

Molecular bases of cancer: new tools for diagnosis, prognosis and treatment

Rosario Perona Abellón

Instituto de Investigaciones Biomédicas del CSIC UAM
C./ Arturo Duperier, 4, 28029 Madrid.
E-mail: rperona@iib.uam.es

Over the last 25 years, cancer research has produced an enormous amount of information that has allowed us to begin to understand the general rules that govern the transformation of normal cells into malignant ones. One of the consequences of this knowledge has been the development of new strategies for drug design based on new targets whose activity is altered in tumoral cells. The discovery of new tumor markers can help to improve diagnosis, prognosis and treatment.

The Spanish Association for Cancer Research (ASEICA) brings together scientists working in different fields of cancer research. Our association is very interested in making new information related to cancer available to professionals working in different fields of oncology. In order to accomplish this, we plan to publish a series of articles that will cover, firstly, current knowledge of the molecular bases of cell transformation and, secondly, background information on molecular biology. We hope that this information will help readers to understand how this knowledge can be translated into new tools to help the cancer patient.

MOLECULAR BASIS OF CELL TRANSFORMATION

Tumorigenesis in humans is a multistep process that is the result of genetic alterations that contribute to the progressive transformation of normal cells into cancer cells. The basic concept in cancer is that tumorigenic cells have defects in the regulatory circuits that control cell proliferation and homeostasis.

Although current information indicates that these circuits represent a complicated network, all cancer cells invariably show changes in a discrete and common group of pathways that dictate malignant growth. Tumor cells become independent of external signals for growth by developing autonomous signals for positive regulation of these pathways. They are insensitive to external signals that inhibit growth or induce differentiation. Tumors develop ways to escape from the induction of apoptosis, and transformed cells acquire genetic instability that generates further genetic alterations. Tumor cells are characterized by induction of angiogenesis and metastatic potential and by the acquisition of unlimited replicative potential (fig. 1).

The order in which these alterations occur in the generation of tumor cells can vary from one tumor cell to another; as a consequence, the relative importance of these changes is very much dependent on the tissue of origin.

MOLECULAR BIOLOGY TECHNIQUES USED IN ONCOLOGY

The word «cancer» covers more than 100 different diseases, and physicians and cancer researches have learned that the more information that can be obtained in order to perform an accurate estimate of the disease, the more beneficial it will be for the patient. Sometimes, drugs that work for one patient do not work the same for another patient, even if they have been diagno-

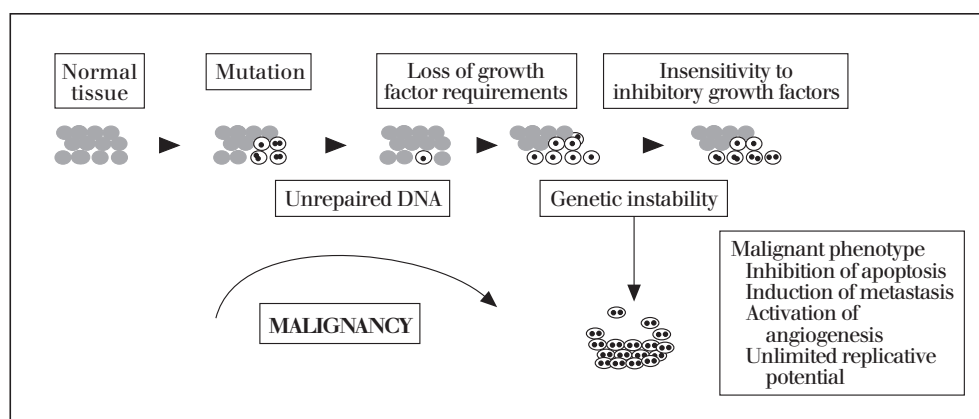


Fig. 1. Pathways that dictate malignant growth.

sed as having the same type of tumor. Additionally, with the information available at present, it is difficult to know if a patient will develop resistance to a certain drug, which makes it difficult to improve clinical results. With all the new information available and all the new tools generated by molecular biologists, it will become possible to draw up molecular profiles for individual patients and develop individualized therapy. Some mutations have been related to resistance to specific drugs, and polymorphisms can be associated with a different prognosis. Finally, mutations in proteins such as p53 have a different role in drug response depending on the specific site of the mutation in the cellular gene. Now researchers have developed a powerful tool to monitor simultaneously the expression of thousand of genes within a tumor, based on microarray technology. This technology has already been applied to the molecular classification of acute leukemias. It has been demonstrated that it is possible to distinguish acute myeloid leukemia from acute lymphoblastic leukemia without previous biological knowledge. The technique can also be used to predict the response of the tumor to different drugs, thus making chemotherapy more rational.

In order to make information on this technology available to clinical oncologists, we will begin in this article by establishing the structural bases of the biological material used in these techniques: DNA or RNA.

The genomes of all organisms are composed of DNA (deoxyribonucleic acid), and that of some viruses is composed of RNA (ribonucleic acid). Both types of molecules are polymers of linear unbranched chains of monomeric units called nucleotides. Each nucleotide has three components: a phosphate group, a sugar and a base. If the molecule is DNA, the sugar is 2'-deoxyribose, and if it is RNA, the sugar is ribose. For DNA the bases are adenine (A), cytosine (C), guanine (G) and thymine (T), and for RNA the fourth base is uracil (U). Nucleotides are linked to one another by phosphodiester bonds to form a DNA or RNA polymer which might be several millions of nucleotides in length in the case of DNA and no more than a few thousands for RNA. In the cells, DNA is double stranded and two polynucleotides are wound around one another to form a double helix. The base components of the nucleotides form hydrogen bonds that hold together the double helix in the two strands. The base pairing is A with T and C with G, making the two strands complementary. RNA is usually single stranded, although base pairing can occur between some parts of the single chain (fig. 2).

The biological information contained in the genome is encoded in the sequence of nucleotides of the DNA and RNA molecules into units called genes. This information is read by proteins with enzymatic activity that attach to the DNA and start a series of biochemical reactions that are known as gene expression. This

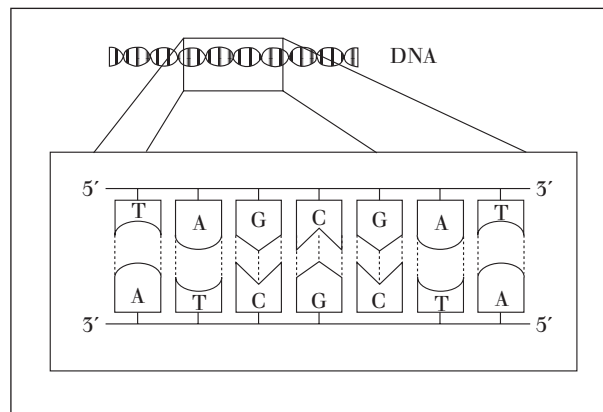


Fig. 2. Complementarity of bases in DNA.

process has two steps, transcription and translation (fig. 3). Transcription produces a copy of RNA, while translation, synthesizes a protein whose aminoacid composition is determined via the genetic code, by the nucleotide sequence of the transcribed RNA molecule. Each time that a cell divides, a copy of the DNA is made. This replication of the DNA has to be extremely accurate in order to avoid the introduction of mutations in the genome. Sometimes, mutations are installed in the DNA sequence due either to physical or chemical mutagens or to replication errors. A complicated network of DNA repair enzymatic complexes repairs the DNA, and defects in expression of these genes can contribute to the generation of tumor phenotype.

The human genome is physically divided in two components:

1. The mitochondrial genome, which is a circular DNA molecule of 16,569 base pairs located in the mitochondria.
2. The nuclear genome, which comprises approximately 3×10^9 base pairs of DNA.

Until the complete sequence of the DNA is known, the general agreement is that it contains between 80,000 and 100,000 genes.

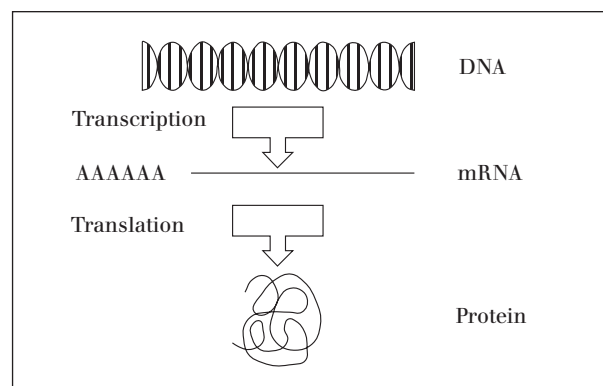


Fig. 3. Flux of genetic information.

Other types of elements can also be found within the human genome:

1. Pseudogenes, are basically a non-functional copy of a gene that contains a mutation that negatively affects the reading of the information.

2. Repetitive elements, which appear in many places of the genome, belong to four main groups:

a) LINE: long interspersed nuclear elements.

b) SINE: short interspersed nuclear elements.

c) LTR: long terminal repeats.

d) DNA transposons.

3. Microsatellites contain a short motif repeated in tandem. Many microsatellites are polymorphic, and the number of repeats varies among individuals. These elements are useful markers of the genome.

4. Nongenic, nonrepetitive single copy DNA of unknown function makes up more than 50% of the genome.

Genes are transcribed into an RNA molecule called messenger RNA, which is translated into a protein. Two other types of RNA, ribosomal RNA and transfer RNA, are not translated and have different functions in the cell.

Each gene has different parts:

1. ORF (open reading frame) is the part of the gene that codifies for a protein. Each triplet of nucleotides is a codon specific for a different aminoacid following the rules of the genetic code. The ORFs are read in a 5-3' direction along the messenger RNA. Each ORF starts in an initiation codon and finishes in a termination codon. The part of the messenger RNA located upstream (right to the 5' of the ORF) of the ORF is called the leader sequence and after ORF is the non translated sequence.

2. Many eucaryotic genes are discontinuous and are divided into regions called exons (the part of the gene that codifies for a protein) and introns (the non-codifying region). Introns are processed and eliminated from the primary or immature RNA, producing a mature RNA molecule.

References

- Genomes TA. Brown. Biosh Scientific Publishers, 1999.
Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JA. Molecular biology of the cell. Garland Publishing Inc, 1994.
DeVita V, Hellman S, Rosenberg SA. Cancer, principles & practice in oncology. Philadelphia: JB Lippincott Company, 1995.