5 GENE REGULATION IN EUKARYOTES

The control of gene expression in eukaryotes is much more complex and less well defined than tbe prokaryotes and viruses. The eukaryotic cell also contains greater amounts of genetic information, and its DNA is complexed with a wide variety of proteins to form chromatin. They also have more than one chromosome contained in a nuclear envelope. Translation occurs in the cytoplasm. In multicellular organisms, tissue and cell-specific gene products are restricted, even though each cell contains a full genetic complement. In comparison to bacteria and viruses, the eukaryotic organisms are not amenable to mutagenesis and experimentation. Thus, our understanding of the regulatory mechanisms in eukaryotes is not as precise as that of the prokaryotes and bacteria. Many speculations have been made and many models have been proposed to explain regulation of gene expression in eukaryotes.

One of the popular early models was proposed by R.J. Britten and E.H. Davidson (Fig. 15.8). This model proposes an integrated regulation of sets of structural genes by means of moderately repetitive regulator genes. It takes into account the observed interspersion of single, copy DNA .sequences and repetitive DNA sequences. See Figure 15.8 carefully.



Flg.15.8: The **Britten-Davidson** model of regulation *of* **gene expression** in **eukaryotes** showing two **variations** of **integrated** regulation. (a) **represents** a system **based on** redundancy of **'receptor'** genes, and in (b) **system** based on redundancy of **integrator** gene **is shown.** S_1, S_2, and S_3 are three sensor **genes** that respond to three **different** signals, that **may** be honuoue-receptor complexes. The diagrams **represent** the **events** that occur **after sensor** genes **have** triggered **the transcription** of **their respective integrator** gents (I, I_1, I, **or** I_A, I_1, I_C). The **products** of Integrator **genes** - 'activator RNAs' **diffuse from their site** of **synthesis** (Integrator genes) to **their sites** *of* action **(receptor** genes). **With the** binding **of activator** RNAs to **their** respeclive receptor genes somehow **triggers** the **transcription** of the **contiguous** producer genes P_A, P_B, P_C. **Depending on** which **integrator gene(s)** is **activated** by **sensor gene(s),** one, two or **all three** of **the** producer **(structural) genes** may be turned on. **From:** R.J. **Britten** and EII. **Davidson. 1969.** Gene **Regulation** for **Higher** Cells **:** A **Theory.** Science, **Vol.** 165, pp 349-357.

In the Britten-Davidson model, the specific ***Sensor genes*** represent sequence-specific binding sites, that respond to **specific** signals such as hormone-receptor protein complexes, When the sensor genes receive the right signal, they activate the transcription of the adjacent ***integrator genes.*** The products **of integrator** gene then interact in a sequence-specific manner with the ***receptor genes.*** According to **Britten** and Davidson, the integrator **gene** products were ***activator RNAs*** that interact directly with the receptor **genes** to trigger the transcription of the contiguous ***producer genes.*** The producer genes are considered analogous to the structural genes in **prokaryotic operons.** They also **point** that, it would make no difference whether the active integrator **gene-**products are RNA molecules or proteins.

When either **the** receptor genes (see Fig. **15.8a)** or the integrator genes (see Fig. **15.8b)** are **redundant,** various combinations **of** the producer genes can be turned on in response to different signals. And if both the integrator genes and the receptor genes are redundant, complex integrated circuits of gene expression can easily be devised, although testing the validity of such models is far more difficult.

The most attractive feature **of** this model is that it provides a plausible reason **for** observed **pattern** of interspersion of moderately repetitive sequences and single-copy DNA sequences. **There** are strong evidences showing that most structural genes (producer genes) are actually **single** copy DNA sequences. According to the above model, the adjacent **moderately repetitive** DNA **sequences** contain various kinds of regulator genes (sensor, integrator and receptor genes).

Further studies **comparing** the **complexity** of heterogeneous nuclear RNA **(hnRNA)** populations and mRNA **populations** in different types of cells have shown that **hnRNA** populations are **usually more complex** (as they contain more distinct sequences). **than mRNA populations. These** findings suggest that considerable regulation occur; post-transcriptionally during RNA processing, that is, in-the **hnRNA → mRNA** stage.

Based on these observations, Davidson and Britten proposed a second model, the Davidson-Britten model (see Fig. 15.9). This model proposes that gene expression is regulated at the level of RNA processing. According to this model, most of the structural genes are located in *constitutive transcription units,* which are transcribed at basal levels in all cells. These constitutive transcripts are processed, only in cells that contain the appropriate *integrating regulatory transcripts (IRTs).* The *IRTs* are transcribed in a cell-specific manner, and these must be present before the *constitutive transcripts (CTs)* of the structural genes can be processed into mRNAs.

These 'integrating regulatory transcripts' contain repetitive sequences that interact with different structural gene transcripts like the repetitive 'integrator' genes interacted with different 'receptor' genes in the original Britten-Davidson model. The main difference is that the regulation occurs post-tmnscriptionally during RNA in the new Davidson-Britten model, rather than transcriptionally as in the original model.

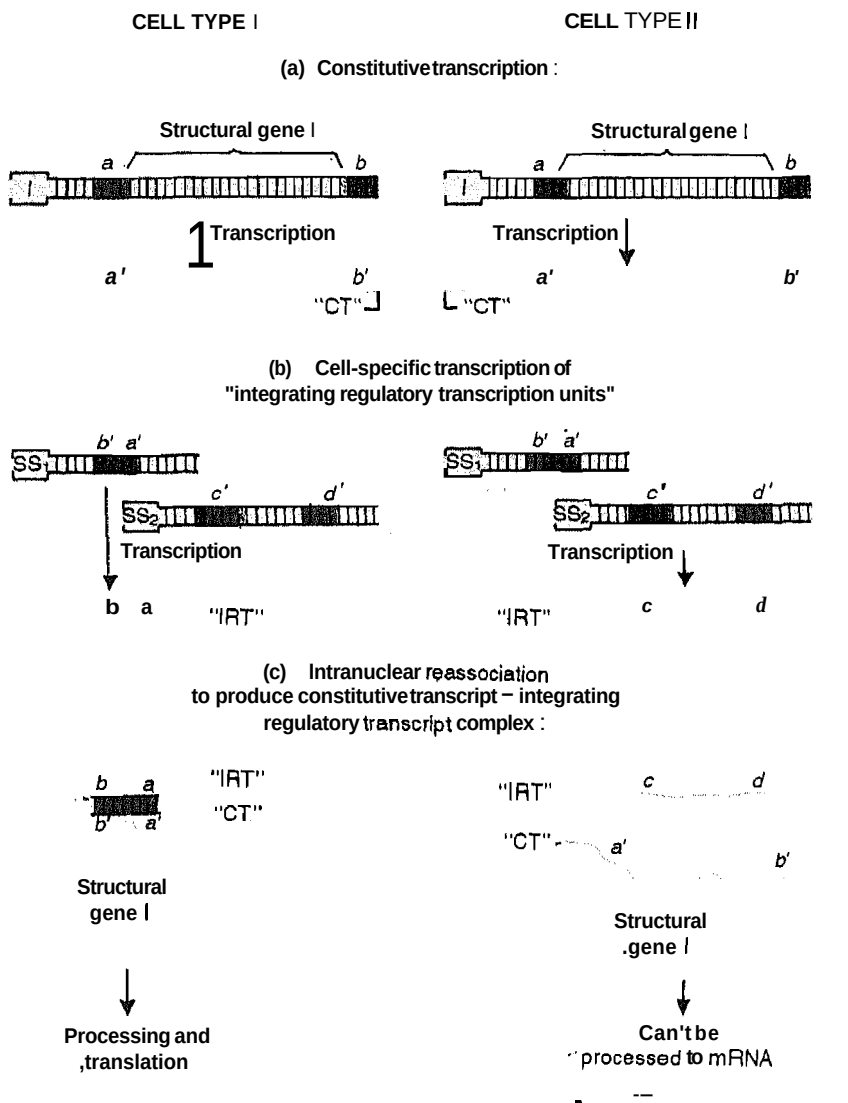


Fig.15.9: **The** new **Davidson-Britten** model for **the regulation** of gene expression. It is **based on** RNA **processing level in eukaryotes. 3 The majority of the** structural genes are **assumed** to **be located** In **'constitutive transcription units'. These are transcribed continually** in **all cells. The letter** I **denotes** the **transcription initiation site. The letters** a, a', b **and** so on **represent middle-repetitive** sequences **that are** proposed to be **involved** in **regulating gene expression. The sequences** a', ***b'***, ***c'*** and **d' are complementary to sequences** a, b, ***c*** and d **respectively. (b) Structural** gene **expression is regulated** by **repetitive** RNA **sequences transcribed** in a **tissue** or **cell-specific manner as components of "integrating regulatory transcription units (IRTUs).** The transcription of IRTU Is **controlled** by **nucleoprotein 'sensors'** (*SS*), **which** respond to **specific external** signals. **Different** tissues or **cell types transcribe different IRTUs.** (c) A give **structural gene constitutive transcript (*CT*) can** be processed **into mRNA and** thus **expressed** only If It **forms a complex with 'integrating** regulatory **transcript' (*IRT*). Thus,** lor **the expression of a particular structural** gene **in** any **given cells,** the **presence** of **appropriate nuclear**

IRT is required. Different types of cells will have *IRTs* carrying overlapping populations of repetitive sequences, so that they process some of the same structural gene transcripts and some different gene transcripts. The complex *IRT* populations of different cells are, however, not illustrated in this diagram. (Based on E R. Davidson and R.J. Britten, 1979. Science. Vol. 204 : 1052-1059).

SAQ 2

Match the items given in column I with those of column II. Write your answer in the space provided.

I		II	
i) Long-term gene regulation	[ ]	a)	several genes having similar base pair sequence of a short segment.
ii) Antennapcdia	[ ]	b)	respond to certain hormone-receptor proteins
iii) Homoeobox	[ ]	c)	development and differentiation
iv) Sensor genes	[ ]	d)	regulation of gene expression at the level of RNA processing
v) Davidson-Britten model	[ ]	e)	homoeotic mutation

15.6 SUMMARY

In this unit you have learnt that:

- Eukaryotes being more complex also contain more genetic material. They are assumed to have a more complex system of regulation.
- The discovery of operon system in bacteria prompted a search for similar control systems in eukaryotes. No operons however, have been found in eukaryotes, although the genes for related functions in some biochemical pathways arc closely linked or are contiguous. More often the genes that are coordinately regulated are dispersed throughout the genome.
- The .regulation in highcr eukaryotes can be either short-term or long-term. In higher eukaryotes one of the well studied systems of short-term gene regulation is the control of enzyme synthesis by steroid hormones. The specificity of hormone action is caused by the specific array of receptors in the genome of each hormone. The long-term regulation of gene expression refers to the regulation during the development and differentiation of organisms.
- One of the most popular models of regulation of gene expression in eukaryotes is the Davidson-Britten model. According to this model, regulation occurs at the level of RNA processing.

15.7 TERMINAL QUESTIONS

1) Why is gene regulation assumed to be more complex in a multicellular eukaryote than in a prokaryote? Why is the study of this phenomenon in eukaryotes more difficult?
2) Arc operons more common in bacteria or in higher organisms?
3) How do steroid hormones induce transcription of cukaryotic genes?
4) Why is the model proposed by Britten and Davidson more acceptable than the operon model for explaining regulation in cells of higher animals?

15.8 ANSWERS

Self-assessment Questions

1) i)
   iv)
   vii) .
   viii)
2) i) c
   ii) e
   iii) a
   iv) b
   v) d

Terminal Questions

1) Hint: One, they contain greater amount of genetic information. Two, their DNA is associated with various proteins, and it is present in chromosomes enclosed in nuclear envelope. In addition, various kinds of tissues/cells have cell-specific gene products that are associated with the particular genes.
   You may also refer to Section 15.2 and Section 15.5.

2) Operons are common in prokaryotes. Gene clusters that resemble operons (actually are not operons) exist in several microorganisms, e.g. *Neurospora.* No operon is known in higher eukaryotes.

3) Steroid hormones pass through the cell to the nucleus where they combine with transcriptional activator proteins, and stimulate transcription.

4) In prokaryotes, the structural genes specifying the enzymes in a metabolic pathway are usually arranged as groups of contiguous genes. This facilitates regulation by operon mechanism. In higher eukaryotes, such genes are usually not in clusters and are often unlinked. The complex patterns of gene expression during development in higher animals almost certainly require complex integrated controls' that can governexpression. Genomes of higher eukaryotes contain single copy DNA sequences (structural genes) that are interspersed with middle-repetitive