RAMSADAY COLLEGE

Content Type: PDF

Department: BOTANY

Teacher: Dr. Arnab Kumar De

Caption: e-Material of BOT-A-CC-4-10-TH

GENETICS

7. Structural organisation of Gene:

7.3. Overlapping gene.

Overlapping genes

Definition: An overlapping gene is a gene whose expressible nucleotide sequence partially overlaps with the expressible nucleotide sequence of another gene. In this way, a nucleotide sequence may make a contribution to the function of one or more gene products.

Overprinting refers to a type of overlap in which all or part of the sequence of one gene is read in an alternate reading frame from another gene at the same locus. Overprinting has been hypothesized as a mechanism for *de novo* emergence of new genes from existing sequences, either older genes or previously non-coding regions of the genome. Overprinted genes are particularly common features of the genomic organization of viruses, likely to greatly increase the number of potential expressible genes from a small set of viral genetic information. We will now an example of overlapping genes in φX 174 bacteriophage.

Overlapping Genes in φX174:

The bacteriophage φX 174 is an extremely small icosahedral virus containing a single stranded DNA molecule about 5,400 nucleotides long. Nine genes A, B, C, D, E, J, F, G and H have been identified on φX DNA and they code for 9 specific proteins.

The combined weight of these proteins is 250,000 Daltons. A genome of about 5,400 bases has a maximum coding capacity for about 1800 amino acid residues with a combined weight of about 200,000.

Obviously, the total mass of proteins coded for is significantly greater than is expected from the amount of DNA contained in φX genome. The works of Sanger and Coulson (1975), Barrelled (1976) and Weisbeek (1977) have revealed several mechanisms underlying the compact genetic organisation in φXDNA.

The authors have employed the techniques of restriction mapping using several different enzymes such as Hind II, Hae III, Hpa II and Alu I. By applying Sanger and Coulson's 'plus and minus' technique for determining DNA, sequences, Barrell et al., have found the sequences of genes D, E and J and in fact the entire sequence of φX174 DNA. Weisbeek have sequenced genes A and B.

The termination codon of gene D overlaps the initiation codon for gene J by one nucleotide. Further, the mutation amb 6 determined genetically to lie in gene J actually lies 179 nucleotides before the initiating codon for J. Gene J therefore lies in another gene.

The location of gene E has been identified from two amber mutations am^3 and am^6. Both the mutations lie within the sequence of gene D. Analysis of the sequence around the amber mutations has shown that the sequence of gene E (273 nucleotides long) overlaps the sequence of gene D (456 nucleotides long); that both genes are translated in two reading frames in two different phases.

The sequence for gene E is displaced one base to the right from that of gene D. By identifying the initiation and termination codons of gene D, it is concluded that the sequence

of gene E lies in the latter part of gene D, and specifies a protein of about 10,000 Daltons. Thus gene E overlaps a portion of gene D; the sequence in the overlap region codes for two proteins in two phases (Fig. 22.12).

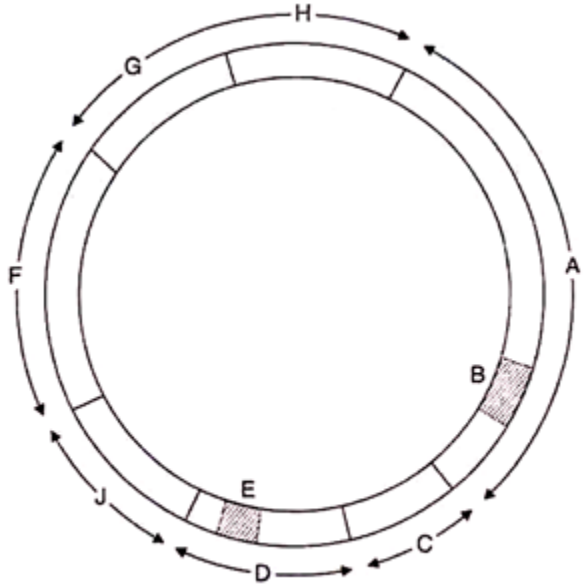

Fig. 22.12 The genetic map of ϕX174 showing overlapping genes.

The genes A and B have been characterised by Weisbeek (1977) by mapping positions of several mutations by a marker rescue technique. All the mutations in gene B have been identified within the sequence of gene A.

However, the nonsense mutations in gene A do not impair the function of gene B; similarly nonsense mutations in gene B do not affect the activity of gene A; the genes A and B belong to different complementation groups.

The study of amber mutations in gene A which results in synthesis of shorter chains of A and A* proteins has proved that gene B is completely contained within gene A and is translated in two different reading frames: one which leads to synthesis of A and A * protein, the other for synthesis of B protein. Smith (1977) working in Sanger's lab have shown that gene A not only overlaps gene B but even extends beyond gene B.

Overlapping genes have also been identified in the single-stranded DNA virus G4 which is closely related to and has the same order of genes as φX174. In G4 also gene A overlaps gene D, and gene E overlaps gene D. It also has a gene if containing portions of sequence of gene A and C. The last 86 nucleotides of gene A and the first 89 nucleotides of gene C together constitute the sequence of gene K.