Exposing the MYtH about base excision repair and human inherited disease

Jeremy P. Cheadle* and Julian R. Sampson

Institute of Medical Genetics, University of Wales College of Medicine, Heath Park, Cardiff CF14 4XN, UK

Received June 19, 2003; Accepted July 29, 2003

Base excision repair (BER) protects against damage to DNA from reactive oxygen species, methylation, deamination, hydroxylation and other by-products of cellular metabolism. Until last year, inherited deficiencies in the BER pathway had not been causally linked with any human genetic disorder. An apparent explanation was functional redundancy between proteins in this and other pathways. However, it was recently discovered that biallelic mutations in the BER DNA glycosylase MYH lead to an autosomal recessive syndrome of adenomatous colorectal polyposis and very high colorectal cancer risk. We review the molecular mechanism of tumourigenesis in MYH polyposis, the preliminary delineation of the MYH polyposis phenotype and the functional overlap of MYH with other repair proteins.

THE BASE EXCISION REPAIR PATHWAY

The base excision repair (BER) pathway plays a significant role in the repair of mutations caused by reactive oxygen species (ROS) that are generated during aerobic metabolism (1). BER is a multi-step process that involves the sequential activity of several proteins. DNA glycosylases initiate this repair pathway by recognizing and removing a damaged or improper base by hydrolysing the N-glycosidic bond. To date, 10 DNA glycosylases have been characterized and cloned in humans, and each excises an overlapping subset of oxidized, deaminated, alkylated or mismatched bases (2). To complete the repair process, the apurinic/apyrimidinic (AP) site is further processed by an incision step, DNA synthesis, an excision step, and DNA ligation through either the short or long-patch BER pathways. Despite the critical nature of these functions, knockout mouse models of individual glycosylases appeared to be phenotypically normal. Partial redundancy between glycosylases and an overlap with transcription coupled repair was proposed as a likely explanation (3). Furthermore, although inherited deficiencies involving components of the nucleotide excision repair, mismatch repair and recombinational repair pathway had all been linked to specific human genetic disorders, no inherited disorder of BER had been identified (3).

ESTABLISHED COLORECTAL CANCER GENES

Inherited factors are thought to play a major role in at least 15% of colorectal cancers (CRC), but established CRC predisposition genes account for only a minority of these (4). Familial adenomatous polyposis (FAP; MIM 175100) is an autosomal dominant disorder associated with the development of hundreds or thousands of colorectal adenomas, some of which progress to cancer. There are a number of associated extracolonic manifestations which include congenital hypertrophy of the retinal pigment epithelium (CHRPE), upper gastrointestinal tumours, desmoid tumours, hepatoblastoma, epidermoid skin cysts and benign osteoid tumours (Gardner’s syndrome) and cerebellar medulloblastoma (Turcot syndrome). FAP is caused by inherited mutations within the adenomatous polyposis coli (APC) gene that acts as a gatekeeper regulating proliferation of colonic cells (5). Tumours develop in patients with FAP after somatic inactivation of the wild-type APC allele in accordance with Knudson’s ‘2-hit hypothesis’, and it has recently been suggested that different combinations of APC mutation confer different growth advantages in colorectal tumours (6–8). Attenuated FAP (AFAP) is associated with smaller numbers of adenomas and is caused by germline mutations in the extreme 5’ or 3’ ends of APC or in the alternatively spliced region of exon 9 (5). Tumour development in at least some cases of AFAP appears to require somatic second and third hits of the wild-type and attenuated APC alleles (9,10).

Hereditary non-polyposis CRC (HNPCC; MIM 114500) is an autosomal dominant disorder characterized by early-onset CRC (in the absence of florid polyposis) and other extracolonic cancers, notably endometrial cancer and cancers of the stomach, small bowel, ureter and renal pelvis. HNPCC is caused by inherited deficiencies in the mismatch repair (MMR) pathway (11). Germline mutations are most frequently found in...
**MSH2** and **MLH1**, and cause a high degree of somatic microsatellite instability (MSI) in the associated colorectal tumours. Mutations in **MSH6** are less frequent and are associated with less marked MSI. For mismatch recognition, the MSH2 protein forms a heterodimer with MSH6 or MSH3 depending on whether base-base mispairs (MSH2/MSH6) or insertion-deletion loops (MSH2/MSH3 and/or MSH2/MSH6) are repaired (11). Tumour development in HNPCC requires somatic inactivation of the wild-type MMR allele, again in accordance with Knudson’s ‘2-hit’ hypothesis (11).

**ADENOMATOUS COLORECTAL POLYPOSIS AND INHERITED MUTATIONS OF MYH**

Al-Tassan et al. (12) investigated a British family (family N) in which three siblings were affected by multiple colorectal adenomas and carcinoma. Sequencing of the entire **APC** open reading frame (ORF) in constitutional DNA samples from two of the affected siblings, together with haplotype and expression analyses, excluded an inherited **APC** gene defect. Assessment for MSI in DNA extracted from 11 tumours from family N also excluded an underlying MMR defect. However, the pattern of somatic **APC** mutations in tumours from family N provided a clue as to the underlying genetic defect. Sequencing of the **APC** ORF in each of the 11 tumours revealed 18 somatic mutations, 15 of which were G:C→T:A transversions (12). This class of mutations accounts for only some 10% of reported somatic **APC** mutations, with frameshift mutations and loss of heterozygosity being the more usual classes of mutations leading to somatic inactivation of **APC** in colorectal tumours (8,12). Comparison of the findings in family N with a database of over 800 somatic **APC** mutations from sporadic and FAP-associated colorectal tumours, confirmed that the excess of G:C→T:A transversions in family N was highly significant (P = 10^{-12}).

8-Oxo-7,8-dihydro2′-deoxyguanosine (8-oxoG) is the most stable product of oxidative DNA damage (13) and readily mispairs with adenines (14), leading to G:C→T:A mutations in repair-deficient bacteria and yeast (15–18). In *Escherichia coli*, three enzymes help protect cells against the mutagenic effects of guanine oxidation (16). MutM DNA glycosylase removes the oxidized base from 8-oxoG:C base pairs in duplex DNA, MutY DNA glycosylase excises A residues misincorporated opposite unrepaired 8-oxoG during replication, and MutT, an 8-oxo-dGTPase, prevents the incorporation of 8-oxo-dGMP into nascent DNA (Fig. 1). Homologues of *mutM*, *mutY* and...
mutations. They also identified biallelic mutation of MYH in 8/107 APC-mutation-negative FAP-like cases (>100 colorectal adenomas).

The mean age at diagnosis of the 25 unrelated MYH polyposis cases reported by Sampson et al. (25) was 46 years (median 48 years, range 13–65 years; Table 1). Nine were specified as having over 100 adenomas (one had over 400), 11 had 10–100 adenomas, and in five, the adenomas were ‘multiple, too many to count’, ‘numerous’ or ‘throughout the colon’. Twelve of the 25 index cases (48%) had colorectal cancer diagnosed at a mean age of 49.7 years. The index cases had a total of 64 siblings of whom 17 (27%) were known to be affected by colorectal polyposis, consistent with autosomal recessive transmission.

Among the 14 cases with biallelic MYH mutations identified by Sieber et al. (26), the age at diagnosis ranged from 30 to 70 years, the total number of polyps ranged from 18 to 1000, and 50% had colorectal cancer. Both Sampson et al. (25) and Sieber et al. (26) reported the identification of microadenomas in the background colorectal mucosa in patients with biallelic MYH mutations. Previously, this feature had been considered pathognomonic of FAP (27). Duodenal adenomas have also been noted in several patients (as found in patients with FAP due to germline mutation of APC) and possible CHRPE noted in one. No other frequent phenotypic manifestations of biallelic MYH mutation have been reported to date.

Together, these data indicate that the colorectal phenotype of MYH polyposis may closely resemble AFAP (<100 adenomas), or FAP (100–1000 adenomas), but not severe FAP (>1000 adenomas; Table 1). We propose that this may reflect the number of somatic mutations required for initiation of adenoma development. In FAP, adenoma development requires only a single somatic APC mutation. Families with biallelic MYH mutations may be more comparable to patients with AFAP who develop smaller numbers of adenomas that require two somatic APC mutations for initiation (9,10). By contrast, most patients with HNPCC develop only one or a few adenomas or carcinomas whose initiation requires somatic inactivation of a wild-type MMR allele and two somatic APC mutations in the target cell.

### Table 1. Phenotypic characteristics of 39 apparently unrelated cases with adenomatous colorectal polyposis and biallelic mutations of MYH

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Sampson et al. (25), n = 25</th>
<th>Sieber et al. (26), n = 14</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of adenomas</td>
<td>14–400+</td>
<td>18–1000 (&lt;1000)</td>
</tr>
<tr>
<td>Mean age at presentation</td>
<td>46 years (range 13–65 years)</td>
<td>51.3 years (range 30–70 years)</td>
</tr>
<tr>
<td>Cancer at diagnosis</td>
<td>12/25 (48%)</td>
<td>7/14 (50%)</td>
</tr>
<tr>
<td>Mean age at cancer diagnosis</td>
<td>49.7 years</td>
<td>—</td>
</tr>
<tr>
<td>Excess of other cancers</td>
<td>None noted</td>
<td>None noted</td>
</tr>
<tr>
<td>Known duodenal adenomas</td>
<td>1/25</td>
<td>2/14</td>
</tr>
</tbody>
</table>

Cases reported by Sampson et al. (25) were identified from a cohort of patients with at least 10 adenomas and those reported by Sieber et al. (26) from a cohort of patients with at least three adenomas.
SOMATIC G:C → T:A MUTATIONS PRIMARILY OCCUR AT GAA SEQUENCES IN APC

Analyses of MYH-deficient colorectal tumours have shown that the two bases immediately 3' to the somatically mutated G are almost always AA and this preponderance of G:C → T:A mutations at GAA sequences is highly significant (12,22). This sequence specificity occurs irrespective of the nature of the germline MYH mutations (including very early truncating mutations which are predicted to result in a complete absence of functional MYH) (22). Recently, Chmiel et al. (28) demonstrated that wild-type MutY has a 3-fold decrease in adenine glycosylase activity on a GAA containing duplex as compared with non-GAA containing duplex. These data indicate that the GAA sequence specificity may not reflect a direct function of MYH but, more likely, improper recognition/repair by compensatory glycosylases.

A possible explanation for the apparently specific predisposition to colonic tumours in patients with MYH defects is the high level of oxidative damage affecting this organ (29). However, the prevalence of GAA target sites in APC as compared with other key genes involved in tumourigenesis may also be a factor. APC has a total of 216 GAA sites in which G:C → T:A mutations could lead to termination codons, whereas TP53, PTCH, RB1, NF1 and VHL (that are frequently mutated during tumourigenesis in the brain/breast/lung, skin, retina, Schwann cells and kidney) have only 12, 34, 61, 139 and eight sites, respectively. Although the prevalence of GAA sites in APC and high levels of oxidative damage in the gut may contribute towards the colorectal phenotype seen in association with MYH deficiency, the reasons underlying this phenotype specificity remain unclear.

THE SPECTRUM OF MYH MUTATIONS AND DIAGNOSTIC ISSUES

To-date, six truncating (252delG, Y90X, Q324X, 1103delC, E466X and 1419delC), four missense (W117R, Y165C, V232F and G382D), one in-frame insertion (137insIW) and two putative splice site mutations (347 - 1G → A and 891 + 3A → C) have been reported in MYH in patients with colorectal polyposis (12,22,25,26) (Fig. 2). Although there has been some bias in the ascertainment of MYH-deficient cases, by far the most common mutations among Caucasians are Y165C (36 mutations, 53%) and G382D (22 mutations, 32%) (25,26) (Fig. 3). Specific mutations in MYH have been identified in different ethnic populations and diagnostic screening strategies will have to be optimised accordingly. For example, E466X accounts for all (8/8) mutations so far reported in Indian cases (25).
FUNCTIONAL DOMAINS AND BIOCHEMISTRY OF mutY/MYH

Although only limited data is available on MYH due to the difficulty of expressing the protein (30–34), Parker et al. (35) have shown that MYH interacts with AP endonuclease, PCNA, and RPA, suggesting a role in long patch BER and Boldogh et al. (36) have shown an association with the replication foci, suggesting a role in replication-coupled repair. More extensive structural and biochemical information is available on MutY. The N-terminal domain of MutY (Met-1 to Lys-225) contains the catalytic region (37) and shares several motifs with other BER glycosylases, including the helix–hairpin–helix (HhH), pseudo HhH and the iron–sulfur cluster loop motif (38). MutY contains a C-terminal domain that is not found in the BER glycosylase superfamily, with sequence and structural homology to MutT (39) and the C-terminal domain of MYH correspondingly shares homology with MTH1 (40). NMR and biochemical studies have suggested that the C-terminal domain plays a role in 8-oxoG recognition (39,41,42).

MutY is a monofunctional BER glycosylase that is capable of removing adenine from 8-oxoG:A, G:A, and C:A mispairs (43). Like all DNA–nucleotide-modifying enzymes, MutY has to recognise and access chemical adducts on DNA bases hidden within the double helix of DNA (44–46). These enzymes expose their targets by rotating the phosphodiester bonds surrounding the nucleotide, causing the target base to be flipped out of the DNA helix (47–52). Crystallographic studies on the catalytic core of MutY have revealed an active site binding pocket for the extruded adenine (38), indicating that MutY uses a base-flipping mechanism, with compression of the DNA intrastrand phosphate distance by the HhH and pseudo-HhH motifs. However, MutY is unique among BER enzymes in recognizing a mismatch between a damaged 8-oxoG and a normal adenine while exclusively catalysing the removal of the undamaged base (42,53,54). On the basis of NMR studies, a double base flipping mechanism for MutY has recently been proposed in which both adenine and 8-oxoG are flipped from the helix during the repair process (55).

FUNCTIONAL OVERLAP WITH OTHER REPAIR PATHWAYS

Ni et al. (56) demonstrated that Msh2p/Msh6p in Saccharomyces cerevisiae bind to 8-oxoG:A mismatches and repair 8-oxoG lesions. However, because S. cerevisiae does not contain MutY and MutT homologs, it was initially unclear whether MSH2/MSH6 played a similar role in other organisms. This was recently resolved by Mazurek et al. (57), who showed that human MSH2/MSH6 were activated upon recognition of 8-oxoG. Furthermore, Gu et al. (58) demonstrated that MYH interacts with the MSH2/MSH6 heterodimer via MSH6, and MSH2/MSH6 stimulates the DNA binding and glycosylase activities of MYH with an 8-oxoG:A mismatch. Because both MYH and MSH6 interact with PCNA and co-localize to the replication foci, PCNA may act as a co-ordinator of both repair pathways (35,36,58–61). Therefore, MYH-mediated BER may co-operate with MMR in protecting against the mutagenic effects of 8-oxoG. Other repair pathways have also been implicated in the repair of 8-oxoG: the Cockayne syndrome B gene product may be required for general genome repair (62), and BRCA1 and BRCA2 required for transcription coupled repair (63).
CONCLUDING REMARKS

Although MYH polyposis appears to be transmitted as an autosomal recessive trait, the risk of colorectal adenoma and carcinoma in heterozygotes has not been determined. Neither is it clear whether inherited mutations in other components of the 8-oxoG repair system, or other BER glycosylases, predispose to tumours in humans. Whatever the outcome of investigations to address these questions, the myth that inherited defects of BER lack phenotypic consequences has now been firmly dispelled.

ACKNOWLEDGEMENTS

We thank Professor Sheila David for helpful comments and Dr Nada Al-Tassan for help with preparing the figures. J.P.C. and J.R.S. are supported by Cancer Research UK, Tenovus, CETIC (W.D.A.), KEF and the Wales Gene Park.

REFERENCES


