GENERAL CYTOGENETICS

J. SYBENGA

GENERAL CYTOGENETICS

J. SYBENGA

University of Agriculture
Wageningen

1972

NORTH-HOLLAND PUBLISHING COMPANY – AMSTERDAM • LONDON
AMERICAN ELSEVIER PUBLISHING CO., INC. – NEW YORK

PREFACE

The field of cytogenetics is wide and diversified. The different specializations (Drosophila polytene cytology, human karyotyping, plant cytogenetics, etc.) seem to have only few contact points and tend to be treated as separate entities. Yet they all have one common basis: the chromosome, its behaviour and the genetic consequences of this behaviour. It is gradually becoming clear that the chromosomes of all higher organisms, in composition and behaviour are fundamentally equal. Studying chromosomes in plants is quite relevant for understanding chromosomes in man and vice versa. It is important, therefore, to try to treat the entire science of cytogenetics as one unit. While such an approach may already be difficult with a simpler subject, it is truly complicated with this particularly wide field, certainly when the book must be kept within pocket size. There are a few ways out: one can generalize and consider only the major aspects in a simplified way that makes them understandable to the uninitiated student without further help. Or one can make the coverage wider and introduce some detail, hoping that the student has a sound basic knowledge of genetics and can count on someone to help him out when he gets baffled by the complexities of cytogenetic thought. The latter alternative was chosen, but is should be understood that the coverage still had to be far from complete.

The book contains a reasonable amount of systematically arranged information that may be used for reading assignments accompanying a cytogenetics course, and may perhaps replace extensive note-taking during lectures.

Because of the suitability of plant chromosomes for experimental manipulation, their study has contributed considerably to the understanding of chromosome behaviour. It is for this reason that many examples have been taken from plant cytogenetics. This may perhaps help preventing unnecessary "rediscovery" in humans and other animals of phenomena that for decades have been known to occur in plants.

CONTENTS

Chapter 1

BACKGROUND

1.1. The material basis

1.1.1. *Self-reproduction*

An essential characteristic of living material is its capacity for self-reproduction. In its absence, life, once originated, would not escape extinction: there are numerous internal and external processes that carry out an efficient break-down. This self-reproduction must be extremely exact as the vital processes of even the simplest forms of life are so complicated that slight deviations from an established pattern, proven to be efficient, will almost certainly result in a reduction of fitness.

All living material, from virus to complex multicellular organism contains a single basic substance that combines two essential functions: (1) exact replication (self-reproduction) and (2) strict regulation of the vital processes. The two functions are carried out in sequence, not simultaneously. This basic substance is usually DNA (deoxyribonucleic acid). In some cases (plant viruses for instance) it is the related substance RNA (ribonucleic acid). The exact replication of the basic regulating substance is the root of the self-reproduction of the species.

1.1.2. *DNA, RNA*

DNA and RNA consist of long molecular chains of *nucleotides*, which are the mono-phosphate esters of *nucleosides*. Nucleosides contain a pentose sugar (deoxyribose in DNA and ribose in RNA) and one of four nitrogenous bases: adenine, guanine (both purines), thymine (only in DNA and substituted by uracil in RNA)

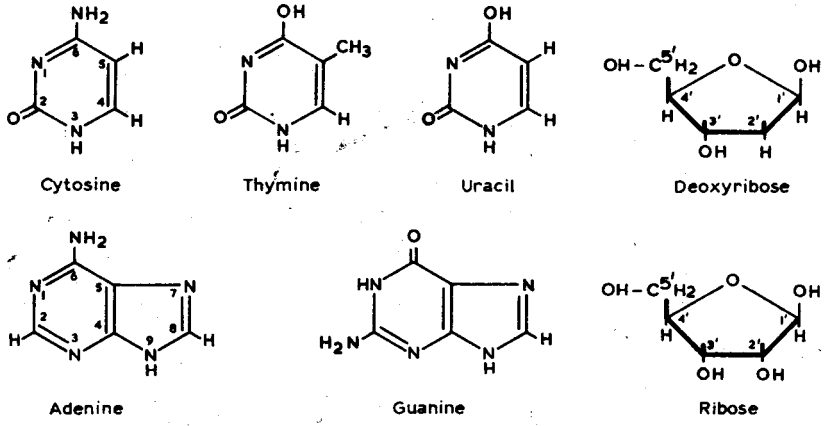

Fig. 1.1. The components of DNA and RNA. DNA contains the pyrimidines cytosine and thymidine, the purines adenine and guanine and the pentose sugar deoxyribose. RNA contains the pyrimidines cytosine and uracil, the purines adenine and guanine and the pentose sugar ribose.

and cytosine (all three pyrimidines) (see fig. 1.1). Thus only four types of nucleosides occur: adenosine, guanosine, thymidine and cytidine in DNA; adenosine, guanosine, uridine and cytidine in RNA. As a consequence, DNA and RNA contain only four types of nucleotide. DNA and RNA principally differ only in respect to the pentose sugar and one of the four nitrogenous bases.

1.1.3. *Replication*

Polymerization of the nucleotides into the DNA (or RNA) chains occurs serially. The nucleotide to be built in is in triphosphate form, with its phosphate groups attached to the 5′C atom of the sugar. In the (enzymatic) process of attachment two of the three phosphate groups are removed and the remaining one binds to the 3′C atom of the previously attached nucleotide of the growing chain (fig. 1.2). Thus the backbone of the DNA (and RNA) macromolecules is formed by successive (deoxy)ribose and phosphate groups, with the nitrogenous bases sticking out from the sides. The chain is not symmetrical: one end is the 3′C atom of the sugar, the other end the phosphate group at the 5′C atom: a

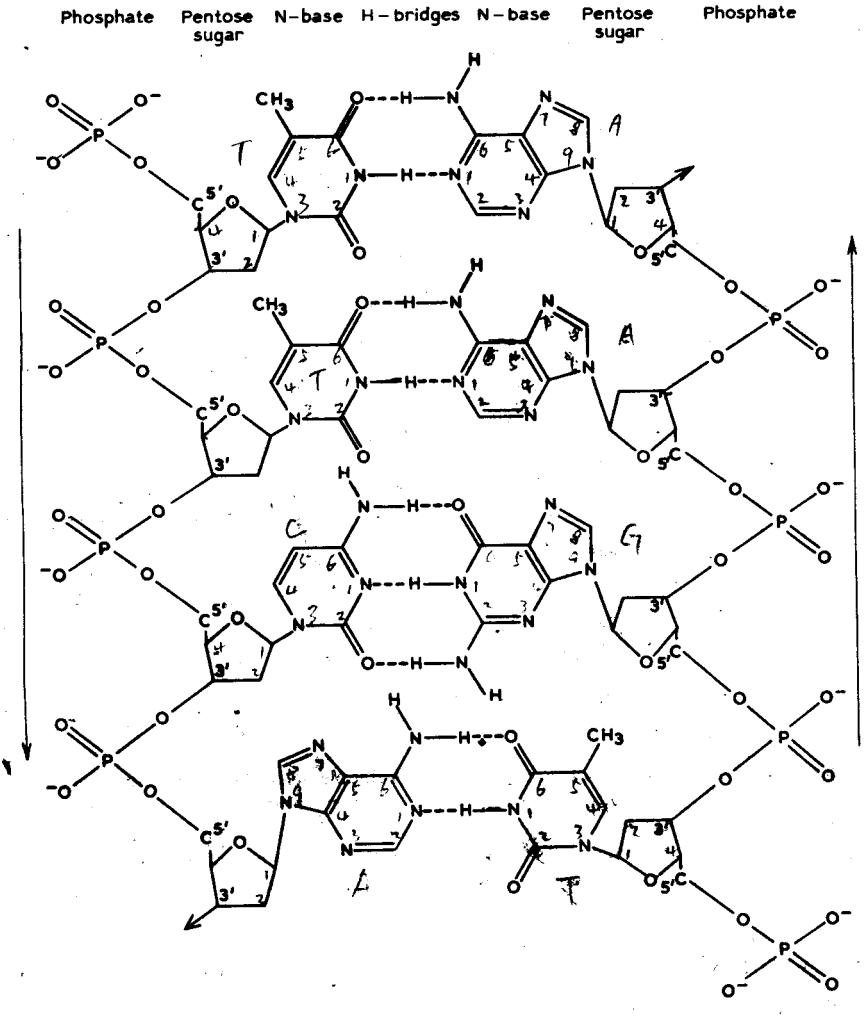

Fig.1.2. A fragment of a DNA double chain. The backbones of the two chains are formed by alternating sugar and phosphate groups. The polarities of the two chains is opposite. The nitrogenous bases are attached to the sugars and the complementary bases of the two chains are connected by hydrogen bridges. There are two between thymidine and adenine and three between cytosine and guanine. The double chain has the form of a helix with one revolution per 10 nucleotides.

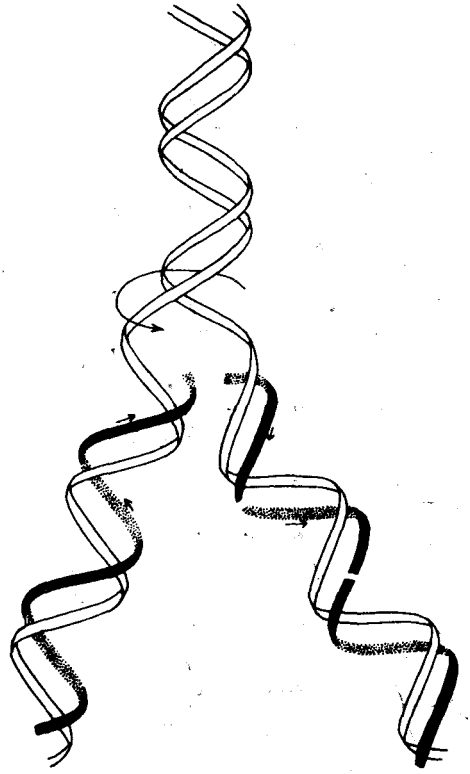

Fig.1.3. Semi-conservative replication of the DNA double helix. After unwinding, one of the two original chains directly forms a new, complementary chain. The other first forms fragments (Okazaki fragments) in the opposite direction, i.e. the same direction in respect to polarity. The fragments are later linked together by ligases. The two daughter helices both contain one old and one new strand and are completely equivalent.

DNA (and RNA) molecule has a definite *polarity* which has important consequences.

When DNA re(du)plicates, the new chain is laid down alongside an old chain, which functions as a template. There is a positive–negative relation between the two chains, in the sense that the nucleotides of the new chain are not identical to those of the old chain, but of specific complementary types. Opposite to thymine in the old chain, adenine is built into the new chain and vice versa.

Opposite to cytosine comes guanine (and vice versa). For stereochemical reasons other combinations are excluded. Hydrogen bonds form between the bases of the old and the new chains: two between thymine and adenine and three between cytosine and guanine. The orientation of the new chain is such that its polarity is reversed compared to that of the old chain. The two chains normally remain together: DNA is *double stranded* (fig.1.2). Before new synthesis can take place, the two strands must be separated. It might be expected that both strands are synthesized continuously starting from one initiation point. Then, because of the opposite polarity of the two strands, synthesis in respect to polarity would be into one direction in one strand and into the other direction in the other strand. There are strong indications that this is not the case: one strand is synthesized continuously, but the other strand is synthesized in sections (Okazaki fragments) in the same direction in respect to polarity, i.e. into the *opposite* direction when considered from the initiation point. The fragments are subsequently (enzymatically) linked together by ligases (fig.1.3).

Some important aspects of the structure of DNA are:

(1) DNA normally is double stranded;
(2) The two strands (the old and the new one) are not identical but complementary;
(3) The two strands have opposite polarity;
(4) In the total DNA there are equal quantities of adenine and thymine, and of guanine and cytidine, on a molecular basis. This is independent of the sequence or of the relative frequencies of nucleotides in each chain.

1.1.4. *Transcription*

In a comparable fashion but under different conditions and effected by different enzymes the DNA functions as a template for the formation of chains of RNA (transcription). After completion these RNA chains are released from the DNA without forming H-bonds. The RNA macro-molecule formed on the DNA is called *messenger* RNA (mRNA). It is carried to the cytoplasm where it is used as a template for the formation of polypeptide chains. This polypeptide synthesis is carried out by small bodies,

the ribosomes, that usually act in groups (polysomes), moving along the mRNA molecules in the process of polypeptide formation. The RNA moiety of the ribosomes (rRNA) consists of at least one small and two large RNA molecules that are transcribed on specific segments of the DNA. In the ribosome they are combined with proteins. The ribosomes pick up a third type of RNA: the *transfer* RNA (tRNA), that consists of much shorter chains, folded in a specific manner and often containing special base-types. These tRNAs are formed on their own DNA segments. The tRNA molecules have two recognition sites: on one an amino acid is attached enzymatically by a specific synthetase. The other site consists of a group of three nucleotides. Type and sequence of these nucleotides appear to be specific for the amino acid attached to the other site. Thus, for each amino acid there exists (at least) one specific tRNA characterized by a distinct group of three nucleotides. After having picked up a tRNA molecule with an amino acid, the ribosome first attaches the tRNA to the mRNA: the three specific tRNA nucleotides are fitted on a complementary set of three nucleotides in the mRNA. The first is placed on a specific initial site and the second follows on the next three nucleotides of the mRNA. The ribosome must select the correct tRNA for each site. For fitting the tRNA nucleotide triplets onto the mRNA triplets the same rules are followed as with transcription: adenine on uracil (which in RNA takes the place of thymine in DNA) etc. The row of amino acids is threaded together to a polypeptide chain and then the tRNA molecules are released from the mRNA and from the amino acids. Now both RNA forms are available for a new cycle, or are broken down. The sequence of amino acids in the polypeptide thus depends on the nucleotide sequence in the RNA, and consequently on that in the DNA, three nucleotides corresponding to one amino acid. This *triplet* of nucleotides in the mRNA contains the code for a specific amino acid and is named *codon*. Some codons have specific functions, such as interpunction. One amino acid may correspond to more than one codon: the code is degenerated. The codons UAA, UAG and UGA indicate the end of the polypeptide chain, the codons AUG and GUG the beginning.

For the replication of DNA both strands are used. For transcription only one is available: this prevents that two different polypeptide chains are derived from one DNA double strand. Which one is read off is determined by specific base sequences.

1.1.5. *Mutation*

Some polypeptide chains are built into structural proteins, but others are used to make enzymes. Small changes in the DNA, resulting from loss or from doubling of single nucleotides, can alter the entire pattern of transcription, since always three consecutive nucleotides form one codon. When the reading frame is shifted, all following codons are changed. A polypeptide may be formed, but it may not be expected to be functional as part of an enzyme. When no alternative intact DNA chain is available, a functional enzyme is not formed and the character conditioned by this enzyme is not expressed. Such a heritable change in a character is called a mutation, in this case a "frame shift" mutation.

Other changes in the DNA only locally affecting the reading frame such as inversion of a small segment or translocation from the original site to another location very nearby, or the replacement of one nucleotide by another, all will yield minor effects if the corresponding polypeptide segment is of minor importance. Whenever amino acid sequences in corresponding polypeptides have been studied in related species, small variations without great consequences have been detected. Occasionally, however, even simple amino acid substitutions, the consequence of single nucleotide replacements, have been found to have drastic effects. Both, harmless variations and drastic changes due to single amino acid substitution (sickle-cell anaemia for instance) have been found in haemoglobin and other components of the blood of man.

All such DNA changes (mutations) are transmitted with each cycle of DNA replication: they are heritable. In nature, mutations occur "spontaneously" (cause unknown) in a low frequency. They can be induced experimentally in high frequencies by chemicals (such as ethyl-methane-sulphonate, EMS, and other alkylating agents), by ionizing radiations, and by ultraviolet radiation. A high spontaneous mutation frequency would thoroughly disturb the

biochemical system of an organism. Therefore all forms of life have mechanisms that protect their DNA against a high mutation rate, and other mechanisms that can repair mutational damage. A few mutations, however, may result in a favourable mutant *phenotype* (the actual appearance of an individual, determined by a heritable component, the *genotype* and environmental factors). The individual in which this mutation is expressed, may have a reproductive advantage relative to the non-mutant individuals, or it may be able to maintain itself under unusual conditions. Mutations are the most important means of living material to improve itself even when it is at the expense of a large number of failures. Therefore, for no form or life is it desirable that mutation is precluded completely.

1.1.6. *Gene, cistron, operon, factor, allele*

For a long time geneticists have used the term "gene" for the hypothetical basic unit of function, of mutation and of recombination (see chapter 2). It has become clear that no single unit would simultaneously satisfy the criteria of these three categories. The unit of mutation may be a small group of atoms in one nucleotide. The unit of recombination, at least in lower organisms may be one nucleotide. The unit of function, however, can be quite large, consisting of a segment of DNA corresponding to at least one polypeptide chain, often to several.

Since 1961 when Benzer published his studies on complementation of different mutations of a single function in bacteriophage the term *cistron* has been in use as a unit of function. The term refers to specific differences in complementation when two mutations occur in the same DNA double strand (cis) and when the two occur in different double strands (trans) in the same cell. The cistron perhaps corresponds with one polypeptide chain. As complementation phenomena are more complicated than originally understood, the value of the experimental cistron is somewhat doubtful, but as a concept of functional unit it has survived.

Another unit of function, or rather of regulation is the *operon* (Jacob and Monod): a number of DNA segments coding for different enzymes, usually related to a single metabolic function in