Seminars in Colon and Rectal Surgery 27 (2016) 172–175

Contents lists available at ScienceDirect

Seminars in Colon and Rectal Surgery

journal homepage: www.elsevier.com/locate/yscrs

Molecular genetics of colorectal cancer

James Church, MB,ChB, FRACS, FACS, FACG*

Department of Colorectal Surgery, Center for Hereditary Colorectal Neoplasia, Digestive Disease and Surgery Institute, Cleveland Clinic, 9500 Euclid Ave, A30, Cleveland, OH 44195

A B S T R A C T

Colorectal cancer is the end result of an accumulation of destabilizing mutations and other genetic events, which occur in clones of colonocytes over many years. While each colorectal cancer is genetically unique, there are at least three distinct mechanisms by which the process occurs. The commonest is chromosomal instability, producing microsatellite stable, aneuploid cancers. The second is DNA promoter methylation that underlies CpG island methylation phenotype cancers, either microsatellite stable or unstable, and the third is loss of DNA mismatch repair, causing microsatellite unstable, diploid cancers. Cancers arising from these three mechanisms are biologically different and the differences have clinical implications.

Introduction

Colorectal neoplasia results from an accumulation of genetic abnormalities in the nuclei of clones of colonocytes that allow the cells to escape the normal controls on cellular growth, death, and differentiation. The causes of these genetic abnormalities include chance events that happen during cell division, elements of lifestyle, personal characteristics, and inheritance. Each colorectal cancer is genetically unique, with hundreds of different mutated genes. However, the number of driver genes, genes which when mutated drive carcinogenesis forward, is limited.[1] Despite the genetic heterogeneity displayed in colorectal cancers there are three main routes through which colorectal cancer develops.[2] These three routes produce cancers of different biology. It is important to distinguish between the routes, because the differences in biology have different clinical implications. The aim of this article is to provide an understandable account of the ways in which colorectal cancers develop at a molecular level, so that the clinical implications of differences in biology can be understood.

The basic building blocks

A gene is a sequence of nucleotide bases in DNA that codes for a specific protein. The genetic code is based on triplets of nucleotide bases, called codons, which encode a specific amino acid. A protein is a sequence of amino acids that performs a specific function in a cell. The nucleotide code is redundant as there are many more codons than amino acids. Thus, the actual DNA sequence can vary considerably yet still produce a consistent functioning protein. However, when changes in the DNA sequence result in a gene that encodes an abnormally functioning protein, a genetic disease may arise. This is an underlying process in cancer.

Put simply, cancer is uncontrolled cell growth that results from an alteration in the normal checks and balances of cellular homeostasis. In normal cellular processes, there is a balance between the proteins produced by tumor suppressor genes and proteins produced by proto-oncogenes. Inactivation of tumor suppressor genes or inappropriate activation of proto-oncogenes produces uncontrolled acceleration of cell growth that can ultimately lead to neoplasia. The most common way these genes are inactivated or activated is by mutation. A mutation is a permanent structural change in the DNA sequence. Some mutations are harmless due to the redundancy of the code, however, a deleterious mutation is one associated with a harmful change in the structure of the protein, which causes abnormal function. While a mutation is the most common way in which gene function is lost or disordered, other processes contribute to cancer development.

Epigenetic changes

An epigenetic change in gene function is brought about by an event external to the gene sequence; the gene maintains its normal sequence and yet its function is changed. The process of DNA methylation, by which methyl groups are added to the CpG dinucleotide, is a common way of controlling gene expression (the production of protein by the gene) for various normal cellular

* Corresponding author.
E-mail address: churchj@ccf.org

http://dx.doi.org/10.1053/j.scrs.2016.04.013
1043-1489/

processes. Multiple CpG clusters, or islands, tend to exist in gene promoter regions. The promoter is the span of DNA proximal to the coding region of the gene that controls it transcription. When CpG islands in the promoter region become methylated, gene expression is shut down. There is no protein produced. This is a normal way of gene expression regulation. But when this happens in a tumor suppressor gene, the loss of that tumor suppressor protein plays an important role in carcinogenesis.[3] Conversely, loss of methylation of a gene that is normally methylated can inappropriately increase expression of that gene. When this happens to an oncogene, neoplasia may result.[4] Cancers that arise via promoter methylation are termed CpG Island Methylator Phenotype or CIMP-positive cancers.

Chromosomal changes

DNA is organized into threads of nucleic acids and proteins called chromosomes. During cell division, chromosomes are pulled apart and reorganized into the daughter cells. When this process goes wrong, pieces of a chromosome, or sometimes whole chromosomes, can be deleted, transposed, or misplaced. If a segment of a chromosome is lost, all the genes on that segment are also lost. This is apparent as loss of heterozygosity (LOH), where there is only one copy of the gene present.[5] Normally, there are two copies of each gene (on the two different alleles, one from each parent) which may be different (heterogyzous). If one of a pair of different (heterozygous) alleles is lost by a chromosomal event, the apparent heterozygosity is lost. LOH is associated with aneuploid tumors (an abnormal number of chromosomes) and is a feature of chromosomal unstable colorectal cancer. Chromosomal instability in colon cancer normally follows an initial *APC* gene mutation.

MicroRNA

RNA is a single stranded chain of nucleotide bases that carries the genetic code from the DNA to the ribosome where it is translated and the protein is made. The main types of RNA include messenger RNA (mRNA) that takes the genetic code to the ribosome, transfer RNA (tRNA) that delivers the appropriate amino acid to the growing protein molecule, and ribosomal RNA (rRNA), the catalytic part of the ribosome. RNA also plays an important role in the regulation of gene expression, through microRNAs (miRNA). miRNAs are very short segments of RNA that do not code for any protein, but that interact with normal RNA and change gene expression by altering gene interactions. miRNA provides another layer in the complex biology of genetic control over protein function, and represents another opportunity for abnormalities to influence gene expression, and plays a role in carcinogenesis.[6]

The importance of DNA repair

Maintenance of the fidelity of the DNA is a highly conserved aspect of cell biology. There are multiple redundant DNA repair systems designed to prevent mistakes in replication from becoming permanent mutations in the daughter cells. Abnormal DNA repair is one of the primary causes of mutations. In addition to DNA repair function, there are multiple other cellular check points that avoid propagation of a mutation. DNA that is too distorted for repair is destroyed by the process of apoptosis or programmed cell death.[7] The key player in the coordination of cell cycle arrest, DNA repair, and apoptosis is *TP53*.[8] Mechanisms of DNA repair involved in colorectal cancer are shown in Table 1, along with the syndromes that occur when there is a germline mutation that inactivates these systems. When there is a germline mutation in a DNA repair gene, the stage is set for a "mutator phenotype," where the consequences of failed repair are thousands of mutations in multiple genes.[9–11] This leads to accelerated carcinogenesis with variable phenotypes that is characteristic of the syndromes in Table 1.

Signal transduction pathways

Although there are multiple genes mutated in colorectal cancers, there are only a few that drive the process of cancer. These genes are called driver genes, while mutated genes that do not influence the development or progression of cancer are referred to as passenger genes. Vogelstein and Kinzler suggest that only three driver gene mutations are necessary to produce colorectal cancer.[4] The first is a "gateway gene" that leads to a neoplasia "breakthrough" phase. The second driver gene leads to a neoplasia "expansion" phase, and the third enables invasion and metastasis. The driver genes vary in particular cancers, but most belong to a limited number of key signal transduction pathways, which must be inactivated for neoplasia to progress. In colorectal cancer these pathways are the wnt/wingless pathway (driver genes include *APC* and *CTNNB1*/β catenin), the EGF pathway (drivers include *KRAS* and *BRAF*), and the TGF β pathway (*SMAD4*), and p53 mediated cell cycle arrest, apoptosis and DNA repair.[12]

Mechanisms of colon carcinogenesis

There are three main molecular mechanisms by which colorectal cancer develops chromosomal instability, microsatellite instability, and methylation. Each of these results in distinct clinical phenotypes is discussed below.

Chromosomal instability (CIN) is represented by the classic "Vogelgram,"[12] beginning with an inactivating mutation in the gateway driver gene *APC* and progressing through other driver mutations in *KRAS*, *SMAD4*, and *TP53*. The *APC* mutation predisposes the cell to chromosomal instability, which leads to LOH and aneuploidy. Genetic testing of chromosomal unstable cancers shows them to be CIMP-negative and microsatellite stable (MSS).[13]

Microsatellites are short repeating nucleotide base sequences that are prone to slippage during DNA replication that creates loop mismatches. Unrepaired mismatches are apparent as "unstable" microsatellites, where the length of a particular microsatellite in a tumor is different compared to that in a normal cell. This is a feature of defective DNA mismatch repair such as that seen in Lynch syndrome, and in cancers that have lost mismatch repair

Table 1
DNA repair mechanisms and colorectal cancer.

Mechanism	Purpose	Genes	Syndromes
DNA mismatch repair	Repairs DNA mismatches	*MLH1*, *MSH2*, *PMS2*, *MSH6*	Lynch syndrome
Base excision repair	Repairs oxidative damage	*MUTYH*, *NTHL1*	MYH-associated polyposis, NTHL1-associated polyposis
DNA proofreading	Prevents mistakes in replication	*POLD1*, *POLE*	Polymerase proofreading associated polyposis
TP53 control	Coordinates cell cycle arrest, DNA repair and apoptosis	*TP53*	Li Fraumeni Syndrome