Published in final edited form as:

Fam Cancer. 2011 June ; 10(2): 297–301. doi:10.1007/s10689-011-9427-0.

Mutation deep within an intron of *MSH2* causes Lynch syndrome

Mark Clendenning,

Familial Cancer Laboratory, Queensland Institute of Medical Research, 300 Herston Road, Herston, QLD, Australia

Daniel D. Buchanan,

Familial Cancer Laboratory, Queensland Institute of Medical Research, 300 Herston Road, Herston, QLD, Australia

Michael D. Walsh,

Familial Cancer Laboratory, Queensland Institute of Medical Research, 300 Herston Road, Herston, QLD, Australia

Belinda Nagler,

Familial Cancer Laboratory, Queensland Institute of Medical Research, 300 Herston Road, Herston, QLD, Australia

Christophe Rosty,

Familial Cancer Laboratory, Queensland Institute of Medical Research, 300 Herston Road, Herston, QLD, Australia

Bryony Thompson,

Molecular Cancer Epidemiology, Queensland Institute of Medical Research, Herston, QLD, Australia

Amanda B. Spurdle,

Molecular Cancer Epidemiology, Queensland Institute of Medical Research, Herston, QLD, Australia

John L. Hopper,

Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, Melbourne School of Population Health, The University of Melbourne, Parkville, VIC, Australia

Mark A. Jenkins, and

Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, Melbourne School of Population Health, The University of Melbourne, Parkville, VIC, Australia

Joanne P. Young

Familial Cancer Laboratory, Queensland Institute of Medical Research, 300 Herston Road, Herston, QLD, Australia

Mark Clendenning: mark.clendenning@qimr.edu.au

Correspondence to: Mark Clendenning, mark.clendenning@qimr.edu.au.

Electronic supplementary material The online version of this article (doi:10.1007/s10689-011-9427-0) contains supplementary material, which is available to authorized users.

Abstract

Lynch syndrome, a heritable form of cancer predisposition, is caused by germline mutations within genes of the DNA mismatch repair family, and can be rapidly identified in young onset cancer patients through the detection of loss of expression of at least one of these genes in tumour samples. To date, such causative mutations have only been identified within exonic and splice site regions. Though this approach has been successful in the majority of families, a considerable number remain in which no mutation has been found. To address this situation, we used an alternative mutation discovery procedure which involved haplotype analysis of the locus containing the gene lost in the tumour and delineation of segregating haplotypes, followed by an investigation of splicing aberrations to uncover cryptic splice sites which lay outside the genomic regions routinely examined for mutations. In this report, we show that an intronic mutation 478 bp upstream of exon 2 in the *MSH2* gene causes Lynch syndrome through creation of a novel splice donor site with subsequent pseudoexon activation, thus highlighting the need for more extensive sequencing approaches in families where routine procedures fail to find a mutation.

Keywords

Lynch syndrome; *MSH2*; Novel mutation mechanism

Introduction

Colorectal cancer (CRC) is the 3rd most common form of cancer in the United States and has the third highest morbidity [1]. Heritable forms of this disease account for approximately 20% of cases and of these around 20% are either confirmed or suspected carriers of a mutation within one of the four primary DNA mismatch repair (MMR) genes (*MLH1*, *MSH2*, *MSH6* or *PMS2*). Individuals are suspected of carrying a Lynch syndrome (MIM120435) associated mutation if they have a strong family history of colorectal cancer such that they meet either the Amsterdam [2] or the Bethesda criteria [3]. Lynch syndrome is an autosomal dominantly-inherited cancer predisposition with high penetrance, and therefore the identification of all relatives carrying the family mutation is of paramount importance in the prevention of colorectal and other spectrum cancers. In a suspected Lynch syndrome family, an individual's tumour is assessed for microsatellite instability (MSI) and immunohistochemical (IHC) absence of the mismatch repair proteins, both of which are hallmarks of the Lynch syndrome phenotype. To enable appropriate diagnosis, counselling and surveillance, the MMR gene indicated by the IHC MMR deficiency profile [4], is screened for causative mutations by Sanger sequencing of exonic and flanking splice site regions and by MLPA (multiplex ligation-dependant PCR amplification) testing for large exonic deletions and duplications.

Using these selective sequencing approaches, causative mutations have been identified in the majority of CRC cases suspected of Lynch syndrome, however there remains a substantial proportion of families (where affected individuals demonstrate MMR IHC deficiency and/or MSI in the presence of a family history that meets the Amsterdam II criteria) for which a mutation has not been identified. Recently, the identification of a novel genetic mechanism for the inactivation of the *MSH2* gene has been described whereby

deletions of the 3' end of the *EPCAM* gene creates an *EPCAM/MSH2* fusion transcript resulting in methylation of the *MSH2* promoter and transcriptional silencing [5], however these do not account for all the families in which CRC cases have a tumour displaying loss of staining for the MSH2 and MSH6 heterodimer proteins but have no identifiable mutation.

MMR gene mutation detection by both research and clinical laboratories has focused on routine sequence analysis of the transcribed regions such that only short stretches of intronic sequence flanking the exons are screened. This practice is reflected by an analysis of the Leiden Open Variation Database (www.lovd.nl) where 90% of the variants identified within introns of the *MSH2* gene ($n = 761$) occur within 20 bp of the exon, and the majority of those seen deeper within the intron have come from only two studies which have measured non-deleterious tagging SNPs across the genomic region [6, 7].

In this study we describe a family with Lynch syndrome caused by a novel *MSH2* mutation situated deep within intron 1 (4,521 bp downstream from exon 1 and 478 bp upstream of exon 2), demonstrating the benefit of a more extensive sequencing approach for such families.

Methods and results

Family X (Fig. 1; Table 1) was recruited through the Australasian Colorectal Cancer Family Study (ACCFs) [8], under the study recruitment guidelines of a family meeting the Amsterdam II criteria [2]. Colorectal tumour material was available for MMR IHC staining and MSI testing [4] for individuals 1, 9 and 12. All three demonstrated microsatellite instability and/or tumour-specific loss of MSH2 and MSH6 proteins (Table 1). Sanger sequencing and MLPA testing for mutations [9] in individual 9 identified no pathogenic mutations or variants of uncertain clinical significance within *MSH2* or the *EPCAM* gene. Testing for mutations in *MSH6*, the heterodimeric partner of *MSH2*, also identified no pathogenic or unclassified variants.

A panel of seven previously described [10] microsatellite markers spanning 4.2 Mb across chromosome 2p, inclusive of the *MSH2*, *MSH6* and *EPCAM* genes, was used to identify a disease-segregating haplotype and hence establish that a genetic connection existed between the tumour IHC MMR deficiency and affected members of the family. Affected individuals 1 and 9, who are first cousins once removed, demonstrated a shared haplotype across the *MSH2* locus (Fig. 1) and thus provided evidence against possible somatic inactivation of MMR. This conserved haplotype was also seen in individual 10 who is currently unaffected, but it has not previously been observed amongst a set of 118 healthy Caucasian controls [10].

In the absence of a coding mutation or gene re-arrangement, we hypothesised that the shared haplotype in Family X harboured a variant causing aberrant regulation or splicing of the *MSH2* gene. To address this hypothesis, immortalised lymphoblastoid cell lines were generated following the principles of Sugden and Mark [11]. Cell lines were cultured in the presence of 100 µg/ml of cycloheximide to eliminate nonsense mediated decay, and RNA was subsequently extracted using TRIzol[®] reagent (Invitrogen). To detect aberrant splicing,

complementary DNA (cDNA) was generated, using the ImProm-IITM Reverse Transcription System (Promega), for individuals 9 and 10 of Family X and from three blood donor controls. This cDNA was then used to amplify the *MSH2* gene in four overlapping fragments (Supplementary Table 1) that were resolved on a 2% agarose gel.

An aberrant band approximately 100 bp larger than the wild type band was seen in individuals 9 and 10 for the amplicon spanning exons 1 through 4 (Fig. 2a). It can be seen from lanes 1 and 3, which distinguish between cycloheximide treated and untreated cell lines, that nonsense mediated decay is involved in breaking down some of the aberrant splice form, however a significant level of transcript still remains in the untreated cells (lane 3). This aberrant banding was not observed in unrelated controls, and no other patient specific aberrations could be seen in the remaining amplicons (Supplementary Figure 1).

To characterise the nucleotide sequence of the larger aberrant product, the band was excised and DNA extracted from the agarose gel, using Qiagen's QIAquick gel extraction kit. Sanger sequencing of the DNA fragment revealed an insertion of 75 nucleotides between the end of exon 1 and the beginning of exon 2 (Fig. 2b), the sequence of which mapped to deep within intron 1 of *MSH2* (c.212–553_c.212–479). Primers were generated (Supplementary Table 1) to amplify across the identified inserted region of intron 1. Subsequent germline sequencing revealed a T > G nucleotide change at intronic position c.212–478 for individuals 1, 2, 9 and 10 but not for individual 11, concordant with the segregating haplotype (Fig. 1). This T > G transversion results in the formation of a canonical splice donor site at the 3' end of the inserted sequence (NNSPLICE 0.9, score = 1), which together with an existing predicted cryptic acceptor site at the 5' end (NNSPLICE 0.9, score = 0.8) accounted for the observed insertion and aberrant transcript (Fig. 3). The inserted sequence is in-frame with the wild type transcript, however it contains a stop codon at the 3' end and is predicted to result in a truncated protein of just 94 amino acids. The variant is not recorded in dbSNP nor was it seen amongst a set of 100 healthy blood donor controls.

Conclusion

This study has identified a novel mutation in the *MSH2* locus and provides evidence for a mechanism of disease causation for rare intronic variants far removed from the traditional splice sites. This type of mutation mechanism has been previously reported for cystic fibrosis [12], but has not been previously described in Lynch syndrome. We have not excluded the possibility that the mutation identified here is family specific, however, additional intronic mutations located outside of the current screening regions may be present in families without an identified mutation. An analysis of the intronic sequences within *MSH2* revealed additional putative splice donor ($n = 163$) or splice acceptor ($n = 284$) sites (NNSPLICE 0.9, score = 0.75) which could potentially be involved in cryptic splicing given an upstream or downstream mutation which creates a complementary splice acceptor or splice donor site, respectively.

These current findings highlight the success of additional molecular studies in families with suspected Lynch syndrome and demonstrate the potential benefit of utilising more extensive

sequence analysis methods to identify causative mutations in suspected Lynch syndrome families where routine methods have failed to find a causative mutation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank individual participants in the study who made this work possible and the contribution of the Jeremy Jass Memorial Pathology Collection. This work was funded by the National Cancer Institute, under RFA no. CA-95-011, and through a cooperative agreement with the Australasian Colorectal Cancer Family Registry (U01 CA097735). The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute nor does mention of trade names, commercial products, or organisations imply endorsement by the US Government. JY is a Cancer Council Queensland Senior Research Fellow.

References

1. Jemal A, Siegel R, Xu J, Ward E. Cancer statistics. *CA Cancer J Clin*. 2010 doi:[10.3322/caac.20073](https://doi.org/10.3322/caac.20073).
2. Vasen HF, Watson P, Mecklin JP, Lynch HT. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. *Gastroenterology*. 1999; 116(6):1453–1456. [PubMed: 10348829]
3. Umar A, Boland CR, Terdiman JP, et al. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *J Natl Cancer Inst*. 2004; 96(4): 261–268. [PubMed: 14970275]
4. Lindor NM, Burgart LJ, Leontovich O, et al. Immunohistochemistry versus microsatellite instability testing in phenotyping colorectal tumors. *J Clin Oncol*. 2002; 20(4):1043–1048. [PubMed: 11844828]
5. 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(1):112–117. doi:[10.1038/ng.283](https://doi.org/10.1038/ng.283). [PubMed: 19098912]
6. Mann A, Hogdall E, Ramus SJ, et al. Mismatch repair gene polymorphisms and survival in invasive ovarian cancer patients. *Eur J Cancer*. 2008; 44(15):2259–2265. doi:[10.1016/j.ejca.2008.07.010](https://doi.org/10.1016/j.ejca.2008.07.010). [PubMed: 18723338]
7. Schafmayer C, Buch S, Egberts JH, et al. Genetic investigation of DNA-repair pathway genes PMS2, MLH1, MSH2, MSH6, MUTYH, OGG1 and MTH1 in sporadic colon cancer. *Int J Cancer*. 2007; 121(3):555–558. doi:[10.1002/ijc.22735](https://doi.org/10.1002/ijc.22735). [PubMed: 17417778]
8. Newcomb PA, Baron J, Cotterchio M, et al. Colon Cancer Family Registry: an international resource for studies of the genetic epidemiology of colon cancer. *Cancer Epidemiol Biomark Prev*. 2007; 16(11):2331–2343. doi:[10.1158/1055-9965.EPI-07-0648](https://doi.org/10.1158/1055-9965.EPI-07-0648).
9. Walsh MD, Buchanan DD, Cummings MC, et al. Lynch syndrome- associated breast cancers: clinicopathologic characteristics of a case series from the colon cancer family registry. *Clin Cancer Res*. 16(7):2214–2224. doi:[10.1158/1078-0432.CCR-09-3058](https://doi.org/10.1158/1078-0432.CCR-09-3058). [PubMed: 20215533]
10. Clendenning M, Baze ME, Sun S, et al. Origins and prevalence of the American Founder Mutation of MSH2. *Cancer Res*. 2008; 68(7):2145–2153. doi:[10.1158/0008-5472.CAN-07-6599](https://doi.org/10.1158/0008-5472.CAN-07-6599). [PubMed: 18381419]
11. Sugden B, Mark W. Clonal transformation of adult human leukocytes by Epstein-Barr virus. *J Virol*. 1977; 23(3):503–508. [PubMed: 197258]
12. Chillon M, Dork T, Casals T, et al. A novel donor splice site in intron 11 of the CFTR gene, created by mutation 1811 + 1.6kbA- >G, produces a new exon: high frequency in Spanish cystic fibrosis chromosomes and association with severe phenotype. *Am J Hum Genet*. 1995; 56(3):623–629. [PubMed: 7534040]

Fig. 1. Pedigree and genotyping data for Family X. Disease associated haplotypes are shown in blue with the associated mutation status depicted in the chromatograms below. Lynch spectrum cancers or polyps are displayed beneath each individual along with the age at onset. Complete clinical data is provided in Table 1

Fig. 2.

Aberrant cDNA banding and sequencing chromatogram from the 5' end of the *MSH2* transcript. **a** Amplicon of *MSH2* Exon 1 through 4 for patient 9 (*lanes 1* and *3*) and an unrelated, unaffected control sample (*lanes 2* and *4*). *Lanes 1* and *2* show products obtained from cycloheximide treated LCLs while products in *lanes 3* and *4* were untreated. **b** Sequencing chromatogram from patient 9 showing the start of the heterozygous sequence indicating the presence of an inserted sequence. The lower trace shows sequence obtained from a gel extraction of the upper aberrant band from panel (**a**)

Fig. 3. Genomic overview of the transcriptional aberration identified in Family X. **a** Sequence of the genomic region preceding exon 2. Intronic sequences are depicted in lower case with splice sites *underlined*. The beginning of exon 2 is highlighted *red* and the newly transcribed exon is highlighted *yellow*. The causative mutated base is shown in *red* below the wild type sequence. Nucleotides are all numbered relative to the first base of exon 2. **b** Schematic representation of the effect of c.212-478T > G on the MSH2 mRNA

Table 1

Clinical attributes of Family X

Individual	Age at last follow-up	Segregating haplotype	Mutation	Phenotype (Age)	IHC	MSI ^a
1	23	Yes	Yes	Colon (19)	MSH2/MSH6 deficient	7/9
2	40	Yes	Yes	Polyp (36)	–	–
5	38 ^b	–	–	Breast (35)	–	–
				Intestine (37)	–	–
6	64	–	–	None reported	–	–
9	42	Yes	Yes	Colon (35)	MSH2/MSH6 deficient	–
10	34	Yes	Yes	None reported	–	–
11	36	No	No	Polyp (35)	–	–
12	45 ^b	–	Yes	Bladder (44)	–	–
				Colon (45)	MSH2/MSH6 deficient	6/8
13	60	No	No	Polyp (52)	–	–
31	–	–	–	Endometrium (47)	–	–

– Not available/tested or not applicable

^aMarkers unstable/markers tested

^bAge at death