The Origin, Development and Present Status of the Concept of the Gene: A Short Historical Account of the Discoveries[†]

Petter Portin*

Laboratory of Genetics, Department of Biology, University of Turku, FIN-20014 Turku, Finland

Abstract: The classical view of the gene prevailing during the 1910's and 1930's comprehended the gene as the indivisible unit of genetic transmission, genetic recombination, gene mutation and gene function. The discovery of intragenic recombination in the early 1940's led to the neoclassical concept of the gene, which prevailed until the 1970's. In this view the gene or cistron, as it was now called, was divided into its constituent parts, the mutons and recons materially identified as nucleotides. Each cistron was believed to be responsible for the synthesis of one single mRNA and concurrently for one single polypeptide. The discoveries of DNA technology, beginning in the early 1970's, have led to the second revolution in the concept of the gene in which none of the classical or neoclassical criteria for the definition of the gene hold strictly true. These are the discoveries concerning gene repetition and overlapping, movable genes, complex promoters, multiple polyadenylation sites, polyprotein genes, editing of the primary transcript and gene nesting. Thus, despite the fact that our comprehension of the structure and organization of the genetic material has greatly increased, we are left with a rather abstract, open and general concept of the gene.

The gene is operationally defined on the basis of four genetic phenomena: genetic transmission, genetic recombination, gene mutation, and gene function. These criteria of definition are interdependent; we cannot for example observe gene function or gene mutation without transmission, while on the other hand, we cannot observe transmission without gene function.

According to the so called classical view of the gene, which prevailed during the 1910's and 1930's, all four criteria led to one and the same unit. According to the classical view, the gene was the smallest indivisible unit of transmission, recombination, mutation and gene function.

The classical view of the gene begins with the work of Gregor Mendel [1], in which he explained definitively the transmission of genes -or elements as he called these units of inheritance - and their independent assortment. The gene as the unit of transmission means that each gamete includes one unit of each gene. The term "gene" was coined by Wilhelm Johannsen [2]. He wished this unit of heredity to be free of any hypotheses regarding its physical or chemical nature; i.e. the genes could be treated as calculating units.

The actual formulation of the classical concept of the gene must be attributed to the American Thomas Hunt Morgan and his school, which included Calvin Blackman Bridges, Herman Joseph Muller and Alfred Henry Sturtevant. They created the chromosome theory of inheritance, according to which the genes reside in the chromosomes like beads on a string. The chromosome theory of inheritance begins, however, already in the works of Walter S. Sutton and Theodore Boveri in 1903 [3, 4]. They called attention to the fact that the Mendelian rules of inheritance were explained by the behaviour of chromosomes in meiosis. Already earlier Boveri [5] had demonstrated the individuality of chromosomes, and in 1904 [6] he showed that chromosomes preserved their individuality during cell division. Both these characteristics of chromosomes are naturally necessary properties of the hereditary material.

The chromosome theory of inheritance developed as a precise theory due to the work of the Morgan school. They observed [7, 8] that the number of linkage groups in *Drosophila*

[†]This work was supported by The Academy of Finland.

* Address correspondence to this author at the Laboratory of Genetics, Department of Biology, University of Turku, FIN-20014 Turku, Finland; telephone: 358-2-333 5570; telefax: 358-2-333 6680; e-mail: petter.portin@utu.fi

1389-2029/00 $25.00+.00

melanogaster was the same as the haploid number of chromosomes of that species. Sturtevant [9] was able to map six sex-linked genes of *D. melanogaster* into a linear order, and called attention to the fact that the linear structure of the linkage group corresponded to that of the chromosome. In fact, there exists an epistemological correspondence between the concepts of the linkage group and the chromosome. These facts were, however, only indirect evidence in favour of the chromosome theory of inheritance. The first direct evidence was obtained by Bridges [10], who was able to show that a certain abnormal behaviour of sex-linked genes of *D. melanogaster*, namely non-disjunction, corresponded to the analogous non-disjunction of the sex chromosomes. Further direct evidence was gained when Herman J. Muller and Theophilus S. Painter [11] and Theodosius Dobzhansky [12] demonstrated that the X-ray-induced structural changes of the chromosomes were associated with corresponding changes in the linkage relations of the genes. This was the first step towards the physical mapping of genes, which Bridges [13 - 15] then carried much further when he was able to map genes on the salivary gland giant chromosomes of *D. melanogaster*, in the best case to an accuracy of a single chromosome band. Already earlier Painter [16] had discovered the value of salivary gland chromosomes in genetics when he showed that structural changes in *Drosophila* linkage groups can be correlated with changes in the sequence of transverse discs or bands in these chromosomes.

The gene as the unit of mutation became apparent from the fact that the mutant alleles of a single gene were mutually exclusive. Thus the gene was believed to alter in the event of mutation as a single unit. This view gained further support when Muller [17, 18] showed that X-rays increase the number of mutations in linear proportion to the amount of radiation. Since it was known that radiation was quantated, it was believed that the genetic material was likewise quantated, the gene being the atom of genetics.

The gene as the unit of gene function is defined on the basis of the complementation test. In that test, if the heterozygote a/b is phenotypically mutant, *a* and *b* are alleles of the same gene. If, on the contrary, *a* and *b* complement each other, in such a way that the heterozygote is phenotypically of wild type, a and b are mutations of different genes, and thus the genotype will be written as *a +/+ b*.

The nature of gene function was substantially specified by George W. Beadle and Edward Tatum [19] and Adrian M. Srb and Norman H. Horowitz [20] when they showed, using *Neurospora crassa* (Ascomycetes) as their experimental organism, that genes control the synthesis of enzymes, and in particular that each individual gene is responsible for the synthesis of one single enzyme. This one gene - one enzyme hypothesis was the culmination of the classical view of the gene.

THE BREAK UP OF THE CLASSICAL VIEW OF THE GENE

The classical concept of the gene started to break down as soon as it had been completely formulated. C. Peter Oliver [21] and Edward B. Lewis [22] namely observed in *Drosophila melanogaster* the phenomenon of intragenic recombination. Mutations which on the basis of the complementation test were alleles, recombined with a very low frequency. Thus the atom of genetics was not indivisible. Melvin M. Green and Kathleen Green [23] were able to map mutations of the *lozenge* locus of *D. melanogaster* into linear order. Joseph A. Roper [24] and Guido Pontecorvo [25] carried the analysis even further. They observed in the ascomycete *Aspergillus nidulans* intragenic recombination within genes which were known to control the synthesis of one single enzyme. David M. Bonner [26] and Norman H. Gilles [27] observed the same in *Neurospora crassa*. Robert H. Pritchard [28] was the first to demonstrate with microbic fungi that mutations within a single gene could be mapped into linear order by means of recombination.

THE NEOCLASSICAL CONCEPT OF THE GENE

Our comprehension of the nature of the genetic material became more accurate when Oswald Avery *et al*. [29] demonstrated that the substance causing transformation in bacteria was DNA. Transformation had been discovered by Frederic Griffith [30]. He observed that killed bacterial cells injected into mice were able to transform genetically different living bacteria into their own kind. Martin H. Dawson and Richard Sia [31] and Lionel Alloway [32] demonstrated transformation *in vitro* in a cell-free extract. Avery's group [29] were able to isolate from this extract the substance which was responsible for transformation. That substance was DNA. Thus genes consisted of DNA. Further support for the DNA theory of inheritance was gained when Alfred D. Hershey and Martha Chase [33] demonstrated that DNA alone was responsible for the multiplication of bacteriophages.

From the point of view of the conceptual framework of genetics, the experiments of Seymour Benzer [34 - 37] involving the genetical fine structure of the bacteriophage T4 rII-region turned out to be of fundamental importance. With the aid of a selective technique he was able to map hundreds of mutations of that region into a linear order. The gene as the unit of function was not indivisible; within the gene, independently mutating mutation sites existed, which could be separated from one anothers by means of genetic recombination. Benzer created a new terminology. He called the unit of genetic function the cistron. The cistron is operationally defined with the aid of the *cis-trans* test [Fig. **1**). In the *cis-trans* test *cis*- and *trans*-heterozygotes are compared. In the *cis*-heterozygote the mutations are in the same chromosome but in the *trans*-heterozygote in homologous chromosomes. Thus the genotype of the *cis*-heterozygote is designated as *a b*/+ + and that of the *trans*-heterozygote as *a* +/+ *b*. If the *cis*-heterozygote is of a wild type phenotype and the *trans*-heterozygote is mutant, *a* and *b* are mutations of the same cistron. If, however, both *cis*- and *trans*-heterozygotes are phenotypically of a wild type, *a* and *b* are mutations of different cistrons. The cistron is a synonyme of the gene, but this term should be used only when it is based on *cis-trans* test or biochemical evidence, which will be dealt with later.

The smallest unit of recombination Benzer called the recon, and the smallest unit of mutation the muton. The recon is the smallest unit of genetic material which can be separated from other such units by means of genetic recombination but which cannot be divided further. A muton is the smallest unit of genetic material a change in which is sufficient to cause a mutant phenotype. Benzer also called attention to the correspondence between the linear internal structure of the cistron and the linear structure of the DNA molecule. As is well known, the latter fact had been demonstrated by James D. Watson and Francis H. C. Crick [38].

Alexander Dounce [39] and George Gamow [40] independently presented the so-called colinearity hypothesis, according to which the linear structure of DNA determines the linear primary structure of a polypeptide. The colinearity hypothesis was shown to be true by Anan Sarabhai *et al*. [41] and Charles Yanofsky *et al*. [42, 43], by comparing the genetic map of the T4 phage coat protein gene and the corresponding primary structure of the polypeptide and also by comparing the genetic map of the tryptophane synthetase gene of *Escherichia coli* with the corresponding primary structure of the polypeptide. Further, Yanofsky's group were able to show that the material counterpart of a cistron was that part of the DNA molecule which coded information for the synthesis of a single polypeptide. This, they demonstrated by

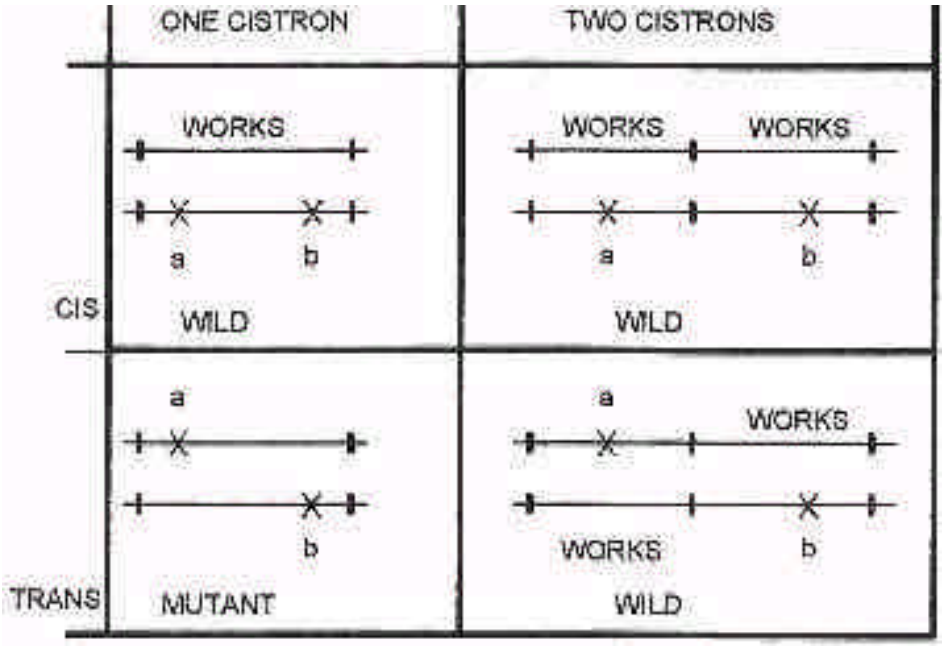

Fig. (1). The principle of the *cis-trans* test. If mutations a and b belong to the same cistron, the phenotypes of the *cis*- and *trans*-heterozygotes are different. If, however, the mutations belong to different cistrons the *cis*- and *trans*-heterozygotes are phenotypically similar.

showing that all the mutations of the *E. coli* tryptophan synthethase A-protein were mutations of the A-cistron, and likewise the mutations of the B-protein of the B-cistron. Finally, Yanofsky's group were able to show by studying the A-protein of the tryptophan synthethase that the material counterpart of both recon and muton was one single nucleotide pair in the structure of DNA [44, 45]. Thus the cornerstone of the neoclassical view of the gene became the one cistron - one polypeptide hypothesis, which replaced the old one gene - one enzyme hypothesis.

Alfred Hershey *et al.* [46] substantially elucidated the role of RNA in protein synthesis by showing that in connection with phage infection the synthesis of RNA in the bacterial cell greatly increased. Elliot Volkin and Lasarus Astrachan [47] measured the base constitution of the RNA involved in phage infection, and demonstrated that it was virus-specific. Cedric Davern and Matthew Meselson [48] and Monica Riley *et al.* [49] for their part demonstrated that the RNA involved in phage infection was unstable, unlike the ribosomal RNA, which was stable. Thus, the hypothesis ripened that protein synthesis was conducted by the unstable RNA and not the ribosomal RNA.

The time was thus ripen to present direct evidence of the role of RNA in protein synthesis. This was done by J. Brenner *et al*. [50], when they showed with the aid of heavy carbon isotopes and ultra-centrifugation that in phage infection it was indeed the RNA of the phage and not the bacterium which was responsible for the synthesis of the phage coat protein. Francois Jacob and Jacques Monod [51] had proposed the name messenger RNA (mRNA) for this protein-synthesis-conducting RNA.

Thus, the neoclassical view of the gene culminated in a theory according to which one gene or cistron controls the synthesis of one messenger RNA molecule, which in turn controls the synthesis of one polypeptide.

THE BREAK DOWN OF THE NEOCLASSICAL GENE CONCEPT AND THE MODERN CONCEPT OF THE GENE

The break down of the neoclassical concept of the gene started in the beginning of the 1970's, with the new discoveries of gene technology and molecular biology. These were the discoveries of repeated genes, interrupted genes and alternative splicing, the special case of immunoglobulin genes, overlapping genes, movable genes, complex promoters, multiple polyadenylation sites, polyprotein genes, the editing of messenger RNA and nested genes. Such observations have led to a situation where none of the classical or the neoclassical criteria of the definition of the gene hold strictly true. We therefore have to adopt a new, open, general and abstract concept of the gene, despite the fact that our comprehension of the nature and organization of the genetic material has greatly increased.

Stuart Linn and Werner Arber [52] and Matthew Meselson and Robert Yuan [53] found specific restriction endonucleases in bacteria, which act when the latter defend themselves agains the attact of bacteriopahges; thus these enzymes restrict the host range of the bacteriophages. Harry Smith and K. W. Wilcox [54] were able to purify these enzymes, and Thomas Kelly and Hamilton Smith [55], Kathleen Danna and Daniel Nathans [56] and Philip Sharp *et al*. [57], determined their mode of action. These enzymes cut DNA molecules each at a specific site. These observations made it possible to isolate genes, to clone them and analyze their biochemical structure in great detail. Following the action of restriction endonucleases, there often arise so-called cohesive ends in the DNA molecules [58], which tend to join together. By this means it is possible for example to join together DNA from any eukaryotic organism and that from the bacterial plasmids. Such recombinant DNA molecules were first constructed by David Jackson *et al*. [59], Peter Lobban and Armin Kaiser [60] and Stanley Cohen *et al.* [61]. Cloned DNA molecules can be physically mapped, using the cutting points of the restriction endonucleases as markers, [62] and sequensed by means of sophisticated biochemical methods [63, 64].

REPEATED GENES

Michael Waring and Roy Britten [65] and Britten and David Kohne [66 - 68] were the first to observe repeated DNA sequences in many organisms by means of the reassociation kinetics of DNA. The first observations of repeated structural genes concerned amphibian ribosomal RNA genes [69 - 71] and sea urchin histone genes [72].

The genes of ribosomal RNA are repeated in several tandem copies. Each one consists one transcription unit, but the gene cluster is usually transmitted from one generation to the next as a single unit. Thus the units of transmission and transcription are not always the same. Likewise the histone genes have been observed to be repeated in

such tandem repeats in many higher eukaryotic organisms [73].

INTERRUPTED GENES AND ALTERNATIVE SPLICING

The first observations of interrupted (split) genes, i.e. genes in which there exist noncoding intron sequences between the coding exon sequences, were made in animal viruses in 1977 [74 - 76]. These observations were based on physical mapping of the inner structure of the genes by means of so-called Southern blotting [62]. In this method, DNA fragments created by restriction endonucleases are separated according to their size in electrophoresis.

Almost at the same time as the virus observations, split genes were also found in eukaryotic organisms [77 - 85]. Since then, split genes have been found in all eukaryotic organisms investigated and their viruses, and it can be said that interrupted genes are the rule rather than the exception in the organization of the genes of multicellular eukaryotic organisms. In unicellular eukaryotic organisms they are found to a lesser extent. However, there are exceptions among multicellular organisms too, of which the evolutionarily old histone genes are worth mentioning [86].

In terms of the concept of the gene, interrupted genes constitute an interesting case in two respect. Firstly, the existence of introns between the coding exons shows that there exists no one-to-one colinear relation between the gene and the polypeptide; coding sequences are interrupted by noncoding DNA. Secondly, when interrupted genes produce messenger RNA, there occurs in certain genes tissue and stage-specific alternative splicing. The interrupted gene produces primary transcription product a heterogenous nuclear RNA molecule, in which both exons and introns are represented. Introns, however, are removed from the primary transcript during the processing of messenger RNA in specific splicing reactions. Splicing is usually constitutive, which means that all exons are joined together in the order in which they occur in the heterogenous nuclear RNA. In many genes, however, alternative splicing has also been observed, in which the exons may be combined in some other way (Fig. **2**). For example, some exon or exons may be skipped in the splicing reaction. The primary order of the exons is not, however, altered even in alternative splicing. Thus, alternative splicing makes it possible for a single gene to produce more than one messenger RNA molecule, which contradicts the basic conceptual framework of the neoclassical view of the gene.

Alternative splicing was first observed in animal viruses [87 - 95]. The first observations of alternative splicing in the genes of eukaryotes concerned murine immunoglobulin genes [96 - 99]. Since then, alternative splicing has been observed in hundreds of genes in various eukaryotic organisms, man included [see 100 for review].

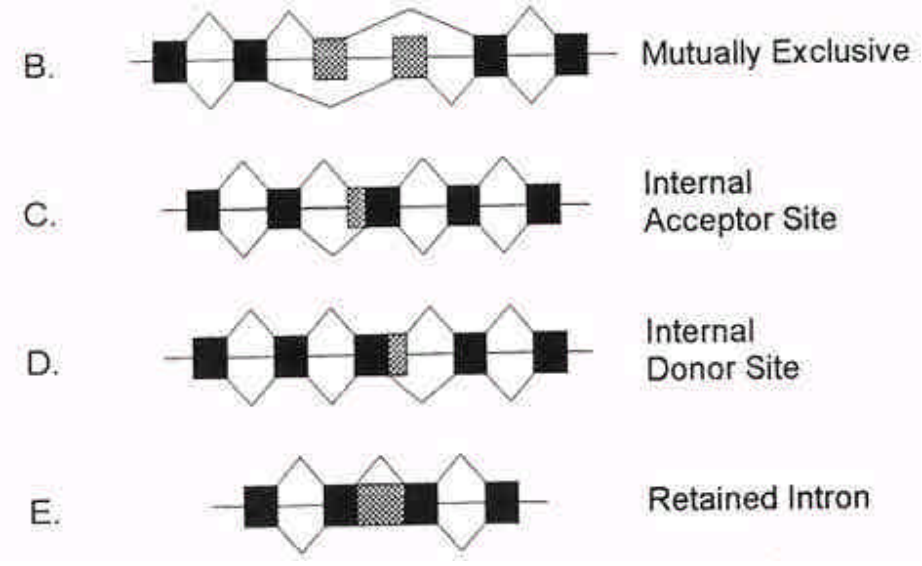

Fig. (2). Patterns of alternative splicing. Constitutive exons (black), alternative exons (dotted), and introns (heavy solid lines) are spliced according to different pathways (fine solid lines) [149].

The tissue specificity of alternative splicing was first shown in the fibrinogene genes of rat and man [101]. The first observations of developmental stage specificity concerned the alcohol dehydrogenase gene of *Drosophila melanogaster* [102]. The first demonstration, that alternative splicing was both tissue and stage-specific concerned the trompomyosin gene of *D. melanogaster* and rat [103, 104]. The tissue and stage specificity of alternative splicing naturally constitutes a previously unknown and effective mechanism of gene regulation.

IMMUNOGLOBULIN GENES

The enormous versatility of antibodies was for a long time a difficult problem in genetics. How was it possible that in the genome there was room for the codes of millions of different antibodies? The matter was solved when it became clear that the functional genes of immunoglobulins mature by means of somatic recombination from a few units in the germ line during the maturation of immune cells.

Each antibody molecule is a tetramere, consisting of two identical light chains and two identical heavy chains. Each chain consists of a constant and a variable region. In the genome of the germ line there are many genes for the variable region and a few genes for the constant region. In somatic recombination these can be combined during the maturation of the functional antibody gene into several thousands of different combinations whereby millions of different antibodies are formed. This phenomenon was first demonstrated by Nobumichi Hozumi and Susumu Tonegawa [105] [see also 106]. The observation was subsequently confirmed in several laboratories using different methods.

The immunoglobulin genes, which can be called assembled genes [107], do not fit any classical or neoclassical definition of the gene, since the genetic unit in the germ line and in the mature immune cell is completely different.

OVERLAPPING GENES

The first observations of overlapping genes were made in the bacteriophages ΦX 174 [108] and G4 [109], and in the animal virus SV40 [110, 111]. In the bacteriophage ΦX 174 several genes overlapped, encoding different proteins read from the same DNA strand but in different reading frames. In G4 the situation was even more complicated. In that phage the same DNA strand encoded as many as three different proteins, the messenger RNA's of which were transcribed overlappingly in all three possible reading frames. In these phages, however, the genes overlapped for only a few codons. In the SV40 virus, on the other hand, the genes overlapped for as many as 122 nucleotides.

Since these early findings, overlapping genes have also been found in eukaryotic multicellular organisms such as *Drosophila melanogaster* [112], mouse [113], and rat [114]. Overlapping genes can be located on the same DNA strand or on opposite strands.

Neither the classical nor the neoclassical view of the gene encompassed the possibility of overlapping genes, since genes were believed to reside on the chromosome always in tandem. The evolution of overlapping genes is also a difficult problem which, however, will not be dealt with here.

MOVABLE GENES

Movable genes are DNA elements which can move from one location to another in the genome of an organism. Already in the 1940's Barbara McClintock [115, 116] explained certain variegated phenotypes of maize by means of movable genes, which she called 'control elements'. At the time, however, these elements appeared so odd that nobody really knew what to think of them. Nowadays movable genes have been found in virtually all organisms and their molecular nature is quite well known [see 117 - 119 for reviews]. Consequently movable genes have became one of most important discoveries of genetics, and Barbara McClintock was awarded the Nobel prize for physiology or medicine in 1983 at the age of 81 years.

In addition to the ability of movable genes to move in the chromosomal complement from one location to another, they can also move from one individual to another and even, to a certain extent, from one species to another [see 120 for review]. Movable genes thus constitute an important evolutionary factor.

The existence of movable genes shows that the hypothesis of a fixed location of the gene in the chromosome, adopted by both the classical and neoclassical view, does not necessarily hold true.

COMPLEX PROMOTERS

Promoters are DNA sequences on the 5' side of the gene on which the RNA polymerase fastens when transcription begins. In all groups of organisms alternative promoters have been shown for many genes. These alternative promoters have been classified into six classes by Ueli Schibler and Filipe Sierra [121] (Fig. **3**). Certain types of alternative promoters make it possible for transcription to start from different points of the gene in different cases, and for the transcripts to have initiation codons at different positions of the chromosome. Thus, it is possible for a single gene in this case too to produce more than one type of messenger RNA molecules, encoding more than one polypeptide. This is again against the basic conceptual framework of the neoclassical view of the gene.

In higher eukaryotes alternative promoters are typically tissue and/or stage specific, as exemplified by the alcohol dehydrogenase gene of *Drosophila melanogaster* [102], the murine alpha-amylase gene [122, 123], and the aldolase gene of rat and man [124 - 126].

According to whether the unit of transcription is controlled by one or several promoters, simple and complex transcription units are distinguished.

MULTIPLE POLYADENYLATION SITES

During the maturation of messenger RNA, about 200 adenosine nucleotides are added in a polyadenylation reaction at the 3' end. These are not coded by the corresponding gene. In certain cases there are multiple alternative polyadenylation sites in the primary transcript. This was first observed in adenoviruses [127 - 131]. In cellular genes many alternative polyadenylation sites have also been found [see 132 for review]. Alternative polyadenylation sites usually involve the untranslated trailer sequence in the messenger RNA, but they can also involve translated sequences, and in this case they can affect the structure of the encoded protein. Thus multiple polyadenylation sites are one mechanism whereby a single gene can control the synthesis of more than one polypeptide.

POLYPROTEIN GENES

In viruses very many genes encode for one single large polypeptide, which, however, after translation is cleaved enzymatically into smaller subunits. Such polyprotein genes are also known in multicellular eukaryotes. Such are for example the neuropeptide genes of mammals [see 133, 134 for reviews] and the proline-rich proteins of salivary glands [135]. Thus, polyprotein genes contradict the hypothesis adopted by the neoclassical view of the gene, that each gene encodes for a single polypeptide.

EDITING OF MESSENGER RNA

In trypanosomes, and in certain plant RNA encoded by mitocondrial DNA, a peculiar

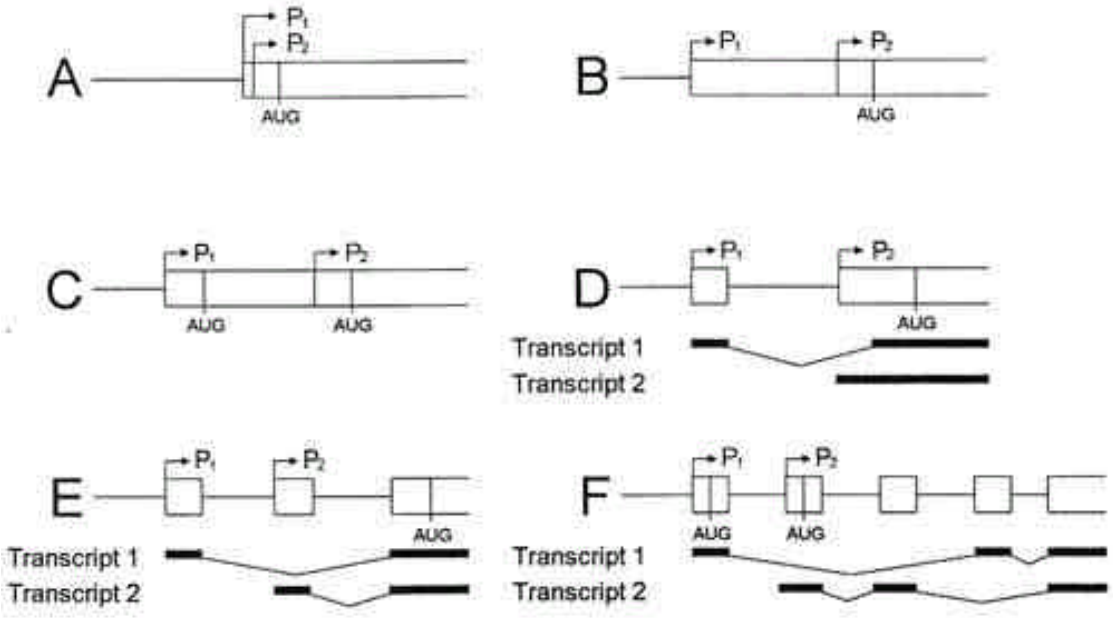

Fig. (3). Various organizations of complex promoters [121].

phenomenon called RNA editing has been observed. In this editing of messenger RNA, uracil nucleotides are removed and cytocine nucleotides are replaced by uracil after transcription. The information for this process comes either from the gene involved or alternatively from outside it. In the latter case the information source is called guide RNA [see 136 for review].

Although probably a rare phenomenon, editing of RNA is contrary to the classical and neoclassical concept of the gene, since in this case messenger RNA can retrieve information from outside the gene by which it was encoded.

NESTED GENES

Nested genes, i.e. a situation in which one gene resides within an intron of another gene, was first demonstrated in the *Gart* locus of *Drosophila melanogaster* by Steven Henikoff *et al*. [137]. In this particular case the nested genes were on opposite strands of DNA. Chun-nan Chen *et al.* [138] in turn demonstrated that in the large intron of the *dunce* locus of *D. melanogaster* there were actually two other genes residing, of which one was the known *Sgs-4* gene. In this case the nested genes were on the same strand of DNA. Barbara Levinson *et al.* [139] were the first to demonstrate nested genes in man. The 22nd intron of the factor VIII gene included another gene in its opposite strand. The large intron of human neurofibromatosis gene includes a total of three other transcription units in two opposed orientation [140].

The existence of nested genes is in contradiction to the central hypothesis adopted by both the classical and neoclassical gene concept, that genes are located in linear order on the chromosome.

ENHANCERS AND THEIR RELATION TO GENETIC COMPLEMENTATION

At the 5' and 3' end of the gene, enhancers are located, which respond to the signals mediated by the proteins regulating the function of the gene. Enhancers can also be located within the introns. The regulative effect of the enhancers is either positive or negative. In the latter case they are often called silencers [for reviews concerning enhancers and silencers, see for example 141, 142].

The relation of enhancers to genetic complementation is interesting. Denote an enhancer by *E* and the transcription unit regulated by it by *g*. In the *cis-trans* test, the $E^- g^-/E^+ g^+$ *cis*-heterozygote is phenotypically wild, whereas the $E^- g^+/E^+ g^-$ *trans-heterozygote* is phenotypically mutant. Thus the *cis-trans* test gives a positive result. This means that we cannot on the basis of a genetic test alone distinguish between an enhancer and the transcription unit regulated by it; biochemical evidence is needed. Thus, by definition, the regulatory elements of a transcription unit, such as enhancers, have to be included in the gene itself.

The first enhancers were demonstrated in the SV40-virus [143, 144]. The first enhancers in nucleated cells were demonstrated in the immunoglobulin heavy chain gene [145, 146].

THE PRESENT CONCEPT OF THE GENE

The examples presented above show that none of the classical or neoclassical criteria of the definition of the gene hold strictly true. The current view of the gene is of necessity an abstract, general, and open one, despite the fact that our comprehension of the structure and organization of the genetic material has greatly increased. Simply, our comprehension has outgrown the classical and neoclassical terminology. Open concepts, with large reference potential, are, however, very usefull in science in general, as pointed out by Richard M. Burian [147]. In fact it should be stressed that our comprehension of the very concept of gene has always been abstract and open as indicated already by Wilhelm Johannsen [2].

Due to the openess of the concept of the gene, it takes different meanings depending on the context. Maxime Singer and Paul Berg [148] have pointed out that many different definitions of the gene are possible. If we want to adopt a molecular definition, they suggest the following definition: "A eukaryotic gene is a combination of DNA segments that together constitute an expressible unit. Expression leads to the formation of one or more specific functional gene products that may be either RNA molecules or polypeptides. Each gene includes one or more DNA segments that regulate the transcription of the gene and thus its expression." (p. 622). Thus the segments of a gene include [1] a transcription unit, which includes the coding sequences, the introns, the flanking sequences - the leader and trailer sequences, and [2] the regulatory sequences, which flank the transcription unit and which are necessary for its specific function.

ADDENDUM

An extensive review on the early development of the concept of the gene has been published by Elof A. Carlson [150], and on the recent development by myself [151] and Carlson [152].

REFERENCES

[1] Mendel, G. (1866) *Ver. naturf. Ver. Brünn*, **4**, 3-44.

[2] Johannsen, W. (1909) Elemente der exakten Erblichkeitslehre. Jena: Gustav Fischer, pp. 143-144.

[3] Sutton, W.S. (1903) *Biol. Bull.*, **4**, 231-251.

[4] Boveri, T. (1903) *Ver. Dtsch. zool. Ges.*, **13**, 10-33.

[5] Boveri, T. (1902) *Ver. phys. med. Gesellsch. Würzburg, N.F.*, **35**, 60-90.

[6] Boveri, T. (1904) Ergebnisse über die Konstitution der chromatischen Substanz des Zellkerns. Jena: Gustav Fischer, pp. 1-22.

[7] Morgan, T.H., Sturtevant, A.H., Muller H.J. and Bridges, C.B. (1915) The mechanism of Mendelian Heredity. New York: Henry Holt, pp. 1-8.

[8] Morgan, T.H. (1919) The Physical Basis on Heredity. New Haven: Yale University Press, pp. 133-138.

[9] Sturtevant, A.H. (1913) *J.Exp. Zool.*, **14**, 43-59.

[10] Bridges, C.B. (1916) *Genetics*, **1** , 1-52 and 107-163.

[11] Muller, H.J. and Painter, T.S. (1929) *Am. Nat.*, **63**, 193-200.

[12] Dobzhansky, T. (1929) *Biol. Zentralbl.*, **49**, 408-419.

[13] Bridges, C.B. (1935) *J. Hered.*, **26**, 60-64.

[14] Bridges, C.B. (1937) *Cytological Fuiji Jubileum Vol*., 745-755.

[15] Bridges, C.B. (1938) *J. Hered.*, **29**, 11-13.

[16] Painter, T.S. (1934) *Genetics*, **19**, 175-188.

[17] Muller, H.J. (1927) *Science*, **66**, 84-87.

[18] Muller, H.J. (1928) *Zeits. Vererbunsl. Suppl.*, **1**, 234-260.

[19] Beadle, G.W. and Tatum, E.L. (1941) *Proc. Natl. Acad. Sci. USA*, **27**, 499-506.

[20] Srb, A.M. and Horowitz, N.H. (1944) *J. Biol. Chem.*, **154**, 129-139.

[21] Oliver, C.P. (1940) *Proc. Natl. Acad. Sci. USA*, **26**, 452-454.

[22] Lewis, E.B. (1941) *Ibid.*, **27**, 31-34.

[23] Green, M.M. and Green, K.C. (1949) *Ibid.*, **35**, 586-591.

[24] Roper, J.A. (1950) *Nature*, **166**, 956.

[25] Pontecorvo, G. (1952) *Adv. Enzym.*, **13**, 121-149.

[26] Bonner, D.M. (1950) *Genetics*, **35**, 655-656.

[27] Giles, N.H. (1952) *Cold Spring Harbor Symp. Quant. Biol.*, **16**, 283-313.

[28] Pritchard, R.H. (1955) *Heredity*, **9**, 343-371.

[29] Avery, O.T., MacLeod, C.M. and McCarty, M. (1944) *J. Exp. Med.*, **79**, 137-158.

[30] Griffith, F. (1928) *J. Hyg. Cambridge*, **27**, 113-159.

[31] Dawson, M.H. and Sia, R.H.P. (1931) *J. Exp. Med.*, **54**, 681-700.

[32] Alloway,J.L. (1932) *Ibid.*, **55**, 91-99.

[33] Hershey, A.D. and Chase, M. (1952) *J. Gen. Physiol.*, **36**, 39-56.

[34] Benzer, S. (1955) *Proc. Natl. Acad. Sci. USA*, **41**, 344-354.

[35] Benzer, S. (1957) in McElroy, W.D. and Glass, B. (eds.) The Chemical Basis of Heredity. Baltimore: Johns Hopkins Press, pp. 70-93.

[36] Benzer, S. (1959) *Proc. Natl. Acad. Sci. USA*, **45**, 1607-1620.

[37] Benzer, S. (1961) *Ibid.*, **47**, 403-415.

[38] Watson, J.D. and Crick, F.H.C. (1953) *Nature*, **171**, 737-738.

[39] Dounce, A.L. (1952) *Enzymologia*, **15**, 503-507.

[40] George Gamow, G. (1954) *Nature*, **173**, 318.

[41] Sarabhai, A.S., Stretton, A.O.W., Brenner, S. and Bolleh, A. (1964) *Ibid.*, **201**, 13-17.

[42] Yanofsky, C., Carlton, B.S., Guest, J.R., Helinski, D.R. and Henning, U. (1964) *Proc. Natl. Acad. Sci. USA*, **51**, 266-272.

[43] Yanofsky, C., Drapeau, G.R., Guest, J.R. and Carlton, B.C. (1967) *Ibid.*, **57**, 296-298.

[44] Crawford, I.P. and Yanofsky, C. (1958) *Ibid.*, **44**, 1161-1170.

[45] Yanofsky, C. and Crawford, I.P. (1959) *Ibid.*, **45**, 1016-1026.

[46] Hershey, A.D., Dixon, J. and Chase, M. (1953) *J. Gen. Physiol.*, **36**, 777-789.

[47] Volkin, E. and Astrachan, L. (1956) *Virology*, **2**, 149-161.

[48] Davern, C.I. and Meselson, M. (1960) *J. Mol. Biol.*, **2**, 153-160.

[49] Riley, M., Pardee, A., Jacob, F. and Monod, J. (1960) *Ibid*, **2**, 216-225.

[50] Brenner, J., Jacob, F. and Meselson, M. (1961) *Nature*, **190**, 576-580.

[51] Jacob, F. and Monod, J. (1961) *J. Mol. Biol.*, **3**, 318-356.

[52] Linn, S. and Arber, W. (1968) *Proc. Natl. Acad. Sci. USA*, **59**, 1300-1306.

[53] Meselson, M. and Yuan, R. (1968) *Nature*, **217**, 1110-1114.

[54] Smith, H.O. and Wilcox, K.W. (1970) *J. Mol. Biol.*, **51**, 393-409.

[55] Kelly, T.J. Jr. and Smith, H.D. (1970) *Ibid.*, **51**, 393-409.

[56] Danna, K. and Nathans, D. (1971) *Proc. Natl. Acad. Sci. USA*, **68**, 2913-2917.

[57] Sharp, P.A., Sugden, B. and Sambrook, J. (1973) *Biochemistry*, **12**, 3055-3062.

[58] Merz, J.E. and Davis, R.W. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 3370-3374.

[59] Jackson, D., Symons, R. and Berg, P. (1972) *Ibid.*, **69**, 2904-2909.

[60] Lobban, P. and Kaiser, A.D. (1973) *J. Mol. Biol.*, **78**, 453-471.

[61] Cohen, S., Chang, A., Boyer, H. and Helling, R. (1973) *Proc. Natl. Acad. Sci. USA*, **70**, 3240-3244.

[62] Southern, E.M. (1975) *J. Mol. Biol.*, **98**, 503-517.

[63] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.

[64] Maxam, A.M. and Gilbert, W. (1977) *Ibid.*, **74**, 560-564.

[65] Waring, M. and Britten, R.J. (1966) *Science*, **154**, 791-794.

[66] Britten, R.J. and Kohne, D.E. (1967) *Carnegie Inst. Washington Yearbook*, **65**, 78-106.

[67] Britten, R.J. and Kohne, D.E. (1968a) *Science*, **161**, 529-540.

[68] Britten, R.J. and Kohne, D.E.(1968b) *Carnegie Inst. Washington Yearbook*, **66**, 73-88.

[69] Miller, O.L.,Jr. and Beatty, B.R. (1969a) *Science*, **164**, 955-957.

[70] Miller, O.L., Jr. and Beatty, B.R. (1969b) *Genetics*, **61**, 133-143.

[71] Miller, O.L., Jr. and Beatty, B.R. (1969c) *J. Cell Physiol.*, **74** (suppl. 1), 225-232.

[72] Weinberg, E.S., Birnsteil, M.L., Purdom, I.F. and Williamson, R. (1972) *Nature*, **240**, 225-228.

[73] Lewin, B. (1980a) Gene Expression, vol. 2. Eukaryotic Chromosomes, 2nd ed. New York: John Wiley and Sons, pp. 706, 909-912.

[74] Berget, S.M., Berk, A.J., Harrison, T. and Sharp, P.A. (1978) *Cold Spring Harbor Symp. Quant. Biol.*, **42**, 523-529.

[75] Broker, T.R., Chow, L.T., Dunn, A.R., Gelinas, R.E., Hassel, J.A., Klessig, D.F., Lewis, J.B., Roberts, R.J. and Zain, B.S. (1978) *Ibid.*, **42**, 531-553.

[76] Westphal, H. and Lai, S.-P. (1978) *Ibid.*, **42**, 555-558.

[77] Breathnach, R., Mandell, J.-L. and Chambon, P. (1977) *Nature*, **270**, 314-319.

[78] Jeffreys, A.J. and Flavell, R.A. (1977) *Cell*, **12**, 1097-1108.

[79] Tilghman, S.M., Tiermeier, D.C., Seidman, J.G., Peterlin, B.M., Sullivan, M., Maizel, J.V. and Leder, P. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 725-729.

[80] Konkel, D., Tilghman, S. and Leder, P. (1978) *Cell*, **15**, 1125-1132.

[81] Lomedico, P., Rosenthal, N., Efstratiadis, A., Gilbert, W., Kolodner, R. and Tizard, R. (1979) *Ibid.*, **18**, 545-558.

[82] Bell, G.I., Pictet, R.L., Ruttner, W., Cordell, B., Tischer, E. and Goodman, H.M. (1980) *Nature*, **284**, 26-32.

[83] Yamada, Y., Avvedimento, V.E., Mudryj, M., Ohkubo, H., Vogeli, G., Irani, M., Pastan, I. and de Crombrugghe, B. (1980) *Cell*, **22**, 887-892.

[84] Wozney, J., Hanahan, D., Morimoto, R., Boedther, H. and Doty, P. (1981a) *Proc. Natl. Acad. Sci. USA*, **78**, 712-716.

[85] Wozney, J., Hanahan, D., Tate, V., Boedther, H. and Doty, P. (1981b) *Nature*, **294**, 129-135.

[86] Kedes, L.H. (1979) *Annu. Rev. Biochem.*, **48**, 837-870.

[87] Berk, A.J. and Sharp, P.A. (1978a) *Proc. Natl. Acad. Sci. USA*, **75**, 1274-1278.

[88] Chosh, P.K., Reddy, V.B., Swinscoe, J.,. Choudary, P.V., Lebowitz, P. and Weissman, S.M. (1978a) *J. Biol. Chem.*, **253**, 3643-3647.

[89] Chosh, P.K., Reddy, V.B., Swinscoe, J., Leibowitz, P. and Weissman, S. M. (1978b) *J. Mol. Biol.*, **126**, 813-846.

[90] Crawford, L.V., Cole, C.N., Smith, A.E., Paucha, E., Tegtmeyer, P., Rundell, K. and Berg, P.(1978) *Proc. Natl. Acad. Sci. USA*, **75**, 1274-1278.

[91] Lai, C.-J., Dhar, R. and Khoury, G. (1978) *Cell*, **14**, 971-982.

[92] Paucha, E., Mellor, A., Harvey, R., Smith, A.E., Hewick, R.M. and Waterfield, M.D.(1978) *Proc. Natl. Acad. Sci. USA*, **75**, 2165-2169.

[93] Canaani, D., Kahana, C., Mukamel, A. and Croner, Y. (1979) *Ibid.*, **76**, 3078-3082.

[94] Horowitz, M., Bratosin, S. and Aloni, Y. (1978) *Nucl. Acids Res.*, **5**, 4663-4676.

[95] Flavell, A.J., Cowie, A., Legon, S. and Kamen, R. (1979) *Cell*, **16**, 357-373.

[96] Sakano, H., Rogers, J.H., Huppi, K., Brack, C., Traunecker, A., Maki, R., Wall, R. and Tonegawa, S. (1979a) *Nature*, **277**, 627-632.

[97] Early, P.W., Davis, M.M., Kaback, D.B., Davidson, N. and Hood, L. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 857-861.

[98] Early, P., Rogers, J., Davis, M., Calame, K., Bond, M., Wall, R. and Hood, L.E. (1980) *Cell*, **20**, 313-319.

[99] Moore, K.W., Rogers, J., Hunkapiller, T., Early, P., Nottenburg, C., Weissman, I., Bazin, H., Wall, R. and Hood, L.E. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 1800-1804.

[100] Benjamin Lewin, B. (1980b) *Cell*, **22**, 324-326.

[101] Crabtree, G.R. and Kant, J.A. (1983) *Ibid.*, **31**, 159-166.

[102] Benyajati, C., Spoerel, N., Haymerle, H. and Ashburner, M. (1983) *Ibid.*, **33**, 125-133.

[103] Basi, G.S., Boardman, M. and Storti, R.V. (1984) *Mol. Cell. Biol.*, **12**, 2828-2836.

[104] Medford, R.M., Nguen, H.T., Destree, A.T., Summers, E. and Nadal-Ginard, B. (1984) *Cell*, **38**, 409-421.

[105] Hozumi, N. and Tonegawa, S. (1976) *Proc. Natl. Acad. Sci. USA*, **73**, 3628-3632.

[106] Tonegawa, S., Brack, C., Hozumi, N. and Pirrotta, V. (1978*) Cold Spring Harbor Symp. Quant. Biol.*, **42**, 921-931.

[107] Dillon, L.S. (1987) The Gene: Its Structure, Function, and Evolution. New York: Plenum Press, p. 12.

[108] Sanger, F., Air, G.M., Barrell, B.G., Brown, N.L., Coulson, A.R., Fiddes, J.C., Hutchison, C.A., Slocombe, P.M. and Smith, M. (1977) *Nature*, **265**, 687-695.

[109] Shaw, D.C., Walker, J.E., Northrop, F.D., Barrell, B.G., Godson, G.N. and Fiddes, J.C. (1978) *Ibid.*, **272**, 510-514.

[110] Contreras, R., Rogiers, R., van de Voorde, A. and Fiers, W. (1977) *Cell*, **12**, 529-538.

[111] Fiers, W., Contreras, R., Haegeman, G., Rogiers, R.., van de Voorde, A., van Heuverswyn, H., van Herreweghe, J., Volckaert, G. and Ysebaert, M. (1978) *Nature*, **273**, 113-210.

[112] Spencer, C.A., Gietz, R.D. and Hodgetts, R.B. (1986) *Ibid.*, **322**, 275-279.

[113] Williams, T. and Mine Fried, M. (1986) *Ibid.*, **322**, 275-279.

[114] Eveleth, D.D. and Marsh, J.L. (1987) *Mol. & Gen. Genet.*, **209**, 290-298.

[115] McClintock, B. (1947) *Carnegie Inst. Washington Yearbook*, **46**, 146-152.

[116] McClintock, B. (1948) *Ibid.*, **47**, 155-169.

[117] Cohen, S.N. and Shapiro, J.A. (1980) *Sci. Am.*, **242**, 36-45.

[118] Fedoroff, N.V. (1984) *Ibid.*, **250**, 64-74.

[119] Berg, D.E. and Howe, M.M. (1989) eds. Mobile DNA. Washington, D.C.: Am. Soc. Microbiol.

[120] Ajioka, J.W. and Hartl, D.L. (1989) in Berg, D.E. and Howe, M.M. eds. Mobile DNA, Washington, D.C.: Am. Soc. Microbiol., pp. 939-958.

[121] Schibler, U. and Sierra, F. (1987) *Annu. Rev. Genet.*, **21**, 237-257.

[122] Sierra, F., Pittet, A.-C. and Schibler, U. (1986) *Mol. Cell. Biol.*, **6**, 4067-4076.

[123] Young, R.A., Hagenbüchle, O. and Schibler, U. (1981) *Cell*, **23**, 451-458.

[124] Joh, K., Arai, Y., Mukai, T. and Hori, K. (1986) *J. Mol. Biol.*, **190**, 401-410.

[125] Schweighaffer, F., Maire, P., Tuil, D., Gautron, S., Daegelen, D., Bachner, L. and Kahn, A. (1986) *J. Biol. Chem.*, **261**, 10271-10276.

[126] Maire, P., Gautron, S., Hakim, V., Gregori, C., Mennecier, F. and Kahn, A. (1987) *J. Mol. Biol.*, **197**, 425-437.

[127] Nevins, J.R. and Darnell, J.E. (1978) *Cell*, **15**, 1477-1493.

[128] Klessig, D.F. (1977) *Ibid.*, **12** , 9-21.

[129] Wilson, M.C. and James E. Darnell, J.E. (1981) *J. Mol. Biol.*, **148**, 231-251.

[130] Ziff, E. and W. Fraser, N.W. (1978) *J. Virol.*, **25**, 897-906.

[131] McGrogan, M. and Raskas, J.H. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 625-629.

[132] Leff, S.E., Rosenfeld, M.G. and Evans, R.M. (1986) *Annu. Rev. Biochem.*, **55**, 1091-1117.

[133] Scheller, R.H., Kaldany, R.-R., Kreiner, T., Mahon, A.C., Nambu, J.R., Schaefer, M. and Taussig, R. (1984) *Science*, **225**, 1300-1308.

[134] Douglass, J., Civelli, O. and Herbert, E. (1984) *Annu. Rev. Biochem.*, **53**, 665-715.

[135] Lyons, K.M., Azen, E.A., Goodman, P.A. and Smithies, O. (1988) *Genetics*, **120**, 255-265.

[136] Weissmann, C., Cattaneo, R. and Billeter, M.A. (1990) *Nature*, **343**, 697-699.

[137] Henikoff, S., Keene, M.A., Fechtel, K. and Fristrom, J.W. (1986) *Cell*, **44**, 33-42.

[138] Chen, C., Malone, T., Beckendorf, S.K.and Davis, R.I. (1987) *Nature*, **329**, 721-724.

[139] Levinson, B., Kenwrick, S., Lakich, D., Hammonds, G. Jr. and Gitschier, J. (1990) *Genomics*, **7**, 1-11.

[140] Xu, G., O'Connell, P., Viskochil, D., Cawthon, R., Robertson, M., Culver, M., Dunn, D., Stevens, J. Ray Gesteland, White, R. and Weiss, R. (1990) *Cell*, **62**, 599-608.

[141] Serfling, E., Jasin, M. and Schaffner, W. (1985) *Trends Genet.*, **1**, 224-230.

[142] Schöler, H.J., Hatzopoulos, A.K. and Schlokat, U. (1988) Enhancers and trans-acting factors in Kahl, G. ed. Architecture of Eukaryotic Genes. Winheim: VCH Verlagsgesellschaft, pp. 89-129.

[143] Banerji, J., Rusconi, S. and Schaffner, W. (1981) *Cell*, **27**, 299-308.

[144] Moreau, P., Hen, R., Wasylyk, B., Everrett, R., Gaub, M.P. and Chambon, P. (1981) *Nucl. Acids Res.*, **9**, 6047-6069.

[145] Gillies, S.D., Morrison, S.L., Oi, V.T. and Tonegawa, S. (1983) *Cell*, **33**, 717-728.

[146] Banerji, J., Olson, L. and Schaffner, W. (1983) *Ibid.*, **33**, 729-240.

[147] Richard M. Burian, R.M. (1985) in Deprew, D.J. and Weber, B.H., eds. Evolution at a Crossroads. Cambridge, Massachusetts, London, England: Bradford Books, pp. 21-42.

[148] Singer, M. and Berg, P. (1991) Genes and Genomes: A Changing Perspective. Mill Valley: University Science Books, p. 622.

[149] Andreadis, A., Gallego, M.E. and Nadal-Ginard, B. (1987) *Annu. Rev. Cell Biol.*, **3**, 207-242.

[150] Carlson, E.A. (1966) The Gene: A Critical History. W.B. Saunders, Philadelphia and London.

[151] Portin, P. (1993) *Quart. Rev. Biol.*, **68**, 173-223.

[152] Carlson, E.A. (1991) *J. Hum. Genet.*, **49**, 475-487.