The Journal of the International Federation of Clinical Chemistry and Laboratory Medicine

eJIFCC

Analytical challenges in the genetic diagnosis of Lynch syndrome – difficult detection of germ-line mutations in sequences surrounding homopolymers

Castillejo M.I., Castillejo A., Barbera V.M., Soto J.L.

Molecular Genetics Laboratory, Elche University Hospital, Elche, Spain

ARTICLE INFO

Corresponding author:
José Luis Soto
Molecular Genetics Laboratory
Elche University Hospital
Camino Almazara nº 11
03203 Elche. Spain
Phone: (+34) 966 616185
E-mail: soto_jos@gva.es

Key words:
diagnosis, Lynch syndrome, analytical challenges

Competing interests:
The authors declare no competing interests.

Ackowledgements:
We thank the Hereditary Cancer Program of the Valencian Region.

ABSTRACT

A genetic diagnosis is essential in families with a suspicion of Lynch syndrome, as it allows the use of proper and specific surveillance programs for high-risk individuals who carry a pathogenic mutation. The prediction and prevention schemes reduce the impact of cancer in high-risk families in a cost-effective manner. Genetic tests for LS are well standardized and broadly used, although there remain some specific difficulties that need to be addressed to reach an optimal diagnosis. In this report, we addressed the problem raised by the detection of mutations at intronic-splicing consensus sites located near mononucleotide repeats. A standard procedure was applied for LS diagnosis in all cases. PCR and Sanger sequencing results of the whole coding sequences and intron–exon boundaries of the *MSH2* gene were analyzed. Moreover, we designed quality-control procedures to verify the attainment of the intended quality of results regarding sequences located in complex contexts. We found eight families with point mutations at intron 5 of the *MSH2* gene located near the BAT26 mononucleotide marker, which could be missed in a regular diagnostic process. Four families had the c.942+2T>A mutation, and the remaining four families had the c.942+3A>T mutation. In conclusion, the detection of pathogenic

mutations located near microsatellite sequences is especially difficult and requires the implementation of specific quality controls to optimize diagnostic methods.

❖❖❖❖❖

INTRODUCTION

Lynch syndrome (LS) (MIN No: 120435) is an autosomal dominant hereditary condition that predisposes to colorectal, endometrial, and other tumors. The syndrome is caused by germ-line mutations in one of the mismatch repair (MMR) genes: *MLH1*, *MLH2*, *MSH6*, or *PMS2*[1]. A genetic diagnosis is essential in families with a suspicion of having LS, as it allows the use of proper and specific surveillance programs for high-risk individuals who carry a pathogenic mutation. Thus, high risk individuals are advised to stay within the normal weight range and refrain from smoking since a high BMI and smoking increase the risk of developing adenomas and colorectal cancer in Lynch syndrome. Regular colonoscopy leads to a reduction of colorectal cancer-related mortality. Hysterectomy and bilateral oophorectomy largely prevents the development of endometrial and ovarian cancer and is an option to be discussed with mutation carriers who have completed their families especially after the age of 40 years[2].

The prediction and prevention schemes reduce the impact of cancer in high-risk families in a cost-effective manner.

In general, genetic tests for LS are well standardized and broadly used, although there remain some specific difficulties that need to be addressed to reach an optimal diagnosis. In addition to the postanalytical limitations in the interpretation of the clinical significance of some genetic variants, there are other analytical challenges, such as the difficult study of the *PMS2* gene because of the high number and homology with several pseudogenes, or the detection of variants located in the proximity of a homopolymer sequence in the MMR genes.

In this report, we addressed the problem raised by the detection of mutations located at intronic-splicing consensus sites, near mononucleotide repeats. In this particular sequence context, the detection of mutations is a real analytical challenge.

METHODS

Our laboratory performs genetic testing for the diagnosis of LS covering a population of over five million people in the southeast of Spain. These genetic tests are requested from the five genetic counseling units of the Public Health Hereditary Cancer Program of the Comunidad Valenciana[3].

The present study was conducted in compliance with the ethical principles for medical research involving human subjects of the Declaration of Helsinki. Informed consent was obtained from all subjects, and the study received the approval of the Ethics Committee of the Elche University Hospital.

A standard procedure was applied for the diagnosis of LS in all cases. Fulfillment of the revised Bethesda Guidelines or loss of expression of MMR genes during universal screening for colorectal and endometrial tumors is required for referral to the Genetic Counseling in Cancer Units[4,5]. Before gene mutation analysis, the tumors of the probands were studied for microsatellite instability and MMR protein expression (MLH1, MSH2, MSH6, and PMS2), to confirm MMR implication and select the target gene/s for mutation analysis.

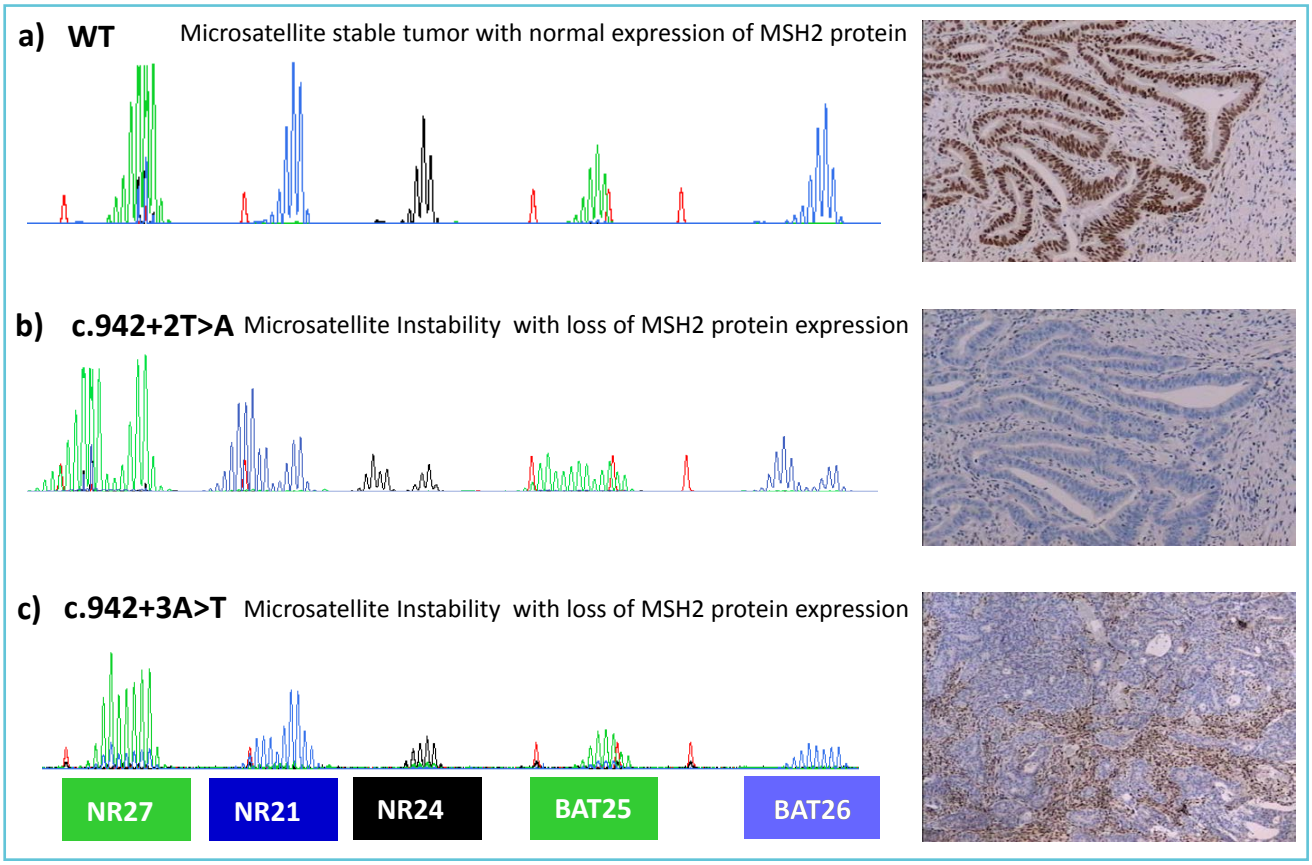
Mutation testing using the probands' blood DNA was then performed to assess the causative germ-line mutation in their families. PCR and Sanger sequencing results of the whole coding

sequences and intron–exon boundaries of the *MSH2* gene were analyzed. The PCR primers and conditions used were reported by Wahlberg et al.[5]. The suspicious genetic variants detected were confirmed by an independent sequence analysis of both DNA strands. The clinical significance of the variants was assessed according to the *InSiGHT Variant Interpretation Committee: Mismatch Repair Gene Variant Classification Criteria, Version 1.9 August 2013* (http://insight-group.org/criteria/)[7]. Pathogenic mutations and variants of unknown clinical significance were deposited in the *InSiGHT* database (*International Society for Gastrointestinal Hereditary Tumors*: http://www.insight-group.org). Once a causative mutation of LS was detected, the patient received the corresponding genetic counseling and genetic predictive tests were offered to at-risk relatives. Genetic predictive tests are usually performed using PCR and sequencing of both strands of the amplicon that contains the mutation detected in the family.

RESULTS AND DISCUSSION

After 10 years of testing experience of genetic diagnosis for LS, we have found eight families with mutations in homopolymer surrounding sequences. All tumors analyzed in these families showed high microsatellite instability with positive results for the five mononucleotide markers analyzed, as well as loss of inmunohistochemical

Figure 1 **Microsatellite Instability analysis and loss of expression of MSH2 protein by Immunohistochemistry in colorectal tumors**



a) wild-type control, b) and c) cases with c.942+2T>A and c.942+3A>T mutations, respectively. NR27, NR21, NR24, BAT25 and BAT26: microsatellite markers.

Table 1 Clinical data of probands and families with pathogenic mutations located near the BAT26 marker

Id.	Sex	FH	Variant	Protein	Proband´s neoplasms	Ages	Relatives* +/-
1	F	AM II	c.942+2T>A	p.Val265_Gln314del	EC, ***CRC***	41, 43	6/13
2	F	BG			***CRC***	33	5/0
3	F	BG			***OC***	45	3/2
4	M	AM II			***CRC***	51	0/12
5	F	AM II	c.942+3A>T		***EC***	42	0/2
6	F	BG			***CRC***, GC	50, 50	0/2
7	F	AM I			***CRC***, BC	48, 49	0/0
8	M	BG			***CRC***	47	0/2

**Relatives: number of predictive tests performed to date in the family with positive/negative results.*
Sex: F, female; M, male.
FH, family history; AM II, Amsterdam Criteria II; AM I, Amsterdam Criteria I; BG, Bethesda Guidelines.
Neoplasms: BC, breast cancer; CRC, colorectal cancer; EC, endometrial cancer; GC, gastric cancer; OC, ovarian cancer.
In bold italic letters: tumors in which microsatellite instability and MMR protein immunohistochemistry were detected.

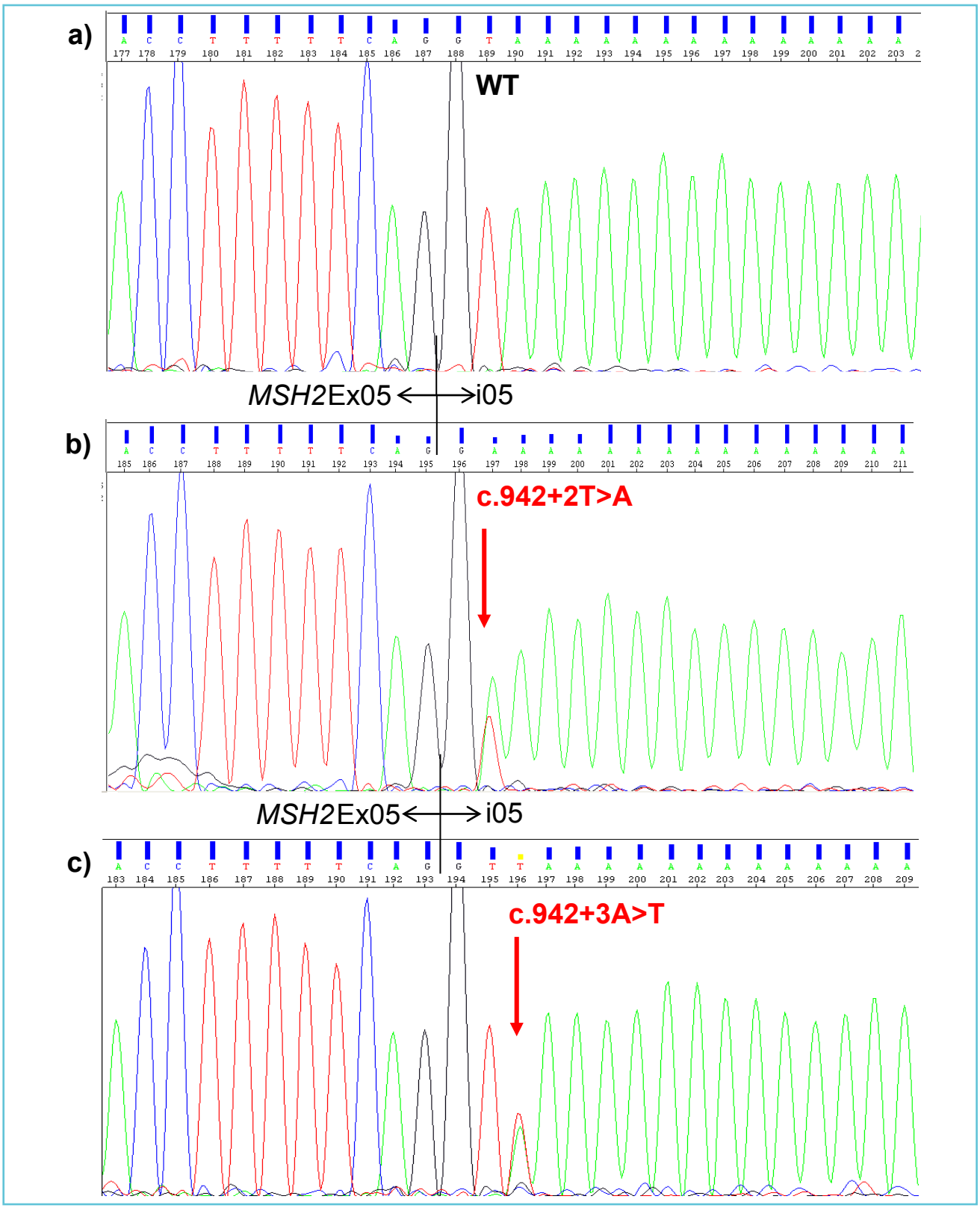
expression of the MSH2 and MSH6 proteins (Figure 1, Table 1).

Four of these families had the c.942+2T>A mutation, whereas the remaining four families had the c.942+3A>T mutation. Both at intron 5 of *MSH2* gene just by BAT26 mononucleotide marker [$(A)_{26}$] (Figure 2, Table 1). These mutations had the same effect at the protein level, i.e., exon 5 skipping (p.Val265_Gln314del). Consequently, important functional domains of the protein were affected, such as MutS II and III, which are connector and lever domains, respectively. These domains play different roles in holding the DNA that is to be repaired. Therefore, both mutations were pathogenic and causative of LS.

The special sequence context of these mutations hampers their detection. In our series, these mutations represent about 18% (8/45) of *MSH2* point mutations and 5% (8/148) of all point mutations detected in the four MMR genes. In addition, and to date, 47 at-risk relatives from these families have also been tested, which led to the identification of 14 mutation-carrier individuals. These high-risk individuals are currently benefiting from a specific surveillance and monitoring program aimed at minimizing the impact of cancer[2].

Up to nine intronic large homopolymer sequences (over 10mer long) are located in the proximity of exons and around splice sites in the MMR genes. To date, pathogenic mutations at eight out of these nine sites have been described in the *InSiGHT* database (Table 2). The splice sites are conserved and essential for exon definition and appropriate splicing. Mutations in those consensus positions generate aberrant transcripts and loss of protein function and are, consequently, pathogenic.

Figure 2 Sanger forward sequence of the MSH2 gene: Ex05–i05 boundary



a) wild-type sequence, b) c.942+2T>A mutation, and c) c.942+3A>T mutation

Table 2 Mononucleotide repeats (>10mer long) located in the proximity of consensus splicing sites of the MMR genes and mutations detected at those sites

Gene	Intron-exon boundaries	Homo polymer	Mutation	Class*	# families Our Lab	# families InSiGHT
MLH1	i04-E05	$(T)_{21}$	c.1039-2A>G	4	0	2
			c.1039-2A>T	4	0	1
			c.1039-1G>A	5	0	4
MSH2	i01-E02	$(T)_{13}$	c.212-2A>G	4	0	2
			c.212-2A>G	5	0	7
	E05-i05	$(A)_{26}$	c.942+2T>A	5	4	6
			c.942+3A>T	5	4	161
MSH6	i06-E07	$(T)_{13}$	c.3556+3_3556+13del	3	0	2
PMS2	i04-E05	$(T)_{13}$	None	-	-	-

**Class: variant classification according to their clinical significance (InSiGHT database): 5, pathogenic; 4, probably pathogenic; 3, unknown.*

It is important to note that mutations that occur in the proximity of a large mononucleotide repeat can be detected only by sequencing of the DNA chain that contains the variant in the 5' side to the repeat. Confirmation by sequencing of the complementary chain is unfeasible; for this reason, special care is needed to detect and confirm these variants.

Currently, the vast majority of tests used for the genetic diagnosis of hereditary cancer syndromes in Europe are laboratory-developed tests (LDT). As stated by the international standard ISO 15189:2012(E) (Medical laboratories, requirements for quality and competence), the laboratory should design quality-control procedures that verify the attainment of the intended quality of the results. Standard operating procedures that include the approaches that are necessary to overcome these specific difficulties are mandatory. For the analysis of probands by Sanger sequencing, validated PCR and sequencing conditions for all amplicons that are needed to cover the regions of interest are required. The use of visual inspection, in addition to the bioinformatics tools used for sequencing analysis, is highly recommended. When there is reason to suspect the presence of mutations, a double check by a second experienced observer and a confirmatory analysis using the same DNA sample are required. A positive-control sample should also be tested in parallel. When next-generation sequencing (NGS) platforms are used for non-validated diagnostic testing, confirmation by Sanger sequencing is compulsory. Furthermore, for predictive testing of at-risk individuals, at least two independent PCR-Sanger sequencing experiments that include positive

and negative controls of the genetic variants in question are recommended.

In conclusion, the detection of pathogenic mutations located near microsatellite sequences is especially difficult and requires the implementation of specific quality controls to optimize diagnostic methods.

REFERENCES

1. Rustgi AK. The genetics of hereditary colon cancer. Genes Dev. 2007, 15;21(20): p. 2525-38.

2. Vasen HF, et al. Revised guidelines for the clinical management of Lynch syndrome (HNPCC): recommendations by a group of European experts.Gut. 2013,62(6): p. 812-23.

3. Cuevas-Cuerda D, Salas-Trejo D. Evaluation after five years of the cancer genetic counselling programme of Valencian Community (Eastern Spain). Fam Cancer. 2014, 13(2): p. 301-9.

4. Umar A, et al. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. J Natl Cancer Inst. 2004, 18;96(4): p. 261-8.

5. Pérez-Carbonell L, et al. Comparison between universal molecular screening for Lynch syndrome and revised Bethesda guidelines in a large population-based cohort of patients with colorectal cancer. Gut. 2012, 61(6): p. 865-72.

6. Wahlberg SS, et al. Evaluation of microsatellite instability and Immunohistochemistry for the prediction of germ-line MSH2 and MLH1 mutations in hereditary nonpolyposis colon cancer families. Cancer Research 2002, 62: p. 3485-92.

7. Thompson BA, et al. Application of a 5-tiered scheme for standardized classification of 2,360 unique mismatch repair gene variants in the InSiGHT locus-specific database. Nat Genet. 2014, 46(2): p. 107-15.