



The molecular genetics of colorectal cancer

Iain Ewing,¹ Joanna J Hurley,² Eleni Josephides,³ Andrew Millar¹

¹Department of Gastroenterology, North Middlesex University Hospital, London, UK
²Institute of Molecular Genetics, Cardiff University
³Department of Gastroenterology, Queen's Hospital Romford, London, UK

Correspondence to

Dr Iain Ewing, Department of Gastroenterology, North Middlesex University Hospital, London N18 1QX, UK; iainewing@doctors.org.uk

Received 11 April 2013
 Revised 31 July 2013
 Accepted 4 August 2013
 Published Online First
 21 August 2013

ABSTRACT

Colorectal cancer is a common but heterogeneous disease, which arises through the accumulation of genetic mutations. Knowledge of the molecular basis of colorectal cancer has advanced at a rapid pace in recent years, reflecting progress made in the field of genomic medicine. Targeted therapies have come into mainstream use, and the exciting prospect of treatment regimens tailored to the mutation profile of individual tumours is beginning to emerge. In order to understand the development and application of the next generation of colorectal cancer treatments, it is important that gastroenterologists have a working knowledge of the pathological mechanisms that drive the disease. This review examines our current understanding of the molecular genetics of colorectal carcinogenesis.

INTRODUCTION

Colorectal cancer (CRC) is the third most common malignancy diagnosed in the UK, with over 40 000 new cases identified annually, and a lifetime risk approaching 1 in 15 for men and 1 in 19 for women.¹ CRC does not represent a single pathological entity, rather a heterogeneous group of diseases arising through various molecular pathways that influence individual susceptibility to cancer, and have the potential to determine responsiveness or resistance to antitumour agents.^{2 3}

Competency 2.C. of the 2010 Gastroenterology curriculum (box 1) states that UK trainees must have an appreciation for the pathology of CRC, including awareness of the molecular genetics of colorectal carcinogenesis, the adenoma-carcinoma sequence and the range of predisposing inherited and acquired conditions. This review summarises current concepts of the molecular basis of CRC, using examples of CRC

Box 1 Gastroenterology curriculum 2010

Competency 2.C. Intestinal disorders: large intestinal tumours

- ▶ Knows the pathology of benign and malignant tumours of the colon and rectum
- ▶ Has awareness of the molecular genetics of colorectal carcinogenesis and the adenoma-carcinoma sequence
- ▶ Knows the range of predisposing conditions, including inherited syndromes and acquired colonic diseases

syndromes caused by germline mutations to illustrate the effects of acquired somatic mutations. Therapeutic targets within the signalling pathways that drive CRC tumourigenesis are also explored.

THE ADENOMA CARCINOMA SEQUENCE

CRC arises as a result of the accumulation of genetic and epigenetic mutations, which transform normal glandular epithelial cells into benign neoplasms (adenomas) and subsequently into invasive carcinomas.^{3 4} Progression of tubulovillous and tubular adenomas has long been recognised, but there is also evidence that serrated adenomas have the potential for malignant transformation.⁴ This represents an alternative pathway for carcinogenesis, where a subset of hyperplastic polyps progress to serrated adenomas and ultimately a smaller proportion to carcinomas.⁴

The National Polyp Study⁵ carried out in the USA between 1980 and 1990 provided proof of the concept of malignant transformation of colorectal adenomas to adenocarcinoma, and the evidence base



Open Access
 Scan to access more
 free content

To cite: Ewing I, Hurley JJ, Josephides E, et al. *Frontline Gastroenterology* 2014;**5**: 26–30.

Table 1 Examples of genomic instability in CRC

Nature of genomic instability	Example gene mutations	Syndrome associated with germline mutation	Comments
Chromosomal instability	Loss of function mutation of <i>APC</i> gene	Familial adenomatous polyposis	Somatic <i>APC</i> mutation found in 85% of sporadic CRC ⁴
Microsatellite instability	Mismatch repair genes <i>MLH1</i> , <i>MSH2</i> , <i>MSH6</i> and <i>PMS2</i>	Lynch syndrome	Somatic inactivation of mismatch repair genes found in 15% sporadic CRC ³
DNA base excision repair defect	<i>MYH</i> gene	<i>MYH</i> -associated polyposis	No recognised somatic equivalent

to support the effectiveness of removing adenomatous polyps at colonoscopy. In this study, the incidence of CRC in a cohort of 1418 patients who had undergone colonoscopy with polypectomy was compared with reference groups including two cohorts in whom colonic polyps were left in situ. Colonoscopic polypectomy resulted in a significantly lower than expected incidence of CRC.

GENOMIC INSTABILITY

Loss of genomic stability facilitates the acquisition of multiple mutations that drive the development of CRC.³ Genomic instability can take a number of forms (table 1), including chromosomal instability (CIN), microsatellite instability, aberrant DNA methylation and DNA repair defects.^{3 4} Genome-wide analysis of gene mutations in CRCs has identified acquired somatic mutations in several 100 genes, and an average of 80 mutations in any single CRC, highlighting the heterogeneity of the disease.⁶ Table 2 highlights some important genes implicated in CRC tumourigenesis.

CHROMOSOMAL INSTABILITY

CIN, defined as the presence of structural aberrations or changes in chromosome copy number, is found in up to 85% of CRCs.⁴ Loss of function of tumour-suppressor genes, including *APC*, whose normal function is to oppose tumorigenesis, has been implicated in the development of CIN.⁷ *APC* regulates spindle microtubules, and is required to detect misaligned chromosomes during mitosis. The significance of this role in maintaining mitotic fidelity is highlighted by the CIN observed in CRC bearing *APC* mutations.⁷

Loss of function of the *APC* gene is further illustrated by the autosomal dominant condition, familial adenomatous polyposis (FAP), in which hundreds to thousands of adenomatous colonic polyps develop, leading to almost 100% lifetime risk of developing CRC in the absence of pre-emptive colectomy.^{12 13} FAP is the consequence of a germline mutation in the *APC* gene, which gives rise to a non-functional truncated protein, leading to accumulation of β -catenin and unregulated expression of a number of genes that drive colorectal tumorigenesis.¹⁴

Table 2 Examples of gene mutations implicated in CRC

Gene or group of genes	Description	Mechanism for mutation increasing CRC risk	Notes
<i>APC</i>	Tumour suppressor gene	Inactivating mutation causes loss of regulation of spindle microtubules during mitosis ⁷	<i>APC</i> mutations cause chromosomal instability
<i>TP53</i>	Tumour suppressor gene	Inactivating mutation causes loss of regulation of cell-cycle arrest and cell death ³	Inactivation may coincide with malignant transformation of adenomas ³
<i>RAS</i>	Oncogene	Activating mutations drive cell growth through MAPK pathway ⁴	<i>KRAS</i> mutation occurs as early event in adenoma-carcinoma sequence: concordance of primary tumour and metastases ⁸
<i>BRAF</i>	Oncogene	Activating mutations drive cell growth through MAPK pathway ⁴	
<i>PIK3CA</i>	Oncogene	Activating mutation upregulates PI3 K pathway, enhancing prostaglandin E2 synthesis and inhibiting apoptosis ⁹	Aspirin is a novel therapeutic agent for mutated <i>PIK3CA</i> tumours ⁹
<i>MLH1</i> , <i>MSH2</i> , <i>MSH6</i> , <i>PMS2</i>	MMR genes	Inactivating mutation impairs ability to repair strand slippage within nucleotide repeats ³	MMR gene mutations cause microsatellite instability
<i>EPCAM</i>	Codes for transmembrane glycoprotein epithelial cell adhesion molecule	Deletion of 3' end of <i>EPCAM</i> leads to epigenetic silencing of <i>MSH2</i> ¹⁰	Novel cause of Lynch syndrome
<i>MYH</i>	Base excision repair gene	Germline inactivating mutation of <i>MYH</i> leads to somatic mutation of <i>APC</i> ¹¹	Somatic mutations of <i>MYH</i> not described

CRC, colorectal cancer; MAPK, mitogen-associated protein kinase; MMR, mismatch repair.

MICROSATELLITE INSTABILITY

Microsatellites are mononucleotide or dinucleotide repeats found throughout the entire genome. Their repetitive nature makes them vulnerable to transcription errors during replication. Microsatellite unstable tumours are distinct from those with CIN as they display a normal karyotype.⁴ The mechanism of tumourigenesis in microsatellite instability involves inactivation of genes responsible for DNA mismatch repair (MMR) through somatic mutation or aberrant methylation.⁴ This loss of MMR gene function and resulting inability to repair strand slippage within nucleotide repeats changes the size of microsatellites.³ This is of particular importance if the microsatellite lies within the coding region of a gene as it may lead to altered gene function or a change in the protein product of gene expression.¹² Somatic inactivation of MMR genes is found in approximately 15% of cases of sporadic CRC. There are associations with older age, female sex and proximal distribution of these tumours.³

Germline mutations of MMR genes are responsible for Lynch syndrome, or hereditary non-polyposis colorectal cancer (HNPCC), which carries a lifetime risk of CRC of about 80%.^{3 13} Mutations leading to loss of function have been identified in four genes involved in MMR: *MLH1*, *MSH2* (accounting for the majority of cases), *MSH6* and *PMS2*.^{12 13} Lynch syndrome is the most common hereditary CRC syndrome accounting for 2%–3% of all cases.^{12 10} Acceleration of the adenoma to carcinoma sequence is seen relative to sporadic CRC, with cancers evident at a median age of around 45 years.^{3 12} Affected individuals are also at increased risk of developing extra-colonic malignancy, in particular, endometrial and ovarian cancers.¹⁴ The inheritance of Lynch syndrome is autosomal dominant. Affected individuals carry a germline mutation in a single copy of a MMR gene. This alone is not thought to account for the observed increased risk of CRC, which occurs only when somatic mutation has affected the remaining wild-type parental allele.³

Germline deletion mutations in the *EPCAM* gene have recently been identified as a novel cause of Lynch syndrome. The mechanism is disruption of the 3' end of *EPCAM*, which leads to epigenetic silencing of the neighbouring *MSH2* MMR gene.¹⁰

The correct identification of patients with Lynch syndrome is clinically relevant as it allows for targeted CRC surveillance for the index case and family members. A definitive molecular diagnosis can be made by germline mutation analysis of the four DNA MMR genes implicated in the pathogenesis of Lynch syndrome.⁴ This process is expensive and a more pragmatic approach is to test for loss of MMR gene products by immunohistochemistry and for microsatellite instability (MSI) using PCR.⁴

ABERRANT DNA METHYLATION

Aberrant methylation of DNA is a further mechanism of gene silencing in patients with CRC that can lead to loss of MMR function.^{3 6} Methylated cytosine is incorporated in the normal genome, representing a fifth DNA base. It occurs outside of exons within CpG dinucleotides, but is largely absent from CpG-rich islands in the promoter regions of approximately half of all genes.^{3 15} In the CRC genome, there is aberrant methylation within promoter-associated CpG islands, leading to silencing of gene expression.^{3 15} Hypermethylation of promoters containing CpG islands is known as the CpG island methylator phenotype (CIMP).⁴ This phenomenon is observed in about 15% of CRCs, most of which show loss of *MLH1* expression resulting in MMR deficiency and microsatellite instability.^{3 16}

DNA BASE EXCISION REPAIR GENES

The *MYH* gene is a base excision repair gene, responsible for repairing DNA damaged by reactive oxygen species.^{12 17} Polyposis develops in the presence of germline mutation of both *MYH* alleles.³ The resulting clinical syndrome, *MYH*-associated polyposis is therefore autosomal recessive.¹⁷ The mechanism of disease following germline inactivation of *MYH* is via subsequent somatic mutation of the *APC* gene causing CIN.¹¹ The risk of CRC approaches 100% by age 60.^{3 12} Thus far only germline inactivating mutations of *MYH* are recognised, with no somatic equivalent. The diagnosis should be suspected in individuals with greater than 15 colonic adenomas, and can be confirmed by genetic testing.³

TUMOUR SUPPRESSOR GENES

Somatic mutations resulting in loss of function of the *APC* gene are the most commonly observed tumour suppressor gene defects in sporadic CRC.^{3 4 14} Other important examples include loss of *TP53* and *TGFβ* function.

TP53 is a key tumour suppressor gene that is mutated in about half of all CRC.⁴ The wild-type p53 protein has a regulatory role in mediating cell-cycle arrest and cell death.³ Inactivation of the *TP53* gene often coincides with malignant transformation of adenomas.^{3 4} The detection of *TP53* mutation currently does not have any prognostic or clinical significance.⁴

Transforming growth factor β (TGFβ) signalling is an important tumour suppressor pathway. Deregulation of this pathway is a frequent observation in CRC, mediated by inactivating mutations of receptor genes (*TGFBR1*, *TGFBR2*) or postreceptor signalling pathway genes (*SMAD2*, *SMAD4*).⁴ Mutation of the TGFβ receptor genes commonly occurs in association with malignant transformation, and is seen in tumours with microsatellite instability.^{3 4} *SMAD4*

deletion has been shown to be associated with malignant transformation in murine models, and loss of expression correlates with lymph node metastases and possibly prognosis in human CRC.⁴

Juvenile polyposis syndrome (JPS) is a rare autosomal dominant disease. It carries an increased risk of development of gastrointestinal cancers. There is some discrepancy in the reported lifetime risk of developing CRC, perhaps reflecting the rarity of the condition. One relatively large registry reported that 14% of patients developed gastrointestinal cancer either by the time of diagnosis or during surveillance.¹⁸ A number of germline mutations, ultimately leading to downregulation of TGF β signaling, have been reported, including inactivating mutations of *SMAD4*.¹²

ONCOGENES

Fundamental cellular activities including differentiation, proliferation and apoptosis, are mediated through intracellular signalling pathways. Oncogenic mutation of genes responsible for controlling these pathways can lead to loss of cellular regulation and subsequent development of invasive, immortal cancer cells.⁴ Examples of such pathways exhibiting oncogenic mutations in CRC include the epidermal growth factor receptor (EGFR), mitogen-associated protein kinase (MAPK) pathway and the phosphatidylinositol 3-kinase (PI3 K) pathway.^{3 4 19}

EGFR activation triggers an intracellular phosphorylation cascade through downstream effectors RAS and BRAF, amplified through the MAPK pathway to promote cell growth.⁴ RAS and BRAF are implicated as oncogenes in a number of human cancers. Activating mutations promoting CRC have been identified in both genes.¹⁹ Mutations in *KRAS* are found in about 40% of CRCs, occurring as a relatively early event in the adenoma-carcinoma sequence.^{4 8} This is clinically relevant as there is concordance between the *KRAS* mutation status of primary tumour and metastases. Genetic analysis of tissue from the colorectal lesion can therefore predict response to targeted therapy in metastatic disease. This has been shown in trials of cetuximab, an immunoglobulin G1 monoclonal antibody against EGFR, which reduces the risk of progression of metastatic CRC, an effect limited to patients with *KRAS* wild-type tumours.⁸

Mutations of the *PIK3CA* gene, leading to upregulation of PI3 K signalling, are present in approximately 15%–20% of CRCs. Resulting enhanced prostaglandin E2 synthesis inhibits apoptosis of CRC cells.⁹ Aspirin may block the PI3 K pathway. Use of aspirin after diagnosis of CRC has been shown to significantly increase survival among patients with mutated *PIK3CA* tumours, in contrast to those with wild-type *PIK3CA*, who do not benefit.⁹ These findings suggest a role for the use of *PIK3CA* mutation status as a biomarker for targeted adjuvant therapy.

CONCLUSIONS

CRC are complex and heterogeneous solid tumours, exhibiting multiple genetic mutations. Individuals may carry predisposing germline mutations and accumulate further somatic mutations at various stages in the transition from normal mucosa, through adenomatous polyp, to invasive cancer.

Enhanced understanding of the molecular basis of CRC has led to new insights into the pathogenesis of familial forms of the disease, and how these relate to the accumulation of somatic mutations in sporadic tumours. Genetic testing for patients at high-risk of germline mutations, for example, testing for MSI in Lynch syndrome, has led to targeted surveillance for CRC that extends to at-risk family members.

Genetic biomarkers that predict response to treatment are beginning to come into routine practice in the management of CRC. *KRAS*-mutational testing to guide anti-EGFR therapy is one of the first examples of individualised targeted cancer therapy, and illustrates how molecular analysis of CRC tissue can improve outcomes by directing therapy to the most appropriate patients.⁴

The introduction of targeted therapies has already had an impact on the management of metastatic CRC. Cetuximab, an EGFR-blocking monoclonal antibody, reduces progression of *KRAS* wild-type metastatic CRC, and can lead to more curative resections of liver metastases.²⁰

The pace of recent advances in our understanding of the molecular basis of CRC and the success of the first wave of targeted treatments provides an optimistic outlook for the future management of CRC. The armoury of specific drugs designed to inhibit oncogenes and signalling pathways is expanding.⁴ There is a real hope that the evolving application of molecular techniques to diagnosis, risk-stratification and management of CRC will translate to reduced disease burden in the future.

Multiple choice questions

1. A 64-year-old man undergoes colonoscopy as part of the national CRC screening programme. An exophytic adenocarcinoma is found in the ascending colon. Staging CT of the chest, abdomen and pelvis is performed, which demonstrates the primary tumour and multiple liver lesions. Following review in the CRC multidisciplinary meeting, the diagnosis is confirmed as CRC with liver metastases that are not amenable to surgical resection.

When considering cetuximab in the adjuvant treatment of metastatic CRC, which of the following statements is most accurate?

- i. *KRAS* mutation testing should be performed on samples from both primary and metastatic tumours

- ii. Cetuximab should be considered alongside conventional chemotherapy
- iii. Cetuximab should be considered if the *KRAS* gene is wild-type on genetic testing of the primary tumour
- iv. Cetuximab should be considered if the *KRAS* gene is mutated on genetic testing of the primary tumour
- v. Cetuximab should be considered if the *KRAS* gene is mutated on genetic testing of both primary and metastatic tumours

Answer: iii) Cetuximab should be considered if the *KRAS* gene is wild-type on genetic testing of the primary tumour.

Mutations of *KRAS* occur relatively early in the adenoma to carcinoma sequence, and there is good concordance in the mutation status of primary and metastatic disease.⁴ Genetic testing of *KRAS* can therefore be performed on the primary tumour only. Cetuximab has been shown to reduce the risk of progression of metastatic CRC, but this effect is limited to patients with *KRAS* wild-type tumours.⁸

2. A 45-year-old woman undergoes colonoscopy because of a family history of CRC affecting her father aged 43 and brother aged 47, and endometrial cancer in a paternal aunt. At colonoscopy, she is found to have four adenomatous polyps and an adenocarcinoma of the sigmoid colon. Subsequent genetic testing confirms the diagnosis of Lynch syndrome (HNPCC).

Which of the following statements is the most accurate?

- i. The inheritance of this condition is autosomal recessive
- ii. Genetic testing will confirm multiple germline mutations of mismatch repair genes
- iii. The median age of development of CRC is 45
- iv. The genetic defect is an example of CIN
- v. Affected individuals are usually found to have hundreds of colonic polyps

Answer: iii) The median age of development of CRC is 45

Lynch syndrome is autosomal dominant, caused by germline mutation of a single mismatch repair gene. Large numbers of colorectal polyps are not characteristic of this condition. Affected individuals are at increased risk of extracolonic malignancy, particularly ovarian and endometrial cancers. The genetic defect is an example of microsatellite rather than CIN. The median age of development of CRC is 45.

Contributors IE wrote the article and prepared the manuscript. JH collaborated on the content of the review and editing of the manuscript. EJ contributed text on inherited polyposis syndromes. AM is supervising senior author.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non

Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>

REFERENCES

- 1 Cancer Research UK, Bowel cancer incidence statistics. 2011.
- 2 Yamauchi M, Lochhead P, Morikawa T, *et al.* Colorectal cancer: a tale of two sides or a continuum? *Gut* 2012;61:794–7.
- 3 Markowitz SD, Bertagnolli MM. Molecular origins of cancer: Molecular basis of colorectal cancer. *N Engl J Med* 2009;361:2449–60.
- 4 Pritchard CC, Grady WM. Colorectal cancer molecular biology moves into clinical practice. *Gut* 2011;60:116–29.
- 5 Winawer SJ, Zauber AG, Ho MN, *et al.* Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. *N Engl J Med* 1993;329:1977–81.
- 6 Casey G, Conti D, Haile R, *et al.* Next generation sequencing and a new era of medicine. *Gut* 2011;62:920–32.
- 7 Caldwell CM, Kaplan KB. The role of APC in mitosis and in chromosome instability. *Adv Exp Med Biol* 2009;656:51–64.
- 8 Van Cutsem E, Kohne CH, Hitre E, *et al.* Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. *N Engl J Med* 2009;360:1408–17.
- 9 Liao X, Lochhead P, Nishihara R, *et al.* Aspirin use, tumor PIK3CA mutation, and colorectal-cancer survival. *N Engl J Med* 2012;367:1596–606.
- 10 Ligtenberg MJ, Kuiper RP, Chan TL, *et al.* Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1. *Nat Genet* 2009;41:112–17.
- 11 Colebatch A, Hitchins M, Williams R, *et al.* The role of MYH and microsatellite instability in the development of sporadic colorectal cancer. *Br J Cancer* 2006;95:1239–43.
- 12 Kastrinos F, Syngal S. Inherited colorectal cancer syndromes. *Cancer J* 2011;17:405–15.
- 13 Woods MO, Younghusband HB, Parfrey PS, *et al.* The genetic basis of colorectal cancer in a population-based incident cohort with a high rate of familial disease. *Gut* 2010;59:1369–77.
- 14 Al-Sohaily S, Biankin A, Leong R, *et al.* Molecular pathways in colorectal cancer. *J Gastroenterol Hepatol* 2012;27:1423–31.
- 15 Issa JP. CpG island methylator phenotype in cancer. *Nat Rev Cancer* 2004;4:988–93.
- 16 Donehower LA, Creighton CJ, Schultz N. MLH1-silenced and non-silenced subgroups of hypermutated colorectal carcinomas have distinct mutational landscapes. *J Pathol* 2013;229:99–110.
- 17 Nieuwenhuis MH, Vogt S, Jones N, *et al.* Evidence for accelerated colorectal adenoma–carcinoma progression in MUTYH-associated polyposis? *Gut* 2012;61:734–8.
- 18 Latchford AR, Neale K, Phillips RK, *et al.* Juvenile polyposis syndrome: a study of genotype, phenotype, and long-term outcome. *Dis Colon Rectum* 2012;55:1038–43.
- 19 Ogino S, Chan AT, Fuchs CS, *et al.* Molecular pathological epidemiology of colorectal neoplasia: an emerging transdisciplinary and interdisciplinary field. *Gut* 2011;60:397–411.
- 20 Arnold D, Seufferlein T. Targeted treatments in colorectal cancer: state of the art and future perspectives. *Gut* 2010;59:838–58.