ACTA ONCOLÓGICA BRASILEIRA Acta Oncol. Bras. Vol. 24 - nº 01, 568-571, jan-fev-mar, 2004

MOLECULAR GENETICS OF COLORECTAL CANCER

WALTER F. BODMER[1], JENNY WILDING[1] AND NICOLA S. FEARNHEAD[1,2]

Cancer is essentially a somatic evolutionary process involving a series of mutations or changes in gene expression ,each of which confers a further advantage to the outgrowing tumour. This results in a clonal process of outgrowth which can sometimes be visualised using X-linked genetic markers for which, in females , only one of the two X chromosomes is expressed in any given cell. Thus tumours should then be seen as patches of tissue expressing either only one or the other of a pair of X-linked markers. The normal patch size in epithelial tissue, however, limits this approach to tumour visualisation for carcinomas(Novelli et al. 2002). The observation of functionally relevant somatic mutations in tumours now provides the most direct evidence for tumour clonality. The description of the series of genetic changes in a tumour , which is the most direct definition of its evolution , is now possible through the development of the extraordinarily powerful techniques of molecular genetics. Colorectal tumours have been a particularly good model for this analysis because of the accessibility of colonic tumours at various stages of their progression and because of the range of inherited predispositions to colorectal cancer which is found in Human populations.

Familial adenomatous polyposis (FAP), other polyposis syndromes and hereditary non-polyposis colorectal cancer (HNPCC) account for about 5% of all colorectal cancers (CRC).Other less well defined hereditary factors may then contribute to an estimated further 20% of CRC.

Familial adenomatous polyposis (FAP) was first clearly described as a dominantly inherited Mendelian trait by Lockhart-Mummery in 1925. The first clue to localising the position of the APC gene came from identification of a patient with colorectal polyposis and mental retardation who had a deletion of the chromosomal band 5q21. Linkage analysis of families with FAP led to the mapping of the APC gene to 5q21 and , eventually , to its cloning.. It consists of 8535 base pairs that encode a 2843 amino acid multidomain protein(see Fearnhead,Britton et al. 2001 for a review). Germline mutations in the APC gene have been demonstrated in most FAP patients. The vast majority are nonsense or frameshift mutations that result in a truncated protein product with abnormal function. As expected from Knudsen's two-hit hypothesis, colorectal tumours from FAP patients nearly all harbour either additional somatic APC mutations or loss of heterozygosity at the APC locus, in addition to the original germline mutation. The type of germline APC mutation in FAP appears to determine the nature of the second somatic hit to APC. If the germline mutation occurs between codons 1194 and 1392, then there is strong selection for allelic loss of APC as the second hit in the development of a colorectal adenoma. If the germline mutation lies outside this region, the second hit in tumorigenesis is most likely to produce a truncating mutation in the somatic mutation cluster region (MCR) between codons 1286 and 1513.

Up to a third of all FAP-causing mutations occur at APC codons 1061 and especially 1309 (Beroud and Soussi 1996), although this figure is probably artificially high due to reporting bias. Other germline mutations are spread fairly evenly between codons 200 and 1600, with mutations occurring only rarely beyond codon 1600. The risk of developing specific manifestations of

1. Cancer Research U.K., Weatherall Institute of Molecular Medicine, John Radcliffe Hospital

2. Department of Colorectal Surgery, John Radcliffe Hospital, Oxford

Proofs and reprints to: 1 - Oxford University - Hertford College, Catte Street - Oxford OX1 3BW - UK

Acta Oncol. Bras. Vol. 24 - nº 01, 568-571, jan-fev-mar, 2004

FAP is often correlated to the position of the inherited APC mutation. A particularly severe phenotype is seen in patients with mutations between codons 1250 and 1464, and especially at codon 1309. Attenuated polyposis, in contrast, is usually attributed to mutations at the extreme 5' or 3' ends of the APC gene..

Knudsen's two hit hypothesis predicted that those genes , such as APC , which give rise to inherited cancer predispositions would also be genes in which somatic mutations in cancers would be found. These would be predicted by loss of heterozygosity around , for example , the APC gene. This was first observed by Solomon et al.(1987) and led eventually to the observation that mutations in the APC gene were most probably the commonest early event in colorectal tumorigenesis.

The APC protein has many well characterized functional domains and interacts with numerous other proteins. It is involved in a wide variety of cellular processes including migration, adhesion, proliferation, and even perhaps aspects of chromosome stability and cytoskeletal organization.

The intriguing interdependence between the first and second hits to the APC gene in both FAP and sporadic cancers suggests that there is strong selection for one truncating mutation within the MCR of APC. This has led to the search for mutations in other genes involved in the APC pathway and these have been found , especially in the E-Cadherin and b-catenin genes. In the former case, loss of gene expresion may often be due to promoter hypermethylation.The evidence suggests that one of the main selective advantages of truncated APC protein is loss of regulation of b-catenin turnover (Polakis 1999).

HNPCC is caused by mutations in one of the mismatch repair genes, most commonly hMLH1 and hMSH2 in about equal numbers .The condition is inherited as an autosomal dominant with an 80% lifetime risk of CRC, and an increased risk in certain other cancers , including especially endometrial cancer in women.

Individuals at risk in HNPCC kindreds are heterozygous for mutations in the mismatch repair genes and so their normal cells do not have an elevated mutation rate . Loss of DNA mismatch repair function in HNPCC therefore requires both the germline mutation and a somatic hit, so that the cell loses its ability to correct errors made during DNA replication. The most vulnerable areas to loss of mismatch repair mechanisms are poly-oligo tracts, and base pair repeats known as microsatellites. Disruption of these sequences is seen in over 90% of CRC arising in HNPCC patients, a phenomenon known as replication error (RER+) or microsatellite instability (MSI).

MSI is also seen in about 15% of sporadic CRC. However in this case most of the changes in gene expression are due to epigenetic silencing of the hMLH1 gene by promoter hypermethylation. Thus, in addition to deletions and inactivating mutations, epigenetic events are now recognised as an important mechanism for gene silencing. Specifically, hypermethylation of CpG islands in the promoter region of a gene often impairs the ability of transactivating factors to bind and initiate gene transcription. Promotor hypermethlyation has now been shown to play a part in the silencing of a number of genes , including for examplle CDKN2A (p16) in some colorectal cancer cell lines and patients (Burri, Shaw et al. 2001). Recently it has been shown in our laboratory(N.Wong personal communication) that the homeobox gene, CDX1, which is thought to play a critical role in the control of the differentiation of colorectal epithelial cells,is also often silenced by promoter hypermethylation in colorectal tumours. Normal colorectal epithelium expresses CDX1 ,and so these results suggest that switching off this gene may be selected for as a mechanism for counteracting the differentiation controls acting on the epithelial cells.

Although RER- and RER+ tumours share early initiating events (involving, in particular, APC or b-catenin mutations), tumours with loss of mismatch repair probably develop along a different genetic pathway to sporadic RER- tumours. The same environmental constraints exist for both pathways and it is therefore likely that, although different sets of genes may be targeted by mutation, there will be overlap of the signaling pathways affected. For example, inactivation of transcription growth factor b (TGFb) signalling may occur via mutation within the poly A tract in the TGFb receptor II (TGFbRII) in RER+ tumours whilst mutation of the downstream signalling molecule SMAD4 may be the most efficient means to inactivate this pathway in RER- tumours (Woodford-Richens, Rowan et al. 2001)..

There has been much debate regarding the requirement for an increased mutation rate or genomic instabi-

Acta Oncol. Bras. Vol. 24 - nº 01, 568-571, jan-fev-mar, 2004

lity in the generation of cancers to account for the large number of genetic alterations present in most tumours. Two main forms of genetic instability occur in colorectal cancer. Approximately 85% of tumours are characterised by global chromosome instability resulting in anueploidy and frequent LOH. The remaining 15% of CRC harbour bi-allelic inactivation of one of the mismatch repair genes and typically retain a near diploid karyotype but exhibit features of deficient DNA mismatch repair in the form of microsatellite instability (MSI) and an increased mutation rate.

It was initially assumed that such an increased tolerance for DNA damage would allow accelerated tumour development and that this would be the main reason for selection of a mismatch repair mutation in RER+ tumours or mutations promoting chromosomal instability in replication error negative (RER-) tumours. However, such mutations would not confer any relative survival advantage to the cell in which they first occur. Recent evidence suggests that some genes implicated in genetic instability also play key roles in the regulation of apoptosis. The attendant increase in mutation rate may thus be a by-product of the primary selection against apoptosis and confer a long-term selective advantage (Tomlinson, Novelli et al. 1996).

The histological progression of colorectal cancer from adenoma to carcinoma , which was first described by Morsen and colleagues, implies a parallel genetic pathway (Figure 1).

The earliest genetic change associated with adenomatous polyps is most frequently mutation and/or loss of the APC gene. The exact sequence of commonly aquired genetic changes accumulated subsequent to inactivation of APC is variable. K-ras mutations are found in approximately 50% of colorectal cancers and are thought to be relatively early events that correlate histologically with early to late adenomas, while p53 muations mostly occur relatively late in the adenoma to carcinoma sequence..

Disruption of the TGFbIIR/SMAD4 pathway and mutations in mismatch repair genes (e.g. hMLH1, hMSH2) and cyclin dependent kinase inhibitors (e.g.

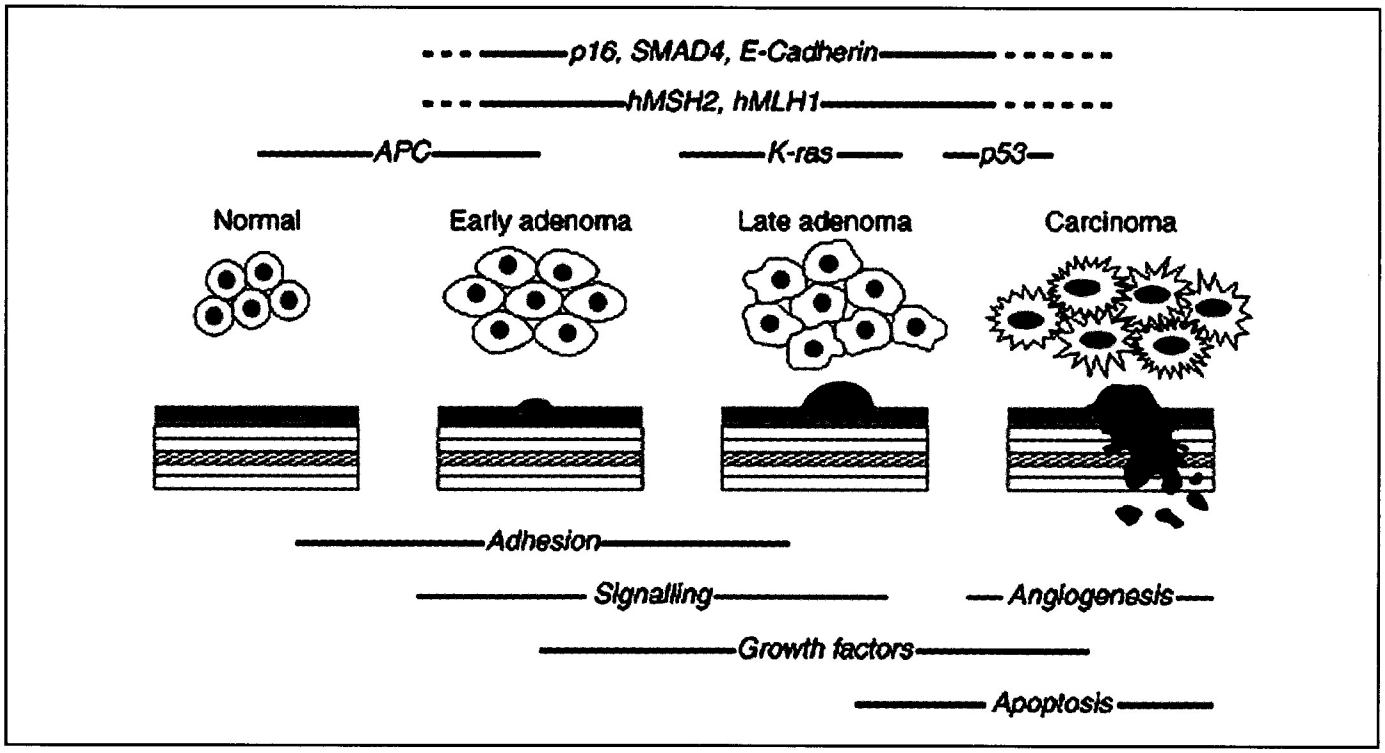

Figure 1: Basic outline of the adenoma to carcinoma sequence. The temporal order in which key genes may be affected is shown above the histological stages of disease during which they are thought to occur. Broken lines are used where the order of accumulation of genetic events is uncertain. Functional pathways affected are indicated at the bottom of the diagram.

Acta Oncol. Bras. Vol. 24 - nº 01, 568-571, jan-fev-mar, 2004

CDKN2A) have all been identified as key factors in the development and progression of colorectal cancer. However, the temporal order of disruption of these genetic pathways as they relate to histologic progression remains uncertain, and is most probably often different in different tumours. Each tumour is, after all, the result of an independent somatic evolutionary process involving a series of genetic or epigenetic changes each of which gives the tumour a further growth advantage.

Apart from the classical autosomal dominant diseases, the genetic mechanisms responsible for inherited predisposition to CRC are likely to include less deleterious mutations in the same genes responsible for FAP and HNPCC. The APC variant I1307K in Ashkenazi Jews and the rarer E1317Q in Caucasoids have been described in patients with multiple adenomas or a carcinoma developing at a young age. These mutations result in APC proteins with amino acid substitutions in functionally critical areas, thus apparently conferring an advantage with respect to tumour function. Incomplete penetrance of such mutations results in increased risk, of perhaps between 20% and 50%, of developing CRC, as distinct to the 70% to 100% risk associated with the clearly inherited syndrome.

Analagous variants in other candidate genes involved in the progression of the adenoma-carcinoma sequence may also carry an increased risk of developing colorectal tumours. Such genetic variability probably contributes substantially to multifactorial disease inheritance, accounting, for example, for much of the upto 20% of hereditary component in CRC beyond the clear cut Mendelian factors.

The Min mouse has a mutation in the mouse equivalent of the APC gene and so this, together with other genetic changes, can be used to construct suitable mouse models for studying therapeutic mechanisms for colorectal cancer, in particular with monoclonal antibodies. In our laboratorywe are working with a monoclonal antibody to a distinctive determinant of CEA which is appropriately expressed in CEA transgenic Min mice (Willkinson et al 2001). There is suggestive evidence that monoclonal antibody therapy works synergenistically with chemotherapy through an antibody dependent NK killing mechanism, and this can now be investigated using these mouse models.

There has long been an interest in the possibility of using identification of rare tumour cells in the blood as a means of early detection of cancers, or at least initially of monitoring recurrence after treatment. We have now shown in our laboratory that it should be possible, using some of our monoclonal antibodies, to detect one tumour cell in several million white cells and then to obtain some characterisation of such identified and isolated cells. This can be achieved either with further monoclonal antibody characterisation or using some of the newly developing techniques for genomic amplification of very small amounts of material even obtaining balanced amplification of the DNA from a single cell scraped from a microscope slide.

References

Beroud, C. and T. Soussi (1996). "APC gene: database of germline and somatic mutations in human tumors and cell lines." Nucleic Acids Res 24(1): 121-4.

Burri, N., P. Shaw, et al. (2001). "Methylation Silencing and Mutations of the p14ARF and p16INK4a Genes in Colon Cancer". Laboratory Investigation 81(2):217-229.

Fearnhead, N. S., M. P. Britton, et al. (2001). "The ABC of APC." Hum Mol Genet 10(7): 721-733.

Novelli , M. R.,Cossu,A.,Tomlinson, I.,(2003) Patch size in human colon confounds tyhe assessment of tumour clonality using X chromosome-linked markers. Proc.Natl. acad. Sci. USA, 100: 3311-3314.

Polakis, P. (1999). "The oncogenic activation of beta-catenin." Curr Opin Genet Dev 9(1): 15-21.

Solomon E, Voss R, Hall V et al Chromosome 5 allele loss in human colorectal carcinomas. Nature 1987, 328 (6131), 626-630.

Tomlinson, I.P., M.R. Novelli, W.F. Bodmer (1996). "The mutation rate and cancer". Proc Natl Acad Sciences U S A 93(25): 14800-14803.

Wilkinson R W, Ross E L ,Poulsom R etal(2001) Antibody targeting studies in a transgenic murine model of spontaneous colorectal tumours ProcNatl Acad Sci 98 : 10256-10260

Woodford-Richens, K.L., A.J. Rowan, et al. (2001). "SMAD4 mutations in colorectal cancer probably occur before chromosomal instability, but after divergence of the microsatellite instability pathway. " Proc Natl Acad Sci U S A 98(17):9719-9723.