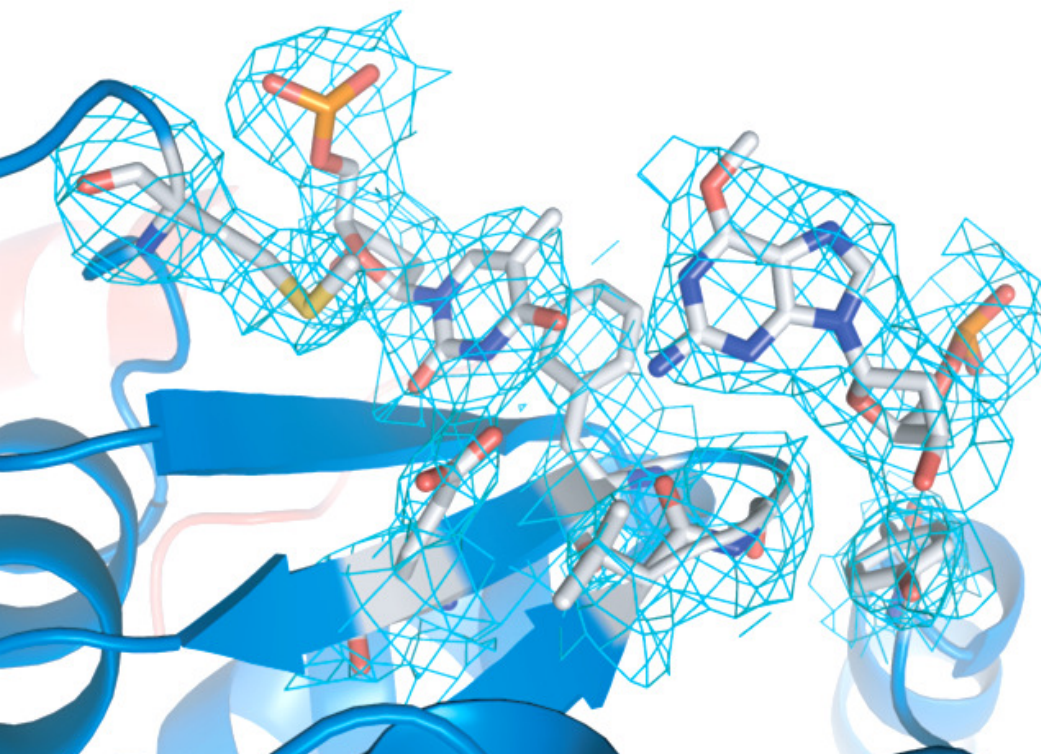




Universidad de Salamanca
Departamento de Medicina -Unidad de Medicina Molecular
Centro de Investigación del Cáncer

Estudio de mutaciones en los genes reparadores en pacientes con síndrome de Lynch



Laura Macías Álvarez
2011

El Dr. D Rogelio González Sarmiento, Catedrático del Departamento de Medicina de la Universidad de Salamanca y el Dr. D. Juan Jesús Cruz Hernández, Catedrático del Departamento de Medicina de la Universidad de Salamanca

CERTIFICAN:

Que el trabajo “Estudio de mutaciones en los genes reparadores en pacientes con síndrome de Lynch” que presenta la licenciada en Biología Doña Laura Macías Álvarez, ha sido realizado bajo nuestra dirección en la Unidad de Medicina Molecular del Departamento de Medicina y, reúne, a nuestro juicio, originalidad y contenidos suficientes para ser presentado ante el comité evaluador correspondiente y optar al título de “Doctor en Biología” por la Universidad de Salamanca.

Y para que así conste, y a los efectos oportunos, expiden el presente certificado en Salamanca,

Fdo. Dr. D. Rogelio González Sarmiento

Fdo. Dr. D. Juan Jesús Cruz Hernández

El presente trabajo de tesis doctoral ha sido financiado por la Gerencia Regional de Castilla y
León, expediente 216/A/08

A Benito, Juan, Sinda y Fina.

Agradecimientos

Le agradezco a mis directores de tesis: el Prof. Dr. D. Rogelio González Sarmiento, por permitirme entrar en el mundo de la ciencia y el Prof. Dr. D. Juan Jesús Cruz Hernández, por su importante papel en el Servicio de Oncología Médica del Hospital Universitario de Salamanca.

Le agradezco al Servicio de Anatomía Patológica por su aporte en este trabajo.

A mis compañeros y amigos de laboratorio que son un compendio de profesionalidad y, ante todo, de amistad y compañerismo.

En especial, quiero mencionar a Eva, su paciencia, cariño y buenas enseñanzas será algunas de las cosas que me dejarán mejor recuerdo de esta etapa.

A Nieves y a Estrella que se ocupan de nosotros con mimo aunque lleguemos a última hora cargados de problemas.

A María José y Luis, que ya saben lo agradecida que estoy por su apoyo diario. A Nacho, me trajeron un hermano en vez de un perro, pero el cambio no ha sido tan malo.

Índice

Glosario

Introducción

1. Epidemiología del cáncer

1.2. Epidemiología del cáncer colorrectal

2. Etiología del cáncer colorrectal

2.1. Factores genéticos

- 2.1.1. Poliposis adenomatosa familiar (PAF)
- 2.1.2. Poliposis adenomatosa familiar atenuada (PAFA)
- 2.1.3. Síndrome de Turcot
- 2.1.4. Síndrome de Gardner
- 2.1.5. Síndrome de Peutz Jeghers
- 2.1.6. Poliposis juvenil familiar
- 2.1.7. Síndrome PTEN-Hamartoma-Tumor
- 2.1.8. Síndrome de poliposis mixta hereditaria

3. Cáncer hereditario de colon no polipósico (CCHNP) o Síndrome de Lynch

3.1. El diagnóstico del síndrome de Lynch

- A) Criterios clínicos del síndrome de Lynch: Ámsterdam, Bethesda y modificaciones
- B) Estudio de inestabilidad de microsátélites

3.2. Bases moleculares del síndrome de Lynch

- Mecanismos de reparación del DNA
- Sistema de reparación de errores por despareamiento (Mismatch Repair System, MMR)
- Sistema reparador de errores por despareamiento en organismos procariontes
- Sistema reparador de errores por despareamiento en organismos eucariotas

4. Efecto de las mutaciones en los genes implicados en el sistema de reparación de errores por despareamiento en el DNA

4.1. Tipos de mutaciones en los genes implicados en el sistema reparador

5. microRNAs

- 5.1. Nomenclatura de los microRNAs
- 5.2. Biogénesis de los microRNAs
- 5.3. microRNAs y cáncer

Hipótesis de trabajo y objetivos

Pacientes y métodos

1. Pacientes

2. Manipulación de ácidos nucleicos

2.1. Extracción de DNA a partir de muestras de sangre periférica

- 2.2. Extracción de RNA a partir de muestras de sangre periférica
- 2.3. Extracción de DNA a partir de tejido tumoral parafinado

3. Análisis de mutaciones puntuales en los genes MLH1, MSH2, MSH6

- 3.1. Amplificación por PCR de los genes MLH1, MSH2, MSH6.
- 3.2. Electroforesis en gel de agarosa.
- 3.3. Análisis de los fragmentos amplificados por PCR mediante CSGE (Electroforesis en gel sensible a la conformación)

4. Análisis de grandes reordenamientos genómicos

- 4.1. MLPA (Multiplex dependent Probe Amplification)
- 4.2. PCR a tiempo real cuantitativa.
- 4.3. PCR de transcripción inversa (RT-PCR)

5. Análisis de inmunohistoquímica

6. Análisis de microsatélites

7. Discriminación alélica mediante sondas Taqman

8. Análisis bioinformáticos

- 8.1. Análisis de conservación entre especies
CLUSTAL W
- 8.2. Análisis posible patogenicidad de las mutaciones
PolyPhen (Polymorphism Phenotyping)
SNP3D
SIFT (Sorting intoleran from tolerant)
- 8.3. Análisis de sitios ESE
ESE Finder
RESCUE-ESE

9. Identificación de microRNAs de unión a los genes MLH1, MSH2 y MSH6

miRBase: La base de datos de los microRNA
TargetScan

Resultados

1. Análisis de mutaciones en los genes MLH1, MSH2 y MSH6:

- 1.1. Estudio del gen MLH1
 - 1.1.1. Mutaciones puntuales patogénicas.
 - 1.1.2. Polimorfismos
 - 1.1.3. Variantes de significado desconocido identificadas en el gen MLH1

- 1.2. Estudio del gen MSH2:
 - 1.2.1. Mutaciones patogénicas
 - 1.2.2. Polimorfismos
 - 1.2.3. Variantes de significado desconocido
- 1.3. Estudio del gen MSH6:
 - 1.3.1. Mutaciones patogénicas
 - 1.3.2. Polimorfismos

2. Grandes reordenamientos genómicos en los genes MLH1, MSH2 y MSH6

3. Análisis del gen MLH1 en probandos con inmunohistoquímica negativa para MLH1

4. Genotipado de polimorfismos en el gen MLH1

5. Estudio de microRNAs en los genes MLH1, MSH2, MSH6

Discusión

1. Estudio del gen MLH1

- 1.1 Variantes de significado desconocido en el gen MLH1
 - 1.1.1. Variantes de significado desconocido localizadas en el dominio de dimerización
 - 1.1.2. Variantes de significado desconocido en el extremo 3'UTR del gen MLH1
- 1.2. Análisis del gen MLH1 en Cáncer colorrectal esporádico con inmunohistoquímica negativa para MLH1.

2. Estudio del gen MSH2

- 2.1. Variantes de significado desconocido en el gen MSH2
 - 2.1.1. Variantes de significado localizadas en el dominio de unión al DNA.
 - 2.1.2 Variantes de significado desconocido localizadas en el dominio de dimerización de la proteína MSH2

3. Estudio del gen MSH6

4. Estudio de grandes reordenamientos genómicos.

5. Estudio bioinformático de miRNAs asociados a genes implicados en el síndrome de Lynch.

Bibliografía

Glosario

APC: “*Adenomatous poliposis coli*”. Gen supresor de tumores. Codifica una proteína que interviene en la ruta de señalización Wnt. Está implicada también en los procesos de migración y adhesión celular, activación transcripcional y apoptosis. Los defectos en este gen provocan poliposis adenomatosa familiar (PAF).

- **APS:** Persulfato de amonio.
- **ATP:** Adenosín trifosfato.
- **BER:** “*Base excision repair*”. Ruta de reparación de DNA por escisión de bases.
- **BLAST:** Programa informático capaz de comparar una secuencia problema con una gran cantidad de secuencias presentes en una base de datos y seleccionar las más parecidas.
- **BMPR1:** “*Bone morphogenetic protein (BMP) receptor*”. Los receptores BMP son una familia de kinasas serina/treonina transmembranales que se clasifican en tipo I y tipo II. BMPR1 pertenece a las de tipo I. Mutaciones en el gen que codifica la proteína aparecen asociadas a la poliposis juvenil familiar.
- **CCR:** Cáncer colorrectal.
- **CCHNP:** Cáncer de colon hereditario no polipósico.
- **CIAA:** cloroformo:alcohol isoamílico 24:1
- **CSGE:** Electroforesis en gel sensible a la conformación.
- **CHPRE:** Hipertrofia congénita del epitelio pigmentario de la retina.
- **DNA:** Ácido desoxirribonucleico.
- **DNA-helicasa II:** Pertenece a una familia de enzimas que rompen los puentes de hidrógeno entre las bases nitrogenadas en los procesos de replicación, transcripción, recombinación y reparación del DNA.
- **DNA-ligasa:** Enzima que forma enlaces covalentes entre el extremo 5' de una cadena polinucleotídica y el extremo 3' de otra cadena polinucleotídica.
- **dNTP's:** Desoxinucleótidos trifosfatos.
- **EDTA:** ácido etilendiaminotetracético.
- **EXOI, EXOV, EXOVII:** Son exonucleasas, enzimas que hidrolizan los enlaces fosfodiéster finales de los ácidos nucleicos
- **FASTA:** Formato de fichero informático basado en texto, utilizado para representar secuencias de ácidos nucleicos o de proteínas en el que los pares de bases o los aminoácidos se representan usando códigos de una única letra. El formato permite incluir nombres y comentarios que preceden a las secuencias.
- **HGMD:** “*Human Gene mutation database*”. Base de datos que agrupa mutaciones de diversos genes.
- **HR:** “*Homologous recombination*”. Reparación del DNA por recombinación homóloga.

- **LKB11/STK11:** Miembro de la familia serina/treonina kinanas. Regula la polaridad celular y actúa como supresor de tumores. Las mutaciones en este gen se han asociado con el síndrome de Peutz-Jeghers.
- **MLH1:** “*MutL homolog 1*”. Proteína implicada en la ruta de reparación de desapareamientos del DNA en organismos eucariotas.
- **MLPA:** “*Multiplex-Ligation Probe Amplification*”. Técnica de amplificación por PCR multiplex dependiente de ligación que permite estudiar grandes reordenamientos genómicos.
- **MMR:** “*Mismatch repair*”. Ruta de reparación de desapareamientos en el DNA.
- **mRNA:** RNA mensajero.
- **MSH2:** “*MutS homolog 2*”. Proteína implicada en la ruta de reparación de desapareamientos del DNA en organismos eucariotas.
- **MSH3:** “*MutS homolog 3*”. Proteína implicada en la ruta de reparación de desapareamientos del DNA en organismos eucariotas.
- **MSH6:** “*MutS homolog 6*”. Proteína implicada en la ruta de reparación de desapareamientos del DNA en organismos eucariotas.
- **MutL:** Proteína implicada en la ruta de reparación de desapareamientos del DNA en organismos procariotas.
- **MutS:** Proteína implicada en la ruta de reparación de desapareamientos del DNA en organismos procariotas.
- **MutS α y MutS β :** Complejos heterodiméricos formados por las proteínas implicadas (MutS homolog 2,3,6) en la ruta de reparación del DNA en organismos eucarióticos.
- **MYH:** Homólogo MutY. Es una enzima implicada en la ruta de reparación del DNA por escisión de bases (BER).
- **NER:** “*Nucleotide excision repair*”. Reparación del DNA por escisión de nucleótidos.
- **NHEJ:** “*Non-homologous end joining*”. Reparación del DNA por unión de extremos no homólogos.
- **p53:** (*TP53 “Tumor protein 53”*). Proteína supresora de tumores que regula genes que inducen la parada del ciclo celular, apoptosis, senescencia y reparación de DNA. Mutaciones en este gen dan lugar al Síndrome de Li-Fraumeni.
- **PAF:** Poliposis adenomatosa familiar.
- **PAFA:** Poliposis adenomatosa familiar atenuada.
- **PCNA:** “*Proliferating Cell Nuclear Antigen*”. El antígeno nuclear de proliferación celular es una proteína nuclear sintetizada en la fase G1 temprana y en la fase S del ciclo celular. Esta proteína se localiza en el núcleo y favorece la síntesis de ADN, ya que es un cofactor de la ADN polimerasa.
- **PCR:** “*Polymerase chain reaction*”. Reacción en cadena de la polimerasa.

- **PMS1:** "*MutL homologous*". Proteína implicada en la ruta de reparación de desapareamientos del DNA en organismos eucariotas.
- **PMS2:** "*Postmeiotic Segregation Increased 2*". Proteína implicada en la ruta de reparación de desapareamientos del DNA en organismos eucariotas.
- **PTEN/MMAC1:** "*Phosphatase and tensin homolog*". Fosfatasa que funciona inhibiendo diversas dianas y funciona como supresor de tumores.
- **RECJ:** Exonucleasa que degrada el DNA de cadena sencilla en sentido 5'-3' y participa en los procesos de recombinación homóloga y reparación de DNA en *Escherichia coli*.
- **RFC:** "*Replication factor C*". Es un complejo proteico requerido para la replicación del DNA.
- **RNA:** Ácido ribonucleico.
- **RPA:** "*Replication protein A*". Proteína de unión a ssDNA que participa en la reparación de DNA.
- **SDS:** dodecil sulfato sódico.
- **SMAD4:** "*SMAD family member 4*". Proteína miembro de la familia SMAD e implicada en la poliposis juvenil.
- **SNP:** "*Single nucleotide polymorphism*". Variación en la secuencia de DNA que afecta a un solo nucleótido.
- **SSB:** "*single-stranded DNA binding proteins*". Proteínas con capacidad de unión a DNA de cadena sencilla.
- **ssDNA:** DNA de cadena sencilla (monocatenario).
- **TBE:** Tris 0.044 M, ácido bórico 0.044 M, EDTA 1.0mM pH= 8,3.
- **TEMED:** N, N, N, N' Tetrametilendiamina.
- **TGF-β:** "*Transforming growth factor beta*". Esta ruta de señalización está implicada en numerosos procesos celulares tanto en organismos adultos como en el desarrollo embrionario incluyendo crecimiento embrionario, diferenciación celular, apoptosis, homeostasis celular y otras funciones celulares.
- **UTR:** "*Untranslated region*". Zonas que flanquean la región codificante de cada gen y no se traducen a proteína.

Introducción

1. Epidemiología del cáncer

El cáncer es la primera causa de muerte en los países desarrollados y la segunda en aquellos países en vías de desarrollo. Los casos de cáncer se incrementan en los países desarrollados por el crecimiento de la población y su envejecimiento; así como la adopción de estilos de vida asociados al desarrollo de cáncer como el sedentarismo, una dieta rica en grasas y el tabaco (Jemal y col., 2011).

La estimación global determinó que se diagnosticaron 12.7 millones de nuevos casos de cáncer en 2008 y el 56% de los casos se produjeron en países desarrollados. El tipo de cáncer con mayor incidencia global son el de mama, en mujeres, y el de pulmón, en varones (Jemal y col., 2011).

La mortalidad estimada para ese mismo año se estableció en 7.6 millones de muertes por cáncer y el 64% de ellas se produjeron en países desarrollados (Jemal y col., 2011). El cáncer de próstata y el de mama son los tipos de tumores con mayor mortalidad en el mundo (Jemal y col., 2011).

En Europa, se estimaron 3.191.600 casos de cáncer diagnosticados en 2006. Para ambos sexos, el tipo de cáncer más frecuente es el de mama, seguido del cáncer colorrectal (CCR), de pulmón y próstata. Específicamente, el cáncer de pulmón, el cáncer colorrectal y el de próstata son los tipos de cáncer que con mayor frecuencia se presentan en varones; mientras que, en mujeres, lo son el de mama, el CCR y de pulmón (Ferlay y col., 2007).

En 2006, se estimaron 1.7 millones de muertes causadas por cáncer en Europa (Ferlay y col., 2007). El cáncer de pulmón fue es la causa mayoritaria de muerte, seguido del CCR y del cáncer de mama (Ferlay y col., 2007).

En España, el cáncer es la segunda causa de muerte (26.9% de todas las defunciones) por detrás de los accidentes cardiovasculares (INE, 2008). El envejecimiento de la población española, el incremento en la incidencia de tumores malignos y la mejora en la supervivencia de los enfermos de cáncer debido a los avances diagnósticos y terapéuticos, ha supuesto un aumento significativo de la prevalencia del cáncer en la población española (López-Abente y col., 2004).

1.1 Epidemiología del cáncer colorrectal:

El cáncer colorrectal, a nivel mundial, es el tercer tipo de tumor más diagnosticado en varones y el segundo en mujeres. En 2008, se contabilizaron 1.2 millones de nuevos casos de este tumor y se estimaron 608.700 muertes por su causa (Jemal y col., 2011). Los ratios más altos de incidencia de CCR

se localizan en Nueva Zelanda, Europa y Norte América y los más bajos en África, en el centro y sur de Asia.

Se ha observado que el número casos de CCR se incrementa rápidamente en varias áreas del mundo históricamente identificadas como de bajo riesgo de CCR como, por ejemplo, España, el este de Asia o el este de Europa; mientras que las zonas de alto riesgo (Estados Unidos, Canadá o Australia) la incidencia de CCR empieza a declinar. Estas observaciones aparecen relacionadas con los cambios en los patrones de estilo de vida y alimentación o al aumento y la eficacia de los programas de prevención, detección y extirpación temprana de pólipos precancerosos (Jemal y col, 2011).

En Europa, según los datos de la IARC (Globocan) para el 2008, la incidencia del cáncer colorrectal ocupa el tercer puesto por detrás del cáncer de próstata y pulmón en varones; mientras que en mujeres, es la segunda causa de cáncer por detrás del cáncer de mama. La tasa de mortalidad para ese mismo año ocupa el segundo puesto para ambos sexos (Ferlay y col., 2007).

El cáncer colorrectal en España se sitúa como la segunda causa de muerte por detrás del cáncer de pulmón en varones y el de mama en mujeres (datos obtenidos del Centro Nacional de Epidemiología, Instituto de la Salud Carlos III, año 2007). En Castilla y León, la mortalidad causada por el CCR en el año 2007 tiene una media superior a española en casi todas sus provincias y en ambos sexos (datos obtenidos del Centro Nacional de Epidemiología, Instituto de la Salud Carlos III, año 2007).

2. Etiología del cáncer colorrectal:

Está establecido que el CCR se desarrolla como consecuencia de la acumulación progresiva de alteraciones genéticas y epigenéticas que provocan la transformación de células normales del epitelio colónico en células tumorales. Este proceso de tumorigénesis es largo y puede tardar entre 8 y 12 años (Jemal y col., 2010).

Se ha observado que las tasas de incidencia del cáncer colorrectal varían en relación con las diferencias raciales y los distintos estilos de vida dependiendo de la zona geográfica y las costumbres de cada comunidad. Así, algunos grupos de inmigrantes aumentan o disminuyen el riesgo asociado a su comunidad original cuando se adaptan al estilo de vida en sus países de acogida, asimilando el riesgo de estos últimos (Hobbs y col. 2001). Los factores de riesgo más importantes en el desarrollo de cáncer colorrectal son la edad, los factores genéticos y diversos factores ambientales. Los factores ambientales tienen gran importancia en la aparición y el desarrollo del cáncer colorrectal; entre ellos el estilo de vida, la alimentación y el consumo de alcohol y tabaco (tabla 1).

Tabla 1: Factores ambientales.

Alcohol	<p>El consumo de más de 30g/día de alcohol aumenta el riesgo de CCR. (Khan y col. 2010; Mizoue y col., 2008)</p> <p>Uno de cada cuatro casos en varones con CCR está provocado por consumos de más de 23 g/día de alcohol. (Khan y col. 2010; Paskett y col., 2007)</p> <p>Tomar más de siete bebidas alcohólicas por día incrementa el riesgo CCR en un 72%. (Khan y col. 2010; Akhter y col., 2007)</p> <p>El consumo de alcohol aparece relacionado con CCR distal. (Khan y col. 2010; Moskal y col., 2007)</p> <p>Los periodos de abstinencia parecen disminuir el riesgo. (Khan y col. 2010; Liang y col., 2009)</p>
Tabaco	<p>Los fumadores habituales presentan un incremento de la incidencia de CCR más que los no fumadores. (Khan y col. 2010; Botteri y col., 2008)</p> <p>El incremento del riesgo viene definido por la cantidad de cigarrillos o paquetes fumados al día. (Khan y col. 2010; Mizoue y col., 2008)</p> <p>Este efecto se hace más visible después de 30 años de hábito. (Khan y col. 2010; Mizoue y col., 2008)</p> <p>La edad de diagnóstico de CCR se adelanta en fumadores frente a los no fumadores. (Khan y col. 2010; Sarebo y col., 2006)</p> <p>Fumadores con 40 años de hábito presentan los tumores de colon y recto con mutación en el gen APC. (Khan y col. 2010; Paskett y col., 2005)</p> <p>Los fumadores habituales desarrollan con mayor frecuencia pólipos hiperplásicos en zonas distales del colon. (Khan y col. 2010; Coyle y col., 2009)</p>
Dieta	<p>La ingesta de pescado y ácidos omega-3 del pescado puede disminuir el CCR. (Khan y col. 2010; Goodherman y col., 1997)</p> <p>Hay una fuerte asociación entre el consumo de carne y la mortalidad por CCR. (Khan y col. 2010; Bingham y col., 1996)</p> <p>En varones, el consumo de carne aumenta el riesgo. (Khan y col. 2010; Larsson y col., 2005)</p> <p>El consumo diario de café disminuye el riesgo de CCR. (Khan y col. 2010; Larsson y col., 2005)</p> <p>El consumo de marisco aparece asociado con el desarrollo de CCR. (Khan y col. 2010; Chao y col., 2005)</p> <p>La leche aparece como protector frente al CCR. (Khan y col. 2010; Chao y col., 2005)</p> <p>El alto consumo de huevos y, en general, comidas ricas en calorías aumentan el riesgo. (Khan y col. 2010; Chao y col., 2005)</p>
Obesidad	<p>La obesidad abdominal está asociada con el incremento de adenocarcinomas colorrectales. (Khan y col. 2010, Thygesen y col., 2008)</p> <p>El 29.5% de todos los cánceres de colon son atribuibles a un índice de masa corporal (IMC) mayor de 22.5. (Khan y col. 2010; Larsson y col., 2005)</p> <p>Cada incremento de 5 unidades en el IMC produce el aumento del riesgo de CCR. (Khan y col. 2010; Kim y col., 2007; Tamakoshi y col. 2004)</p>

Son diversos los estudios que asocian la **obesidad** con el aumento del riesgo de cáncer y, en particular, del cáncer de colon y recto (Khan y col. 2010, Thygesen y col., 2008; Larsson y col., 2005; Tamakoshi y col. 2004). En relación estrecha con la obesidad está la **dieta**. Es difícil establecer exactamente en qué cantidad y cuáles son los alimentos beneficiosos para la prevención del cáncer y los que aumentan la probabilidad de desarrollar la enfermedad. En general, se considera que tanto los aceites vegetales, como el pescado, la fruta y la verdura son beneficiosos; mientras que la carne, los alimentos ricos en grasas saturadas e, incluso, el marisco producen el efecto contrario (Khan y col. 2010; Goodherman y col., 1997; Bingham y col., 1996; Chao y col., 2005).

El **alcohol** es un factor de riesgo directo puesto que su consumo en más de 30 g/día aumenta el riesgo de cáncer de colon y recto. El consumo de alcohol parece estar relacionado con la aparición de tumores en zonas distales del colon. El efecto del alcohol sobre la predisposición a padecer cáncer de colon no es irreversible y parece que el riesgo disminuye durante los periodos de abstinencia (Khan y col. 2010; Mizoue y col., 2008; Paskett y col., 2007; Moskal y col., 2007; Liang y col., 2009).

Fumar también aumenta la incidencia y la mortalidad por cáncer colorrectal. Los fumadores habituales presentan un incremento de la incidencia cáncer colorrectal significativamente más alto que los no fumadores. Además, este incremento viene definido por la cantidad de cigarrillos o paquetes fumados al día, si bien este efecto se hace más visible después de 30 años de hábito. La edad de diagnóstico de CCR se adelanta en fumadores frente a no fumadores. Los efectos del tabaco sobre el desarrollo del tumor empiezan a conocerse. Los fumadores habituales desarrollan con mayor frecuencia pólipos hiperplásicos en zonas distales del colon en comparación con los no fumadores (Botteri y col., 2008; Mizoue y col., 2008; Sarebo y col., 2006; Coyle y col., 2009; Khan y col. 2010).

2.1 Factores genéticos

Entre un 80-85% de todos los casos de cáncer colorrectal diagnosticados anualmente son de tipo esporádico. El 15-20% restante agrupa a individuos afectados con cáncer de colon en los cuales aparece un componente familiar, es decir, tienen familiares de primer o segundo grado con la enfermedad. Sin embargo, tan sólo entre un 1 y el 5% de estos últimos heredan la predisposición a padecer la enfermedad a través de un patrón mendeliano (de la Chapelle y col., 2004).

La línea que separa los casos de cáncer hereditario y los casos familiares no está claramente establecida debido a la dificultad para fijar unos criterios de clasificación. Por lo tanto, exceptuando los síndromes hereditarios claramente definidos, resulta bastante complejo discriminar cáncer esporádico y familiar y, en consecuencia, el porcentaje de cáncer colorrectal considerado familiar posiblemente no esté ajustado a la realidad (la Chapelle y col., 2004). En la tabla 2 mostramos varios síndromes en los cuales desarrollo de tumores colorrectales es un rasgo característico.

Tabla 2: Relación de síndromes que predisponen al desarrollo de cáncer colorrectal. *OMIM™: Online Mendelian Inheritance in Man™.

Síndrome	Nº OMIM*	Gen(es) causante(s)
Síndrome de Lynch o Cáncer de colon hereditario no polipósico (CCHNP)	114500	Genes del sistema reparador de errores del DNA: MLH1, MSH2, MSH6, PMS2 y PMS1
Poliposis adenomatosa familiar (PAF)	175100	APC
Poliposis adenomatosa familiar atenuada (PAFA)	175100	APC (Las alteraciones aparecen en el extremo 3' y 5' de la secuencia del gen)
Poliposis asociada a MYH	604933	MYH
Síndrome de Gardner	175100	APC
Síndrome de Turcot	276350	APC, MLH1 y MSH2
Síndrome de Peutz-Jeghers	175200	STK11/LKB1
Poliposis juvenil familiar	174900	SMAD4 y BRMP1A
Síndrome de PTEN-Hamartoma-Tumor	158350	PTEN/MMAC1
Poliposis mixta hereditaria	601228	Se desconoce el gen

2.1.1 Poliposis adenomatosa familiar (PAF)

La poliposis adenomatosa familiar es un síndrome autosómico dominante de alta penetrancia que afecta por igual a varones y a mujeres. Es una enfermedad de baja incidencia, que representa menos del 1% de los casos de cáncer colorrectal observados (Bisgaard y col., 1994; Varesco y col., 2004, Gryfe y col., 2009), con una incidencia de uno de cada 8000 nacimientos (Gryfe y col., 2009).

El fenotipo de esta enfermedad se caracteriza por la aparición de cientos a miles de pólipos adenomatosos que tapizan el epitelio del colon y el recto. Los primeros pólipos surgen durante la infancia o la adolescencia y, sin tratamiento quirúrgico, alguno de ellos acaba degenerando en adenocarcinomas durante la edad adulta. En principio, las manifestaciones de la enfermedad quedan limitadas a las regiones del colon y el recto. Sin embargo, existen casos en los que se observan pólipos en otras zonas del tracto intestinal, por ejemplo, en el duodeno, en el estómago o en la ampolla de Váter (Nagy y col., 2004).

La PAF también puede presentar manifestaciones extracolónicas: tumores desmoides, hipertrofia congénita del epitelio pigmentario de la retina (Congenital Hypertrophy of Retinal Pigment Epithelium, CHPRE), osteomas, anomalías dentarias, quistes epidermoides y otras neoplasias benignas y malignas de tejidos blandos como el cáncer papilar o folicular de tiroides o el hepatoblastoma en niños (Nagy y col., 2004; Gryfe y col., 2009) (figura 1).

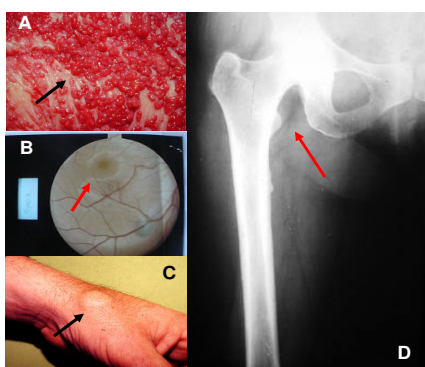


Figura 1: manifestaciones de la PAF. A) Múltiples pólipos colónicos. B) Hiperpigmentación de la retina (CHPRE). C) Quiste epidermoide en la mano. D) Osteoma en el fémur.

La aparición del síndrome es consecuencia de mutaciones germinales en el gen supresor de tumores APC. El 80-90% de las familias con PAF, presentan una mutación en el gen APC en la línea germinal (Gryfe y col. 2009). El 20% de los casos diagnosticados de poliposis adenomatosa familiar presentan una mutación *de novo* en el gen APC (Gryfe y col., 2009).

2.1.2 Poliposis adenomatosa familiar atenuada

La variante atenuada de la poliposis familiar se presenta en un 10% de los pacientes con poliposis adenomatosa. Esta variante difiere de la poliposis clásica en la presencia de menos de 100 pólipos que tapizan la superficie colorrectal. Los pólipos tienden a situarse cerca del ángulo esplénico con una morfología plana, por lo que en un principio esta enfermedad fue descrita como “síndrome hereditario del adenoma plano” (“hereditary flat adenoma syndrome”) (Lynch y col., 1995). Además, se han descrito pólipos gástricos y adenomas duodenales asociados. Por otro lado, son raros los casos de CHPRE y tumores desmoides asociados a la poliposis atenuada, y, además, la media de edad de los pacientes en la cual se manifiesta es mayor que la de los individuos con el fenotipo clásico (Soravia y col., 1998; Knudsen y col. 2003; Gryfe y col., 2009).

Un 10% de las familias con poliposis atenuada presenta mutaciones en el gen APC. Las mutaciones en APC asociadas a esta variante de la enfermedad aparecen localizadas principalmente en tres regiones del gen: el extremo 5', en el exón 9 y en el extremo 3' (Figura 2).

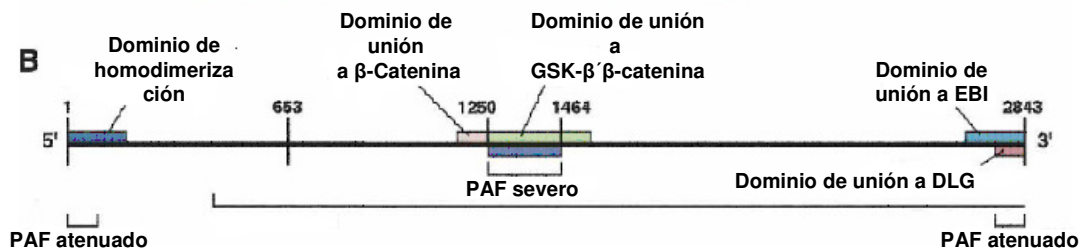


Figura 2: esquema de la proteína APC. Las mutaciones localizadas en el extremo 5', exón 9 y extremo 3' del gen producen una poliposis adenomatosa atenuada.

El 30% de las familias con poliposis atenuada no presentan una mutación en el gen APC, si no que tienen mutaciones en el gen MYH; en estos casos, el síndrome se denomina poliposis asociada a MYH y presenta un patrón de herencia autosómico recesivo (Sieber y col., 2003, Nagy y col., 2004; Gryfe y col., 2009).

2.1.3 Síndrome de Turcot

Este síndrome se caracteriza por el desarrollo de neoplasias en el sistema nervioso central, que incluyen meduloblastomas, astrocitomas y ependimomas. La mayor parte de estos pacientes también desarrollan características típicas de los pacientes afectados de PAF y un 70% de las familias presentan mutaciones en APC. No obstante, hay un grupo de pacientes con este síndrome que presentan

alteraciones en los genes de reparación del DNA, MLH1 y MSH2 en lugar de en el gen APC (Gryfe y col., 2009).

2.1.4 Síndrome de Gardner

El síndrome de Gardner es una variante de la FAP clásica que incluye las mismas manifestaciones extracolónicas. Actualmente, el síndrome de Gardner y la poliposis adenomatosa familiar se consideran diferentes grados de severidad de la misma enfermedad. En ambos síndromes se han encontrado las mismas alteraciones en el gen APC (Gryfe y col., 2009).

2.1.5 Síndrome de Peutz-Jeghers

Es una enfermedad autosómica dominante que se caracteriza por la aparición en la infancia de pólipos hamartomatosos difusos en el intestino y una pigmentación característica en las mucosas de la zona perioral y oral, en la piel de la cara, de los antebrazos, palmas de las manos y los pies, y en la zona perianal (Figura 3). Los pólipos hamartomatosos se distribuyen por todo el tracto gastrointestinal y empiezan a desarrollarse en la primera década de la vida, aunque los síntomas se manifiestan en la segunda y tercera década. Los pacientes afectados tienen una alta predisposición a desarrollar neoplasias gastrointestinales y extra-intestinales. Entre el 50% y el 90% de los pacientes desarrollan cáncer y un 40% de ellos desarrolla el tumor en el colon. El gen responsable del síndrome de Peutz-Jeghers es el gen supresor de tumores LKB11/STK11, el cual se ha encontrado mutado en el 50-70% de los pacientes (Aaltonen y col., 2000; Sánchez-Céspedes M., 2007).



Figura 3: manifestaciones del síndrome de Peutz-Jeghers. A y B) Pigmentación en mucosas y palma de la mano. C) Pólipos hamartomatosos en intestino.

2.1.6 Poliposis juvenil familiar

La poliposis juvenil familiar sigue un modelo de herencia autosómica dominante. Los pacientes afectados presentan múltiples pólipos a lo largo del tracto intestinal y un subgrupo de ellos también manifiesta anomalías congénitas. El adjetivo “juvenil” se refiere, más que a la edad de aparición de los mismos, al tipo de pólipo desarrollado. Estos pólipos tienen un gran potencial de malignización y pueden desarrollar tumores de los cuales un alto porcentaje aparecen en el colon o el recto. Se han descrito mutaciones en los genes supresores de tumores de la vía de señalización TGF- β , SMAD4 y BMPR1, asociados a esta enfermedad (Hizawa y col., 1997; Chow y col., 2008).

2.1.7 Síndrome PTEN-Hamartoma-Tumor

El síndrome de PTEN-Hamartoma-tumor engloba a los síndromes de Cowden, Bannayan-Riley-Ruvalcaba, Proteus y Proteus-like. Todos ellos están causados por mutaciones en el gen PTEN/MMAC1, y se caracterizan por la presencia de múltiples hamartomas que afectan a cualquier zona del cuerpo (Zhou y col., 2003; Nagy y col., 2004).

El síndrome de Cowden es difícil de reconocer, se caracteriza por la presencia de varios hamartomas en varios sitios del cuerpo. Los individuos afectados presentan predisposición a desarrollar diversos cánceres, en particular cáncer de mama, tiroides o en el endometrio.

Los individuos afectados por el síndrome Bannayan-Riley-Ruvalcaba presentan macrocefalia, pólipos hamartomatosos en el intestino y tumores adiposos benignos en la piel (lipomas). Además, antes y después del nacimiento, estos individuos muestran un crecimiento exagerado.

El síndrome Proteus es una enfermedad rara que se caracteriza por el crecimiento desproporcionado de algunas partes del cuerpo, principalmente, el tejido óseo, la piel, el sistema nervioso central, los ojos y el tejido conectivo.

El término Proteus-like se utiliza para describir aquellos que muestran características similares al síndrome Proteus pero que carecen de los criterios de diagnósticos específicos de esta enfermedad, para el síndrome de Cowden y el síndrome Bannayan-Riley-Ruvalcaba (Zhou y col., 2003; Nagy y col., 2004).

2.1.8 Síndrome de poliposis mixta hereditaria

El síndrome de la poliposis mixta hereditaria se caracteriza por la aparición de pólipos adenomatosos, hiperplásicos y hamartomatosos. El riesgo de desarrollar cáncer colorrectal es de un 30% y la edad media de diagnóstico es de 47 años. Se desconoce el gen que provoca la enfermedad (Whitelaw y col., 1997, Rustgy y col., 2007).

3. Cáncer hereditario de colon no polipósico (CCHNP) o Síndrome de Lynch

El síndrome de Lynch (SL) es la forma de cáncer colorrectal hereditario más común, manifestándose en un 2-4% de todos los casos de cáncer colorrectal (la Chapelle y col., 2004; Lynch y col., 2009). Los pacientes diagnosticados con este síndrome desarrollan CCR a una edad media de 45 años, aproximadamente 20 años antes que la población general. Los tumores que desarrollan son, en un

70%, de tipo proximal. Además, el paciente puede presentar cáncer colorrectal múltiple con un desarrollo sincrónico o metacrónico. El tumor asociado al síndrome de Lynch presenta, con frecuencia, una histología característica con células en anillo de sello, infiltración linfocítica peritumoral, diferenciación mucinosa, reacción linfocítica de Crohn o un patrón de crecimiento medular (Lynch y col., 2009).

La enfermedad se caracteriza por el desarrollo de, al menos, un adenoma en el colon o el recto, por el diagnóstico a edades tempranas (antes de los 45 ó 50 años), y por el aumento de probabilidad de que ese adenoma degenera en un tumor maligno con el paso del tiempo (la Chapelle y col. 2004; Lynch y col., 2009).

Existen otros tumores extracolónicos asociados al síndrome de Lynch son los tumores urológicos en varones, los tumores de ovario en mujeres, tumores gástricos, tumores hepatobiliares, tumores pancreáticos y tumores cerebrales. Los tumores urológicos son los más frecuentes de todos los tumores extracolónicos en varones. En concreto, los tumores epiteliales del tracto del sistema urinario (riñón y uréter) (Figura 4). La edad de aparición de los tumores urinarios asociados al síndrome de Lynch se produce entre los 50 y los 70 años (Watson y col., 2008). Los tumores de ovario se presentan a edades tempranas entre los 45 y los 50 años (Watson y col., 2008). Los tumores gástricos y los cerebrales, aparecen con menos frecuencia en estos pacientes con síndrome de Lynch (Watson y col., 2008).

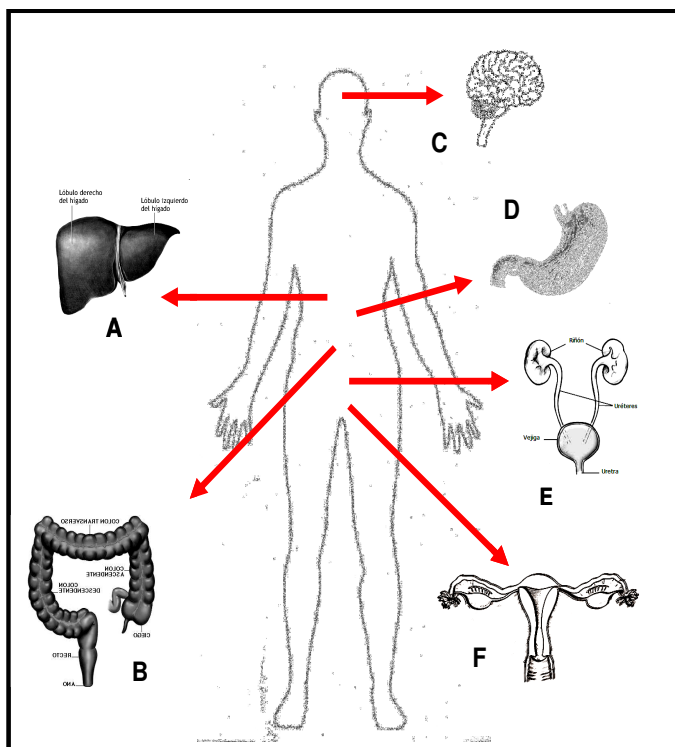


Figura 5: tumores más frecuentes asociados al síndrome de Lynch. A: sistema hepatobiliar y páncreas. B: colon. C: cerebro. D: estómago. E: aparato urinario. F: endometrio y ovario.

La primera descripción de una familia con síndrome de Lynch fue realizada en 1913 por Aldred Scout Warthin, patólogo de la Universidad de Michigan (Warthin y col., 1931). Esta familia, conocida como la familia G (figura 5), presentaba una incidencia inusualmente alta de cáncer colorrectal, gástrico y de endometrio a edades jóvenes. Sesenta años más tarde, Henry Lynch y Anne Krush actualizaron la familia G e identificaron a otras familias con características parecidas (Lynch y Krush, 1971). En 1990 se constituyó un grupo cooperativo, conocido como InSIGHT (<http://www.insight-group.org>), para promover el estudio de esta entidad. La cooperación internacional en esta área del conocimiento ha facilitado el estudio de este síndrome y ha permitido la identificación de las alteraciones genéticas de este síndrome.

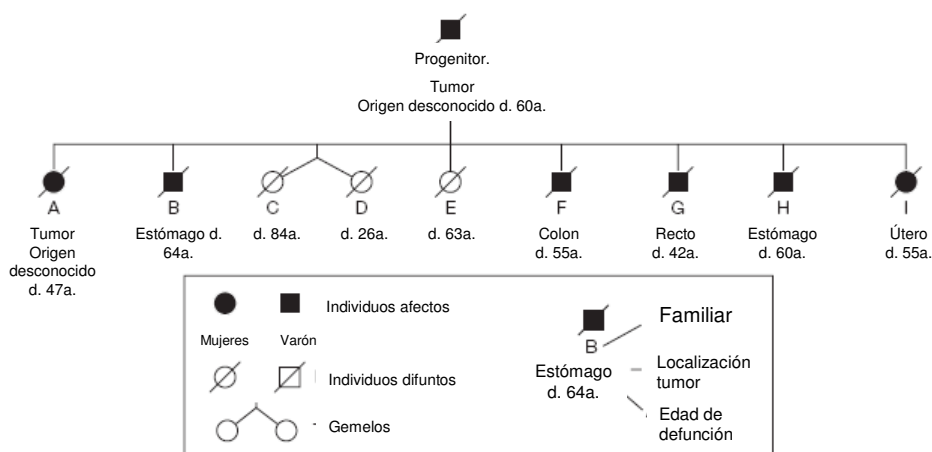


Figura 5: árbol familiar de la familia G (Imagen modificada de Douglas y col., 2005).

El cáncer de colon hereditario no polipósico (CCHNP) es el término más utilizado para identificar este síndrome; sin embargo, no parece el más adecuado ya que la entidad no se limita a la presencia de CCR en la familia. Actualmente, también se le denomina síndrome de Lynch y este término identifica a una entidad de expresión fenotípica variable causada por mutaciones en genes reparadores de los errores de replicación del DNA.

El síndrome de Lynch es una enfermedad hereditaria autosómica dominante en la que la penetrancia de la mutación es de hasta el 80%.

En 1993, Peltomäki y Lindblom identificaron de manera independiente dos *loci* potencialmente implicados en el síndrome de Lynch, uno en el cromosoma 2 y otro en el cromosoma 3 (Lynch y col., 2009; Peltomäki y col., 1993; Lindblom y col., 1993). En estos loci aparecen codificados algunos de los genes que intervienen en la reparación post-replicativa del DNA (*Mismatch Repair genes*, MMR). Durante este mismo periodo, se demostró que los tumores relacionados con síndrome de Lynch tenían unas

características moleculares particulares a las que denominaron de manera genérica “fenotipo mutador”. El fenotipo mutador se refiere, en realidad, a la presencia de alta inestabilidad de microsatélites (MSI) en estos tumores (Lynch y col., 2009; Aaltonen y col., 1993). Los microsatélites son repeticiones de 1 a 6 nucleótidos que se localizan en secuencias no codificantes a lo largo del genoma. El proceso de replicación en estas zonas del genoma tiene tendencia a ser errónea y suele ser corregida por mecanismos de reparación post-replicativos del DNA. La ineficiencia en la actividad reparadora de este mecanismo permite la acumulación de errores en los microsatélites, denominado inestabilidad de microsatélites (Lynch y col., 2009).

3.1 El diagnóstico del síndrome de Lynch:

A) Criterios clínicos del síndrome de Lynch: Ámsterdam, Bethesda y modificaciones.

El diagnóstico del síndrome de Lynch resulta complicado debido a la gran heterogeneidad clínica que presenta este síndrome. Por esta causa, se han desarrollado modelos de criterios mínimos en los que incluir a los pacientes con esta enfermedad. En 1990 se establecieron los primeros criterios, denominados “Criterios de Ámsterdam” (Vasen et al. 1991); sin embargo, estos requisitos resultaron ser demasiado estrictos y excluían a familias con pocos miembros o con manifestaciones extracolónicas de la enfermedad; tampoco incluían las características anatomopatológicas de los tumores colorrectales. Con los años, los criterios de Ámsterdam se han ido modificando, ampliando y corrigiendo con el fin de asegurar un diagnóstico fiable de la enfermedad (Vasen et al., 1999; Benatti y col., 1993; Bellacosa y col., 1996; Lynch y col. 2004) (tabla 3).

En 1996, se propusieron los criterios de Bethesda, que describen prácticamente todas las condiciones clínicas en las que existe sospecha de síndrome de Lynch. Estos nuevos criterios tienen en cuenta también características del tumor, su localización, histología y la edad aparición (Rodríguez-Bigas y col. 1997; Umar y col. 2004).

B) Estudio de inestabilidad de microsatélites:

La inestabilidad de microsatélites (IMS) se encuentra en el 15% de todos los tumores colorrectales. De todos estos tumores con IMS, el 75-80% pertenecen a la categoría de cáncer de colon esporádico y sólo el 20-25% son tumores pertenecientes al síndrome de Lynch. El análisis de la inestabilidad de microsatélites es un método sensible ampliamente utilizado como criterio de inclusión. Además, la técnica inmunohistoquímica (IHQ) permite observar la presencia o ausencia de las proteínas codificadas por los genes implicados en el sistema reparador mediante la tinción del tejido tumoral con anticuerpos específicos para cada proteína con una sensibilidad del 95%.

Tabla 3: criterios clínicos del síndrome de Lynch. (Vasen y col., 1991; Vasen y col., 1999; Umar y col., 2004).

Criterios de Ámsterdam (tienen que cumplirse los 3 criterios)	Criterios de Ámsterdam II
Tres o más familiares con cáncer colorrectal.	Tres o más familiares con cáncer asociados a HNPCC (cáncer colorrectal o cáncer de endometrio, intestino delgado, uréter o pelvis renal).
Al menos dos generaciones afectadas.	Uno de los afectados de la familia debería ser familiar de primer grado de los otros dos.
Al menos un caso de cáncer colorrectal diagnosticado antes de los 50 años de edad.	Dos o más generaciones sucesivas deberían estar afectadas.
	El tumor en uno o más familiares afectados debería ser diagnosticado antes de los 50 años de edad.
Criterios de Bethesda I	Criterios Bethesda II
Incluye a familias que cumplen los criterios de Ámsterdam.	Los criterios son los mismos que para Bethesda I, pero la edad de diagnóstico pasa de 45 a 50 años.
Individuos con dos cánceres asociados al CCHNP, incluyendo cánceres colorrectales sincrónicos y metacrónicos o cánceres extracolónicos asociados al CCHNP.	
Individuos con cáncer colorrectal y un familiar de primer grado con cáncer colorrectal y/o un cáncer extracolónico asociado a CCHNP y/o un adenoma colorrectal; uno de los cánceres diagnosticado antes de los 45 años y el adenoma antes de los 40 años.	
Cáncer colorrectal o endometrial diagnosticado antes de los 45 años.	
Cáncer colorrectal derecho con patrón diferenciado antes de los 45 años.	
Cáncer colorrectal de células en anillo de sello diagnosticado antes de los 45 años.	
Adenomas diagnosticados antes de los 40 años.	
	Bethesda revisado
	Cáncer colorrectal diagnosticado antes de los 50 años.
	Cánceres colorrectales u otros tumores asociados a CCHNP, sincrónicos o metacrónicos, independientemente de la edad del diagnóstico.
	Cáncer colorrectal diagnosticado antes de los 60 años con hallazgos histológicos de linfocitos, reacción linfocitaria tipo Crohn, diferenciación mucinosa/anillo de sello o patrón de crecimiento medular.
	Cáncer colorrectal diagnosticado en dos o más familiares con tumores asociados al CCHNP, independientemente de la edad del diagnóstico.

3.2 Bases moleculares del síndrome de Lynch

3.2.1 Mecanismos de reparación del DNA:

El DNA es una molécula estable en la que, por causa de agentes mutagénicos o errores replicativos, se pueden introducir cambios en su secuencia. Estas alteraciones pueden modificar la información contenida en la molécula. Para minimizar su incidencia e impedir su transmisión de generación en generación, en la célula existen diferentes mecanismos capaces de reparar estas lesiones.

Se han descrito al menos cinco mecanismos de reparación del DNA: reparación por recombinación homóloga (HRR), unión de extremos no homólogos (NHEJ), reparación por escisión de nucleótidos (NER), reparación por escisión de bases (BER) y el mecanismo reparador de desapareamientos (MMR). Todos ellos tienen en común un mecanismo similar en el que aprovechan la redundancia de información que proporciona la estructura en doble hélice del DNA. Estos procesos reparan la hebra dañada copiando la secuencia de la hebra complementaria y “sana”. El proceso básico de los mecanismos de reparación conlleva tres etapas (figura 6):

1. Reconocimiento y eliminación de la porción alterada de la cadena de DNA. Las enzimas encargadas de este paso son nucleasas capaces de hidrolizar los enlaces fosfodiéster que unen los nucleótidos alterados.
2. Reparación de hueco en la cadena de DNA producido por la acción de las nucleasas. Una DNA polimerasa añade los nucleótidos correctos a partir del extremo 3'OH utilizando como molde la hebra complementaria.
3. Una ligasa une el fragmento recién sintetizado a los nucleótidos circundantes de la cadena de DNA.

En general, este tipo de reparación es útil para todo el rango de errores que se produce en la molécula de DNA, que abarca desde cambios de una sola base hasta roturas de doble cadena. Sin embargo, las roturas de doble cadena son más complejas de arreglar y el sistema de reparación por escisión y síntesis a partir de la hebra “sana” no es viable.

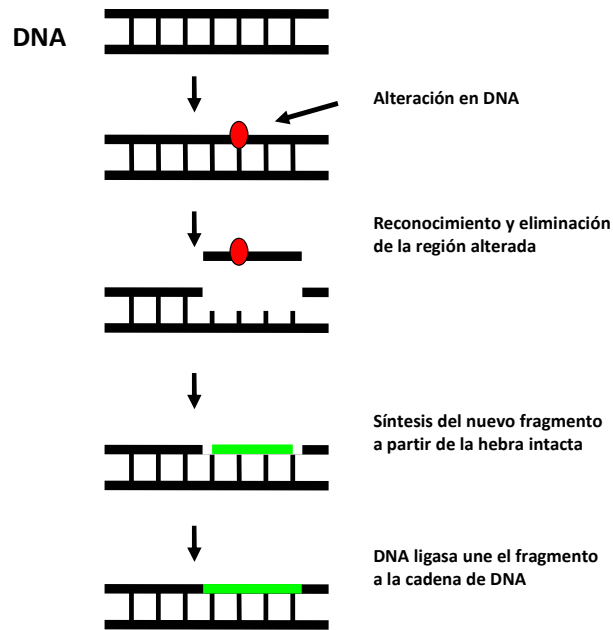


Figura 6: mecanismo general básico de reparación de DNA.

3.2.2 Sistema de reparación de errores por desapareamiento (Mismatch Repair system, MMR)

Durante la replicación del DNA, algunos de los errores cometidos por la polimerasa son corregidos por su propia actividad exonucleasa 3'-5'; sin embargo, esta enzima por sí misma no es capaz de eliminar todos los errores de la nueva cadena de DNA y deben ser reparados por el sistema de reparación de desapareamientos.

El mecanismo general de reparación de errores en DNA (MMR) es un proceso crítico responsable del mantenimiento de la integridad genómica a través de la corrección post-replicativa de errores por apareamientos erróneos de bases y por bucles provocados por pequeñas pérdidas o inserciones de unos pocos nucleótidos. Su inactivación en humanos es la causa de la inestabilidad genética que se asocia con el desarrollo cáncer colorrectal hereditario no polipósico (CCHNP) y una fracción de tumores esporádicos que se producen en algunos tejidos (Dherin y col., 2008)

La reparación de errores post-replicativos no es la única función de este sistema, también participa en la reparación de errores durante la recombinación homóloga entre secuencias casi idénticas, en el aborto del proceso de recombinación (actuando como barrera en la recombinación interespecífica) y en los procesos de respuesta celular a lesiones en el DNA provocados por agentes genotóxicos alquilantes (Edelbrock y col., 2009, Wang y col., 2004).

Este sistema de reparación está muy conservado a través de la evolución y, por eso, los componentes encontrados en estudios iniciales sobre organismos procariotas han podido extrapolarse y confirmarse en diferentes organismos eucariotas (Edelbrock y col., 2009).

3.2.3 Sistema reparador en procariotas:

El sistema de reparación MMR se descubrió inicialmente en *Escherichia coli* y a los genes que codificaban las proteínas encargadas de la reparación se les denominó "Mut" debido a que su inactivación generaba hipermutabilidad en las células. En el mecanismo procariota de la reparación aparecen involucradas principalmente tres proteínas: MutS, MutL y MutH (figura 7).

MutS actúa como un homodímero constituido por dos subunidades de 95 kDa, que tiene la capacidad de detectar varios tipos de errores en el DNA y unirse a ellos con diferente afinidad. Además, esta molécula manifiesta actividad ATPasa y, por ello, se le incluye dentro de la superfamilia de proteínas transportadoras ABC (ATP-binding cassette). Su estructura tridimensional se asemeja a un disco oval con una apertura en el centro con la que interactúa con el DNA (Kunz y col., 2008, Lamers y col., 2000, Obmolova y col., 2000).

Durante la reparación, MutS se une inespecíficamente a la molécula de DNA y se desliza por la doble cadena hasta que tropieza con una lesión. Al reconocer la alteración, MutS activa el proceso de reparación (Hsieh, 2001, Gammie y col., 2007).

La importancia de MutL en este proceso ha sido un misterio durante muchos años al no poder atribuirle ninguna función enzimática. La actividad principal de MutL es la de reclutar a los otros componentes del sistema y es una pieza importante en el ensamblaje del complejo de reparación. MutL activa MutH mediante una reacción dependiente de ATP. Además, MutL interactúa con la DNA-helicasa II y la dirige hacia el error; por último, se une con el homodímero MutS reforzando la unión de éste último a la alteración en la molécula. MutL recluta también a diferentes proteínas encargadas de la escisión de la hebra dañada y de la síntesis de la nueva hebra correcta (Hsieh, 2001, Fukui, 2010).

La diana del sistema reparador es la alteración localizada en la hebra recién sintetizada después de la replicación. En los organismos procariotas, el sistema reparador es capaz de distinguir la hebra madre de la hija a través de la metilación parcial transitoria de secuencias GATC que se da en éstas últimas durante y segundos después de la replicación (Kunz y col., 2008, Fukui, 2010).

MutH surca, en forma de monómero, la hebra hija y reconoce estas secuencias GATC. MutH tiene actividad endonucleasa y, cuando es activada por MutL, origina una pequeña escisión en el sitio

GATC más cercano al lugar donde se localiza el error en la cadena de DNA (Kunz y col., 2008, Fukui, 2010).

El pequeño hueco que deja MutH permite la acción de la DNA-helicasaII (UvrD) y algunas exonucleasas redundantes específicas de DNA de cadena sencilla como son ExoI, ExoV, RecJ y ExoVII (Kunz y col., 2008). Dependiendo de si el hueco producido por MutH se sitúa en el flanco 5' o 3' del error, la escisión del fragmento a reparar se produce en sentido 5'-3' ó 3'-5'. En sentido 5'-3', las encargadas de la escisión son RecJ y ExoVII; mientras que en el sentido opuesto, actúan ExoI y ExoV (Kunz y col., 2008, Fukui, 2010).

Mientras la DNA-helicasaII separa ambas hebras, las exonucleasas hidrolizan el fragmento de DNA que incluye la secuencia GATC de la que partió MutH hasta la zona donde se encuentra la alteración y, rápidamente, una DNA-polimerasaIII re-sintetiza el fragmento ayudada por proteínas de unión a cadena simple (SSB) y DNA-ligasa (Hsieh, 2001, Fukui, 2010).

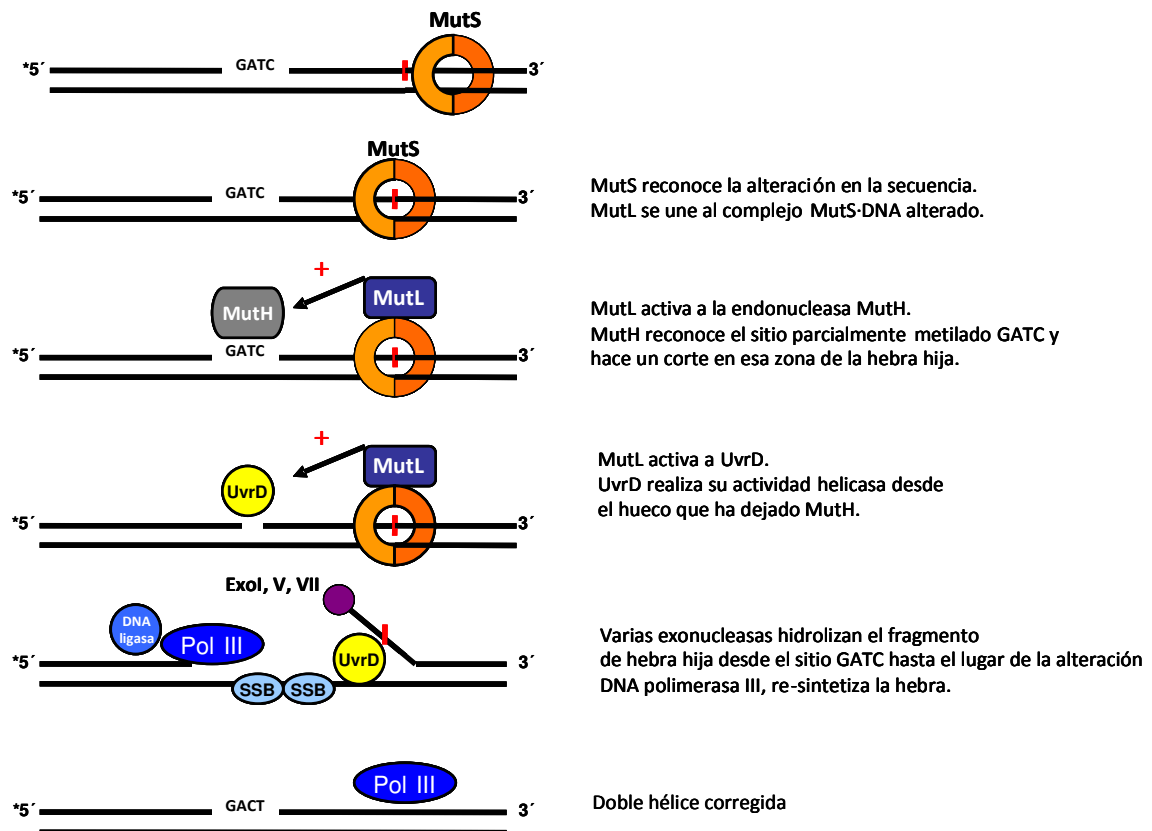


Figura 7: esquema del sistema de reparación de errores en *E.coli*. La marca roja sobre la doble hebra de DNA corresponde con un desapareamiento provocado por una mutación. GATC: secuencia característica de la metilación parcial de la hebra hijas. MutS, MutL y MutH: proteínas implicadas en el sistema reparador de desapareamientos. UvrD: DNA helicasa II. Exo I, V y VII: endonucleasas. Pol III: Polimerasa III. SSB: proteínas de unión a la cadena de DNA simple.

3.2.4 Sistema de reparación en organismos eucariotas:

Aunque el sistema de reparación es un proceso muy bien conservado a lo largo de la evolución, existen diferencias significativas entre el mecanismo procarionta y el eucariota. No obstante, en los organismos eucariotas se han encontrado proteínas homólogas a las proteínas implicadas en el sistema de reparación de los organismos procariontas (tabla 4).

Así, en levaduras se han descrito seis proteínas homólogas a MutS que actúan en forma de heterodímeros. MSH1 es la única de las seis proteínas descritas en levadura que actúa exclusivamente en la mitocondria, mientras que el resto (MSH2-6) tienen su función en el núcleo. MSH2, MSH3 y MSH6 actúan en los procesos de reparación del DNA post-replicativos. MSH4 y MSH5 funcionan como heterodímero durante la meiosis (Kolodner, 1999, Hsieh, 2001; Fukui, 2010).

Tabla 4: correlación de proteínas integrantes del sistema reparador de DNA en organismos procariontas (estudios en *E. coli*) y en organismos eucariotas (levaduras, *S. cerevisiae*, y humanos). NI: No identificada (Tabla modificada de Fukui, 2010).

Función en el sistema reparador	Procariontas (<i>E. coli</i>)	Eucariotas (<i>S. cerevisiae</i>)	Eucariotas (<i>H. sapiens</i>)
Reconocimiento del error	MutS	MSH2, MSH6, MSH3	MSH2, MSH6, MSH3
		MSH1	NI
		MSH4, MSH5	MSH4, MSH5
Fijación al error	MutL	MLH1	MLH1
		PMS1	PMS2
		MLH2, MLH3	PMS1
Incisión de la hebra mutada	MutH	NI	NI
Unión a la hebra monocatenaria	SSB	RPA	RPA
Escisión de la hebra mutada	UvrD (DNA-Helicasa)	NI	NI
	RecJ, EXO I, EXO VII, EXO X	EXO I	EXO I
Síntesis de la hebra reparada	DNA polimerasa III	DNA polimerasa δ	DNA polimerasa δ

En mamíferos se han descrito cinco homólogos de MutS: MSH2, MSH3, MSH4, MSH5 y MSH6. Al igual que en levaduras, estas proteínas ejercen su función en forma de heterodímeros. Los heterodímeros formados por MSH2-MSH6, también denominado MutS α ; y MSH2-MSH3, MutS β , son necesarios para el reconocimiento de la alteración en la secuencia de DNA durante el proceso de reparación. Estos dos complejos proteicos no actúan frente al mismo tipo de error: MutS α , detecta cambios de un nucleótido en la secuencia y bucles en la estructura producidos por pérdidas o inserciones menores de 10 nucleótidos; mientras que MutS β interacciona cuando en la secuencia aparecen

deleciones o inserciones mayores de 10 nucleótidos (Kolodner y col., 1999; Hsieh, 2001; Kunz y col., 2008, Fukui, 2010).

En humanos, el complejo MutSα es el más abundante y activo. Las dos subunidades proteicas que lo forman, muy similares entre sí, constan de cinco dominios funcionales (Lamers y col., 2000; Ollila y col., 2008): Dominio de unión a DNA, dominio conector, “lever domain”, “domain clamp” y el dominio ATPasa. Además, en el extremo C-terminal existe un motivo hélice-giro-hélice con el que se cree que se estabilizan los dominios ATPasa.

El dominio de unión a DNA en MSH2 (figura 8) se localiza en el extremo N-terminal (Malkov y col., 1997; Marti y col, 2002) y se une de manera inespecifica con la hebra en las proximidades de la lesión. En la especificidad de la unión del complejo MutSα al error en la molécula de DNA la proporciona la segunda proteína del heterodímero. (Dufner y col., 2000; Warren y col., 2007).

El dominio conector une el dominio de unión a DNA con el resto de la proteína y es el responsable de las interacciones intramoleculares entre los diferentes dominios (Warren y col., 2007). Los dominios 3 y 4, “lever” y “clamp”, aparecen como mediadores entre los dominios de unión al ATP y de unión al DNA (Ollila y col., 2008).

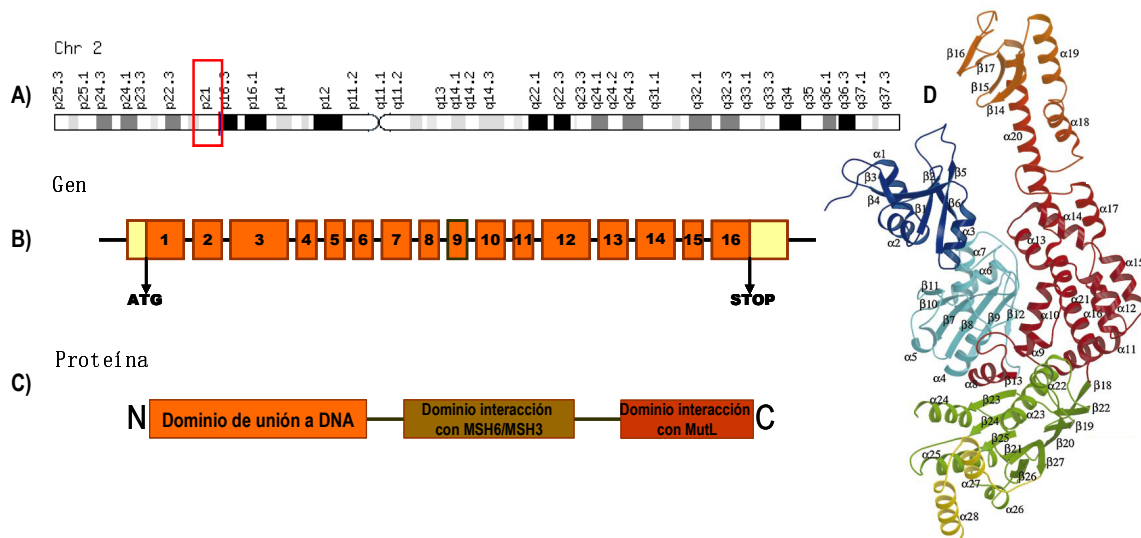


Figura 8: A) El gen MSH2 se localiza en el cromosoma 2p21. B) Esquema de los exones del gen MSH2. C) Esquema simplificado de la estructura de la proteína MSH2. D) Estructura tridimensional de MSH2.

En el extremo C-terminal de MSH2 y MSH6 se localiza el dominio de unión e hidrólisis de ATP. Esta zona de la proteína es crucial en la función del complejo MutSα debido a que modula la conformación de la proteína según su unión a ADP o ATP (Gradia y col., 1997; Warren y col., 2007; Ollila y col., 2008). El intercambio de ADP por ATP en el complejo proteico produce un cambio conformacional que permite el deslizamiento del complejo desde la lesión a lo largo de la hebra de DNA y facilita al

complejo MutS α la búsqueda de las señales que le permitan reconocer a la hebra recién sintetizada (Hsieh, 2001; Kunz y col., 2008). El dominio C-terminal también muestra una estructura en hélice-giro-hélice muy conservada en las proteínas homólogas a MutS. MSH2 tiene la capacidad de unirse a través de esta vía con la proteína MSH6 por el extremo C-terminal y el motivo hélice-giro-hélice parece estar implicado en los cambios en la interacción entre MSH2 y MSH6 (Alani y col., 1996; Marti y col., 2002).

Hasta el momento se han descrito cuatro homólogos de MutL en los organismos eucariotas: MLH1, MLH2, MLH3 y PMS2 (en mamíferos) o PMS1 (en levaduras, plantas y nematodos). Al igual que los homólogos de MutS, estas proteínas actúan como heterodímeros.

En humanos el heterodímero MutL α , formado por la interacción de MLH1 y PMS2, actúa en presencia tanto del complejo MutS α como del complejo MutS β ; mientras que la molécula formada por MLH1 y MLH3, denominada MutL β , interactúa con MutS α (Marsischky y col., 1999).

La proteína MLH1 resulta esencial para el proceso de reparación (figura 9). Consta de un dominio amino-terminal (aproximadamente entre el aminoácido 1 y el 350) muy conservado entre especies, con actividad ATPasa (unión e hidrólisis de ATP), un dominio de unión ("linker") que ocupa la secuencia de aminoácidos desde el residuo 350 hasta el 500 de la proteína y, por último, desde el aminoácido 500 hasta el final de la proteína, la zona carboxi-terminal donde se ubica el dominio de interacción con PMS2, MLH3 y PMS1. El dominio de unión o linker codifica una secuencia de localización nuclear y su actividad parece depender del dominio de interacción situado en la zona carboxi-terminal de la proteína (Mohd y col., 2006).

En el extremo carboxi-terminal de MLH1 se sitúa además una pequeña zona, aparentemente única, denominada dominio homólogo carboxi-terminal o CTH. Este dominio aparece muy conservado en mamíferos, insectos, plantas, hongos y gusanos, pero no parece estar presente en proteínas procariotas MutL o en otros componentes de la familia proteica como son PMS2, MLH3 o PMS1.

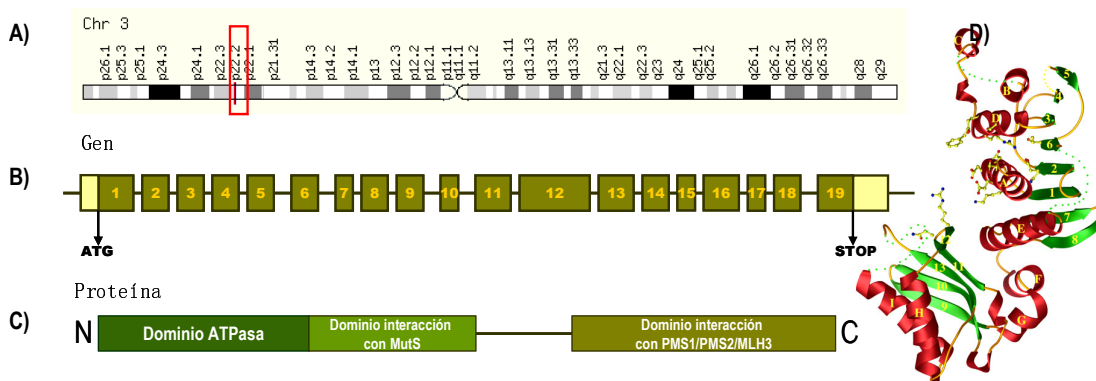


Figura 9: A) El gen MLH1 se localiza en el cromosoma 3p22.2. B) Esquema de los 19 exones del gen MLH1, C) Esquema simplificado de la estructura de la proteína MLH1. D) Estructura tridimensional de MLH1.

Mientras, en procariontas, las zonas GATC semi-metiladas transitoriamente sirven como señalización de la hebra recién sintetizada; en eucariotas no se ha aclarado este mecanismo; sin embargo, se ha postulado que las discontinuidades asociadas a la replicación de la nueva hebra podrían servir como marcador (Kunz y col., 2008). Una vez reconocida la hebra hija, se produce la unión de MutS α o MutS β con el desapareamiento. La interacción de uno u otro heterodímero depende del tipo de error: MutS α , tiene afinidad por los cambios de un único nucleótido o por pérdidas e inserciones de pocos nucleótidos; MutS β , se une con mayor eficiencia a pérdidas e inserciones de más de diez nucleótidos (Kolodner y col., 1999; Kunz y col., 2008; Fukui y col., 2010).

A este complejo formado por MSH2-MSH6 (MutS α) o MSH2-MSH3 (MutS β) y el DNA desapareado se le une uno de los dos heterodímeros MutL α o MutL β . Se ha observado que la molécula formada por MutL α tiene afinidad tanto por MutS α como por MutS β . Al contrario, MutL β parece tener la capacidad de unirse exclusivamente a MutS β (figura 10) (Marsischky y col., 1999; Fukui y col., 2010).

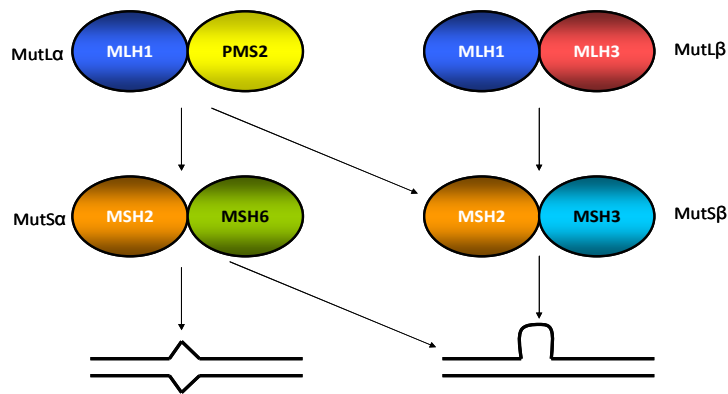


Figura 10: interacciones entre los complejos MutS y los complejos MutL durante el proceso de reparación en humanos. Las interacciones son dependientes del tipo de la secuencia de DNA (imagen modificada de Marsischky y col., 1999)

Si aceptamos la señalización por medio de discontinuidades en la nueva hebra, éstas pueden estar localizadas a ambos lados del error, por lo que la reparación se puede dar tanto en sentido 5´-3´ como en sentido 3´-5´, es decir, en eucariotas el proceso de reparación también sería bidireccional. Sin embargo, en eucariotas no se ha observado que ninguna otra exonucleasa, aparte de EXO1, actúe en este tipo de reparación (Iyer y col., 2006; Fukui y col., 2010).

MutS α activa a EXO1 y forman un complejo que es controlado por la proteína de replicación RPA. Mientras EXO1 escinde nucleótido a nucleótido el fragmento de DNA que contiene el desapareamiento, la proteína RPA se une al DNA monocatenario que se está sintetizando y lo protege. El efecto de RPA se desarrolla en dos planos: reduce la actividad del complejo MutS α -EXO1 e impide la unión de más nucleasas a la hebra de DNA monocatenaria (Iyer y col., 2006; Kunz y col., 2008; Fukui, 2010).

MutL α no interviene en la activación de EXO1, pero sí está implicado en el proceso de escisión. Parece que este heterodímero fortalece la dependencia al error del complejo y es posible que esté implicado en la terminación de este proceso (Kunz y col., 2008; Fukui, 2010).

EXO1 también es necesaria en los procesos de hidrólisis en sentido 3'-5'. MutL α posee actividad endonucleasa latente (esta actividad se encuentra en la subunidad PMS2) que es activada por la interacción de MutS α , PCNA y RFC en un proceso dependiente de ATP. De modo que MutL α , con su actividad endonucleasa, genera un punto de entrada para EXO1 que comienza a degradar la cadena de DNA en sentido 5'-3' (Kunz y col., 2008; Fukui, 2010).

La polimerasa δ junto con PCNA se encarga de la reparación del fragmento que se ha escindido y la DNA ligasa I termina el proceso uniendo los extremos del fragmento recién sintetizado a la hebra de DNA (Kunz y col., 2008; Fukui, 2010).

4. Efecto de las mutaciones en los genes del sistema de reparación de errores en el DNA

Los genes del sistema de reparación de errores tienen un comportamiento similar al de los genes supresores de tumores clásicos; es decir, el cáncer aparece en individuos con una mutación germinal en uno de los alelos cuando el otro alelo se inactiva por cualquier tipo de alteración somática producida en las células del tejido diana (Hemminki y col., 1994 y la Chapelle, 2004). La pérdida de heterocigosidad puede deberse a pérdidas, a mutaciones puntuales o a la metilación de las islas CpG en la región promotora (la Chapelle, 2004).

Las mutaciones en estos genes provocan la aparición del denominado “fenotipo mutador”, es decir, la deficiencia en el sistema reparador provoca un aumento en la frecuencia de mutaciones, sobre todo en secuencias repetitivas de unos pocos nucleótidos, como son las secuencias denominadas microsatélites. La inestabilidad de microsatélites supone un buen marcador de los fallos en el sistema de reparación (la Chapelle, 2004).

4.1 Tipos de mutaciones en los genes implicados en el sistema reparador

El síndrome de Lynch aparece asociado a mutaciones germinales en, al menos, uno de los genes implicados en la vía de reparación de errores del DNA: MLH1, MSH2, MSH6, PMS1 y PMS2 (Marra y col., 1995; Vasen y col., 1996). Las mutaciones aparecen distribuidas a lo largo de toda la secuencia de los genes sin que aparezcan zonas de acumulación de mutaciones (“hot-spots”). La gran

mayoría de los casos de HNPCC se explican por mutaciones en MLH1 o MSH2, mientras que en los otros tres genes, el número de mutaciones descritas y asociadas a la enfermedad se limita a unos pocos casos. Según su efecto sobre la proteína y su función, podemos dividir a las mutaciones en los genes de la ruta de reparación de errores del DNA en tres grupos: patogénicas, polimorfismos (mutaciones neutras) y variantes de significado desconocido.

Las mutaciones patogénicas asociadas al síndrome de Lynch y los polimorfismos detectados en los genes del sistema reparador del DNA aparecen enumerados y clasificados en la base de datos recogida en "The international Collaborative group on Hereditary Non-Poliposis Colorectal Cancer mutation database" (<http://www.insight-group.org>).

Las **mutaciones patogénicas** son aquellas que se pueden considerar como causa principal de la alta susceptibilidad al síndrome de Lynch. El efecto biológico de estas mutaciones resulta en la pérdida de función de la proteína. Existe un amplio rango en el tipo de mutaciones patogénicas y entre ellas se incluyen:

- 1) Las mutaciones de **cambio de marco de lectura** o "**frameshift**" provocan un desfase en la lectura de los tripletes y suelen adelantar la aparición de un codón de parada, produciendo proteínas que son más cortas y que carecen de dominios funcionales de importancia para la actividad de la proteína.
- 2) La mutación **sin sentido** o "**nonsense**", es aquella que, por causa de un cambio en uno o varios nucleótidos dentro de la secuencia, origina un codón de parada prematuro en lugar del aminoácido correspondiente.
- 3) Las mutaciones que afectan a la **maduración de mRNA** también se consideran patogénicas. Son cambios o pérdidas de uno o más nucleótidos en las zonas consenso que intervienen en los procesos de maduración del mRNA.
- 4) Las mutaciones de **cambio de aminoácido** o "**missense**" también pueden resultar perniciosas para la proteína. Éstos son cambios de nucleótido que producen el cambio de aminoácido dentro de la secuencia proteica.
- 5) Las **pérdidas o amplificaciones de grandes fragmentos de un gen** o incluso de la totalidad del mismo. Estos reordenamientos suponen para la proteína la falta o el exceso de dominio importantes y la alteración de su funcionalidad.

Las variantes presentes en la población general con una frecuencia mayor del 1%, al igual que aquellas que en estudios familiares no co-segregan con la enfermedad o aparecen en probandos con otra mutación causante de la enfermedad, se definen como **polimorfismos** o **variantes neutrales**. Dentro de

esta categoría se incluyen **mutaciones silenciosas** o “**silent**” (el cambio de nucleótido en la secuencia codificante del gen no implica un cambio de aminoácido en la proteína), **mutaciones intrónicas**, **cambios en zonas 5´ y 3´ no transcritas** y **mutaciones de cambio de aminoácido** para las cuales se ha estudiado que ni la funcionalidad de la proteína ni la estructura resultan afectadas.

Las **mutaciones de significado desconocido** son aquellas variantes de las cuales se desconoce su repercusión sobre la funcionalidad o la estructura de la proteína. Dentro de este grupo se encuentran mutaciones que cambian el aminoácido, las pérdidas o inserciones que no cambian el marco de lectura y cambios intrónicos. El que una variante de significado desconocido contribuya al fenotipo de la enfermedad o simplemente represente una variante rara en la población, es un problema importante de consecuencias directas para los pacientes de HNPCC. Los portadores de mutaciones patogénicas tienen un riesgo mayor de padecer cáncer de colon y podrían, por tanto, ser monitorizados regularmente para detectar neoplasias a edades tempranas.

Se pueden aplicar varios criterios para evaluar la posible patogenicidad de las variantes de significado desconocido: 1) la aparición *de novo* de la mutación; 2) la segregación de la variante con la enfermedad; 3) la ausencia en individuos control; 4) el cambio de aminoácido que implique un cambio en la polaridad o en el tamaño; 5) la localización del cambio en un dominio dentro de la proteína muy conservado entre especies; 6) el que tenga un efecto en la maduración del mRNA; 7) el que preceda a la pérdida de heterocigosidad, 8) el que suponga la pérdida de proteína en el tumor y 9) la confirmación de su efecto sobre la proteína mediante estudios funcionales (Ou y col., 2007).

La mayoría de los ensayos funcionales están diseñados para medir el efecto biológico o bioquímico de la proteína anómala. Para medir la capacidad de las proteínas para formar complejos proteína-proteína, se realizan ensayos “pull-down” y ensayos de doble-híbrido o la expresión del gen en líneas celulares deficientes en MMR para determinar la estabilización de la unión a la proteína compañera. Otros ensayos se aplican para estudiar el ciclo ADP-ATP que se produce en la formación de dímeros entre la familia de proteínas homólogas MutS (Ou y col., 2007).

En conjunto, MLH1 y MSH2 agrupan respectivamente el 25% y el 49% de todas las mutaciones conocidas en familias con criterios de Ámsterdam (Weber y col., 1997; Wijnen y col., 1997; Devlin y col., 2005). No se han observado diferencias en el efecto sobre el fenotipo de las mutaciones en MLH1 y MSH2, lo que indica que ambas proteínas son de igual importancia en el sistema de reparación.

Las primeras mutaciones en el gen MSH6 asociadas a SL fueron identificadas en 1997. Son menos frecuentes (10% de todas las mutaciones) que las mutaciones en MLH1 y MSH2. Aparecen entre un 2-5% en familias que cumplen criterios de Ámsterdam I y II (Devlin y col., 2008). El fenotipo de los portadores de mutaciones en MSH6 muestra el desarrollo más tardío del tumor colorrectal, en mujeres aumenta el riesgo de aparición de cáncer de endometrio aunque se retrasa la edad de aparición y las familias presentan una penetrancia incompleta (Devlin y col 2008).

La contribución de PMS1 y, PMS2 en el síndrome de Lynch permanece todavía sin definir. Las mutaciones en PMS2 producen la aparición precoz del síndrome de Lynch. Es frecuente que las familias con mutaciones en este gen desarrollen tumores cerebrales a edades tempranas similares a los del síndrome de Turcot (de la Rosa y col., 2000; Nakagawa y col., 2004).

Se han descrito varias mutaciones de cambio de aminoácido y sin sentido (nonsense) en el gen MHL3 en pacientes con CCR, aunque su papel en el desarrollo del síndrome de Lynch ha sido recientemente descartado (Hienonen y col., 2003).

Entre el 30 y el 40% de los casos de familias con criterios de Ámsterdam no tienen evidencia de mutaciones germinales en ninguno de estos genes del sistema reparador post-replicativo de DNA. Cambios o alteraciones en regiones reguladoras no estudiadas o mutaciones en otros genes no identificados, pueden ser la causa de los tumores en estas familias (Papadopoulos y col., 1997; Lindblom y col., 1997).

5. MicroRNAs

Los microRNAs (miRNA) fueron descubiertos por primera vez en 1993 por Victor Ambros, Rosalind Lee y Rhonda Feinbaum durante un estudio sobre el desarrollo de nematodos (Lee y col., 1993; Ambros y col., 2003; Lui y col., 2010). Los microRNA son secuencias de RNA no codificante, cortas y monocatenarias de, aproximadamente, 22 nucleótidos capaces de unirse a secuencias complementarias en las regiones no codificantes localizadas en la región 3' UTR de muchos RNA mensajeros. Su unión a la secuencia complementaria tiene, generalmente, como consecuencia la anulación de actividad del RNA mensajero de modo que los miRNA son considerados un tipo de reguladores post-transcripcionales (Bartel, 2009). Estas secuencias se encuentran conservadas evolutivamente en diferentes especies y en mitocondrias. Hasta el momento se han descrito más de 4000 miRNAs en los estudios multicelulares en eucariotas y más de 700 se han identificado en humanos.

Los miRNAs tienen muchas y diferentes implicaciones biológicas: diferenciación celular, proliferación, apoptosis, en el sistema endocrino, intervienen en la hematopoyesis, en metabolismo lipídico, etc. Modifican los diferentes perfiles de expresión proteica entre los tejidos y parecen ser una de las causas de la diversidad de fenotipos celulares (Lui y col, 2010 y Valeri y col, 2010).

5.1 Nomenclatura de los microRNAs

Para estructurar la enorme cantidad de miRNAs que se han ido identificando; se ha desarrollado una nomenclatura específica. Los miRNAs confirmados experimentalmente se nombran dándoles un número precedido por el prefijo mir- o miR-. El prefijo mir- se aplica a los pre-miRNA, mientras miR- se refiere a la forma madura de la molécula. Los miRNAs con estructuras similares se diferencian con una letra minúscula (Ambros y col., 2003).

5.2 Biogénesis de los microRNA

La mayoría de los microRNA se encuentran codificados en regiones intergénicas o en orientación antisentido dentro de la secuencia exónica de algunos genes y cada uno de estos microRNAs contiene su propio promotor y unidades reguladoras (Bartel, 2004; Rodríguez y col. 2004, Slaby y col, 2009). El estima que el 40% de ellos se localizan en los intrones de genes que codifican proteínas y, con menor frecuencia, en exones. Estos miRNAs pueden compartir el mismo promotor y reguladores de la transcripción de los genes que los contienen (Lagos-Quintana y col., 2001; Bartel, 2004; Slaby y col., 2009).

Otros genes que codifican microRNA muestran un promotor común, de manera que originan los microRNA en unidades policistrónicas de pre-microRNA que luego es procesado. Esta es una estrategia de transcripción común en aquellos miRNAs con homología en la función y en su estructura (Bartel, 2004).

Los promotores de los genes de los microRNAs son similares a los promotores de genes que codifican proteínas, y, por tanto, tienen motivos de unión a proteínas similares a los descritos para las polimerasas (Bartel, 2004).

Por otro lado, durante la maduración de estos miRNA, se observan procesos de edición en al menos el 6% de ellos; hecho que incrementa la diversidad y el alcance de la acción de los mismos (Lui y col., 2010; Budhu y col., 2010).

En el núcleo, la Polimersa II transcribe los microRNA codificados en diferentes partes del genoma al unirse al promotor más cercano a la secuencia del pre-miRNA (Lee y col., 2004). El transcrito

(pri-miRNA) es una secuencia de cien hasta mil nucleótidos protegido por una caperuza en su extremo 5' de adeninas y una cola poli A en su extremo 3'. Se ha observado que algunos de estos pri-miRNA codifican tanto un microRNA como un RNA mensajero (Budhu y col. 2010). Se ha especulado que la Polimerasa III puede transcribir estos microRNAs cuando el microRNA que se transcribe tiene secuencias Alu, tRNA, o unidades promotoras MWIR (Budhu y col., 2010) (figura 11).

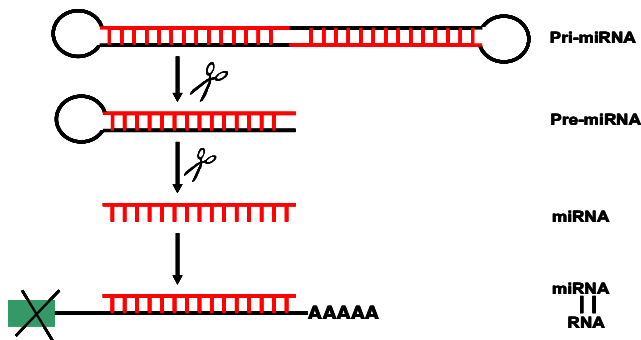


Figura 11: esquema de maduración de microRNA.

Posteriormente a su transcripción, los pri-miRNA son procesados a través de un complejo proteico (complejo proteico microprocesador) hasta pre-miRNA. Estos pri-miRNA contienen al menos una secuencia de un microRNA y hasta 6 miRNA cuando se transcriben en forma policistrónica. Son estructuras en horquilla y bucles de aproximadamente 70 nucleótidos.

5.3 MicroRNA y cáncer

Como ya se ha citado antes, los microRNAs tienen una gran influencia en la regulación biológica. La desregulación de miRNAs puede tener influencias en la carcinogénesis si sus dianas se localizan en genes supresores de tumores u oncogenes. Tanto la sobre-expresión como el silenciamiento de los miRNA se asocian con el desarrollo del cáncer. (Slaby y col., 2009; Faber et al, 2009). La regulación del RNA mensajero a través del miRNA se puede producir por la activación de la transcripción del miRNA o la amplificación del gen que codifica el miRNA o por silenciamiento del miRNA por delección la zona cromosómica del gen que codifica miRNA, por el silenciamiento epigenético o por defectos en la biogénesis del miRNA (Slaby y col, 2009)

Los microRNAs tienen una secuencia específica de unión que se localizan, generalmente, en los extremos 3'UTR de la secuencia del gen, debido a esto, se especula con la idea de que alteraciones en estas secuencias específicas podrían interferir en la formación del complejo miRNA::mRNA (Dong y col., 2011). Se han identificado mutaciones localizadas en el extremo no codificante (UTR) y aunque estas no

alteran la función de la proteína; puede afectar a los niveles de expresión de las proteínas y, por tanto, tener efectos patogénicos (Yu y col, 2007, Dong y col., 2011).

Con respecto al cáncer colorrectal, se están realizando diversos estudios de expresión de microRNAs con la finalidad de utilizarlos como marcadores diagnósticos y pronósticos. Por ejemplo, Landi y colaboradores, en 2008, estudiaron la asociación de la presencia de SNPs en las zonas de unión de miRNA y el riesgo de CCR. Realizaron una selección de 104 genes relacionados con la inflamación, la síntesis de las prostaglandinas y tromboxanos, la obesidad y la resistencia a insulina. Identificaron varios SNPs en algunos de los genes estudiados que indicaban un riesgo en el aumento de CCR a través la alteración de la unión con los microRNA (tabla 5).

Tabla5: SNPS identificados en pacientes con cáncr de colon y recto y microRNAs potencialmente relacionados.

dB SNP ID	Localización	Cambio	miRNA	Efecto	Referencias
rs17281995	3'UTR de CD86	G/C	miR-337	Asociado con aumento del riesgo de CCR en caucásicos	Landi y col., 2008
s1051690	3'UTR de ENSR	G/C	miR-612	Asociado con aumento del riesgo de CCR en caucásicos	Landi y col., 2008
-	Posicion 90 en uc.276	A/G	miR-618	Polimorfismo raro encontrado en pacientes con CCR	Wojcik y col., 2010
rs11077	3'UTR de XPO5	A/C	-	No aparece asociado con la supervivencia.	Boni y col., 2010; Lee y col., 2010
rs7372209	5'UTR de pri-miR-26a-1	C/T	miR-26a-1	Asociación significativa con el ratio de respuesta al tratamiento y el tiempo de progresion.	Boni y col., 2010; Lee y col., 2010
rs1834306	5'UTR de pri-mirR-100	C/T	miR-100	Asociación débil con el tiempo de progresión	Boni y col., 2010; Lee y col., 2010

Hipótesis de trabajo y objetivos

El cáncer de colon no polipósico (síndrome de Lynch) es el síndrome de predisposición al cáncer más frecuente, manifestándose en el 3-5% todos los casos colorrectales.

Es una enfermedad autosómica dominante provocada por alteraciones en alguno de los genes implicados en la ruta de reparación de errores simples de apareamiento en el DNA (sistema MMR): MLH1, MSH2, MSH6, PMS1 ó PMS2.

Los genes MLH1 y MSH2 aparecen más frecuentemente mutados en las familias diagnosticadas con síndrome de Lynch (90%) que en los genes restantes. Una baja proporción (10% de todas las mutaciones) se presentan en el gen MSH6.

La identificación de familias portadoras de mutaciones en los genes de reparación puede ayudar a prevenir la aparición del cáncer y/o detectarlos en fases tempranas, gracias a los protocolos de vigilancia médica y a las medidas preventivas que actualmente se ofrecen a los miembros de las familias con el síndrome de Lynch.

En este trabajo nos propusimos:

1) Realizar la búsqueda de mutaciones en los genes MLH1 y MSH2 en pacientes y familias con sospecha de síndrome de Lynch remitidas a la Unidad de Consejo Genético del Hospital Universitario de Salamanca entre los años 2003 y 2010 con el fin de identificar a las familias portadoras de alteración en alguno de los tres genes.

2) En aquellos pacientes o familias en las que no se detectó mutación puntual, planteamos el análisis de grandes reordenamientos genómicos que nos permitieran identificar las pérdidas o amplificaciones de los genes MLH1, MSH2 y MSH6.

3) Analizar la presencia del polimorfismo c.-93A>G en probandos con sospecha de síndrome de Lynch, individuos con cáncer de colon esporádicos e individuos sanos.

4) Realizar la búsqueda de mutaciones en probandos cuyo análisis de inmunohistoquímica mostraba ausencia de la proteína MLH1 en el tejido tumoral.

5) Identificar mediante herramientas bioinformáticas posibles microRNAs de unión a los genes MLH1, MSH2 y MSH6.

Pacientes y métodos

1. Pacientes

Hemos estudiado 188 probandos pertenecientes a familias con sospecha de síndrome de Lynch remitidas desde diferentes hospitales de la Comunidad de Castilla y León a la Consulta de Consejo Genético del Servicio de Oncología del Hospital Universitario de Salamanca entre los años 2003 y 2010. En todos los casos se obtuvieron muestras de sangre periférica de los probandos previa información y bajo consentimiento informado para los estudios genéticos. La selección de los pacientes se basó en el cumplimiento de al menos uno de los criterios de inclusión en el programa de Consejo Genético del Hospital Universitario de Salamanca establecidos por la Junta de Castilla y León (tabla 6) que son una mezcla de los criterios establecidos internacionales de Amsterdam y Bethesda (tabla 7).

Tabla 6: criterios de la Junta de Castilla y León.

Criterios Programa Consejo Genético	1	Existencia de 3 familiares con cáncer colorrectal.
	2	Existencia de 3 familiares con tumores asociados a síndrome de Lynch.
	3	Un cáncer colorrectal antes de los 50 años.
	4	Dos casos de cáncer colorrectal sincrónico o metacrónico o de otros tumores asociados al cáncer colorrectal.
	5	Un caso de cáncer colorrectal y un familiar de primer grado con tumor asociado al cáncer colorrectal.
	6	Un cáncer de endometrio antes de los 45 años.
	7	Un cáncer de colon derecho con formas histológicas poco diferenciadas (solido/cribiforme), antes de los 45 años.
	8	Pacientes con adenomas colorrectales antes de los 45 años.

Tabla 7: criterios internacionales de Amsterdam y Bethesda. De los criterios de Amsterdam han de cumplirse al menos tres.

Criterios internacionales	AMST	Existencia de tres familiares afectados por cáncer colorrectal.
		Uno debe ser familiar de primer grado de los otros dos
		Deben estar afectados al menos dos generaciones sucesivas
		Uno de los familiares diagnosticado antes de los 50 años.
		Exclusión de PAF
	BETH	Diagnóstico de tumores confirmado histológicamente
		Cáncer colorrectal diagnosticado antes de los 50 años.
		Presencia en el mismo individuo de cáncer colorrectal sincrónico o metacrónico o de otros tumores asociados al cáncer colorrectal
		Cáncer colorrectal antes de los 60 años con presencia de células en anillo de sello o diferenciación mucinosa.
		Cáncer colorrectal con alta inestabilidad de microsatélites
Individuo con cáncer colorrectal y un familiar de primer grado con tumor asociado, uno de ellos diagnosticado antes de los 50 años		

Para los estudios de asociación utilizamos el DNA extraído de la muestra sangre de 145 individuos sanos mayores de sesenta años y 183 individuos con cáncer de colon mayores de 60 años y sin antecedentes familiares de cáncer colorrectal. Las muestras fueron obtenidas bajo consentimiento informado en el Hospital Universitario de Salamanca.

Desde el departamento de Anatomía Patológica del Hospital Universitario de Salamanca obtuvimos doce muestras de tumor fijado en parafina que mostraban la ausencia de la proteína MLH1.

2. Manipulación de ácidos nucleicos

▪ Extracción de DNA de sangre periférica

El DNA de alto peso molecular se extrajo manualmente a partir de muestras de sangre periférica de los pacientes recogidas en tubos con EDTA.

Tras centrifugar las muestras de sangre se aisló la interfase de células leucocitarias, las cuales se resuspendieron en tampón Fornace (0.25M sacarosa; 50mM Tris-HCl pH 7.5; 25mM KCl; 5mM MgCl₂); EDTA (ácido etilendiaminetetraacético) 0.5 pH 8; proteinasa K 20 mg/L y SDS (dodecil sulfato sódico) al 10%. El EDTA es un quelante de iones divalentes que posibilita la inactivación de las nucleasas, la proteinasa K degrada las proteínas, y el SDS es un detergente capaz de romper las membranas celulares. Esta mezcla se incubó a 55°C durante unas 16 horas, tras lo cual se purificó el DNA tratándolo con una mezcla de fenol tamponado y CIAA (cloroformo:alcohol isoamílico 24:1). El DNA contenido en la fase acuosa se limpió de residuos proteicos y detritus celulares mezclándolo con CIAA y centrifugando. Con la adición de etanol absoluto a la fase acuosa, se precipitó el DNA. EL DNA precipitado se lavó con etanol al 70% y, finalmente, se resuspendió en ddH₂O. La muestra de DNA se almacenó en tubos Eppendorf® a – 20° C, con el fin de evitar tanto la degradación progresiva del DNA como la posible contaminación.

La concentración y el grado de contaminación proteica del DNA así obtenido se calcula tras medir la absorbancia a 260 y 280 nm siguiendo la fórmula:

$$\mu\text{g de DNA/ml} = \text{D.O.260} \times (\text{factor de dilución}) \times 50$$

El valor 50 es un factor de corrección introducido, ya que una unidad de densidad óptica con una luz incidente de 260 nm es el valor de absorbancia que tienen 50 µg/mol de DNA.

El cociente D.O.260/D.O.280 se utiliza para determinar el grado de contaminación proteica. Se consideran valores aceptables los comprendidos entre 1.6 y 2.0; siendo el óptimo 1.8. Los valores inferiores a los señalados indican contaminación por proteínas o solventes orgánicos. Los valores superiores indican un exceso de RNA en la muestra.

- **Extracción de RNA a partir de sangre periférica:**

La extracción de RNA se realizó a partir de 1.5 ml sangre periférica de los pacientes recogida en tubos con EDTA siguiendo el protocolo establecidas por el fabricante con el kit QIAamp® RNA Blood MiniKit de QIAGEN. El RNA obtenido se conservó en tubos Eppendorf® de 1.5ml de volumen a -80°C.

- **Extracción de DNA a partir de tejido tumoral parafinado**

A partir de aproximadamente 25mg de tejido tumoral parafinado se extrajo DNA. Primero, la parafina se eliminó mediante tres lavados de la muestra con xileno: se añadió 1200µl de xileno y se agitó la mezcla vigorosamente con vórtex. La muestra se centrifugó a máxima velocidad durante cinco minutos a temperatura ambiente. Se retiró el sobrenadante mediante pipeteo. El xileno se eliminó mediante tres lavados con etanol absoluto: se añadió 1200µl de etanol absoluto y se agitó con vórtex. Se centrifugó a máxima velocidad durante cinco minutos a temperatura ambiente. Posteriormente, se eliminó el sobrenadante mediante pipeteo. La muestra se dejó secar a temperatura ambiente para eliminar los restos de etanol.

La extracción de DNA a partir del tejido tumoral ya desparafinado se realizó siguiendo el protocolo del kit QIAamp® DNA MiniTissue de QIAGEN.

3. Análisis de mutaciones puntuales en los genes MLH1, MSH2 y MSH6:

- **Amplificación por PCR de los genes MLH1, MSH2 y MSH6**

Los exones y las secuencias intrónicas flanqueantes de los genes MLH1, MSH2 y MSH6 de las muestras de DNA fueron amplificadas por medio de la técnica de **PCR** (*Polymerase Chain Reaction*). La amplificación se llevó a cabo en un volumen de 25 µl con los siguientes reactivos:

- 10µl de MasterMix 5'Prime Eppendorf® (contiene Taq Polymerase, KCl, dATP, dGTP, dCTP, dTTP, Mg²⁺ y el detergente Igepal®-CA360).
- 10µl de agua libre de nucleasas.
- 1µl de cada oligonucleótido cebador.

- 1µl de adicional de Mg²⁺.
- Como DNA molde se añadieron 100ng la muestra del paciente obtenido a través del protocolo anteriormente descrito.

Como control negativo se utilizó un tubo de reacción con todos los componentes excepto el DNA molde.

Para la amplificación por PCR de los exones que conforman los genes MLH1, MSH2, y MSH6, respectivamente, se utilizaron varios programas derivados del estándar y adaptados a las exigencias de cada exón.

Programa diseñado para amplificar todos los exones de los genes MLH1, MSH2 y MSH6, excepto los exones 12 y 13 del gen MSH2 (figura 12):

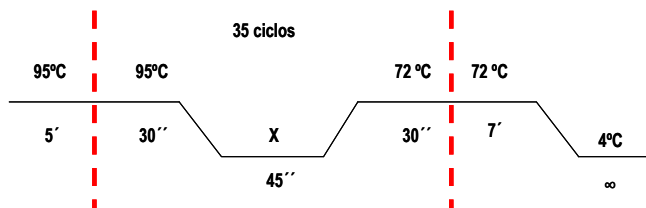


Figura 12: programa de amplificación utilizado para la amplificación de la mayoría de los exones de MLH1, MSH2 y MSH6. X: temperatura de anillamiento para cada par de cebadores que aparecen especificadas en las tablas.

Programa diseñado para amplificar los exones 12 y 13 del gen MSH2 (Figura 13):

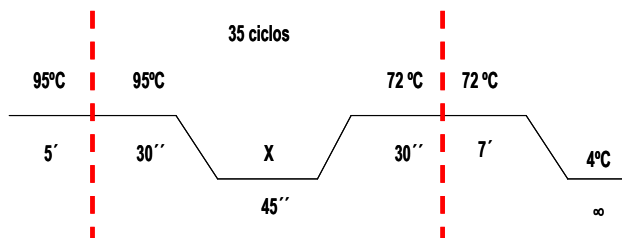


Figura 13: programa de amplificación utilizado para la amplificación de los exones 12 y 13 del gen MSH2. X: temperatura de anillamiento para cada par de cebadores que aparecen especificadas en las tablas.

Los oligonucleótidos utilizados en este trabajo fueron descritos previamente por Hedge y colaboradores (2005), excepto los de los exones 12 y 13 de MSH2, que fueron diseñados en nuestro laboratorio a partir de la secuencia genómica. En las tablas 8, 9 y 10 se recogen los oligonucleótidos utilizados para la amplificación de cada gen.

Tabla 8: oligonucleótidos tomados de Hedge y colaboradores para la amplificación de los exones del gen MLH1. F: oligonucleótido *forward* (sentido); R: oligonucleótido *reverse* (anti-sentido). Secuencia, tamaño del amplicón en pb y temperatura de anillamiento de los oligonucleótidos.

MLH1 Exón	Oligonucleótido (5'-3')	Tamaño amplicón (pb)	Temperatura de anillamiento (°C)
1F	agggtattggctgaaggcac	231	62
1R	gcccgttaagtcgtagccct		
2F	atgtacattagagtagttgcagactgataaatt	221	57
2R	agttccagaacagagaaaggctc		
3F	caagaaaatgggaattcaaagagat	241	55
3R	ctaacaaatgacagacaatgtcatcac		
4F	ccttggtagggtgacagtgg	221	56
4R	caggattactctgagacctaggcaa		
5F	tttccccttgggattagtatctatc	227	53
5R	ccctgaaaacttagaagcaattttattt		
6F	ggacatctgggtttattttcaag	235	56
6R	tgtcaatgtatgagcactagaacaca		
7-8F	gggctctgacatctagtggtgtt	417	56
7-8R	aaaataatgtgatggaatgataaacca		
9F	tctgattctttgtaatgttgagtttg	241	55
9R	cataaaatccctgtgggtgttc		
10F	ctgaggtgattctgactttgtgt	251	59
10R	gaggagagcctgatagaacatctgt		
11F	gtgggcttttctcccct	281	58
11R	ctctcacgtctggccgg		
12F	tttttaatacagactttgctaccaggac	346	55
12R	gttttattacagaataaaggaggtaggctg		
13F	ccaaaatgcaaccacaaaatt	282	58
13R	aacctggcagttgaggcc		
14F	gggtctctagttctgggcctg	271	58
14R	tgctgtgctccctgga		
15F	cccattttgtccaactggtt	203	57
15R	gagagctactatttcagaaacgatcag		
16F	tgggaattcaggctcatttg	292	58
16R	gcacccggctggaaatt		
17F	gcactggagaaatgggattg	221	59
17R	cctccagcacacatgcatg		
18F	agtctgtgatctccglttagaatgag	242	57
18R	ttgtatgaggtcctgtcctagtcct		
19F	catcagccaggacaccagtg	288	58
19R	cggaatacagagaaagaagaacaca		

Tabla 9: oligonucleótidos tomados de Hedge y col. para la amplificación de la región codificante del gen MSH2. F: oligonucleótido *forward* (sentido); R: oligonucleótido *reverse* (anti-sentido). Secuencia, tamaño del amplicón en pb y temperatura de anillamiento de los oligonucleótidos. ^{a, b, c, d} Los oligonucleótidos 5Rb, 12F, 12R, 13F fueron diseñados a partir de la secuencia genómica.

MSH2 Exón	Oligonucleótido (5'-3')	Tamaño amplicón (pb)	Temperatura de anillamiento (°C)
1F	ttcgacatggcggg	285	59
1R	gtccctcccagcacg		
2F	gaagtcagctaatacagtgctga	301	53
2R	aaacacaattaaattctcacattttat		
3F	agagttggattttcctttg	241	57
3R	tcatgtcaattaaagagccttcc		
4F	ttcattttgctttcttctt	316	50
4R	atatgacagaaataccttctaaaaagtcactat		
5F	actggatccagtggtatagaaatctc	285	53
5R	gcttctcagtatatgcaatgaaaca		
5Rb ^a	aaaggtaagggtctgact		55
6F	gcgtagtaaggtttactaatgagc	251	56
6R	catgtggtaactgcaggttaca		
7F	tgagactacgtgcttagtgataaatt	341	53
7R	gcacattgccaagtatatattgtatgag		
8F	tgatgctgtttatctcagtcaaaatt	275	53
8R	aatctacaaacttctaaagtgccct		
9F	gtcttaccattattataggatttgca	217	56
9R	gtatagacaaaagaattccaacctcc		
10F	attgaaaaatggtagtaggtatgatgaa	274	54
10R	cacatcatgttagagcatttaggga		
11F	atatgtttcagtagtacacattgcttca	249	54
11R	tcaaatatcatgattttcttctgtacc		
12F ^b	gggtttgaattcccaaatg	393	46
12R ^c	acaaaacgttaccaccacaa		
13F	tgggggtaatatgggcagta	442	48
13R ^d	tcctctcacaggacagagaca		
14F	tgggcatacctcccaatg	452	55
14R	aataattataacttagaataaggcaattactgat		
15F	tacataaattgctgtcttctcatgc	311	57
15R	aaaaacctcatcttagtgcctgtt		
16F	taattactaatgggacattcacatgtgt	230	55
16R	tacctcattccattactgggatt		

La secuencia flanqueante al exón 5 del gen MSH2 contiene una repetición de 27 adenosinas que imposibilita la correcta amplificación y secuenciación. Para poder analizar la secuencia exónica y la secuencia intrónica se diseñaron un par de cebadores que amplificaban el exón entero. Además se diseñó otro oligonucleótido anti-sentido (*reverse*) que evitaba la repetición de adenosinas problemática.

Tabla 10: oligonucleótidos tomados de Hedge y col. para la amplificación de la región codificante del gen MSH6. **F:** oligonucleótido *forward* (sentido); **R:** oligonucleótido *reverse* (anti-sentido). Secuencia, tamaño del amplicón en pb y temperatura de anillamiento de los oligonucleótidos. ^a El exón 4 del gen está dividido en 7 fragmentos debido a su gran tamaño.

MSH6 Exón	Oligonucleótidos (5'-3')	Tamaño amplicón (pb)	Temperatura de anillamiento (°C)
1F	tgttgattggccactggg	463	65
1R	caaccccctgtgagcctc		
2F	taactgccttaaggaaactgacca	330	60
2R	tcatatagaaaaagtgcctgtctg		
3F	ctggcttgaactgctggat	289	58
3R	cccccttctccccatc		
4.1F ^a	tgacgggtaccattataaagtca	450	58
4.1R ^a	gtattctggttctgatgaaatgctag		
4.2F ^a	gaaggaaacgccctcagc	420	58
4.2R ^a	cagttgccttcatgaataccag		
4.3F ^a	ccacatggatgctcttattgga	420	58
4.3R ^a	tcatctgaaaactgacctatgaaaaact		
4.4F ^a	ttgttgatactcactgggaaagt	420	57
4.4R ^a	ctcctgatcaataaggcatttttg		
4.5F ^a	ctctaggtggtgtcttctacctc	420	57
4.5R ^a	tgagtagcctctcaagatctggaa		
4.6F ^a	cgaagtgtagagcttclaaagaagct	480	57
4.6R ^a	gtcctacagccaattctgttc		
4.7F ^a	agcctcctggaatacctagagaaac	420	58
4.7R ^a	acttattttaggataataacagctggc		
5F	cacttaggctgataaaaccccc	386	57
5R	gtatgttattcctaattgcacaaatgacttt		
6F	aagacaaaagttaagaaactgtactacca	250	56
6R	agaagcaaatatctttatcacatctaaatg		
7F	taacctagaagatgaattatgtaatatgatt	224	53
7R	ttcagataatctctataaaaatagttattgt		
8F	tgagtacttcttatgcatatttact	275	57
8R	aatattagcgatacatgtgctagca		
9F	tgctagcacatgtatcgctaatt	320	56
9R	gcatcatccctcccccttta		
10F	gaagggatgatgcactatgaaaaa	296	52
10R	gtagaaggtagataagaattaaagggttaattt		

- **Electroforesis en gel de agarosa**

El tamaño del producto amplificado y su intensidad se analizaron mediante electroforesis en geles de agarosa preparados al 2% con tampón TBE (Tris 0.044 M, ácido bórico 0.044 M, EDTA 1.0mM pH= 8,3) y teñido con bromuro de etidio. Las muestras migraron en paralelo a un marcador de tamaño de DNA derivado del DNA del fago Φ X-174 cortado por la endonucleasa *Hae III*. Los fragmentos que esta endonucleasa genera en el DNA del fago Φ X-174 se corresponden con los siguientes tamaños: 1.353, 1.078, 0.872, 0.603, 0.310, 0.281, 0.271, 0.234, 0.194, 0.118, 0.072 Kb. La electroforesis se llevó a cabo con un potencial constante de 120 voltios durante 40 minutos.

Para la monitorización de la electroforesis, el DNA se mezcló con dos colorantes incorporados al tampón de carga: el **xileno cianol**, que migra dentro del rango de los fragmentos de 5 Kb en un gel de agarosa al 0.8%, y el **azul de bromofenol**, que migra con los fragmentos de 0.5 Kb.

- **Análisis de los fragmentos amplificados por PCR mediante CSGE (Electroforesis en gel sensible a la conformación)**

Los fragmentos amplificados fueron sometidos a un análisis por CGSE siguiendo la técnica descrita por Orita, con algunas modificaciones (Orita et al., 1989).

Esta técnica requiere una desnaturalización previa de los fragmentos de DNA amplificados por PCR a 95°C y una lenta re-naturalización posterior que permite la formación de heterodúplex (combinación entre una hebra de DNA mutada y otra hebra de DNA no mutada) y homodúplex en el caso de los individuos heterocigotos; y homodúplex en individuos portadores de una mutación en heterocigosis (figura 14). Heterodúplex y homodúplex presentan un patrón de migración diferente en una electroforesis en un gel de acrilamida.

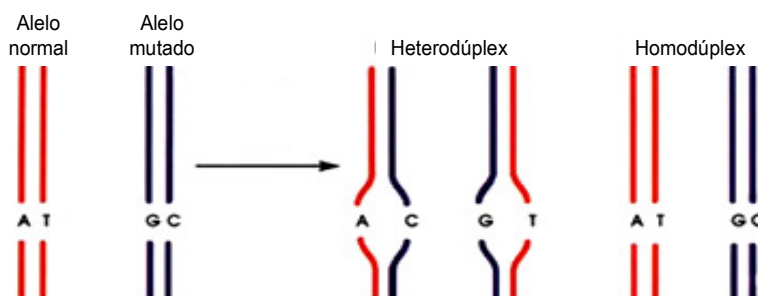


Figura 14: formación de heterodúplex tras la desnaturalización y posterior re-naturalización del producto de PCR en un individuo heterocigoto.

Los geles para electroforesis fueron hechos con MDE™ 2X (AT Biochem, Inc. USA), un polímero de acrilamida modificado derivado del vinilo. Se utilizaron las siguientes cantidades para preparar cada gel:

- Agua destilada: 7.705mL
- Formamida 99%: 5.98mL
- Etilenglicol 99%: 4.025mL
- TBE 10X: 2.415mL
- MDE™ 2X: 20.125mL
- TEMED (N, N, N, N' Tetrametilendiamina): 36.8 µL
- APS 25% (Persulfato de amonio): 138 µL

La electroforesis se llevó a cabo a 180 voltios durante aproximadamente 21 horas (el tiempo varía dependiendo del tamaño del fragmento de PCR).

El gel fue teñido con nitrato de plata. Para ello se utilizó un kit comercial DNA Silver Staining Kit de Amershan Pharmacia®, con un volumen de 250mL, siguiendo las instrucciones del comerciante. Posteriormente, los fragmentos de DNA con patrones diferentes de migración fueron secuenciados.

▪ **Secuenciación automática**

Todos los productos de PCR analizados a través de CGSE y que mostraban un patrón de migración diferente fueron secuenciados para identificar la naturaleza de la variante encontrada.

La muestra amplificada de DNA se purificó con el fin de eliminar cualquier resto de cebadores, dNTPs o iones procedentes de la reacción, que puedan inhibir la reacción de secuenciación. Para la purificación se utilizó el kit comercial "High Pure Product Purification Kit", Roche.

La secuenciación de los productos purificados se llevó a cabo en el secuenciador ABI 377 (Applied Biosystems) en el Servicio Central de Secuenciación de la Universidad de Salamanca y en el secuenciador ABI 3100 (Applied Biosystems) en la Unidad de Genómica y Proteómica del Centro de Investigación del Cáncer (CSIC-Universidad de Salamanca).

Se preparó una muestra con el cebador con sentido y otra con el cebador antisentido a una concentración de 3pmol en un volumen de 2 µL. El DNA se preparó a una concentración de 40 – 60ng para un volumen final de 8µL.

4. Análisis de grandes reordenamientos genómicos

- **MLPA (Multiplex Ligation-dependent Probe Amplification)**

Para el análisis de grandes reordenamientos (pérdidas o amplificaciones) en los genes MLH1 y MSH2 se empleó la técnica de amplificación por PCR multiplex dependiente de ligación (Multiplex Ligation-dependent Probe Amplification, MLPA). La técnica fue descrita por Schouten en 2002 (Schouten y col., 2002). La técnica consta de varias fases: desnaturalización previa del DNA muestra, hibridación de las sondas, ligación y amplificación por PCR multiplex (figura 15).

En esta técnica se utilizan dos sondas diferentes, específicas y adyacentes para cada exón. Cada sonda lleva en su secuencia uno de los oligonucleótidos empleados para la amplificación final de los fragmentos.

En el primer paso de la técnica, cada par de sondas hibridará con el DNA geonómico previamente desnaturalizado de la muestra. Las dos partes de la sonda hibridan con secuencias diana adyacentes. La ligación de las sondas supone el siguiente paso en esta técnica. La secuencia que proporcionan las dos sondas ligadas sirve como molde para la amplificación utilizando cebadores universales. Los productos amplificados se someten a cromatografía capilar y el resultado se visualiza en forma de electroferograma.

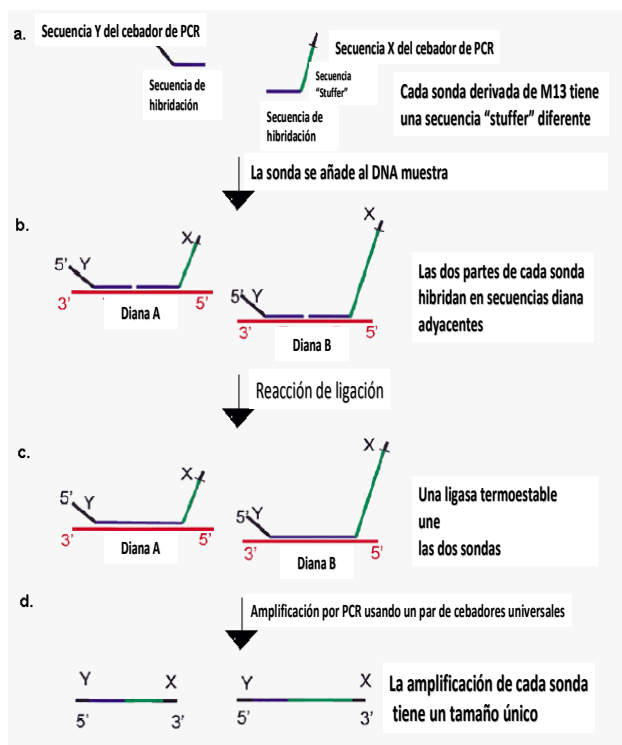


Figura 15: a) esquema de las sondas de MLPA y de las diferentes etapas de la técnica. b) hibridación, c) ligación y d) amplificación.

Los estudios de MLPA se llevaron a cabo con los reactivos SALSA MLPA KIT P003 MLH1/MSH2 y SALSA MLPA KIT P072-B1 MSH6 (MCR-Holland, Ámsterdam, Holanda). El kit para analizar los genes MLH1 y MSH2 contiene una mezcla de 43 sondas para cada uno de los 19 exones de MLH1 y los 16 exones de MSH2; y 7 sondas para otros genes humanos localizados en distintos cromosomas que sirven de control interno del proceso (tabla 11). Además, incluye 4 y 3 fragmentos para el control de la cantidad de DNA, la desnaturalización y la hibridación.

El kit para el análisis de MSH6 contiene 10 sondas para los correspondientes exones del gen. Para algunos exones, existen varias sondas. Además, el kit cuenta con seis sondas que se localizan en el promotor del gen MSH2. Cuatro de estas sondas detectan secuencias del gen EPCAM (o TACSTD1) cuyas deleciones pueden provocar el silenciamiento del gen MSH2. Hay otra sonda diseñada para detectar el exón 1 del gen MSH2 y tres más para el exón 1 del gen MLH1. Por último, incluye también tres sondas capaces de hibridar con el gen MUTYH (tabla 12).

El protocolo aplicado es el recomendado por el fabricante con ligeras modificaciones:

Desnaturalización e hibridación de las sondas: Se diluyeron de 100ng de DNA en ddH₂O hasta un volumen de 5µl. Se desnaturalizaron a 98°C durante 10 minutos y después se atemperaron a 25°C en un termociclador. Se añadió 1,5µl de mezcla de sondas SALSA (SALSA *Probemix*) y 1,5µl de tampón MLPA. Posteriormente, se incubaron un minuto a 95°C e, inmediatamente, a 60°C durante 16-20 horas.

Reacción de ligación: Después de la incubación del DNA con las sondas durante 16 horas a 60°C, se redujo la temperatura del termociclador a 54°C y se añadió la mezcla de ligación que incluye dos tampones, ddH₂O y ligasa 65. Se incubó la muestra con la ligasa a 54°C durante 15 minutos y después se inactivó la enzima calentando la muestra 98°C durante cinco minutos.

Tabla 11: sondas utilizadas en el kit SALSA MLPA P003 MLH1/MSH2, para la detección de grandes reordenamientos en los genes MLH1 y MSH2. La tabla indica la sonda, el exón y el gen donde se localiza, así como la localización cromosómica de las sondas control.

Tamaño (nt)	Sonda SALSA MLPA	Posición cromosómica		
		Referencia	MLH1	MSH2
130	Reference probe 00797-L00463	5q31		
136	Reference probe 00981-L0566	10p12		
142	MLH1 probe 00886-L00474		Exon 1	
148	MSH2 probe 12036-L02162			Exon 1
154	MLH1 probe 01008-L00577		Exon 2	
160	MSH2 probe 00906-L00494			Exon 2
166	MLH1 probe 00888-L00476		Exon 3	
172	MSH2 probe 01029-L00601			Exon 3
178	MLH1 probe 00889-L00477		Exon 4	
184	MSH2 probe 00908-L00496			Exon 4
190	MSH2 probe 11287-L12006			Exon 1
196	Reference probe 3861-L3610	3p24		
202	MLH1 probe 00890-L6064		Exon 5	
211	MSH2 probe 0909-L00497			Exon 5
217	MSH2 probe 13145-L14624			Exón 1
225	MLH1 probe 00891-L14625		Exon 6	
229	MSH2 probe 00910-L00498			Exon 6
238	MLH1 probe 00892-L00480		Exon 7	
247	MSH2 probe 11634-L12398			Exon 7
256	MLH1 probe 00893-L00481		Exon 8	
265	MSH2 probe 00912-L00582			Exon 8
274	MLH1 probe 00894-L00482		Exon 9	
283	Reference probe 00438-L00003	17q21		
292	MSH2 probe 00913-L00583			Exon 9
301	MLH1 probe 00895-L00483		Exon 10	
310	MSH2 probe 11288-L12007			Exon 10
319	MLH1 probe 00896-L00484		Exon 11	
329	MSH2 probe 00915-L00503			Exon 11
337	MLH1 probe 00897-L00485		Exon 12	
346	MSH2 probe 00916-L00504			Exon 12
355	MLH1 probe 00898-L00486		Exon 13	
364	MSH2 probe 01013-L00575			exon 13
375	Reference probe 00681-L11147	4q25		
382	MLH1 probe 00899-L00586		exon 14	
391	MSH2 probe 00918-L00506			Exon 14
401	MLH1 probe 00900-L00488		Exon 15	
409	MSH2 probe 00919-L00585			Exon 15
418	MLH1 probe 01009-L00576		Exon 16	
427	MSH2 probe 01053-L14623			Exon 16
436	MLH1 probe 01030-L00602		Exon 17	
445	MLH1 probe 01031-L00603		Exon 18	
454	MLH1 probe 12094-L12994		Exon 19	
463	Reference probe 00979-L00568	10p14		
472	EPCAM probe 13147-L14404		Región 5' antes del comienzo del gen MSH2	
481	EPCAM probe 13130-L03603		Región 5' antes del comienzo del gen MSH2	
490	Reference probe 04274-L03639	13q12		

Tabla 12: sondas utilizadas en el kit SALSA MLPA P072 MSH6, para la detección de grandes reordenamientos en el gen MSH6. La tabla indica la sonda, el exón y el gen donde se localiza, así como la localización cromosómica de las sondas control.

Tamaño (nt)	Sonda SALSA MLPA	Posición cromosómica				
		Referencia	MLH1	MSH2	MSH6	MUTYH
130	Reference probe 04158-L03513	11q13				
136	EPCAM probe 04249-L03604			Región 5' antes del comienzo del gen MSH2		
142	MSH6 probe 04243-L03598				Exon 7	
148	MLH1 probe 01685-L01265		Exon 1			
154	Reference probe 10051-L10475	8q22				
160	MSH6 probe 04244-L03599				Exon 8	
167	MLH1 probe 01686-L01266		Exon 1			
174	MSH6 probe 04245-L04021				Exon 9	
184	Reference probe 05072-L04472	21q22				
190	EPCAM probe 11983-L12811			Región 5' antes del comienzo del gen MSH2		
196	MSH6 probe 04246-L03601				Exon 10	
202	MSH2 probe 12006-L12810			Región 5' antes del comienzo del gen MSH2		
208	MSH6 probe 06230-L05733				Exon 1	
222	MSH6 probe 12009-L13477				Exon 3	
229	Reference probe 08023-L07804	11q24				
238	MSH6 probe 12010-L12846				Exon 5	
247	MSH6 probe 12011-L12847				Exon 4	
256	Reference probe 06480-L06006	1p13				
265	MSH6 probe 12012-L12848				Exon 2	
274	MSH6 probe 12005-L00909				Exon 5	
281	Reference probe 01633-L01171	11q23				
292	MUTYH probe 04677-L03353					Exon 18
301	MSH6 probe 06229-L05732				Exon 1	
310	EPCAM probe 13131-L03603			Región 5' antes del comienzo del gen MSH2		
319	MLH1 probe 07186-L13476		Exon 1			
328	MSH6 probe 01250-L00798				Exon 1	
337	KCNK12 probe 08663-L13228			Región 3' después del final del gen MSH2		
346	MSH6 probe 04247-L03602				Exon 2	
355	MSH2 probe 02735-L02162			Exon 1		
364	MSH6 probe 01252-L00902				Exon 4	
373	Reference probe 09563-L10017	20p13				
382	MSH6 probe 01253-L00801				Exon 6	
402	MUTYH probe 04323-L01368					Exon 4
412	EPCAM probe 13146-L14626			Región 5' antes del comienzo del gen MSH2		
418	MUTYH probe 03964-L03351					Exon 5a
427	EPCAM probe 13215-L14404			Región 5' antes del comienzo del gen MSH2		
436	Reference probe 09614-L09909	20p12				

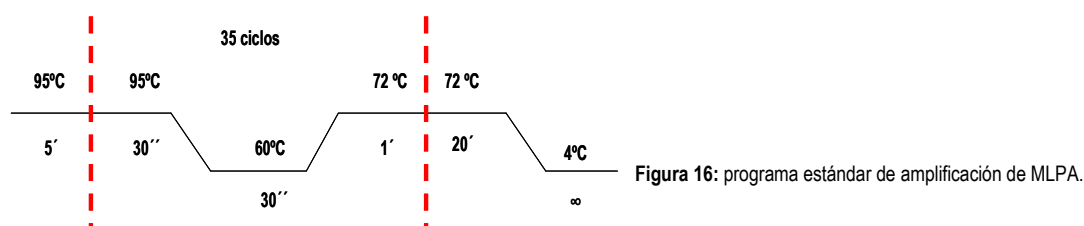
Reacción de PCR: La mezcla de PCR contenía Taq Polimerasa SALSA, tampón de dilución de la enzima, agua destilada y estéril y cebadores universales marcados con fluorocromo:

Sentido: 5' - *GGGTTCCCTAAGGGTTGGA - 3'

Anti-sentido: 5' - GTGCCAGCAAGATCCAATCTAGA - 3'

*Marcado con fluorocromo (FAM).

El programa de PCR utilizado fue el siguiente (figura 16):



Para comprobar la correcta amplificación, se cargaron 6 µl de producto de PCR en un gel de agarosa al 2% teñido con bromuro de etidio y observados bajo luz ultravioleta.

Para visualizar los resultados se preparó cada muestra del siguiente modo: 1 µl del producto de PCR, 0,3 µl de marcador de tamaño marcado con ROX (GeneScan TM-500, Applied Biosystems) y 9 µl de formamida desionizada. El análisis se realizó en el secuenciador ABI Prism 3700 (PE Applied Biosystems) y a través del programa informático Peak Scanner v1.0.

El análisis se realizó mediante la comparación entre un electroferograma procedente de un control y el electroferograma problema. La pérdida de la secuencia de un exón supone la reducción en un 35-55% del área relativa del pico de amplificación. La ganancia de 2 a 3 copias de la secuencia da lugar a un incremento del 30-55% del área relativa del pico de amplificación. La desviación estándar de cada pico debe ser menor del 10% para todas las sondas. Todos los análisis se realizaron por triplicado.

Los resultados se analizaron mediante la comparación del valor del área de los picos generados en el electroferograma por los productos de amplificación de cada muestra y una media de los valores del área de esos mismos picos en muestras control que no presentan cáncer colorrectal.

- **PCR a tiempo real cuantitativa relativa:**

En aquellos pacientes, analizados en primer lugar mediante MLPA, en los cuales identificamos amplificaciones o deleciones en alguno de los genes reparadores; recurrimos a la

confirmación de estas alteraciones mediante el análisis de número de copia del gen mediante PCR a tiempo real cuantitativa comparativa. La tecnología de la PCR en tiempo real permite detectar y evaluar la progresión de la reacción de amplificación en cada uno de sus ciclos. Se fundamenta en la adaptación de un termociclador y un fluorímetro, que permite detectar la fluorescencia emitida por sondas o componentes intercalantes unidos al DNA. La monitorización de esta fluorescencia permite analizar la cinética de la PCR en cada ciclo, determinando el número de copias de fragmento de interés al inicio de la reacción (figura 17). Este valor se puede cuantificar con gran precisión a través del ciclo umbral (Ct). El Ct es el número de ciclos necesarios para que se produzca un aumento de fluorescencia (con respecto a la señal base) suficiente para ser detectado por el lector y es inversamente proporcional a la cantidad inicial de moléculas en la muestra.

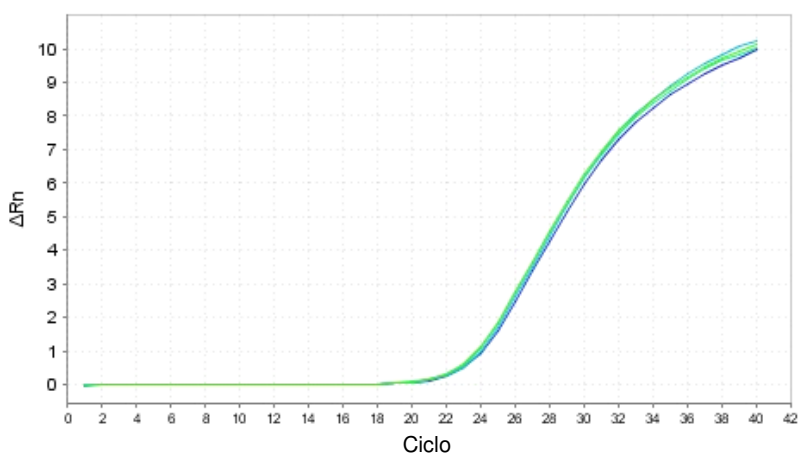


Figura 17: curva de amplificación. Se puede observar el aumento exponencial de fluorescencia a medida que aumenta el número de ciclos.

Por tanto, el uso de la PCR a tiempo real cuantitativa comparativa, permite cuantificar la cantidad de DNA de un determinado fragmento comparando esta amplificación con otra realizada simultáneamente y siempre que ambas pertenezcan a la misma muestra. Para analizar estos cambios relativos se elige como patrón un fragmento de referencia que no varía su carga genética en la patología a estudiar. Comparando el Ct del fragmento de referencia con el fragmento problema de una muestra y el control, se pueden confirmar cambios relativos en la carga genética como pérdida de heterocigosidad, grandes deleciones o amplificaciones. De este modo, las cantidades relativas de los dos productos seguirán siendo las mismas si la cantidad de DNA de partida de ambas muestras es igual; o por el contrario, si hubiera una discrepancia en estos valores iniciales, la cantidad de fragmentos amplificados mostraría diferente cinética y por tanto diferente fluorescencia emitida.

Por estas razones, esta técnica requiere la amplificación de varios fragmentos: el correspondiente a la secuencia de interés, el fragmento de referencia. Además es necesario usar un control.

En estos ensayos utilizamos SYBR Green como agente fluoróforo ya que se une con afinidad al surco menor del DNA bicatenario aumentando su fluorescencia 1000 veces; los oligonucleótidos empleados son los de la tabla. Cada amplificación se realizó por triplicado para minimizar la variabilidad de concentración en las muestras.

La curva de fusión del producto es otro parámetro que permite la caracterización del producto amplificado de acuerdo con su temperatura de fusión (T_m). Esta curva se obtiene tras el fin de la amplificación y se produce como consecuencia del decaimiento de la señal fluorescente debido a la fusión dependiente de la temperatura del producto. La derivada de la curva de fusión nos revela un pico máximo que corresponde con esta temperatura (figura 18).

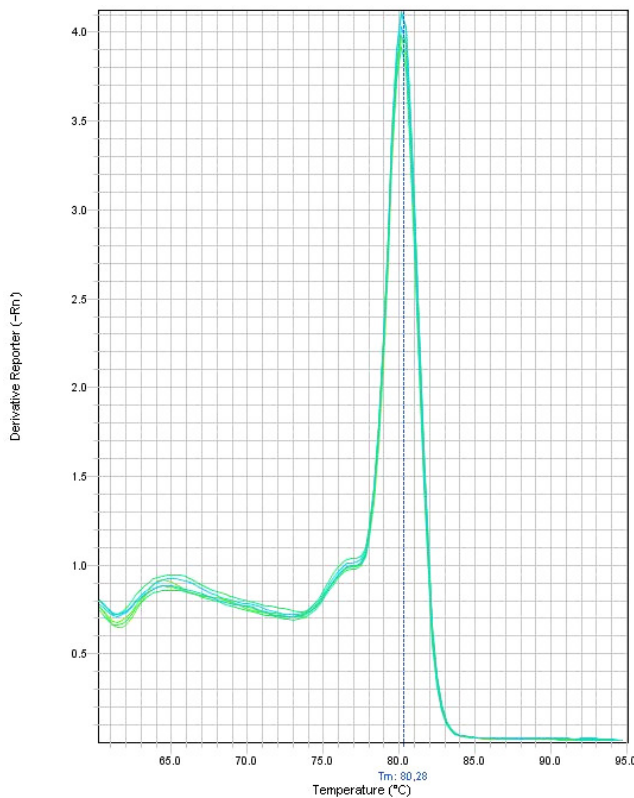


Figura 18: curva de “melting”. En esta gráfica se puede observar la disminución de la fluorescencia (-Rn) en función del aumento de la temperatura.

Para nuestro análisis mediante PCR a tiempo real cuantitativa comparativa Se utilizaron los siguientes controles:

- Control interno que permita comprobar la correcta amplificación: consiste en la amplificación de zonas del genoma que son ajenas a los genes de estudio. En este estudio hemos utilizado dos controles internos: El exón 7 del gen p53 y el exón 1 del gen LEMD3 (tabla 13).
- Muestra de DNA control de un individuo sano.

Tabla 13: cebadores utilizados para amplificar los controles internos de la reacción: p53 exón 7 y LEMD3 exón 1E.

Control interno	Oligonucleótidos 5'-3'
p53 Exón 7	F: aaggcgcaactggcctcatctt R: gaggtggatggtagtag
LEMD3 Exón 1E	F: gcggctgccgggagctagacag R: ggtaggtatgattggagcccg

Los cebadores para el análisis de los exones de interés aparecen especificados en la tabla 14:

Tabla 14: secuencia y temperatura de anillamiento de los cebadores utilizados para la amplificación de los exones 4, 6, y 17 del MLH1.

Amplicón	Oligonucleótidos (5'-3')	Temperatura anillamiento (°C)
MLH1 Exón 17	F: gcactggagaaatgggattg R: cctccagcacacatgcatg	58
MLH1 Exón 4	F: ctttggtaggtgacagtgg R: caggattactctgagacctaggcaa	60
MLH1 Exón 6	F: ggacatctgggtttatttcaag R: tttcaatgtatgacactagaacaca	60

La amplificación se llevó a cabo en un volumen de 20µl cuyos componentes son:

- 10µl de FastStart SyBR Green® Master Rox de Roche.
- 0.5µl de cada cebador.
- 9µl de agua libre de nucleasas.
- 1µl de DNA a una concentración de 15 ng/µl.

Como control de la reacción se mezclaron todos los componentes en un tubo sin DNA.

Las condiciones de PCR utilizadas para amplificar el exón 17 de MLH1 fueron las siguientes (figura 19):

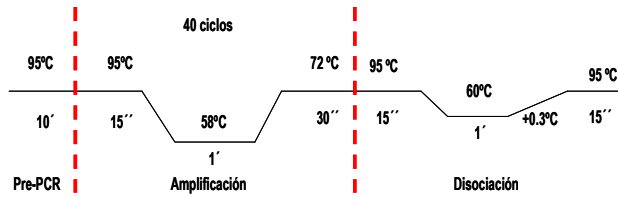


Figura 19 programa de amplificación para el exón 17 de MLH1.

Para la amplificación de los exones 4 y 6 del gen MLH1 utilizamos el siguiente programa (figura 20):

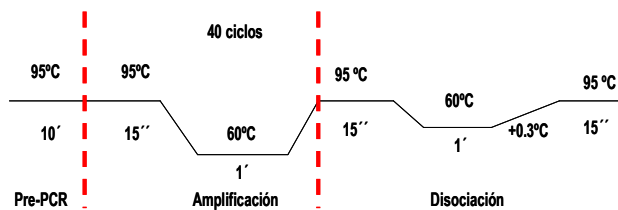


Figura 20: programa de amplificación para los exones 4 y 6.

▪ **PCR de transcripción inversa (RT-PCR):**

Para la síntesis de cDNA se utilizó el kit *ImProm-II™ Reverse Transcription System* (Promega) siguiendo las especificaciones del fabricante. Utilizamos oligo dT, oligonucleótido que se anilla en la cola poliA de los RNAs mensajeros. Para comprobar la correcta síntesis de cDNA, amplificamos por PCR un fragmento del gen de la enzima GAPDH. Este gen se expresa constitutivamente en todos los tejidos (tabla 15).

Tabla 15: secuencia de oligonucleótidos y condiciones de termociclador utilizados para la amplificación del gen de la enzima GAPDH.

	Oligonucleótidos (5'-3')	Condiciones de PCR
GAPDH-F	CCA CCC ATG GCA AAT TCC ATG GCA	55 °C anillamiento 30 ciclos 30" extensión
GAPDH-R	TCT AGA CGG CAG GTC AGG TCC ACC	

5. Análisis de inmunohistoquímica:

El análisis de inmunohistoquímica se realizó en el departamento de Anatomía Patológica del Hospital Universitario de Salamanca.

Las finas láminas de tejido de tumor parafinado se desparafina y se prepara para la tinción con el anticuerpo específico para cada proteína. En la tabla 16 se describen los anticuerpos utilizados para la tinción de las proteínas MLH1, MSH2 y MSH6.

Tabla 16: resumen de los anticuerpos utilizados durante la inmunohistoquímica, la dilución utilizada y la casa comercial.

Anticuerpo	Casa comercial	Dilución
MLH1	BD (Becton Dickinson)	1:20
MSH2	Biocare Medical	Prediluido
MSH6	Biocare Medical	Prediluido

6. Análisis de la inestabilidad de microsatélites (MSI)

El análisis de los microsatélites se realizó sobre 12 muestras de DNA extraído a partir de sangre periférica y DNA de tumor parafinado. Los marcadores incluidos en nuestros análisis son los recomendados por el panel de Bethesda (tabla 17).

Tabla 17: marcadores incluidos en el estudio de MSI. Se incluye la localización en el cromosoma y en el gen correspondiente, número de repeticiones que presenta, el fluorocromo, la secuencia de ambos oligonucleótidos y el tamaño aproximado del producto de PCR. * Oligonucleótido marcado con fluorocromo.

Marcador	Locus	Secuencia	Nº Repeticiones	Oligonucleótidos sentido y antisentido (orientación 5'-3')	Tamaño del amplicón (pb)	Fluorocromo
BAT-25	4q12	Intrón oncogen c-kit	1bp	*TCG-CCT-CCA-AGA-ATG-TAA-GT TCT-GGA-TTT-TAA-CTA-TGG-CTC	115-125pb	TET
BAT-26	2p21-22	Intrón 5 del gen MSH2	1bp	*TGA-CTA-CTT-TTG-ACT-TCA-GCC AAC-CAT-TCA-ACA-TTT-TTA-ACC	110-121pb	HEX
D2S123	2p16	Locus del gen MSH2	2pb	* AAA-CAG-GAT-GCC-TGC-CTT-TA GGA-CTT-TCC-ACC-TAT-GGG-AC	197-227pb	FAM
D5S346	5q21-22	Locus del gen APC	2pb	*ACT-CAC-TCT-AGT-GAT-AAA-TCG-GG AGC-AGA-TAA-GAC-AAG-TAT-TAC-TAG	140-160pb	FAM
D17S250	17q11.2	Locus del cromosoma 17	2pb	*GGA-AGA-ATC-AAA-TAG-ACA-AT GCT-GGC-CAT-ATA-TAT-ATT-TAA-ACC	110-125pb	FAM

La amplificación se realizó:

- 0.08 µl de Gold Taq DNA polimerasa® (Applied Biosystems) 5U/µl,
- 1.0µl de tampón adecuado (10X PCR Buffer II),
- 0.8µl de solución de MgCl₂ 2.5mM, 1.0µl de dNTPs (desoxinucleótidos trifosfato) 25uM,
- 0.3µl y 1.0µl de DNA obtenido por los métodos anteriormente descrito (concentración entre 0.1 y 0.µl/ml).

El volumen final de reacción fue, en todos los casos, de 10 µl que se ajustó con ddH₂O libre de nucleasas. Las reacciones de amplificación se llevaron a cabo en un termociclador de Applied Biosystems, cuyas condiciones generales se muestran en la figura 21.

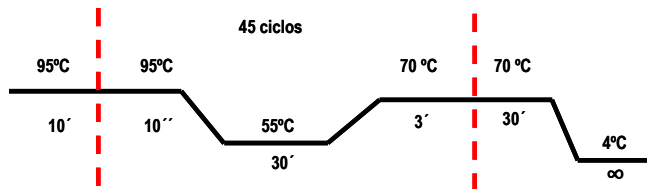


Figura 21: programa utilizado para la amplificación de los marcadores de inestabilidad de microsatélites.

El análisis de los fragmentos amplificados se llevó a cabo resuspendiendo en 1 µL, de producto amplificado 12 µl de formamida desionizada y 0.2 µl de marcador de tamaños ROX (GeneScan™ 350). Esta mezcla se analizó en un secuenciador ABI PRism 3700 (PE Applied Biosystems) y los resultados se visualizaron con el software Peak Scanner v1.0 comparando los electroferogramas procedentes de las muestras de DNA de sangre periférica con los procedentes de tejido tumoral para ver las diferencias de tamaño de secuencia o de pico (figura 22).

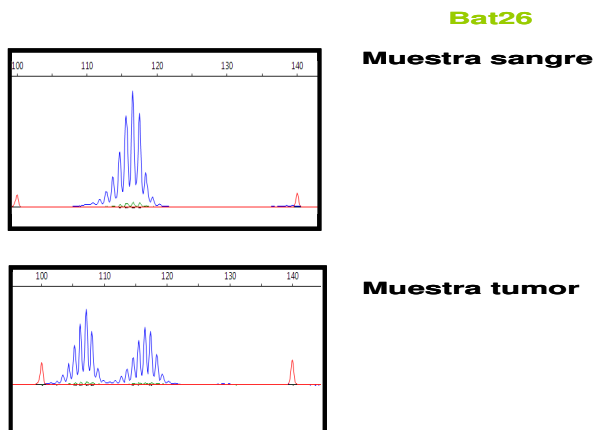


Figura 22: electroferograma de las secuencias de microsatélites. El color azul se corresponde con el marcador BAT26. El color rojo corresponde al marcador ROX. El electroferograma superior muestra el patrón de la secuencia del marcador en DNA y el inferior muestra la secuencia en tumor.

7. Discriminación alélica mediante sondas TaqMan

En la PCR con sondas TaqMan o sondas de hidrólisis, los procesos de amplificación y detección se producen de manera simultánea en el mismo vial cerrado, sin necesidad de ninguna acción posterior. Además, mediante detección por fluorescencia se puede medir durante la amplificación la cantidad de DNA sintetizado en cada momento, ya que la emisión de fluorescencia producida en la reacción es proporcional a la cantidad de DNA formado. Esto permite conocer y registrar en todo momento la cinética de la reacción de amplificación. Los termocicladores para llevar a cabo la PCR con sondas TaqMan incorporan un lector de fluorescencia y están diseñados para poder medir, en cualquier momento, la fluorescencia emitida en cada uno de los viales donde se realice la amplificación.

Para la discriminación alélica hemos empleado sondas específicas marcadas con fluorocromos. Estas sondas de hibridación específica son oligonucleótidos marcados con un fluorocromo donador en el extremo 5' que emite fluorescencia al ser excitado y un aceptor en el extremo 3' que absorbe la fluorescencia liberada por el donador. Para que esto ocurra, las moléculas donadora y aceptora deben estar espacialmente próximas. Además, el espectro de emisión de la primera se ha de solapar con el espectro de absorción de la segunda. En la tabla 18 se describen los fluorocromos más empleados y sus espectros de excitación y emisión. En todos nuestros ensayos de discriminación alélica mediante PCR con sondas TaqMan los fluorocromos empleados fueron VIC y FAM.

Tabla 18: Principales moléculas fluorescentes empleadas como marcadores en la PCR con sondas TaqMan. Las sondas TaqMan que utilizamos en este trabajo están marcadas con FAM y VIC.

Fluorocromo	Máx λ_{abs} (nm)	Máx λ_{em} (nm)
Cascade blue (varios)	374-403	422-430
YOYO-1	491	509
Bodipy	503	512
Fluoresceína (FITC)	494	520
SYBR Green I	497	520
TOTO-1	513	532
FAM	495	535
Luciferina	430	540
VIC	538	554
JOE	525	555
HEX	530	560
Cy3	552	565
POPO-3	534	570
Rodamina	540	570
NED	553	575
TAMRA	560	580
Naranja de acridina	460, 502	526, 650
Cy5	643	667
Quantum Red/Red 670	480, 565	670
Bromuro de etidio	526	605
ROX	580	605
Red 613	480, 565	613
Rojo Texas	596	615
Homodimero de etidio	534	616
Yoduro de propidio	536	617
IRD700	685	705
Cy7	743	767
IRD 800	795	849

Mientras la sonda está intacta, la fluorescencia emitida por el donador es absorbida por el aceptor. Sin embargo, durante la amplificación del DNA diana, la sonda se hibrida con su cadena complementaria. Al desplazarse a lo largo de la cadena, en su acción de síntesis, la DNA polimerasa de *Thermus aquaticus*, que tiene actividad 5' exonucleasa, hidroliza el extremo libre 5' de la sonda, produciéndose la liberación del fluorocromo donador. Como donador y aceptor están, ahora, espacialmente alejados, la fluorescencia emitida por el primero es captada por el lector (figura 23).

El incremento de DNA en cada ciclo se corresponde con un aumento de hibridación de las sondas, lo que conlleva un aumento en la misma proporción de fluorescencia emitida. El empleo de estas sondas garantiza la especificidad de la detección y permite identificar polimorfismos o mutaciones puntuales.

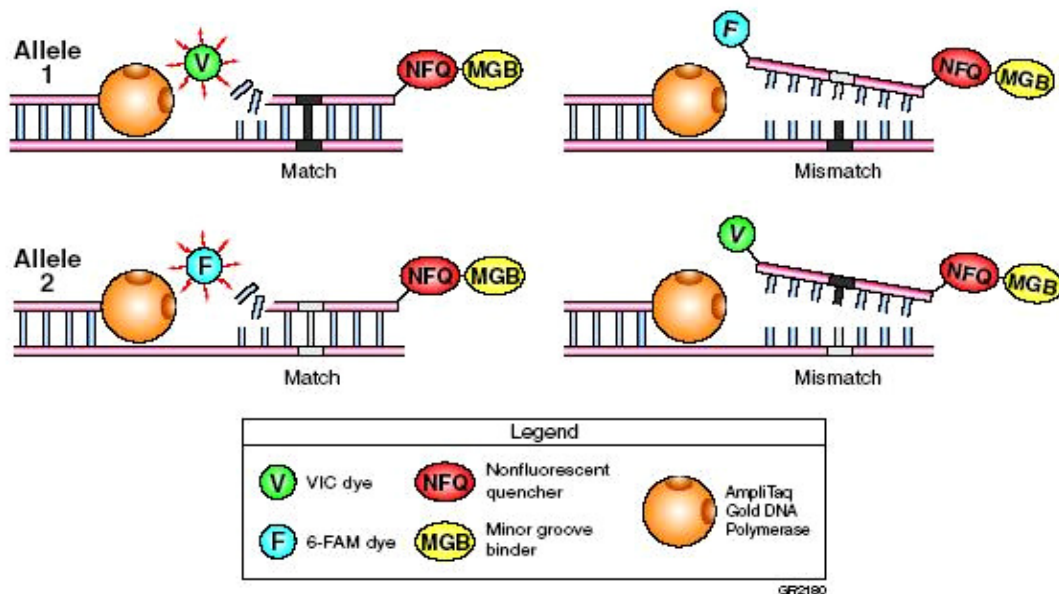


Figura 23. Mecanismo de la PCR con sondas TaqMan.

Esta técnica nos permitió analizar rápida y eficazmente la distribución del polimorfismo c.-93A>G localizada en el promotor del gen MLH1. Utilizamos para ello la sonda C_7535141_1 de Applied Biosystem. Nuestro estudio se realizó en el termociclador StepOnePlus de Applied Biosystems que dispone de varios canales de lectura y permite detectar la emisión de distintos fluorocromos a la vez. De esa manera, se pueden usar varias sondas marcadas con distintos fluorocromos, para identificar los diferentes alelos descritos en cada uno de los genes estudiados.

8. Análisis bioinformáticos

8.1 Análisis de conservación entre especies

- **CLUSTAL W** (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>)

Clustal es una herramienta bioinformática que se utiliza para realizar alineamientos múltiples de secuencias. Existen dos variantes de este programa: Clustal W y Clustal X. En el nuestro trabajo hemos utilizado CLUSTAL W porque esta versión tiene una interfaz gráfica que nos permite trabajar más fácilmente. Este programa acepta un amplio rango de formatos de entrada. Incluyendo NBRF/PIR, FASTA, EMBL/Swissprot. Clustal, GCC/MSF, GCG9 RSF y GDE. El formato de salida puede ser alguno de los siguientes: Clustal, NBRF/PIR, GCG/MSF, PHYLIP, GDE, NEXUS.

Hay tres etapas importantes:

- 1) Hacer un alineamiento por pares.
- 2) Crear un árbol filogenético (o usar un árbol definido por el usuario).
- 3) Usar el árbol filogenético para llevar a cabo el alineamiento.

Los alineamientos de parejas son computados con un método todos contra todos y las similitudes son guardadas en una matriz. Ésta es luego convertida en una matriz de distancias, donde la longitud de la distancia refleja la distancia evolutiva entre cada par de secuencias. Los usuarios pueden alinear las secuencias usando las configuraciones por defecto, pero ocasionalmente es más útil personalizar los parámetros. Los principales parámetros son la penalización por abrir huecos ("gaps") y por extenderlos.

8.2 Análisis posible patogenicidad de las mutaciones

El estudio del comportamiento teórico de las variantes sin clasificar (variantes de significado desconocido) se llevó a cabo mediante el empleo de tres programas informáticos (SIFT, PolyPhen y SNP3D).

- **PolyPhen (Polymorphism Phenotyping)** (<http://www.bork.emblheidelberg.de/PolyPhen>)

Este programa predice el posible impacto de una sustitución aminoacídica en la estructura y función de la proteína humana, empleando consideraciones físicas y comparativas (Stiziel y col., 2003; Nakken y col., 2007).

La secuencia de la proteína se introduce en el programa en formato FASTA. Se lleva a cabo un alineamiento de la misma con las secuencias pertenecientes a las proteínas homólogas en las especies que se encuentran descritas en BLAST. El resultado del alineamiento múltiple es empleado por el software PSIC (Position-Specific Independent Counts) para calcular la matriz del perfil. Los elementos de la matriz son ratios (valores de perfil) logarítmicos que indican la probabilidad de que en una posición determinada se encuentre un aminoácido, frente a que ese aminoácido se localice en cualquier otra posición (frecuencia de fondo). El programa PolyPhen analiza los valores absolutos de la diferencia entre los valores de perfil de las dos variantes alélicas en la posición polimórfica. Un valor muy alto de esta diferencia podría indicar que la sustitución estudiada aparece poco o nada en la familia de proteínas. Polyphen también muestra el número de secuencias alineadas en la posición de búsqueda. Este número indica la credibilidad de los cálculos del valor del perfil.

Los resultados del estudio pueden determinar que el cambio es:

- Probablemente patogénico: existe una gran probabilidad de que afecte a la estructura o a la función de la proteína.
- Posiblemente patogénico: se cree que podría afectar a la función o estructura de la proteína.
- Benigno: es muy probable que carezca de efecto fenotípico.
- Desconocido: en algunos casos, la carencia de datos no permite que el programa pueda hacer una predicción.

- **SNP3D** (<http://www.snp3d.org/>)

Este programa basa su predicción en la hipótesis de que muchos SNPs patogénicos afectan a la función de la proteína disminuyendo la estabilidad proteica. Emplea el software SVM (Support Vector Machina) para encontrar el patrón de diferenciación entre los SNPs asociados a la enfermedad y a los neutrales. La tasa de falsos positivos y negativos es de 15% y 26%, respectivamente (Yue y col., 2006a).

También emplea un segundo modelo que se basa en el análisis de la homología de secuencia en familias de proteínas relacionadas con proteínas humanas. El nivel de conservación de una posición SNP y la probabilidad de observar una variación en una familia hace posible distinguir entre mutaciones patogénicas y neutrales, con una tasa de falsos positivos y negativos del 10% y 20%, respectivamente (Yue y col., 2006)

El resultado positivo para el valor SVM indica que se trata de una variante clasificada como neutral, mientras que el valor negativo la identifica como un cambio patogénico. Cuanto mayor es el valor, más probable es la clasificación. La precisión es significativamente mayor para los valores mayores de

0.5 o menores de -0.5. Para las variantes que afectan más a la función de la proteína que a la estabilidad, el modelo de estabilidad es positivo y el modelo del perfil debe ser negativo. La precisión es significativamente mayor cuando los dos métodos coinciden (Yue y col, 2006b).

- **SIFT (Sorting intoleran from tolerant)** (<http://blocks.fhrc.org/sift/SIFT.html>)

El programa SIFT es una herramienta que clasifica las sustituciones de aminoácidos en una proteína dada y predice si estos cambios provocarán un efecto fenotípico en la proteína.

SIFT se basa en la premisa de que los aminoácidos importantes de una proteína están conservados en la evolución, por lo que los cambios en los mismos deben afectar a la funcionalidad de la proteína. Con una secuencia proteica dada, SIFT escoge proteínas relacionadas y obtiene un alineamiento múltiple de éstas con la proteína a analizar, y basándose en los aminoácidos presentes en cada posición del alineamiento realiza una predicción de las sustituciones que afectarán a la proteína. Las sustituciones en una posición conservada en el alineamiento serán consideradas “no tolerables” para la mayoría de los cambios, mientras que las posiciones que no están conservadas en el alineamiento tolerarán mejor los cambios de aminoácido (Henikoff, 2001).

Se introduce la secuencia a analizar en formato FASTA y el programa primero busca secuencias similares a la introducida por nosotros, después escoge las secuencias estrechamente relacionadas que puedan tener una función similar y luego obtiene el alineamiento de las secuencias escogidas. Por último, calcula las probabilidades normalizadas para todas las posibles sustituciones del alineamiento. Las sustituciones con una puntuación menos de 0.05 serán clasificadas como deletéreas y las puntuaciones mayores o iguales a 0.05, “tolerables” o neutrales.

La precisión en las predicciones de SIFT es del 60-80%, obteniéndose un 30% aproximadamente de falsos positivos. No obstante, la precisión de las predicciones depende del alineamiento que se obtenga en cada caso.

8.3 Análisis de sitios ESE

- **ESE Finder** (<http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home>)

Para saber si las variantes en la secuencia afectan a los sitios consenso de unión de las proteínas SR (*Serine/Arginine-rich proteins*) y se altera la unión de éstas a los sitios ESE (*Exon Splicing Enhancer* o sitios exónicos de potenciación del procesamiento) utilizamos el programa de predicción ESE Finder (Cartegni y col., 2003; Smith y col., 2006). Para ello, se introduce la secuencia normal y la

secuencia con el cambio. El programa valora la presencia de sitios ESE tanto en la secuencia normal como en la mutada que puede provocar el cambio de nucleótido. A nivel general, consideran significativos los cambios exónicos que hacen que desaparezca algún sitio ESE.

El programa estudia el cambio de los sitios consenso de unión de estas proteínas SF2/ASF, SF2/ASF (IgM-BRCA1), SC35, SRp40 y SRp55. Considera los siguientes valores umbral para cada uno de ellas (tabla 19):

Tabla 19: valores umbral para las diferentes proteínas de procesamiento.

Proteína	Valor umbral
SF/ASF	1.956
SF/ASF (IgM-BRCA1)	1.867
SC35	2.383
SRp40	2.67
SRp55	2.676

Como valor umbral se considera al valor mínimo para que una secuencia determinada se considere significativa para la unión de alguna de las proteínas de procesamiento.

El programa también permite obtener una representación gráfica de de los sitios de unión de las proteínas en la secuencia estudiada. La secuencia analizada se representa en el eje X del gráfico. Sólo se representan los valores que superan el umbral. La altura de las barras corresponde a los valores del motivo, la anchura corresponde a la longitud del motivo de unión (6, 7, 8 nucleótidos) y cada una de las proteínas tiene un código de color (figura 24).

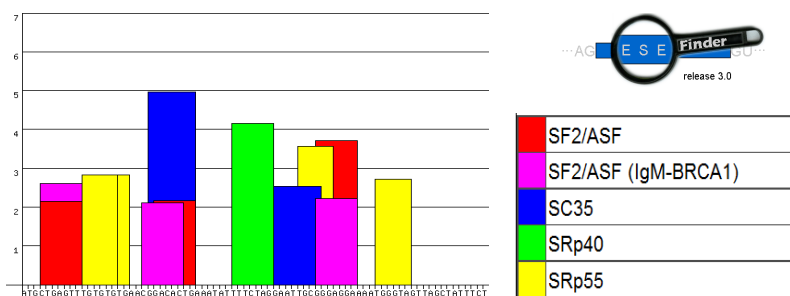


Figura 24: gráfico que muestra los diferentes sitios de reconocimiento de las proteínas de procesamiento en una secuencia de DNA dada.

- RESCUE-ESE (<http://genes.mit.edu/burgelab/rescue-ese/>)

Este programa es un método híbrido computacional y experimental que ha sido desarrollado para identificar secuencia con actividad ESE. En esta aproximación las secuencias

hexaméricas específicas son identificadas como candidato sitios ESE en base a que tienen una mayor frecuencia de presencia en los exones que en intrones y, además, una frecuencia mayor en exones con zonas no canónicas de procesamiento. En este programa se analizan 238 secuencia hexaméricas que son sitios ESE potenciales.

Al igual que en el programa ESE FINDER, RESCUE-ESE permite valorar el efecto de la alteración mediante la comparación de la secuencia normal con la mutada.

Después de las valoraciones, el programa emite una imagen en la que localiza las secuencias ESE dentro de la secuencia de análisis (figura 25).

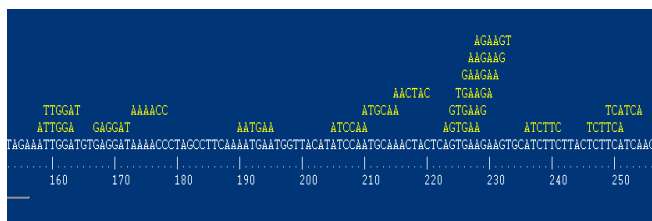


Figura 25: secuencia de un exón del gen MLH1 y los sitios ESE que aparecen en esta secuencia.

Identificación de microRNAs de unión a los genes MLH1, MSH2 y MSH6

En este estudio se utilizó un programa de unión de microRNAs a las secuencias UTR de los genes reparadores y una base de datos que nos permitiera localizar los microRNAs. Para cada gen, se procedió de la siguiente manera:

- 1) Se identificó de la secuencia de UTR 3' según el buscador de genoma UCSC (<http://genome.ucsc.edu>).
- 2) Se identificaron los microRNAs capaces de unirse a estas secuencias del extremo 3'UTR a través de programas: TargetScan, miRBase.
- 3) Los polimorfismos localizados tanto en la secuencia de los microRNAs, como en la secuencia del extremo 3'UTR de los genes se determinaron usando la base de datos de SNPs del sistema NCBI (<http://www.ncbi.nlm.nih.gov/snp>).

- **miRBase: La base de datos de los microRNA.** (<http://www.mirbase.org/>)

El programa miRBase proporciona los siguientes servicios:

- *miRBase database* es una base de datos de las secuencias de los miRNA publicados. Cada entrada de esta base de datos muestra la estructura del transcrito del miRNA teórica con la información de la localización y la secuencia del miRNA maduro. Ambas secuencias, la del transcrito y la del miRNA maduro, se pueden descargar para su uso.
- *miRBase Registry* proporciona localizadores de los genes de los miRNA para nuevos miRNA.
- *miRBase targets database* o *microCosm* es una herramienta para la predicción de dianas de los miRNA.

Esta base de datos está mantenida y ubicada en los servidores de la “Faculty of Life Sciences en la Universidad de Manchester y apoyada económicamente por Biotechnology and Biological Sciences Research Council (BBSRC).

- **TargetScan:** (<http://www.targetscan.org/>)

Este programa predice las dianas biológicas de los microRNAs a través de la búsqueda de la presencia de sitios 8-mer y 7-mer conservados que puedan aparearse con las regiones de unión de los miRNA. Opcionalmente, aquellas regiones no conservadas también pueden ser predichas. También diferencias aquellos sitios que no perfectamente complementarios y que son compensados con la secuencia conservada en el extremo 3' del miRNA. En mamíferos, las predicciones se basan en cálculos en base a los valores de eficiencia de unión de cada sitio. TargetScanHuman considera las uniones con los UTRs en humanos, aunque también puede identificar dianas en zonas codificantes del gen.

Resultados

1. Resultados

Hemos estudiado los genes reparadores MLH1, MSH2 y MSH6 en probandos incluidos en el Programa de Consejo Genético de la Junta de Castilla y León entre los años 2003 y 2010. En la tabla 20 se resume la distribución de probandos según los estudios realizados. Los criterios establecidos por el Programa de Consejo Genético fueron las directrices de inclusión de los probandos en el estudio (descritos en el apartado de Material y Métodos).

Tabla 20: esquema de la distribución de los estudios realizados sobre los probandos incluidos en este trabajo.

Estudio molecular	Nº Pacientes estudiados
Genes MLH1, MSH2 y MSH6 (estudio completo)	29
Sólo genes MLH1 y MSH2	153
Sólo gen MSH6	6
Total	188

Los criterios de inclusión del Programa de Consejo Genético son una mezcla de los criterios establecidos para el diagnóstico clínico del Síndrome de Lynch. Los pacientes utilizados en este estudio pueden ser clasificados en aquellos que, al menos, cumplen un criterio de la Junta de Castilla y León y aquellos que cumplen los criterios de Ámsterdam o de Bethesda. Además, incluidos en el estudio hay probandos sanos cuyas familias cumplen criterios de inclusión pero no hemos tenido acceso a familiares enfermos y otros que no cumplen criterios pero que fueron incluidos en el estudio por motivos compasivos (tabla 21 y 22).

Tabla 21: resumen de número de probandos que cumplen los criterios de Ámsterdam (AMST), Bethesda (BETH), probandos sanos (SANOS) y aquellos que no cumplen criterios (NO).

Criterio	Nº Probandos	Nº de probandos con mutación patogénica
AMST	30	6
BETH	67	6
NO	52	1
SANOS	39	2

Tabla 22: resumen de número de probandos que cumplen los criterios establecidos por la Junta de Castilla y León (C1: existencia de 3 familiares con cáncer colorrectal, C2: existencia de 3 familiares con tumores asociados* al CCHNP, C3: un cáncer colorrectal en un paciente menor de 50 años, C4: dos casos de cáncer colorrectal sincrónico o metacrónico, o de otros tumores asociados al cáncer colorrectal, C5: un caso de cáncer colorrectal y un familiar de primer grado con tumor asociado al CCHNP, uno de ellos antes de los 50 años, C6: un cáncer de endometrio antes de los 45 años, C7: un cáncer de colon derecho, con formas histológicas poco diferenciadas (solido/cribiforme), antes de los 45 años, C8: pacientes con adenomas colorrectales antes de los 45 años), probandos sanos (SANOS) y aquellos que no cumplen criterios (NO). ^AHay probandos que cumplen varios criterios de la Junta de Castilla y León.

Criterios	Nº Probandos ^A	Nº de probandos con mutación patológica
C1	22	1
C2	23	2
C3	33	1
C4	9	2
C5	2	0
C6	1	0
C7	1	0
C8	2	0
C1+C2	0	0
C1+C3	7	2
C1+C4	3	0
C1+C5	1	1
C1+C6	1	0
C1+C7	1	0
C1+C8	0	0
C2+C3	10	1
C2+C4	1	0
C2+C5	0	0
C2+C6	1	0
C2+C7	1	0
C2+C8	1	0
C3+C4	0	0
C3+C5	0	0
C3+C6	1	0
C3+C7	0	0
C3+C8	1	0
C4+C5	0	0
C4+C6	1	1
C4+C7	0	0
C4+C8	0	0
C5+C6	0	0
C5+C7	0	0
C5+C8	0	0
C6+C7	0	0
C6+C8	0	0
C7+C8	0	0
C2+C3+C5	1	1
C2+C3+C7	1	0
NO	24	1
SANOS	39	2

Incluimos en el estudio del gen MSH6 aquellos probandos que presentaban tumores de endometrio y familias con gran carga de tumores asociados a Síndrome de Lynch sin mutaciones en MLH1 o MSH2. De este modo, realizamos la búsqueda de mutaciones en MSH6 en 35 probandos.

2. Análisis de mutaciones en los genes MLH1, MSH2 y MSH6:

Para el rastreo de mutaciones en los genes MLH1, MSH2 y MSH6 hemos analizado un total de 188 probandos de familias no relacionadas. En ellos hemos identificado 61 mutaciones (Tabla 23): el 36.07% localizadas en el gen MLH1, el 39.34% en el gen MSH2 y el 24.59% en el gen MSH6.

Dentro de ellas trece (21.31%) son mutaciones consideradas patogénicas. En conjunto, estas mutaciones germinales patogénicas identificadas engloban tanto mutaciones puntuales como grandes reordenamientos genómicos. Seis de estas mutaciones patogénicas se han identificado en el gen MLH1, seis en el gen MSH2 y una en el gen MSH6. Se consideran mutaciones patogénicas aquellas que provocan un cambio en la fase de lectura o un codón de parada dando lugar, durante la traducción, a una proteína más corta y carente de dominios importantes para su función; mutaciones que provocan un cambio de aminoácido que afecta a la estructura y función de la proteína y aquellas mutaciones intrónicas que producen errores durante la maduración del RNA mensajero.

Son polimorfismos 36 (59.02%) de las 61 alteraciones identificadas: entre ellos se encuentran variantes intrónicas, cambios silenciosos y cambios de aminoácidos. Además, hemos identificado doce (19.67%) variantes de significado desconocido y se corresponden con cambios de aminoácido o cambios en el extremo 3'UTR del gen.

Tabla 23: frecuencias y porcentajes del total de mutaciones identificadas en los tres genes estudiados en este trabajo. Este resumen incluye los grandes reordenamientos genómicos encontrados en este estudio.

Tipo de mutación		Mutaciones encontradas en MLH1	Mutaciones encontradas en MSH2	Mutaciones encontradas en MSH6	Porcentaje sobre el total de mutaciones*
Patogénicas	Cambios de marco de lectura	0	2	1	3 (4.92%)
	Mutaciones sin sentido	3	2	0	5 (8.20%)
	Mutaciones de cambio de aminoácido	1	0	0	1 (1.64%)
	Mutaciones que afectan al procesamiento del mRNA	0	0	0	0
	Grandes reordenamientos	2	2	0	4 (6.56%)
Total mutaciones patogénicas		6	6	1	13 (21.31%)
Polimorfismos	Mutaciones silenciosas	2	3	4	9 (14.75%)
	Cambios de aminoácido	1	0	1	2 (3.28%)
	Variantes intrónicas	7	9	9	25 (40.98%)
Total polimorfismos		10	12	14	36 (59.02%)
Variantes de significado desconocido	Cambios de aminoácido	4	6	0	10 (16.39%)
	Extremo 3'UTR	2	0	0	2 (3.28%)
Total de variantes de significado desconocido		6	6	0	12 (19.67%)
Total de todas las mutaciones		22 (36.07%)	24 (39.34%)	15 (24.59%)	61 (100%)

2. 1. Estudio del gen MLH1:

2.1.1. Mutaciones puntuales patogénicas en el gen MLH1:

El análisis de las regiones codificantes y las regiones intrónicas flanqueantes del gen MLH1 nos ha permitido caracterizar cuatro mutaciones patogénicas (tabla 24). Todas ellas aparecen descritas en diferentes bases de datos y en la bibliografía revisada.

Tres de estas alteraciones (**c.298C>T**, **c.676C>T**, **c.1459C>T**) son mutaciones sin sentido y la mutación **c.350C>T**, es una mutación de cambio de sentido (figura 26).

Tabla 24: mutaciones patogénicas encontradas en el gen MLH1. ^ALa secuencia de referencia utilizada es NM_000429.3 y normas de nomenclatura utilizadas de HSGV.

MLH1	Alteración^A	Exón	Codon	Consecuencia	Referencia bibliográfica	Probandos
p.Arg0Stop	c.298C>T	3	100	Codon parada prematuro	Wang y col., 1999; Peel y col., 2000; Samowitz y col 2001; Ellison y col., 2004; Hampel y col., 2005	1
p.Thr117Met	c.350C>T	4	117	Cambio de treonina por metionina	Shchervakova y col., 1999; Liu y col., 1997; Jäger y col., 2001; Plotz y col., 2006	1
p.Arg226Stop	c.676C>T	8	226	Codon parada prematuro	Molseim y col., 1996; Shin y col., 1998; Lamberti y col., 1999; Otway y col., 2000; Nomura y col., 2000; Caldés y col., 2004	4
p.Arg487Stop	c.1459C>T	13	487	Codon parada prematuro	Leung y col., 1998; Fidalgo y col., 2000; Peel y col., 2000; Caldés y col., 2000.; Lage y col., 2004; Lee y col., 2005	1

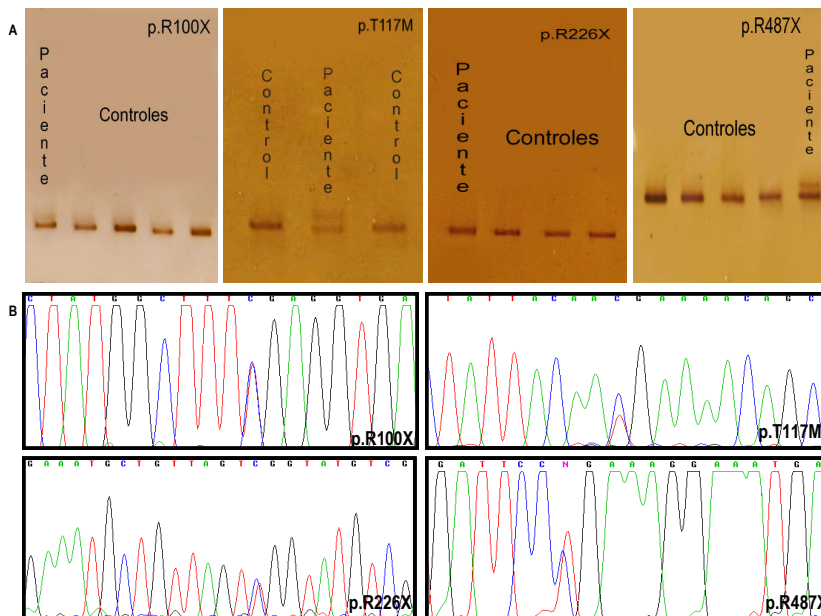


Figura 26: A) Patrones de migración alterados identificados en los pacientes portadores de mutaciones patogénicas en el gen MLH1 mediante CGSE. B) Cromatograma de la secuenciación automática donde se confirmó la alteración.

Las características clínicas de las familias en las cuales se ha identificado aparecen descritas en la tabla 25. Todas salvo una cumplen, al menos, un criterio establecido por el programa de Consejo Genético de la Junta de Castilla y León y los criterios internacionales de Amsterdam o Bethesda. La familia 4 muestra no presenta criterios debido a la escasez de datos de los miembros; sin embargo fue incluida debido motivos compasivos.

Tabla 25: características de las familias con mutaciones patogénicas en MLH1. *AP: antecedentes personales. **AF: antecedentes familiares. Ca: cáncer, pat: familiar por parte de padre; mat: familiar por parte de madre. adc: adenocarcinoma. C1: existencia de 3 familiares con cáncer colorrectal, C2: existencia de 3 familiares con tumores asociados* al CCHNP, C3: un cáncer colorrectal en un paciente menor de 50 años, C4: dos casos de cáncer colorrectal sincrónico o metacrónico, o de otros tumores asociados al cáncer colorrectal, C5: un caso de cáncer colorrectal y un familiar de primer grado con tumor asociado al CCHNP, uno de ellos antes de los 50 años, C6: un cáncer de endometrio antes de los 45 años, C7: un cáncer de colon derecho, con formas histológicas poco diferenciadas (solido/cribiforme), antes de los 45 años, C8: pacientes con adenomas colorrectales antes de los 45 años. AMST: Criterios de Amsterdam. BETH: Bethesda. ^Excluye al probando

MLH1	ID FAMILIA	AP*	AF**	C. Junta de Castilla y León	Criterio Internacional	Familiares portadores / Familiares estudiados ^A
p.Arg100Stop	FAM 4	Ca. colon 51 años	Desconocidos	NO CRITERIO	NO CRITERIO	1/2
p.Thr117Met	FAM 44	Adc. colon derecho 50 años	Varios casos ca.colon en familia matema (2 generaciones)	C1 y C3	AMST	2/5
p.Arg226Stop	FAM 1	Ca.colon 29 años	Padre: ca. vesícula 51 años; tío pat: ca. pulmón 62 años; abuela pat: ca. riñon y ca. matriz 93 años. Hermano: ca. intestino 40 años; hermano: ca. duodeno 40 años;	C3	BETH	2/5
p.Arg226Stop	FAM 92	Ca.colon 83 años	hermano: ca. páncreas 82 años; hermano: ca.colon 64 años; sobrina: ca. piel 30 años; sobrino: ca. recto 40 años.	C2 y C3	AMST	2/7
p.Arg226Stop	FAM178	Ca.colon 55 años	Padre: ca. colon 70 años, 4 familiares ca. colon y estómago. 2 generaciones.	C1 y C5	AMST	0
p.Arg487Stop	FAM 10	Adc.colon 59 años, adc.ciego 62 años, adc. endometrio, adc. duodeno.	Madre: ca. colon 56 años; abuela: ca. ginecológico?	C4	BETH	0/1

2.1.2. Polimorfismos:

Hemos encontrado un total de diez polimorfismos en el gen MLH1 (tabla 26). En concreto, dos mutaciones silenciosas, siete variantes intrónicas y una mutación que produce un cambio de sentido. De ellos, cuatro han sido descritos por primera vez en este trabajo.

Tabla 26: relación de polimorfismos identificados en el gen MLH1.

POLIMORFISMOS MLH1	Exón	Codon	Consecuencia	Referencia bibliográfica	Probandos
c.375A>G	Exón 4	c.375A>G	p.Ala125Ala	Mangold y col., 2005	2
IVS5-51T>C	Intrón 5	c.454-51T>C	/	Nystrom-Lahti y col., 1996	7
IVS5+79A>G	Intrón 5	c.453+79A>G	/	Hudler y col., 2004; Lee y col., 2005; Fredriksson y col., 2006	2
IVS5+92A>T	Intrón 5	c.453+92A>T	/	No descrito	2
IVS6+169C>A	Intrón 6	c.545+169C>A	/	No descrito	3
p.I219V	Exón 8	c.655A>G	p.Ile219Val	Buerstedde y col. 1995; Takashashi y col., 2007	13
p.E234E	Exón 9	c.702G>A	p.Glu234Glu	Curia y col. 1999; Grabowski y col. 2005	1
IVS13+14G>A	Intrón 13	c.1558+14G>A	/	Ghimenti y col. 1999; Furihata y col., 2001; Gómez y col., 2005; Auclair y col., 2006; Fredriksson y col. 2006	8
IVS15+30T>G	Intrón 15	c.1731+30 T>G	/	No descrito	1
IVS15+28A>C	Intrón 15	c.1731+28 A>C	/	No descrito	1

La mutación de cambio de sentido, c.655A>G, provoca en la proteína el cambio del aminoácido isoleucina por una valina. En diferentes trabajos se ha analizado su efecto sobre la proteína concluyendo que no produce ninguna pérdida o deficiencia en la función de MLH1.

Cuatro de los cambios no aparecen descritos en las bases de datos ni en la bibliografía consultada. Todos ellos son variantes intrónicas que, por localizarse en zonas alejadas de las secuencias críticas para la correcta maduración, las hemos clasificado como polimorfismos.

Como ya describimos en material y métodos con los programas ESE FINDER y RESCUE_ESE realizamos un análisis bioinformático de predicción teórica del efecto de los polimorfismos exónicos sobre el procesamiento del RNA mensajero (tabla 27).

Tabla 27: predicciones realizadas mediante los programas ESE FINDER y RESCUE ESE. Sombreados en azul claro y marcados con el signo positivo aparecen las zonas en las que el polimorfismo analizado puede afectar a la maduración del RNA mensajero. = el programa no predice un cambio en ese sitio.

MLH1	ESE FINDER								RESCUE ESE
	5SS_U2	BRANCH SITE	3SS_U2	SF2/ASF	SF2/ASF (IgM-BRCA1)	SC35	SRp40	SRp55	
c.375A>G	=	=	=	=	=	=	=	=	=
c.655A>G	=	=	=	=	=	=	=	=	+
c.702G>A	=	=	=	=	=	=	=	=	+

El programa RESCUE ESE considera que la presencia de las variantes c.655A>G y c.702G>A produce la pérdida de un sitio ESE (figura 27). Estos resultados son orientativos y se requiere una comprobación mediante RT-PCR.

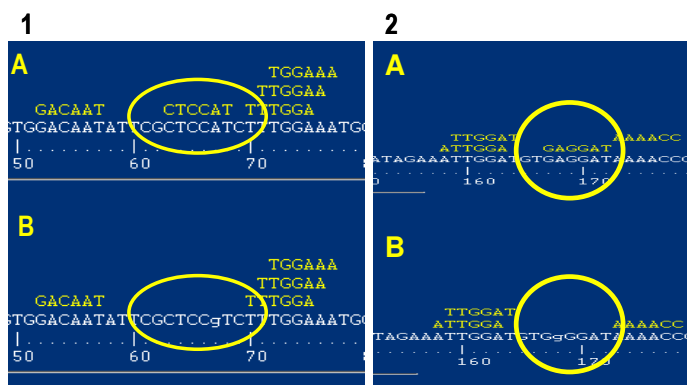


Figura 27: 1) resultado de análisis bioinformático RESCUE ESE del efecto del polimorfismo exónico **c.655A>G** sobre los sitios ESE: 1.A) Sitios ESE presentes en la secuencia normal. 1.B) Sitios ESE presentes en la secuencia mutada. 2) resultado de análisis bioinformático RESCUE ESE del efecto del polimorfismo exónico **c.702G>A** sobre los sitios ESE: 2.A) Sitios ESE presentes en la secuencia normal. 2.B) Sitios ESE presentes en la secuencia mutada.

2.1.3. Variantes de significado desconocido:

En el gen MLH1 hemos encontrado seis variantes sin clasificar (tabla 28). Todas, excepto una, aparecen descritas en la bibliografía y bases de datos consultadas. Entre ellas se encuentran cuatro variantes localizadas en las regiones codificantes del gen y dos localizadas en la región 3'UTR. Una de las dos variantes de significado desconocido localizadas en el extremo 3'UTR (c.*2303C>T) no ha sido descrita hasta el momento.

Tabla 28: variantes de significado desconocido identificadas en el gen MLH1. ^AIn silico: los programas utilizados aparecen descritos en el apartado Material y métodos. ^BCo-ocurrencia: presencia de dos mutaciones, una de ellas patogénica, en un mismo individuo.

MLH1	Exón	Cambio nucleótido	Consecuencia	Dominio	Co-ocurrencia ^B	Estudio In silico ^A			Presencia en población control	Referencia bibliográfica
						Estudio conservación entre especies	Programas de predicción de patogénicidad	Programas de predicción de sitios ESE		
p.G22A	1	c.65G>C	p.Gly22Ala	Dominio ATPasa	No	Aparece en todas las secuencias analizadas	No tolerable	Aparece nuevo sitio de unión para proteína SF2/ASF y pérdida de un sitio ESE localizado en el extremo 3' del exón	0/100	Fearnhead y col., 2004; Woods y col., 2005; Barnetson y col., 2008.
p.R69K	2	c.206G>A	p.Arg69Lys	Dominio ATPasa	No	Aparece en todas las secuencias analizadas, excepto en <i>S. cerevisiae</i>	Posiblemente patogénica y tolerable	La mutación añade dos nuevos sitios ESE al ya existente en la secuencia y la creación de un sitio de unión para la proteína SC35	0/100	Klaus y col., 1998; Ellison y col., 2001; Takahashi y col., 2007.
p.K618A	16	c.1852AAdel GCins	p.Lys618Ala	Dominio de dimerización	No	No aparece conservado en <i>Sacharomyces cerevisiae</i> , <i>Solanum lycopersicum</i> , <i>Trypanosoma brucei</i> y <i>Giardia intestinalis</i> .	No tolerable	La mutación provoca la pérdida de tres sitios ESE.	0/100	Caldés y col., 2002; Fearnhead y col., 2004; Takahashi y col., 2007; Barnetson y col., 2008; Blasi y col., 2010; Castillejos y col., 2011.
p.P640L	17	c.1919C>T	p.Pro640Leu	Dominio de dimerización	No	Aparece en todas las secuencias analizadas, excepto en <i>Giardia intestinalis</i> .	Posiblemente patogénica	La mutación crea un sitio de unión de la proteína SC35.	0/100	Tournier y col. 2008
3'UTR c.*32_34del	/	c.*32_34del	Disminución en la expresión del mRNA en leucemias	Extremo 3'UTR	Si	-	-	-	-	Curia y col., 1999; Murata y col., 2002; Mao y col., 2008
3'UTR c.*2303C>T	/	c.*2303C>T	Desconocida	Extremo 3'UTR	No	-	-	-	-	No descrito

Hemos realizado un rastreo en la población general de las mutaciones p.G22A, p.R69K, p.K618A, p.P640L. La identificación se llevó a cabo mediante PCR, CGSE y secuenciación del DNA de 100 individuos sanos de ambos sexos, mayores de 60 años y sin antecedentes personales de cáncer. En ninguno de los controles analizados apareció alguna de las variantes.

Sobre las mismas variantes realizamos un estudio de conservación de aminoácidos mediante alineación múltiple entre secuencias de diferentes especies con el programa informático CLUSTALW. Esta técnica informática permite saber si el aminoácido que se localiza en esa posición dentro de la secuencia proteica aparece conservado evolutivamente. Hemos utilizado las secuencias de diferentes especies del reino animal y vegetal (figura 28).

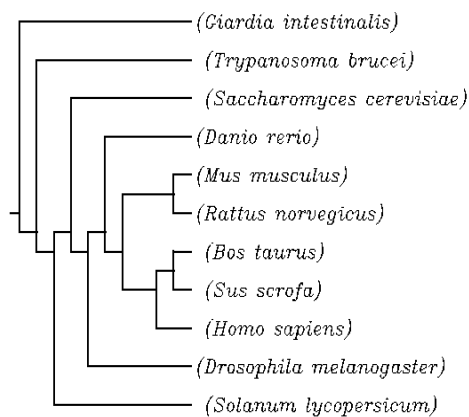


Figura 28: dendograma de la relación entre las especies utilizadas par el estudio de conservación de secuencias. *Homo sapiens* (humano), *Drosophila melanogaster* (mosca de la fruta), *Solanum lycopersicum* (tomate), *Saccharomyces cerevisiae* (levadura), *Giardia intestinalis* (parásito intestinal), *Trypanosoma brucei* (protista), *Bos taurus* (toro), *Mus musculus* (ratón común), *Danio rerio* (pez cebra), *Rattus norvegicus* (rata), *Sus scrofa* (jabali).

En las cuatro variantes analizadas, el aminoácido que presenta la proteína normal en *H. sapiens*, aparece conservado en la mayoría de las secuencias de las especies analizadas (figura 29).





Figura 29: CLUSTALW: comparación de la conservación evolutiva de las variantes de significado desconocido encontradas en el gen MLH1. A) Conservación entre especies del aminoácido glicina en la posición 22. B) Arginina en la posición 69. C) Lisina en la posición 618. D) Prolina en la posición 640.

Con la ayuda de diversos programas de predicción realizamos un estudio teórico global del efecto sobre el procesamiento del RNA mensajero y el efecto sobre la proteína de estas variantes de significado desconocido identificadas en MLH1. Los programas de predicción de alteraciones en la maduración del mRNA utilizados fueron: ESE FINDER Y RESCUE ESE. En la tabla 29 se resumen los resultados obtenidos en el estudio *in silico*.

Tabla 29: predicciones realizadas mediante los programas ESE FINDER y RESCUE ESE. Sombreados en gris y marcados con la cruz aparecen las zonas en las que el polimorfismo analizado puede afectar a la maduración del RNA mensajero. = el programa no predice un cambio en ese sitio.

VSD MLH1	ESE FINDER								RESCUE ESE
	5SS_U2	BRANCH SITE	3SS_U2	SF2/ASF	SF2/ASF (IgM-BRCA1)	SC35	SRp40	SRp55	
G22A	=	+	+	=	+	=	=	=	=
R69K	=	=	=	=	=	+	=	=	+
K618A	=	=	=	=	=	=	=	=	+
P640L	=	=	=	=	=	+	=	=	=

En todas las variantes de significado desconocido los programas encuentran valores significativos. La mutación p.G22A crea un nuevo sitio para la proteína SF2/ASF, además elimina un sitio ESE en el extremo 3' del exón (figura 30).

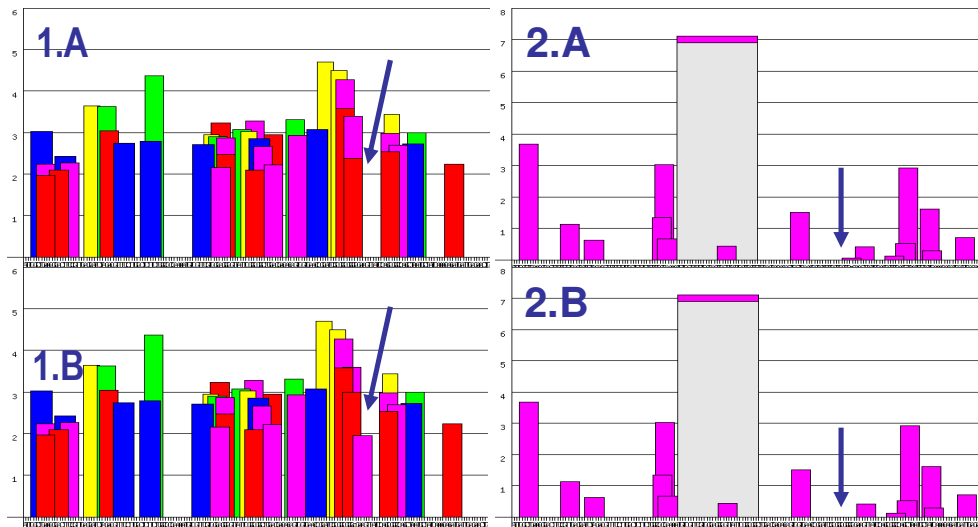


Figura 30: resultado de análisis bioinformático del efecto de la variante de significado desconocido **p.G22A** sobre los sitios ESE. 1: ESE FINDER 1.A) Sitios de unión de las proteínas presentes SR en la secuencia normal. 1.B) Sitios de unión a las proteínas SR presentes en la secuencia mutada. 2: ESE FINDER: 2.A) Sitios canónicos de procesamiento presentes en la secuencia normal. 2.B) Sitios canónicos de procesamiento presentes en la secuencia mutada.

La variante de significado desconocido **p.R69K** crea en la secuencia dos nuevos sitios ESE y un sitio nuevo de unión de la proteína SC35 según las valoraciones de los programas RESCUE ESE y ESE FINDER, respectivamente (figura 31).

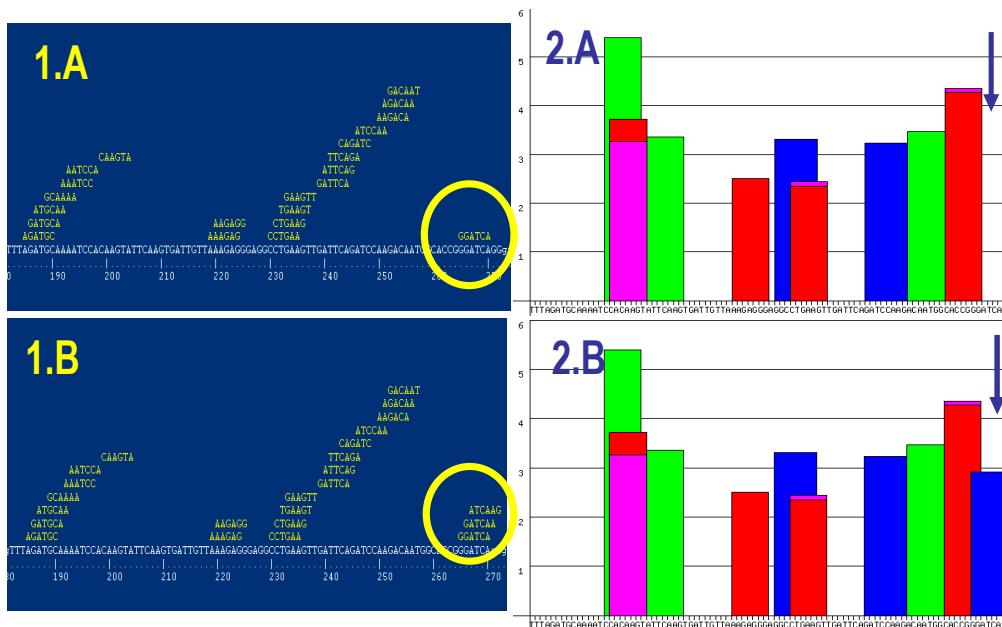


Figura 31: resultado de análisis bioinformático del efecto de la variante de significado desconocido **p.R69K** sobre los sitios ESE. 1: RESCUE ESE 1.A) Sitios ESE presentes en la secuencia normal. 1.B) Sitios ESE presentes en la secuencia mutada. 2: ESE FINDER: 2.A) Sitios de unión de las proteínas SR presentes en la secuencia normal. 2.B) Sitios de unión de las proteínas SR presentes en la secuencia mutada.

El programa RESCUE ESE valora que el cambio p.K618A produce en la secuencia de DNA la pérdida tres sitios ESE. El cambio p.P640L crea un nuevo sitio de unión para la proteína SC35 (figura 32)

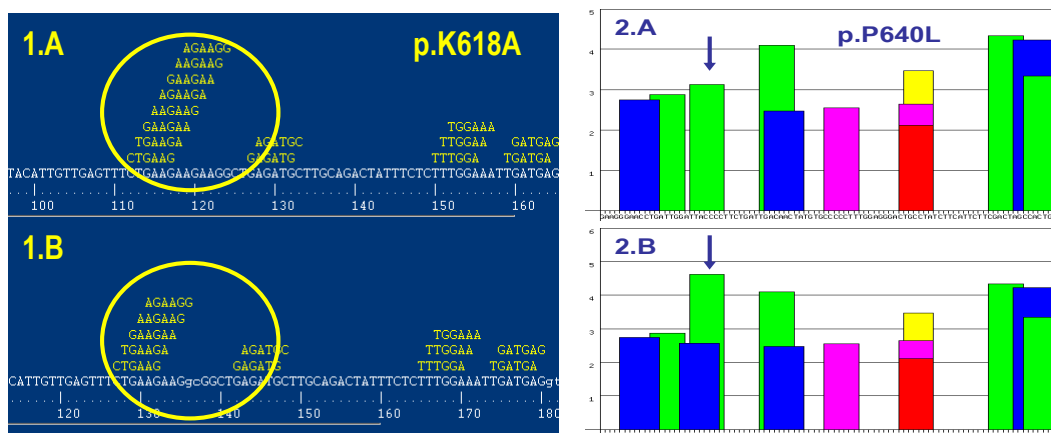


Figura 32: resultado de análisis bioinformático del efecto de las variantes de significado desconocido p.K618A y p.P640L sobre los sitios ESE. 1) Valoración de RESCUE ESE de la mutación p.K618A 1.A) Sitios ESE presentes en la secuencia normal. 1.B) Sitios ESE presentes en la secuencia mutada. 2) Valoración de ESE FINDER de la mutación p.P640L. ESE FINDER: 2.A) Sitios de unión de las proteínas SR presentes en la secuencia normal. 2.B) Sitios de unión de las proteínas SR presentes en la secuencia mutada.

Los programas de predicción del efecto sobre la proteína analizan no sólo la conservación entre especies del aminoácido en una posición, si no que también valoran la estructura del aminoácido y la localización de éste dentro de la proteína. En el estudio hemos utilizado tres programas diferentes: Polyphen, SIFT y SNP3D. En la tabla 30 aparece un resumen de los datos obtenidos.

Tabla 30: patogenicidad teórica obtenida a través de diversos programas informáticos. ^A Los valores calculados por cada uno de los programas. Cada programa utiliza un algoritmo diferente y sitúa un valor umbral diferente.

VSD MLH1	POLYPHEN	Valores ^A	SIFT	Valores ^A	SNP3D	Valores ^A
p.G22A	Benigno	1.477	No tolerable	0.00	No tolerable	-2.59
p.R69K	Posiblemente dañino	1.592	Tolerable	0.42	Tolerable	0.44
p.K618A	Posiblemente dañino	1.989	No tolerable	0.04	No tolerable	-1.10
p.P640L	Probablemente dañino	3.292	No tolerable	0.04	No tolerable	-3.01

Las mutaciones p.K618A y p.P640L aparecen como posiblemente patogénicas en los tres programas informáticos utilizados. La mutación p.G22A es considerada por PolyPhen como benigna mientras que los programas SIFT y SNP3D coinciden en que el cambio no es tolerable por la proteína. La variante de significado desconocido p.R69K es valorada como tolerable por SIFT y SNP3D, mientras que el programa PolyPhen valora el cambio como posiblemente dañino.

2. 2. Estudio del gen MSH2:

2.2.1. Mutaciones patogénicas:

En el gen MSH2 hemos identificado cuatro mutaciones patogénicas (tabla 31). Todas las mutaciones encontradas aparecen descritas en las bases de datos o han sido publicadas por otros grupos.

Tabla 31: mutaciones patogénicas identificadas en el gen MSH2. * Excluye al probando. La secuencia de referencia que hemos utilizado para nombrar las mutaciones es NM_000.251.

MSH2	Alteración	Exón	Codón	Consecuencia	Referencia bibliográfica	Probandos
p.Gln130ValfsX2	c.387_388delTC	3	130	Codón de parada prematuro	Buerstedde y col., 1995; Wagner y col., 2002; Mangold y col., 2005	1
p.Arg389Stop	c.1165C>T	7	389	Codon de parada prematuro	Cunningham y col., 2001; Wagner y col., 2003; Sutter y col., 2004; Pinol y col., 2005; Sanchez-Abajo y col., 2005; Mongical-Artus y col., 2006; Overbeet y col., 2007.	1
p.Arg406Stop	c.1216C>T	7	406	Codón de parada prematuro	Leach y col.; 1993; Liu y col.; 1994; Weber y col.; 1997; Ponz de León y col., 2004; Mangold y col., 2005	1
p.Ile747ArgfsX2	c.2239_2240delTA	14	747	Codón de parada prematuro	Caldés y col. 2000.	1

Estas alteraciones se distribuyen a lo largo de la secuencia codificante del gen y dos consisten en pérdidas de dos nucleótidos que provocan un cambio en la fase de lectura y la aparición de un codón de parada prematuro, c.387_388delTC y c.2239delTA. Las otras dos mutaciones c. 1165C>T y c.1216C>T son un cambio sin sentido que generan un codon de parada prematuro (figura 33).

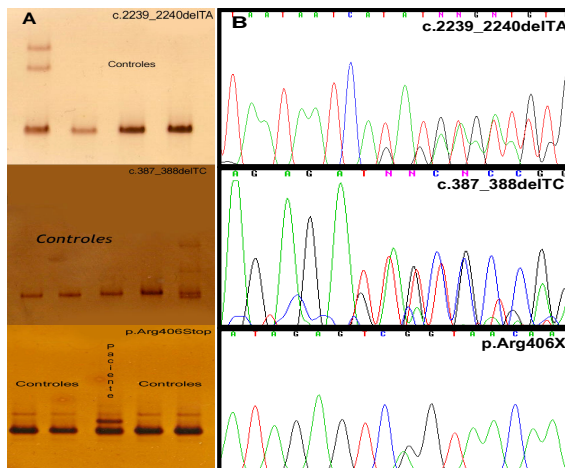


Figura 33: A) Patrones de migración alterados identificados en los pacientes portadores de mutaciones patogénicas en el gen MSH2 mediante CGSE. B) Cromatograma de la secuenciación automática donde se confirmó la alteración.

Las familias en las cuales hemos identificado los cambios cumplen los criterios de Ámsterdam o de Bethesda, excepto en una familia, familia 115, de la cual carecemos de datos y cuyo probando es sano. Esta familia fue incluida en el estudio por motivos compasivos (tabla 32). La familia 116 cumple los criterios de Amsterdam; sin embargo, no existía la posibilidad de obtener muestra de un miembro enfermo.

Tabla 32: Características de las familias portadoras de mutación en MSH2. ^AAP: Antecedentes familiares. ^BAF: Antecedentes familiares. T: tumor, Ca: cáncer. C1: existencia de 3 familiares con cáncer colorrectal, C2: existencia de 3 familiares con tumores asociados* al CCHNP, C3: un cáncer colorrectal en un paciente menor de 50 años, C4: dos casos de cáncer colorrectal sincrónico o metacrónico, o de otros tumores asociados al cáncer colorrectal, C5: un caso de cáncer colorrectal y un familiar de primer grado con tumor asociado al CCHNP, uno de ellos antes de los 50 años, C6: un cáncer de endometrio antes de los 45 años, C7: un cáncer de colon derecho, con formas histológicas poco diferenciadas (solido/cribiforme), antes de los 45 años, C8: pacientes con adenomas colorrectales antes de los 45 años. AMST: criterios de Amsterdam. BETH: criterios de Bethesda.

MSH2	ID FAMILIA	AP ^A	AF ^B	Criterio de la Junta de Castilla y León	Criterio Internacional	Familiares portadores / Familiares estudiados*
p.Gln130ValfsX2	FAM 15	Sano	DESCONOCIDOS	NO CRITERIO	NO CRITERIO	4/4
p.Arg389Stop	FAM 6	Ca. colon 26años.	Hermano: t. cerebral 56 años; padre: ca. colon 68años, tía mat: cáncer?; abuelos mat: ca. Colo? 88 años.	C2, C3 y C5	AMST	1/1
p.Arg406Stop	FAM 116	Sana	(Éxitus) Madre: ca. endometrio 46 años, ca. colon 52 años; tía materna: ca. colon 44 años; abuelo materno: ca. colon?; abuela materno: Ca. mama	NO CRITERIO	NO CRITERIO	0/0
p.Ile747ArgfsX2	FAM 16	Ca. colon 52 años.	Padre: ca.colon 44 años; 3 tíos paternos: ca.colon 35 años y 47 años; tía paterna: ca. útero 55 años.	C1	AMST	3/5

2. 2. 2 Polimorfismos:

En el gen MSH2 hemos identificado 12 polimorfismos que aparecen descritos en la tabla 33. Nueve de los polimorfismos identificados son variantes intrónicas. Su localización dentro del intrón es variable, habiendo variantes alejadas de las zonas consenso implicadas en la maduración del mRNA y otras más cercanas a estas secuencias. Todas aparecen descritas en las bases de datos y la bibliografía consultada, excepto dos que describimos por primera vez en este trabajo.

Tabla 33: relación de polimorfismos identificados en el gen MSH2.

MSH2	Exón	Codon	Consecuencia	Referencia bibliográfica	Probandos
IVS1+9C>G	Intrón 1	c.211+9C>G	/	Palicio y col., 2002; Cederquist y col. 2004; Capellá y col., 2008;	4
p.K113K	Exón 2	c.339G>A	p.Lys113Lys	Tournier y col. 2008.	1
IVS3-39C>A	Intrón 3	c.645-39C>A	/	No descrito	1
IVS6-80A>G	Intrón 6	c.1077-80G>A	/	Bujalkova y col. 2008	2
IVS6-10T>C	Intrón 6	c.1077-10T>C	/	Borresen y col. 1995; Desai y col. 2000; Woods y col. 2005; Rubio del Campo y col. 2008	4
p.A328A	Exón 6	c.984C>T	p.Ala328Ala	Curia y col. 1999; Lee y col. 2005; Yap y col 2008; Tournier y col. 2008.	2
IVS7-12A>C	Intrón 7	c.1277-12A>C	/	No descrito	1
IVS9-9A>T	Intrón 9	c.1511-9A>T	/	Bujalkova y col. 2008	2
IVS10+12G>A	Intrón 10	c.1661+12G>A	/	Wijnen y col. 1994; Wahlberg y col. 1997; Hedge y col. 2005; Apeessos y col. 2005.	2
p.K579K	Exón 11	c.1737A>G	p.Lys579Lys	Sánchez de Abajo y col. 2005; Tournier y col. 2008.	1
IVS11-62G>A	Intrón 11	c.1760-62G>A	/	Tournier y col. 2008	1
IVS12-6T>C	Intrón 12	c.2006-6T>C	/	Rubio del Campo y col. 2008; Capella y col. 2008; Tulupova y col. 2008; Park y col. 2008	4

Las tres mutaciones restantes son consideradas silenciosas puesto que en ningún caso provocan el cambio de aminoácido en la proteína. Sobre ellas realizamos el análisis de sitios ESE a través de los programas ESE FINDER y RESCUE-ESE. Los resultados aparecen descritos en la tabla 34.

Tabla 34: predicciones realizadas mediante los programas ESE FINDER y RESCUE ESE. Sombreados en azul y marcados con el signo positivo aparecen las zonas en las que el polimorfismo analizado puede afectar a la maduración del RNA mensajero. = el programa no predice un cambio en ese sitio.

MSH2	ESE FINDER									RESCUE ESE
	5SS_U2	BRANCH SITE	3SS_U2	SF2/ASF	SF2/ASF (IgM-BRCA1)	SC35	SRp40	SRp55		
c.1737A>G	=	=	=	+	+	=	=	=	+	
c.984C>T	=	=	=	+	+	+	=	=	=	
c.339G>A	=	=	=	+	+	=	=	=	+	

Ambos programas coinciden en la predicción de la alteración de los sitios ESE ocasionada por las variantes c.1737A>G y c.339A>G. Estos dos cambios afectan a los sitios de unión de las proteínas SF2 (SF2/ASF y SF2/ASF (IgM-BRCA1)). El cambio c.1737A>G no supone la pérdida de sitios ESE sino que cambia su secuencia (figura 34). La presencia del cambio c.339A>G añade un nuevo sitio ESE a la secuencia (figura 35). El polimorfismo c.984C>T afecta al sitio de unión de las proteínas SF/ASF y SC35 (figura 36).

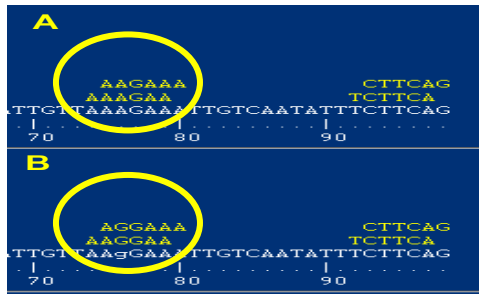


Figura 34: resultado de análisis bioinformático del efecto del polimorfismo exónico **c.1737A>G** sobre los sitios ESE. 1: ESE FINDER: RESCUE ESE: A) Sitios ESE presentes en la secuencia normal. B) Sitios ESE presentes en la secuencia mutada.

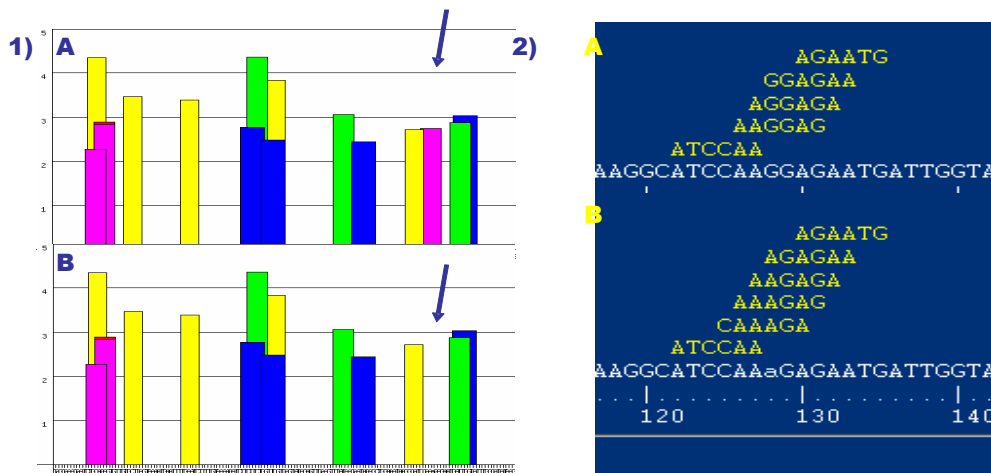


Figura 35: resultado de análisis bioinformático del efecto del polimorfismo exónico **c.339G>A** sobre los sitios ESE. 1: ESE FINDER: 1.A) Sitios de unión de las proteínas SR en la secuencia normal. 1.B) Sitios de unión a las proteínas SR en la secuencia mutada. 2) RESCUE ESE: 2.A) Sitios ESE presentes en la secuencia normal. 2.B) Sitios ESE presentes en la secuencia mutada.

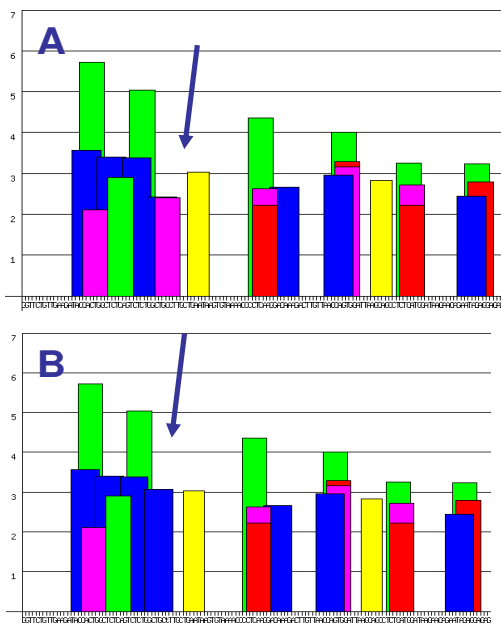


Figura 36: resultado de análisis bioinformático del efecto del polimorfismo exónico **c.984G>A** sobre los sitios ESE. ESE FINDER: A) Sitios de unión de las proteínas SR en la secuencia normal. B) Sitios de unión a las proteínas SR en la secuencia mutada.

2.2.3. Variantes de significado desconocido:

Las variantes de significado desconocido identificadas en el gen MSH2 aparecen descritas en la tabla 35. Hemos encontrado 6 mutaciones de cambio de sentido de las cuales tres aparecen descritas en este trabajo por primera vez. Todas las mutaciones de significado desconocido están distribuidas a lo largo de la secuencia codificante del gen y se sitúan, preferentemente, en dos dominios importantes para la proteína: el dominio de unión al DNA y el dominio de dimerización.

Para la aproximación teórica del significado de estas variantes, hemos seguido el mismo protocolo de análisis que realizamos para las variantes de significado desconocido identificadas en el gen MLH1. En primer lugar, un rastreo de las mutaciones sobre población control en el que no identificamos ninguna de las variantes en ninguno de los controles analizados.

Tabla 35: variantes de significado desconocido identificadas en el gen MSH2. ^A*In silico*: los programas utilizados aparecen descritos en el apartado Material y métodos. ^BCo-ocurrencia: presencia de dos mutaciones, una de ellas patogénica, en un mismo individuo.

MSH2	Exón	Cambio nucleótido	Consecuencia	Dominio	Co-ocurrencia	In silico			Presencia en población control	Referencia bibliográfica
						Conservación entre especies	Programas de predicción de patogenicidad	Programas de predicción de sitios ESE		
p.H46Q	1	c.138C>G	p.His46Glu	Dominio de unión a DNA	No	Aparece en todas las secuencias analizadas excepto en <i>Bos taurus</i>	No tolerable	Altera el sitio de unión de SF/ASF y SRp55	0/100	Barneston y col. 2008, Fearnhead y col., 2004;
p.N127S	3	c.380A>G	p. Asp127Ser	Dominio de unión a DNA	No	Aparece en todas las secuencias analizadas excepto en <i>Schizosaccharomyces pombe</i>	No tolerable	altera el sitio de unión SC35	0/100	Chen-Shtoyerman y col., 2003; Levene y col., 2003.
p.E188Q	3	c.562G>C	p.Glu188Gln	Dominio de unión a DNA	No	Aparece en todas las secuencias analizadas	Posiblemente dañino	Altera el sitio de unión de SRp55	0/100	No descrito
p.K249R	4	c.746A>G	p.Lys249Arg	Dominio de unión a DNA	No	No aparece en <i>Arabidopsis thaliana</i> ni en <i>Schizosaccharomyces pombe</i>	Tolerable	No altera ningún sitio ESE	0/100	No descrito
p.Q377R	7	c.1130A>G	p.Gln377Arg	Dominio de interacción con MSH3/MSH6	No	No aparece en <i>Arabidopsis thaliana</i> ni en <i>Schizosaccharomyces pombe</i>	Tolerable	Altera el sitio de unión de SF/ASF y SRp40	0/100	No descrito
p.P616R	12	c.1847C>G	p.Pro616Arg	Dominio de interacción con MSH3/MSH6	No	Aparece en todas las secuencias analizadas excepto en <i>Schizosaccharomyces pombe</i>	No tolerable	Altera el sitio de unión de SRp55	0/100	No descrito

En segundo lugar, realizamos un estudio de conservación entre especies utilizando el programa CLUSTALW. En todas las variantes, el aminoácido que aparece en la secuencia de la proteína normal se repite en la mayoría de las secuencias aminoacídicas de las especies (Figura 37 y 38).

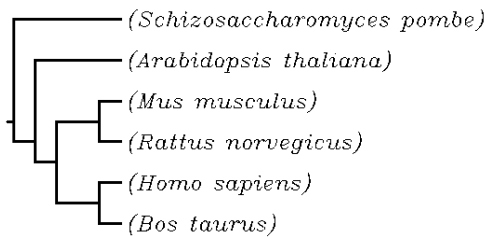


Figura 37: dendograma de la relación entre las especies utilizadas par el estudio de conservación de secuencias. *Homo sapiens*, *Bos taurus*, *Mus musculus*, *Rattus norvegicus*, *Arabidopsis thaliana* y *Saccharomyces pombe*.

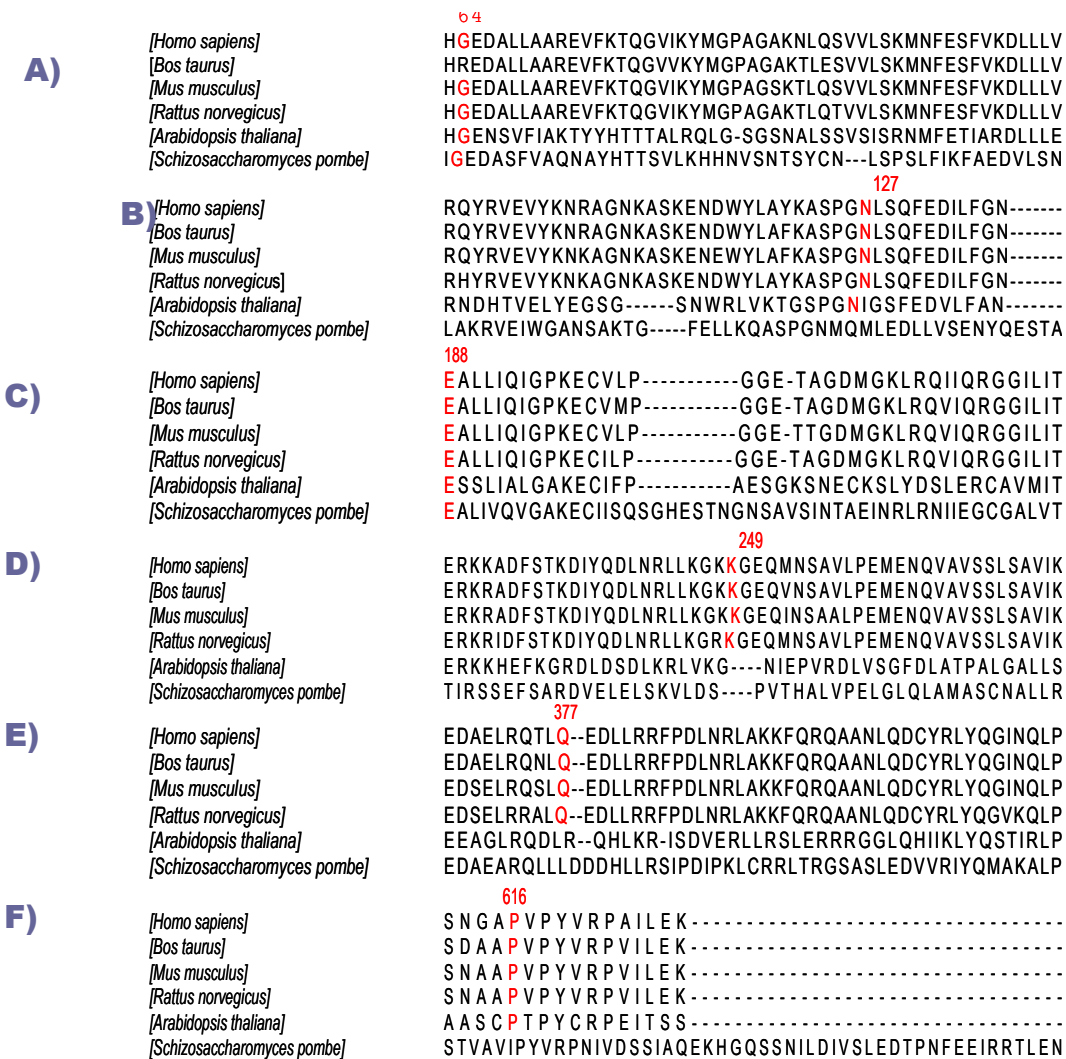


Figura 38: conservación de las variantes de significado desconocido encontradas en el gen MSH2 entre diferentes especies. A) Conservación entre especies del aminoácido glicina en la posición 64, B) Asparragina en la posición 127, C) Glutámico en la posición 188, D) Lisina en la posición 249, E) Ac. Glutámico en la posición 377, F) Prolina en la posición 616.

Por último, el análisis teórico del procesamiento del RNA mensajero mediante ESE FINDER Y RESCUE ESE y el análisis del efecto sobre la proteína mediante PolyPhen, SIFT y SNP3D aparecen resumidos en las tablas 36 y 37:

Tabla 36: predicciones realizadas mediante los programas ESE FINDER y RESCUE ESE. Sombreados en gris y marcados con la cruz aparecen las zonas en las que el polimorfismo analizado puede afectar a la maduración del RNA mensajero. = el programa no predice un cambio en ese sitio.

MSH2	ESE FINDER								RESCUE ESE
	5SS_U2	BRANCH SITE	3SS_U2	SF2/ASF	SF2/ASF (IgM-BRCA1)	SC35	SRp40	SRp55	
H46Q	=	+	+	+	=	=	=	+	=
N127S	=	=	=	=	+	=	=	=	=
E188Q	=	=	+	=	=	=	=	+	+
K249R	=	=	=	=	=	=	=	=	=
Q377R	=	=	=	+	=	=	+	=	+
P616R	=	=	=	=	=	=	=	+	=

Todas las variantes de significado desconocido analizadas con los programas ESE FINDER y RESCUE ESE muestran posibles alteraciones en el mecanismo de maduración del RNA mensajero; solo la mutación p.K249R no parece tener efecto sobre los sitios ESE (figuras 39, 40, 41, 42 y 43).

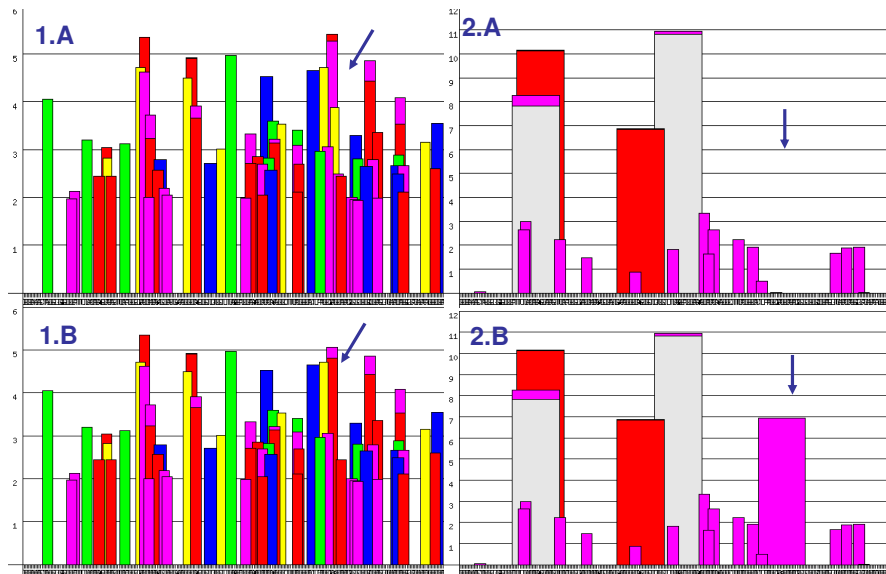


Figura 39: resultado del análisis bioinformático de la variante de significado desconocido p.H46Q sobre los sitios ESE. ESE FINDER: 1.A) sitios de unión de las proteínas SR en la secuencia normal. 1.B) sitios de unión de las proteínas SR en la secuencia mutada. 2.A) Sitios canónicos de procesamiento presentes en la secuencia normal. 2.B) Sitios canónicos de procesamiento presentes en la secuencia mutada.

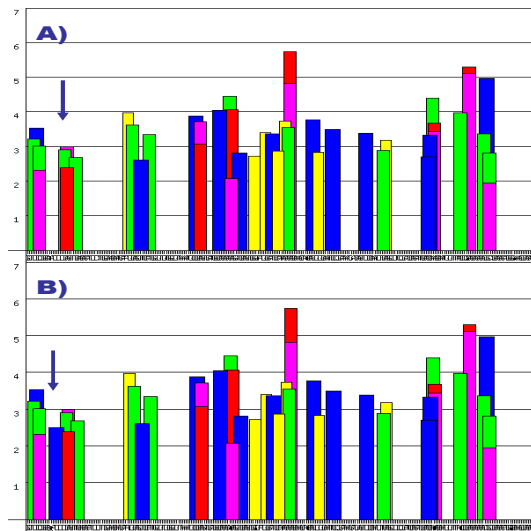


Figura 40: resultado del analisis bioinformático de la variante de significado desconocido p.N127S sobre los sitios ESE. ESE FINDER: A) sitios de unión de las proteínas SR en la secuencia normal. B) sitios de unión de las proteínas SR en la secuencia mutada.

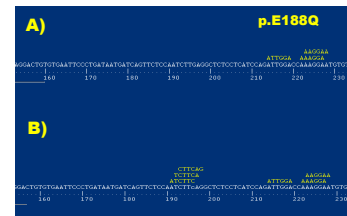
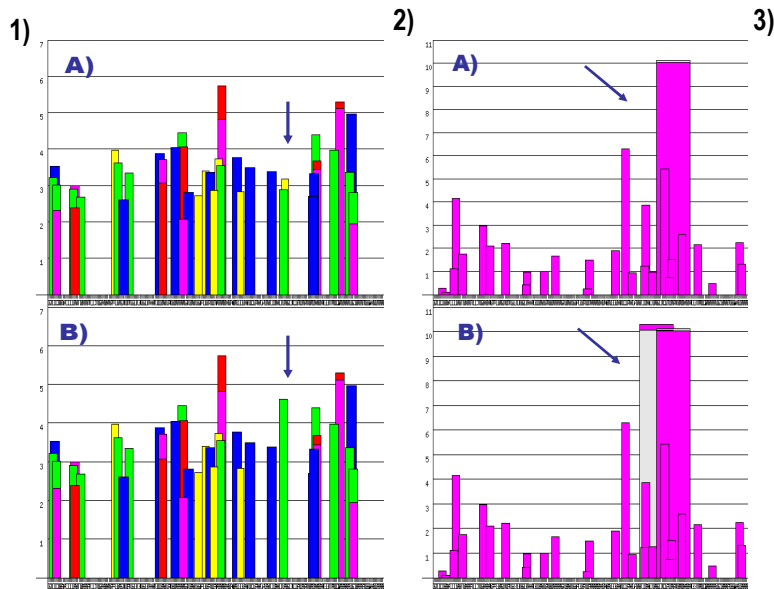


Figura 41: resultado del analisis bioinformático de la variante de significado desconocido p.E188Q sobre los sitios ESE. ESE FINDER: 1.A) sitios de unión de las proteínas SR en la secuencia normal. 1.B) sitios de unión de las proteínas SR en la secuencia mutada. 2.A) Sitios canónicos de procesamiento presentes en la secuencia normal. 2.B) Sitios canónicos de procesamiento presentes en la secuencia mutada. 3) RESCUE ESE: A) sitios ESE presentes en la secuencia normal. B) Sitios ESE presentes en la secuencia mutada.

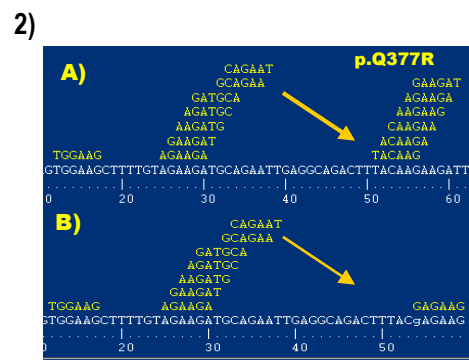
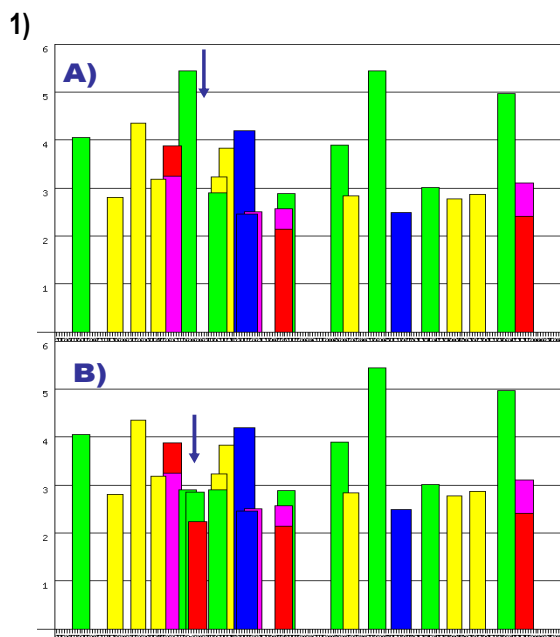


Figura 42: resultado del analisis bioinformático de la variante de significado desconocido p.Q377R sobre los sitios ESE. ESE FINDER: 1.A) sitios de unión de las proteínas SR en la secuencia normal. 1.B) sitios de unión de las proteínas SR en la secuencia mutada. 2.A) Sitios canónicos de procesamiento presentes en la secuencia normal. 2.B) Sitios canónicos de procesamiento presentes en la secuencia mutada.

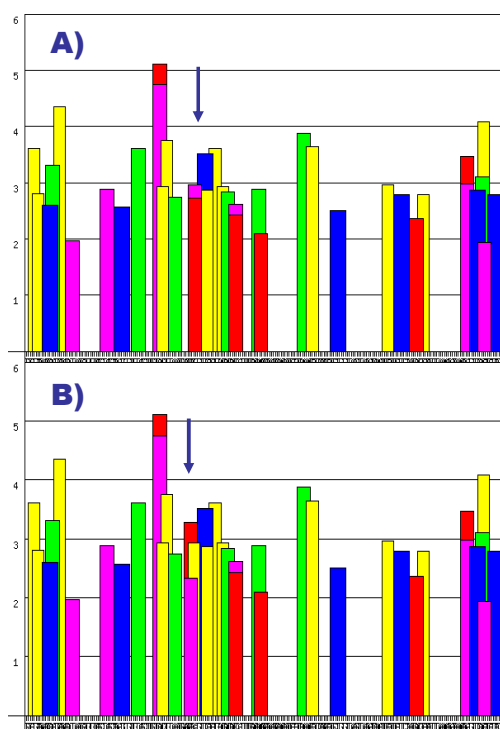


Figura 43: resultado del analisis bioinformático de la variante de significado desconocido p.P616R sobre los sitios ESE. ESE FINDER: 1.A) sitios de unión de las proteínas SR en la secuencia normal. 1.B) sitios de unión de las proteínas SR en la secuencia mutada. 2.A) Sitios canónicos de procesamiento presentes en la secuencia normal. 2.B) Sitios canónicos de procesamiento presentes en la secuencia mutada.

Tabla 37: patogenicidad teórica obtenida a través de diversos programas informáticos. * Los valores calculados por cada uno de los programas. Cada programa utiliza un algoritmo diferente y sitúa un valor umbral diferente.

VSD MSH2	POLYPHEN	Valores*	SIFT	Valores*	SNP3D	Valores*
p.H46Q	Probablemente dañino	3.003	Tolerable	0.33	No tolerable	-1.42
p.N127S	Probablemente dañino	2.303	Tolerable	0.64	No tolerable	-0.05
p.E188Q	Posiblemente dañino	1.893	Tolerable	0.48	Tolerable	0.64
p.K249R	Benigno	0.498	Tolerable	0.46	No tolerable	-0.39
p.Q377R	Benigno	0.233	Tolerable	0.23	Tolerable	2.36
p.P616R	Posiblemente dañino	1.565	Tolerable	0.44	Tolerable	0.98

Las mutaciones p.H46Q, y p.N127S muestran una posible patogenicidad valorada por dos programas de predicción. En el caso de las variantes p.E188Q y p.P616R únicamente el programa PolyPhen las valora como posiblemente dañinas para la proteína. La mutación p.K249R es valorada como no tolerable por la proteína por el programa SNP3D. Los tres programas son unánimes a la hora de valorar como tolerable a la mutación p.Q377R.

2.3. Estudio del gen MSH6:

Hemos analizado el gen MSH6 en 35 probandos que cumplían criterios de Amsterdam o Bethesda sin mutación patogénica en MLH1 y MSH2 y aquellos individuos con cáncer de endometrio o familias con agregaciones de tumores asociados al síndrome de Lynch.

2.3.1. Mutaciones patogénicas:

De las 16 alteraciones identificadas en la secuencia del gen descritas en nuestro trabajo sólo una se considera patogénica. Para la nomenclatura de las mutaciones hemos utilizado la secuencia de referencia NM_000179.2.

La inserción c.1458_1459dupT aparece en un probando con cáncer de endometrio y cáncer de páncreas a los 66 años. Carecemos de información sobre los antecedentes familiares. La mutación se produce en el exón 4 del gen y provoca un codón de parada prematuro en el triplete 486 (figura 44). Esta mutación no aparece en la bibliografía consultada y se describe en este trabajo por primera vez; de modo que, debido al codón prematuro de parada que produce, consideramos esta mutación patogénica.

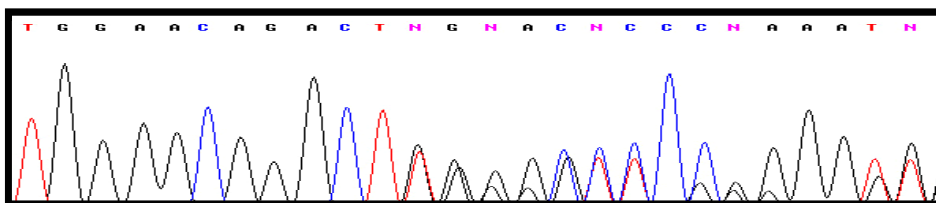


Figura 44: secuencia correspondiente a la mutación identificada en el exón 4 del gen MSH6.

2.3.2. Polimorfismos:

El resto de las mutaciones descritas en el gen MSH6 son consideradas polimorfismos y entre ellas se incluyen cambios intrónicos sin repercusión aparente en la maduración del RNA mensajero (tabla 38). Entre los cambios identificados, también, encontramos varias mutaciones silenciosas (p.Arg62Arg, p.Pro92Pro, p.Asp180Asp, p.Tyr214Tyr) ampliamente descritas en la bibliografía y que aparecen con frecuencia en los probandos analizados. Igualmente, la mutación encontrada en nuestro estudio, p.Gly39Glu; aparece ampliamente descrita y es considerada polimorfismo (Campbell y col., 2008; Smith y col., 2008; Pérez-Carbonero y col. 2009).

Tabla 38: relación de mutaciones encontradas en MSH6. *Probandos en los cuales hemos encontrado la variante.

MSH6	Exón/Intrón	Posición mutación*	Cambio de base	Efecto sobre la proteína	Bibliografía	Frecuencia de aparición*
c.116G>A	1	c.116	G>A	p.Gly39Glu	Campbell y col., 2008; Smith y col., 2008; Pérez-Carbonero y col. 2009	3
c.186G>A	1	c.186	G>A	p.Arg62Arg	Hedge y col., 2005; Charames y col., 2000; Peterlongo y col., 2003; Perez-Carbonero y col. 2009	13
c.276A>G	2	c.276	A>G	p.Pro92Pro	Zidan y col., 2007; Song y col., 2006; Verma y col., 1999; Charames y col., 2000; Hedge y col. 2005; de Abajo y col., 2005; Pérez-Carbonero y col., 2009	3
IVS1+22C>G	Intrón 1	c.260+22	C>G	/	Woods y col., 2005; Kolodner y col., 1999; Plaschke y col., 2000; Peterlongo y col., 2003	7
IVS2+52T>A	Intrón 2	c.457+52	T>A	/	Plaschke y col., 2000; Peterlongo y col., 2003;	6
c.540T>C	3	c.540	T>C	p.Asp180Asp	Wijnen y col., 1999; Plaschke y col., 2000; Kolodner y col., 1999; Pérez-Carbonero y col., 2009	9
IVS7+35InsATCT	Intrón 7	c.3646+35	insATCT	/	Plaschke y col., 2000; Kolodner y col., 1999; Peterlongo y col., 2003; Lucci-Cordisco y col., 2001	1
IVS6-16delT	Intrón 6	c.3557-16	delT	/	NO DESCRITO	1
IVS9+15insACTA	Intrón 9	c.4001+15	dupACTA	/	Peterlongo y col., 2003; Nilbert y col., 2008	1
IVS8+54C>G	Intrón 8	c.3801+54	C>G	/	Kolodner y col., 1999; Verma y col., 1999	5
c.642C>T	4	c.642	C>T	p.Tyr214Tyr	Hendricks y col., 2003; Plaschke y col., 2000; Wijnen y col., 1999; Kolodner y col., 1999; Vahteristo y col., 2005	5
IVS3-56C>T	Intrón 3	c.628-56	C>T	/	Kolodner y col., 1999; Verma y col., 1999; Charames et al. 2000; Plaschke et al. 2000; Vahteristo et al. 2005	1
IVS7+32delCTAT	Intrón 7	c.3646+32	delCTAT	/	Vahteristo et al. 2005; Pérez-Carbonero y col. 2009	6
IVS5+14A>T	Intrón 5	c.3438+14	A>T	/	Plaschke et al. 2000; Kolodner y col., 1999	2

Sobre los polimorfismos exónicos encontrados en el gen MSH6, también realizamos un estudio con el programa ESE FINDER con el fin de predecir una posible alteración del procesamiento del mRNA (tabla 39). En ella se puede observar que los polimorfismos c.276A>G y c.540T>C presentan valores significativos en diferentes proteínas SR y el programa RESCUE ESE corrobora esos resultados. El programa ESE FINDER muestra que el polimorfismo c.186G>A parece afectar a una de las zonas consenso del procesamiento del RNA mensajero.

Tabla 39: predicciones realizadas mediante ESE FINDER y RESCUE ESE. Sombreados en gris y marcados con la cruz aparecen las zonas en las que el polimorfismo analizado puede afectar a la maduración del RNA mensajero. = el programa no predice alteración.

MSH6	ESE FINDER								RESCUE ESE
	5SS_U2	BRANCH SITE	3SS_U2	SF2/ASF	SF2/ASF (IgM-BRCA1)	SC35	SRp40	SRp55	
c.116G>A	=	=	=	=	=	=	=	=	=
c.186G>A	=	=	+	=	=	=	=	=	=
c.276A>G	=	=	=	+	+	+	+	=	+
c.540T>C	=	=	=	+	+	=	=	=	+
c.642C>T	=	=	=	=	=	=	=	=	=

Los resultados indican que las mutaciones c.116G>A y c.642C>T no presentan alteraciones en los sitios ESE. Según la valoración del programa ESE FINDER, la mutación c.186G>A afecta a la secuencia ESE localizada en el extremo 3' del exón; pero el programa RESCUE ESE no predice la alteración de sitios ESE para esta mutación.

ESE FINDER y RESCUE ESE predicen la alteración de los sitios ESE en presencia de los cambios c.276A>G y c.540T>C. La mutación c.276A>G afecta a los sitios de unión de varias proteínas SF2/ASF, SC35 y SRp40. El programa RESCUE ESE valora la presencia de dos sitios ESE (5'AGGAGA3' y 5'GGAGAT3') en la secuencia normal y, cuando presenta el cambio, la secuencia pierde un sitio ESE (5'AGGAGA3') (figura 45). La mutación c.540T>C afecta a uno de los sitios de unión de la proteína SF2/ASF. RESCUE ESE predice cuatro sitios ESE en la secuencia normal. Cuando en la posición 450 se produce el cambio de timina por citosina, los cuatro sitios ESE se pierden y crea tres nuevos (figura 46).

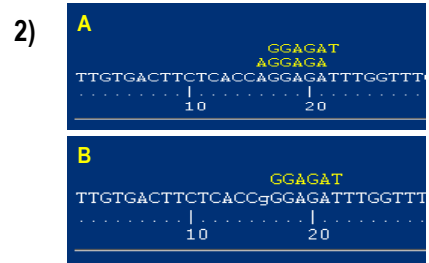
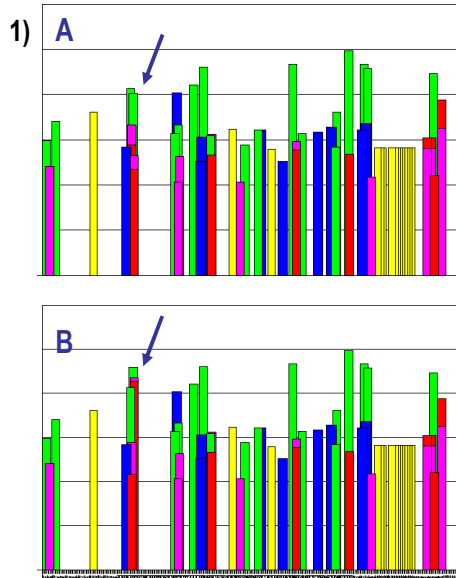


Figura 45: resultado de análisis bioinformático del efecto del polimorfismo exónico **c.276A>C** sobre los sitios ESE. 1: ESE FINDER: 1.A) Sitios de unión de las proteínas SR en la secuencia normal. 1.B) Sitios de unión a las proteínas SR en la secuencia mutada. 2) RESCUE ESE: 2.A) Sitios ESE presentes en la secuencia normal. 2.B) Sitios ESE presentes en la secuencia mutada. 3) Leyenda de la gráfica ESE FINDER.

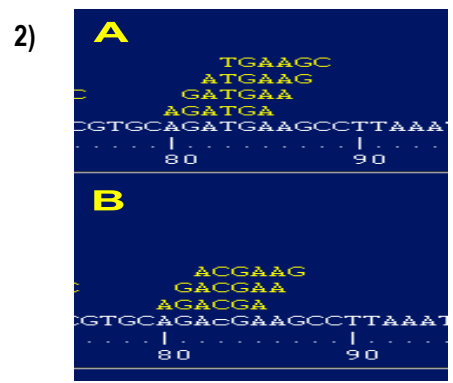
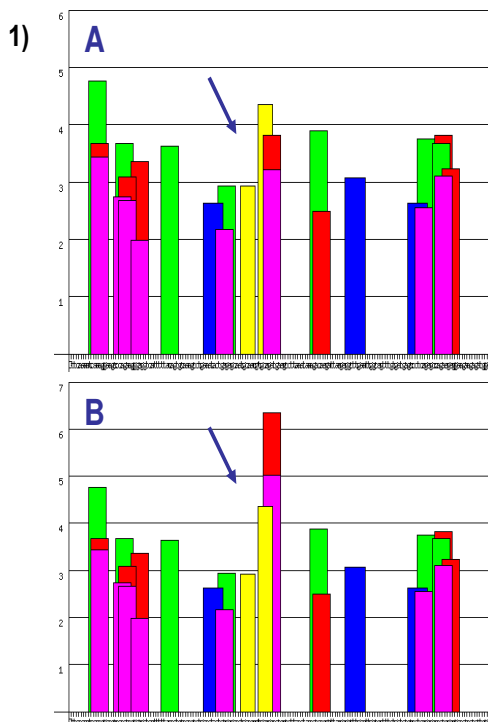


Figura 46: resultado de análisis bioinformático del efecto del polimorfismo exónico **c.540T>C** sobre los sitios ESE. 1: ESE FINDER: 1.A) Sitios de unión de las proteínas SR en la secuencia normal. 1.B) Sitios de unión a las proteínas SR en la secuencia mutada. 2) RESCUE ESE: 2.A) Sitios ESE presentes en la secuencia normal. 2.B) Sitios ESE presentes en la secuencia mutada. 3) Leyenda de la gráfica ESE FINDER.

3. Grandes reordenamientos genómicos en los genes MLH1, MSH2 y MSH6:

Se ha observado que los grandes reordenamientos genómicos suponen una proporción importante del espectro de mutaciones encontrado en los genes del sistema de reparación (Castellví-Bel y col., 2005). En este trabajo hemos realizado el análisis de grandes reordenamientos genómicos en los genes MLH1 y MSH2 en 75 familias que cumplieran los criterios clínicos de Ámsterdam y Bethesda, en las cuales no detectamos ninguna mutación patogénica puntual. El DNA genómico del probando de cada familia fue analizado por triplicado mediante la técnica de MLPA.

Además de los genes MLH1 y MSH2, en el gen MSH6 también se presentan mutaciones germinales que son la causa del síndrome de Lynch y, aunque el porcentaje de reordenamientos genómicos en este gen es menor que en los anteriores, se han descrito hasta un 8% (Castellvi-Bell y col, 2005). En consecuencia, en 14 probandos cuyas familias cumplieran los criterios más estrictos y en los cuales no habíamos encontrado mutación germinal puntual realizamos un análisis por MLPA del gen MSH6.

En los pacientes analizados hemos identificado cuatro alteraciones diferentes (tablas 40 y 41). El probando de la familia 84, con claros criterios de Ámsterdam, presenta en la prueba realizada de MLPA un electroferograma con una pérdida de más del 30% del área de los picos de amplificación de todos los exones correspondientes al gen MLH1, lo que puede ser interpretado como una delección completa del gen en heterocigosis (tabla 40). En la familia 67, el probando muestra un electroferograma en el que la disminución de más del 30% del área se produce en los picos que se corresponden con los exones 4, 5 y 6 del gen MLH1 (tabla 40).

Tabla 40: grandes reordenamientos identificados en los genes MLH1 y MSH2.

ID Familia	Gen	Mutación	Pérdida	Efecto	Referencia bibliográfica	Probandos
FAM 84	MLH1	c.(1-?)_(2272+?)del	Exones del 1 al 19	Pérdida de la proteína	Staaf y col., 2008; van der Klift y col. 2005 di Fiore y col. 2004	1
FAM 67	MLH1	c.(307-?)_(545+?)del	Exones del 4 al 6	La proteína pierde parte de su dominio ATPasa.	di Fiore y col. 2004	1
FAM 184	MSH2	c.(?-68)_211+?del	Extremo 5' y exón 1 del gen MSH2	Pérdida del promotor y el exón 1 del gen	Aparece recogido en base de datos LOVD	1
FAM 187	MSH2	c.943-?_1076+?del	Exón 6	La pérdida del exón produce un codón de parada prematuro.	Guille y col., 2002; Overbeek y col., 2007	1

Tabla 41: Características de las familias portadoras de mutación en MSH2. ^AAP: Antecedentes familiares. ^BAF: Antecedentes familiares. Ca: cáncer. C1: existencia de 3 familiares con cáncer colorrectal, C2: existencia de 3 familiares con tumores asociados* al CCHNP, C3: un cáncer colorrectal en un paciente menor de 50 años, C4: dos casos de cáncer colorrectal sincrónico o metacrónico, o de otros tumores asociados al cáncer colorrectal, C5: un caso de cáncer colorrectal y un familiar de primer grado con tumor asociado al CCHNP, uno de ellos antes de los 50 años, C6: un cáncer de endometrio antes de los 45 años, C7: un cáncer de colon derecho, con formas histológicas poco diferenciadas (solido/cribiforme), antes de los 45 años, C8: pacientes con adenomas colorrectales antes de los 45 años. AMST: Criterios de Amsterdam. BETH: Bethesda.

Gen	Mutación	ID Familia	A.P. ^A	A.F. ^B	Criterio de la Junta de Castilla y León	Criterio Internacional
MLH1	c.(1-?)_(2272+?)del	FAM 84	Ca.colon 57años	Hermano: ca. colon 52 años; Madre: ca. colon 52 años; tío mat: ca. colon metacrónico 50 años; tío mat: ca.colon 50 años y 80 años; abuelo mat: ca. colon 50 años.	C1 y C3	AMST
MLH1	c.(307-?)_(545+?)del	FAM 67	Ca.Colon 44años y 47años.	Abuelo Mat: ca. intestino?; Tío Mat: ca.colon?	C4	BETH
MSH2	c.(?-68)_211+?del	FAM 184	Ca. urotelial 54 años.	Tres casos de ca. urotelial. 2 generaciones, Familiar 2º grado ca. endometrio	C2	BETH
MLH2	c.943-?_1076+?del	FAM 187	C. colon 48 años.	Madre: Ca. Colon 70 años; Abuela mat: Ca. Colon; Tíos mat: Ca. Colon y Ca. Gástrico.	C1 y C3	AMST

La técnica de MLPA es un procedimiento con bajas tasas de reproducibilidad y, por lo tanto, ha de ser comprobado con otras metodologías. En este trabajo hemos utilizado la técnica de PCR cuantitativa a tiempo real para confirmar las posibles alteraciones observadas a través de la técnica de MLPA. De este modo, con el MLPA podemos realizar un rastreo de los posibles reordenamientos genómicos en ambos genes por una técnica rápida que nos permite analizar ambos genes a la vez y por completo y, una vez detectada una alteración, confirmarla más específicamente con otras técnicas (tabla 42).

Tabla 42: resumen de grandes reordenamientos genómicos identificados. Se especifica el tipo de técnica utilizada como método de confirmación.

Gen	Nº alteraciones	Tipo alteración	Consecuencia	Confirmación	A. F.
MLH1	2	c.(1-?)_(2272+?)del c.(307-?)_(545+?)del	Pérdida completa de la secuencia del gen Pérdida de los exones 4 al 6	PCR cuantitativa PCR cuantitativa y RT-PCR	AMST BETH
MSH2	1	c.(?-68)_211+?del	Pérdida de la región promotora del gen y el exón 1	Kits de MLPA independientes	BETH
MSH6	0	-	-	-	-

Así, para confirmar la pérdida del gen completo MLH1 en el paciente, diseñamos un experimento de PCR cuantitativa comparativa a tiempo real. En él, amplificamos el fragmento que coincide con el exón 17 del gen tanto en el DNA del paciente como en el DNA del control. Como control interno de la correcta amplificación en ambas muestras amplificamos, en paralelo, el fragmento correspondiente al exón 7 del gen

p53. Los resultados de esta técnica confirmaron la pérdida del 50% del gen MLH1 en el paciente frente a la amplificación en el control (figura 47).

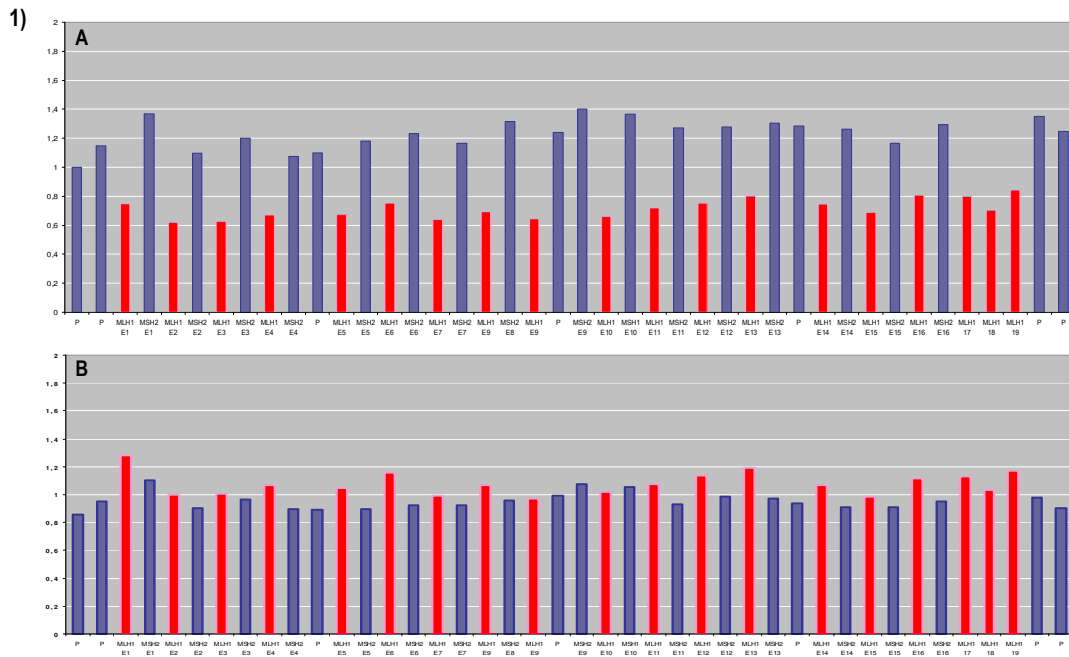
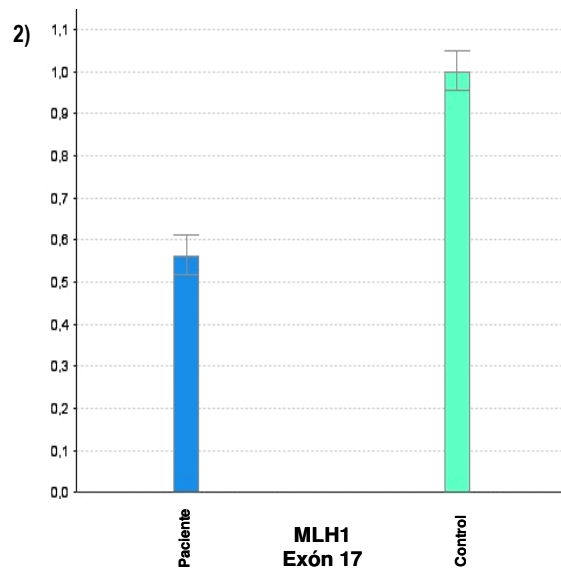


Figura 47: deleción del gen MLH1. 1) resultado del MLPA del probando de la familia 84. El eje X indica el fragmento amplificado. El eje Y muestra los valores normalizados del área de cada pico con respecto al área relativa. Los valores por debajo de 0.6 y por encima de 1.6 son considerados como alterados. A) Electroferograma normalizado del probando. B) Electroferograma normalizado de un control. 2) Confirmación mediante PCR cuantitativa de la pérdida del exón 17 de MLH1. La gráfica muestra los niveles de amplificación del control y del paciente. Se observa una reducción de, aproximadamente, el 50% de la amplificación en la muestra problema.

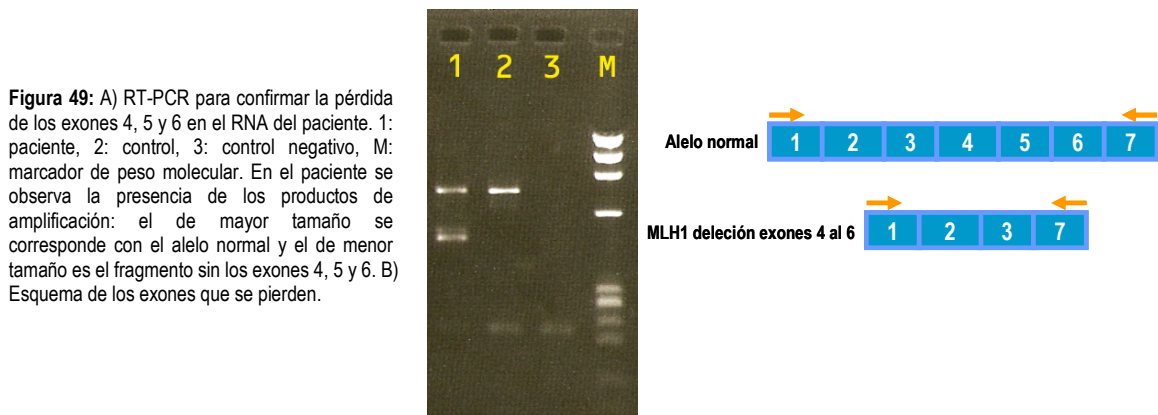


Para confirmar el la pérdida heterocigota de los exones 4 al 6 del gen MLH1, realizamos la misma técnica de PCR cuantitativa. En este caso, utilizamos cebadores diseñados para amplificar el exón 4 y el exón 6 de manera independiente. El gen control que utilizamos fue el exón 1 del gen LEMD3. En ambas amplificaciones observamos que había una reducción del 50% de la intensidad de fluorescencia la reacción de amplificación del DNA del paciente con respecto al DNA control (figura 48).



Figura 48: deleción de los exones 4 al 6 del gen MLH1. A) Electroferograma normalizado del paciente. B) Electroferograma normalizado del control. Comparando ambos se puede observar la disminucióndel área de las barras correspondientes a los exones 4, 5 y 6 en el paciente. C) Confirmación mediante PCR cuantitativa: la gráfica muestra los niveles de expresióndel paciente y el control en dos exones amplificadoss de manera independiente.

Sobre el cDNA, sintetizado a partir de una muestra de RNA del paciente, realizamos una RT-PCR que, de manera complementaria, nos permitiera confirmar la presencia de la alteración en el RNA. Con cebadores sentido y antisentido estudiamos un fragmento de amplificación de aproximadamente 700 pares de bases que incluye desde el exón 1 hasta el exón 7 de MLH1 (figura 49). En la muestra de cDNA del probando obtuvimos dos productos de amplificación uno de 700 pares de bases y otro de aproximadamente 400 pares de bases. La secuenciación de ambos amplicones confirmó que la banda de mayor tamaño se correspondía con la secuencia normal del gen, mientras que la de pequeño tamaño perdía los exones 4, 5 y 6.



En el probando de la familia 184, utilizando el kit diseñado para el gen MSH6 identificamos la pérdida de parte de la zona promotora y el exón 1 del gen MSH2 (figura 50). El probando pertenece a una familia con una amplia agregación de tumores uroteliales, que se incluyen dentro de los tumores asociados al síndrome de Lynch y la presencia de un familiar de segundo grado con tumor de endometrio (tabla 40).

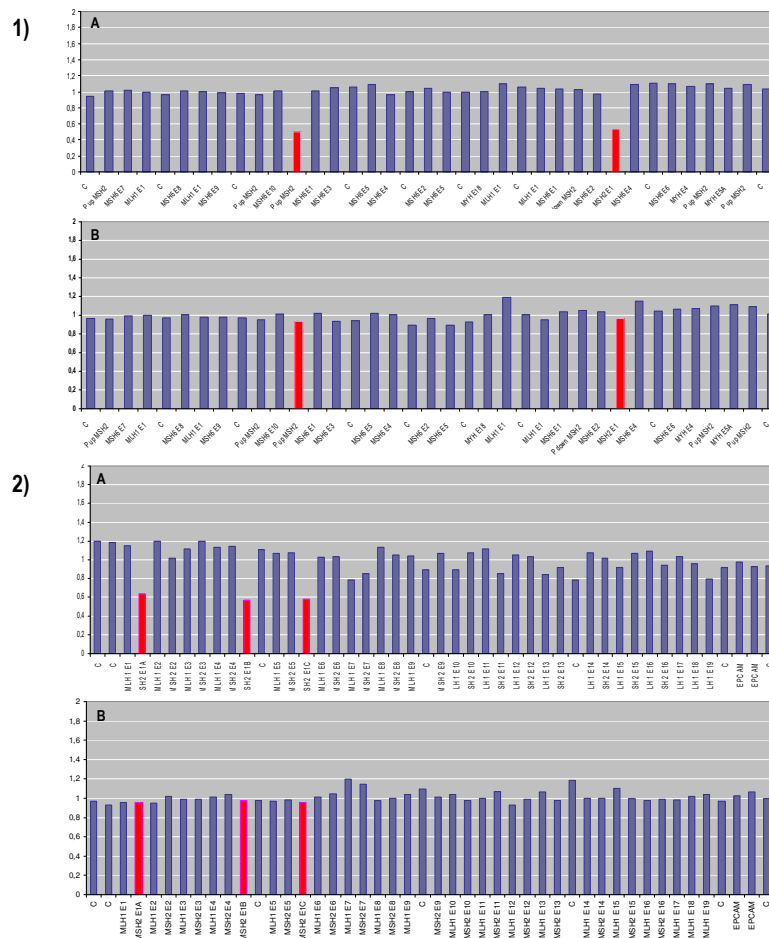


Figura 50: 1) electroferograma normalizado resultado del kit Salsa MLPA p072. A: electroferograma del paciente de la familia 197. B: electroferograma de un control sano. 2) electroferograma normalizado resultado de Kit Salsa MLPA p003 (para el análisis de MLH1 y MSH2). A) electroferograma del paciente 197. B) electroferograma de un control sano.

El resultado del MLPA del probando de la familia 187 muestra la reducción del área correspondiente a la amplificación del exon 6 de MSH2 (figura 51). Esta familia cumple los criterios de Amsterdam: el probando está afectado con cáncer de colon a los 48 años y familiares directos también presentan cáncer de colon.

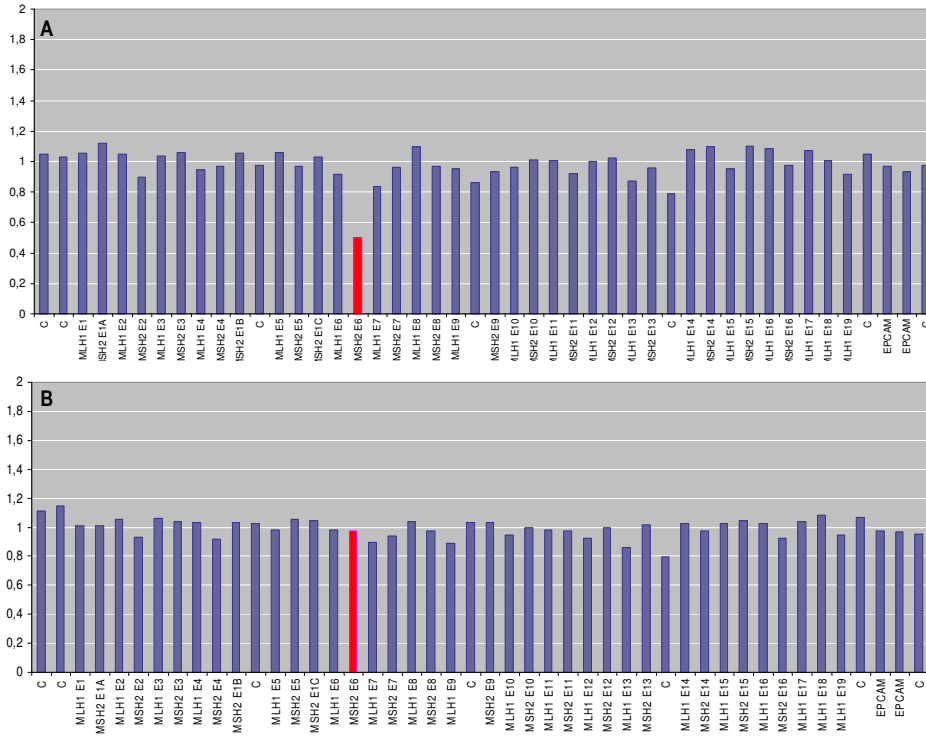


Figura 51: 1 electroferograma normalizado resultado de Kit Salsa MLPA p003 (para el análisis de MLH1 y MSH2. A) electroferograma del paciente 187. B) electroferograma de un control sano.

4. Análisis del gen MLH1 en probandos con inmunohistoquímica negativa para MLH1

Desde el año 2008, en el Departamento de Anatomía Patológica del Hospital Clínico Universitario, se estableció el análisis de la presencia de proteína MLH1, MSH2 y MSH6 en tumores de colon y recto mediante inmunohistoquímica. Doce individuos con cáncer de colon que acudieron a la consulta de Consejo Genético mostraban la pérdida de la proteína MLH1 en el tejido tumoral y de estos doce pacientes contábamos con muestra de sangre y de tumor en nuestro laboratorio.

El análisis del estado de los microsatélites de doce pacientes mostró una alta inestabilidad (figura 52). Sobre la muestra de sangre de estos doce pacientes realizamos la búsqueda de mutaciones puntuales del gen MLH1 mediante amplificación por PCR, análisis por CSGE y secuenciación de los patrones alterados en el gen MLH1 con el fin de descartar una posible mutación patogénica que impidiera la expresión del gen. Ninguno presentaba mutación patogénica alguna y sí, en la mitad de ellos, aparecía el polimorfismo frecuente p.I219V (tabla 43).

Tabla 43: resumen de las técnicas y resultados obtenidos durante el análisis de 12 probandos con análisis de inmunohistoquímica. ^AA.P.: antecedentes personales. ^BA.F.: antecedentes familiares.

ID probando	A.P. ^A	A.F. ^B	IHQ	Estudio molecular MLH1	Análisis Inestabilidad microsatélites
IHQ1	Ca.Colon 78a.	Padre: Estómago 70a.	MLH1 (-)	p.I219V	Alta
IHQ2	Ca. Colon 72a.	Sin antecedentes	MLH1 (-)	p.I219V	Alta
IHQ3	Ca. Colon 74a.	Madre: Ca. Colon	MLH1 (-)	Normal	Alta
IHQ4	Ca. Colon 64a.	Sin antecedentes	MLH1 (-)	Normal	Alta
IHQ5	Ca. Colon 80a.	Sin antecedentes	MLH1 (-)	Normal	Alta
IHQ6	Ca. Colon 79a.	Hija: Oclusión intestinal 54a.	MLH1 (-)	p.I219V	Alta
IHQ7	Ca. Colon 80a.	Sin antecedentes	MLH1 (-)	Normal	Alta
IHQ8	Ca. Colon 71a.	Sin antecedentes	MLH1 (-)	p.I219V	Alta
IHQ9	Ca. Colon 73a.	Hijo: Ca. Colon 45a.	MLH1 (-)	Normal	Alta
IHQ10	Ca. Colon 69a.	Sin antecedentes	MLH1 (-)	p.I219V	Alta
IHQ11	Neo colon 78a.	Padre: ca. Recto; Herman: ca. Tiroides.	MLH1 (-)	Normal	Alta
IHQ12	Ca. Próstata + colon 67a.	Padre: Ca. Próstata 81	MLH1 (-)	p.I219V	Alta

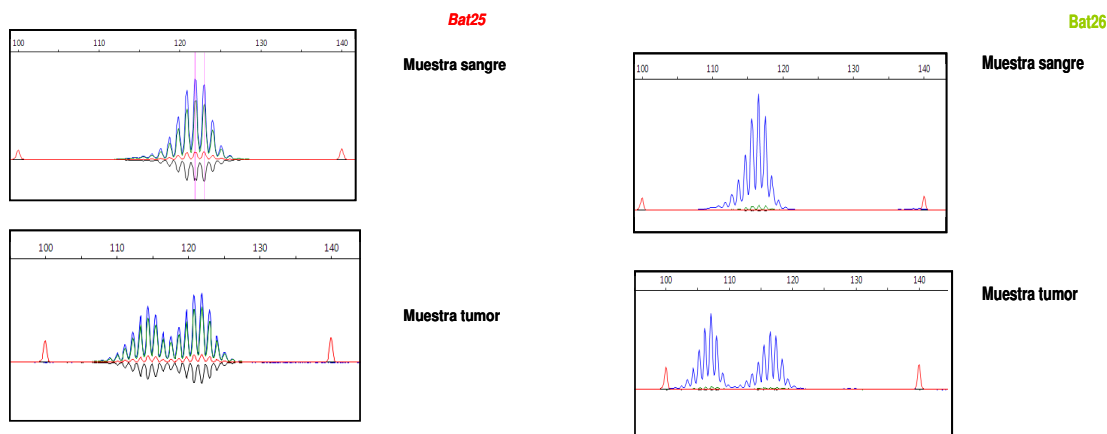


Figura 52: electroferograma que representa dos marcadores utilizados comúnmente para el análisis de la inestabilidad de microsatélite en un paciente: Bat25 y Bat26. En la comparación de la muestra de sangre frente al tejido tumoral se observan cambios en los patrones de amplificación.

6. Genotipado de polimorfismos en el gen MLH1

El polimorfismo de MLH1 c.-93G>A (rs1800734) está localizado en la región promotora del gen, en una zona del promotor que sirve como lugar de unión de factores de transcripción (Mei y col. 2010). Este polimorfismo ha sido asociado con diversos tipos de cáncer como el de pulmón, el de mama, el cáncer de colon o de endometrio (Mei y col, 2010). Se llevó a cabo el análisis de la distribución del genotipo del polimorfismo c.-93A>G en población con sospecha de síndrome de Lynch, población con cáncer de colon y población control. Hemos utilizado un total de 451 individuos distribuidos en los siguientes grupos (tabla 44).

Tabla 44: resumen de los probandos utilizados en el análisis de la distribución del polimorfismo rs1800734. n: n° de muestra.

Grupo	n	Características
Síndrome de Lynch	121	Individuos incluidos en Programa de Consejo Genético. En este grupo se engloban pacientes que cumplen criterios de Amsterdam, Bethesda
Cáncer de colon esporádico	145	Individuos mayores de 60 años con cáncer de colon y sin agregación de tumores en la familia.
Población control	147	Individuos sanos y mayores de 60 años.

En el estudio caso control del polimorfismo rs1800734 de MLH1 el alelo A tiene una baja presencia en la población de cáncer de colon asociado con síndrome de Lynch. La comparación de la presencia de este alelo entre las poblaciones control y cáncer de colon esporádico no muestran diferencias significativas ($p=0.563$, OR (GG): 0.874-1.281; OR (AG,AA): 0.726-1.191). Sin embargo, la comparación de la población de síndrome de Lynch con las poblaciones de controles y esporádicos, muestra diferencias significativas (p (Controles vs. Lynch)=0.040, OR (GG): 0.826-0.990, OR (AG + AA): 1.008-1.960; p (Esporádicos vs. Lynch)= 0.010, OR (GG): 0.659-0.940; OR (AG+AA): 1.086-2.026) (tabla 45 y figura 53).

Tabla 45: distribución de los genotipos del polimorfismo en MLH1 c.-93A>G.

Estudio SNP (rs 1800734)	Distribución de genotipos			HWE	Agrupación de genotipos		Total
	GG (%)	AG (%)	AA (%)		GG (%)	AG, AA (%)	
Controles	85 (57.8)	57 (38.8)	5 (3.4)	1.2	85 (57.8)	62 (42.2)	147
Esporádicos	100 (54.6)	70 (38.3)	13 (7.1)	1.05	100 (54.6)	83 (45.4)	183
Lynch	84 (69.4)	32 (26.4)	5 (4.1)	1.001	84 (69.4)	37 (30.6)	121

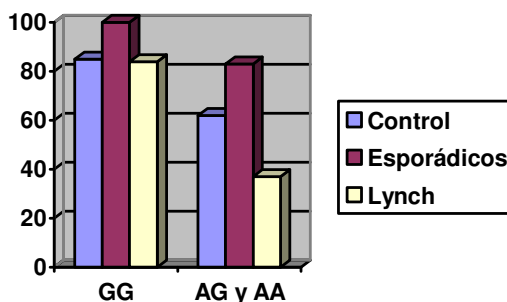


Figura 53: representación gráfica de la distribución de genotipos del SNP rs 1800734.

5. Estudio de microRNAs en los genes MLH1, MSH2 y MSH6:

Los microRNA son pequeñas secuencias de RNA no codificante que participan en la regulación post-transcripcional de muchos genes humanos y, entre ellos, aquellos genes de control del desarrollo, la diferenciación celular y la apoptosis. La actividad de los microRNAs se realiza a través de procesos complejos que implican la unión de las pequeñas moléculas de RNA a la secuencia 3'UTR del gen por complementariedad. La existencia de alteraciones en la secuencia del microRNA o de la secuencia 3'UTR del gen, bien sea por mutaciones puntuales o la pérdida total de la secuencia, puede provocar el desequilibrio en la regulación de los transcritos del gen y ser causante de enfermedad (Liu y col. 2010; Valeri y col. 2010).

En esta parte del trabajo hemos realizado una primera aproximación hacia el conocimiento de la importancia de los microRNA en el desarrollo del Síndrome del Lynch. En primer lugar, utilizando las bases de datos sobre microRNA disponibles en internet, hemos identificado aquellos RNA de pequeño tamaño que tienen la capacidad de unirse complementariamente a diferentes regiones del extremo 3'UTR de MLH1, MSH2 y MSH6. Posteriormente, describimos la relación de polimorfismos localizados en los extremos 3'UTR de estos 3 genes y la posibilidad de que, en esa zona, se uniera un microRNA. Por último, hemos analizado la existencia de polimorfismos en los genes de los microRNAs con capacidad de unión a los genes de estudio.

Los resultados obtenidos muestran que existen varios microRNAs con capacidad para unirse al extremo 3'UTR de los genes reparadores estudiados. Los microRNAs identificados pertenecen a varias familias y se codifican en diferentes cromosomas (tablas 46, 47 y 48).

Tabla 46: resumen de miRNAs con capacidad de unión al extremo 3'UTR de los genes MLH1. Las zonas coloreadas se corresponden con la secuencia complementaria que permite a unión.

miRNA	Secuencia 3'-5' miRNA	Secuencia 5'-3' Extremo 3'UTR de MLH1	Posición cromosoma
hsa-miR-148b	UGUUUCAAGACACU ACGUGACU	AUAUGGUU <u>UUUAUGCACUGU</u>	12: 54731000-54731098
hsa-miR-148a	UGUUUCAAGACAUC ACGUGACU	AUAUGGUU <u>UUUAUGCACUGU</u>	7: 25989539-25989606
hsa-miR-152	GGUUCAAGACAGU ACGUGACU	AUAUGGUU <u>UUUAUGCACUGU</u>	17: 46114527-46114613 [-]
hsa-miR-188-3p	ACGUUUGGGACGU ACACCUC	UGGUUUUU <u>AUGCACUGUGGGAU</u>	X: 49768109-49768194 [+]
hsa-miR-568	CACACAU <u>AUGUA</u> AUUGUA	<u>UGUAUCA</u> AGUGUGA <u>UAUACAAA</u>	3: 114035322-114035416 [-]
hsa-miR-302b	GAUGAUUU <u>GUACCU</u> UCGUGAAU	AACAUAAGUGU <u>GGUAGCAUUA</u>	4: 113569641-113569713 [-]
hsa-miR-372	UGCGAGUU <u>UACAGCG</u> UCGUGAAA	AACAUAAGUGU <u>GGUAGCACUUA</u>	19: 54291144-54291210 [+]
hsa-miR-302d	UGUGAGUU <u>UGUACCU</u> UCGUGAAU	AACAUAAGUGU <u>GGUAGCACUUA</u>	4: 113569160-113569227 [-]
hsa-miR-520a-3p	UGUCAGGU <u>UCCCU</u> UCGUGAAA	AACAUAAGUGU <u>GGUAGCACUUA</u>	19: 54194135-54194219 [+]
hsa-miR-302c	GGUGACUU <u>UGUACCU</u> UCGUGAAU	AACAUAAGUGU <u>GGUAGCACUUA</u>	4: 113569519-113569586 [-]
hsa-miR-520b	GGGAGAUUU <u>UCCCU</u> UCGUGAAA	AACAUAAGUGU <u>GGUAGCACUUA</u>	19: 54204481-54204541 [+]
hsa-miR-520c-3p	UGGGAGAUUU <u>UCCCU</u> UCGUGAAA	AACAUAAGUGU <u>GGUAGCACUUA</u>	19: 54210707-54210793 [+]
hsa-miR-361-5p	CAUGGGGACCUCU AAGACUAUU	GACUU <u>AUACUUGCUUCUGAUAG</u>	X: 85158641-85158712 [-]
hsa-miR-340	UUAGUCAGAGUA ACGAAUAUU	CUUCUGAUAGU <u>AUCCUUUAUAC</u>	5: 179442303-179442397 [-]
hsa-miR-568	CACACAU <u>AUGUA</u> AAUUGUA	UCUGAU <u>GAUUCUUUAUACAC</u>	3: 114035322-114035416 [-]
hsa-miR-374a	GUGAAUAGU CCAACA-UAAUAUU	AUACACAGUGAU <u>UAUUAUAAA</u>	X: 73507121-73507192 [-]
hsa-miR-208b	UGUUUGGAAA CAAGCAGAAUA	UAAAUA <u>AAUAGAUGUCUUAAC</u>	14: 23887196-23887272 [-]

Tabla 47: resumen de miRNAs con capacidad de unión al extremo 3'UTR de los genes MSH2. Las zonas coloreadas se corresponden con la secuencia complementaria que permite a unión.

miRNA	Secuencia 3'-5' miRNA	Secuencia 5'-3' Extremo 3'UTR de MSH2	Posición Cromosoma
hsa-miR-21	AGUUGUAGUCAGACU <u>AUUCGAU</u>	AUGAAGGUAAU <u>UUGAUAAGCUA</u>	17: 57918627-57918698 [+]
hsa-miR-9			
hsa-miR-590-5p	GACGUGAAA <u>UACUUAUUCGAG</u>	AUGAAGGUAAU <u>UUGAUAAGCUA</u>	7: 73605528-73605624 [+]
hsa-miR-889	uguuacc ACAGGC UAUAAUU	UUUUUAU <u>UUGUUAUUA</u>	14: 101514238-101514316 [+]
hsa-miR-559	AAAACCAGGUAAA <u>UAGAAA</u>	UUUUCAAAG <u>AUUAUUAUUG</u>	2: 47604814-47604909 [+]
hsa-miR-548b-5p	cggguUUGGUGU- <u>UAAUGAAA</u> a	auauuuAGUA <u>AUAUUAUCUUUg</u>	6: 119390212-119390308 [-]
hsa-miR-548c-3p	cguuucauu <u>aacUCUAAAAc</u>	aggacauuuca AGAUUUUa	12: 65016289-65016385 [+]
hsa-miR-30c-2	cgACUCUC-ACA- <u>UCCU-ACAAAUGu</u>	aaUGAGAGCUGUAACUG AGGACUGUUUGCa	6: 72086663-72086734 [-]
hsa-miR-30c-1	cgACUCUC-ACA- <u>UCCU-ACAAAUGu</u>	aaUGAGAGCUGUAACUG AGGACUGUUUGCa	1: 41222956-41223044 [+]
hsa-miR-30a	gaAGGUCAGCUCCU- <u>ACAAAUGu</u>	gcUGUAACUGAGGACUGUUUG Ca	6: 72113254-72113324 [-]
hsa-miR-30b	ucgACUCACAUCCU- <u>ACAAAUGu</u>	cugUAACUG-AGGACUGUUUG Ca	8: 135812763-135812850 [-]
hsa-miR-212	cCGGCACUGACCUCUGACAAu	AGCUGUAACUGAGGACUGUUU	17: 1953565-1953674 [-]
hsa-miR-132	gcUGGUACCGACAUCUGACAAu	gaGCUGUAACUGAGGACUGUUu	17: 1953202-1953302 [-]
hsa-miR-935	c-gccaUCGCCUUCGCCAUUGACc	U-GAAA <u>AUGAGAGCUGUAACUGA</u>	19: 54485561-54485561 [+]
hsa-miR-499-3p	ucGUGUCUG-AACGA- <u>CACUACAa</u>	gaCAUAGGCAUAAUAGGUGAU <u>g</u>	20: 33578179-33578300 [+]
hsa-miR-135a-1	agUGUAUCC-UUAUUUUUCGGU <u>Au</u>	ugACAUAGGC <u>AAUAAUAGUGAUg</u>	3: 52328235-52328324 [-]
hsa-miR-135a-2	agUGUAUCC-UUAUUUUUCGGU <u>Au</u>	ugACAUAGGC <u>AAUAAUAGUGAUg</u>	12: 97957590-97957689 [+]
hsa-miR-592	uguaguAGCGUUAACUGUUU	GACUUUU GCAAUUGACAUAG	7: 126698142-126698238 [-]
hsa-miR-545	cgugUGUUUUUAC-AAACGACu	ggcaUA <u>AAUAGUGAUGUGCUGa</u>	X: 73506939-73507044 [-]

Tabla 48: resumen de miRNAs con capacidad de unión al extremo 3'UTR de los genes MSH6. Las zonas coloreadas se corresponden con la secuencia complementaria que permite a unión.

miRNA	Secuencia 3'-5' miRNA	Secuencia 5'-3' Extremo 3'UTR MSH6	Posición Cromosoma
hsa-miR-409-3p	UCCCAAGUGGCUCGUUGUAAG	GGUGGUAUUUCAGA---CAACAUUA	14: 101531637-101531715 [+]
hsa-miR-561	UGAAGUUCUAGAA--UUUGAAAC	ACAUUAUGAUCUAAUAAACUUUA	2: 189162219-189162315 [+]
hsa-miR-142-5p	UCAUCACGAAAGAUGAAAUAC	AUUUAUGAUCUAAUAAACUUUAUU	17: 56408593-56408679 [-]
hsa-miR-330-3p	AGAGACGUCCGGCACCGAAACG	CUGACUACAUUGGAAGCUUUGAG	19: 46142252-46142345 [-]
hsa-miR-888	ACUGACUGUCGAAAACUCUACU	GACUACAUUGGAAGCUUUGAGUU	X: 145076302-145076378 [-]
hsa-miR-1231	CGUCGACAGCGGGUCUGUG	CAAAGGUGGUAUUUCAGACAAC	1: 20177739-20177830 [+]
hsa-miR-632	AGGGUGUCUUCGUCUGUG	CAAAGGUGGUAUUUCAGACAAC	17: 30677128-30677221 [+]

A pesar de la variabilidad de los miRNAs identificados y de la variabilidad de su localización a lo largo de la secuencia de no codificante, no hemos identificado polimorfismos en esas zonas de unión. Excepto para el gen MLH1: el SNP rs2234891 se localiza en la zona de unión del microRNA miR-374a (tablas 49, 50 y 51).

Tabla 49: SNP localizados en la región 3'UTR en los genes MLH1.

SNP MLH1	Localización cromosómica	Alteración	Cambio	miRNA que se une	Alelo 1	Alelo 2
rs 72299848	3:37067178-37067180	-/TTC	c.*30_32del3	/	/	/
rs 56329536	3:37067321-37067324	-/AATA	c.*173_176del4	/	/	/
rs 55711899	3:37067185	-/CTT	c.*35_37del3	/	/	/
rs 2234891	3:37067313	-/GATT	c.*164_167del4	hsa-miR-374a	/	/
rs 1803985	3:37067249	G/T	c.*101G>T	/	/	/

Tabla 50: SNP localizados en la región 3'UTR en los genes MSH2.

SNP MSH2	Localización cromosómica	Alteración	Cambio	miRNA que se une	Alelo 1	Alelo 2
rs66505139	2:47563765	-/T	c.*173delT	/	/	/
rs17225060	2:47563818	A/G	c.*254A>G	/	0.988 A	0.012 G
rs17225053	2:47563733	G/T	c.*169T>G	/	0.988 T	0.012 G

Tabla 51: SNP localizados en la región 3'UTR en los genes MSH6.

SNP MSH6	Localización cromosómica	Alteración	Cambio	miRNA que se une	Alelo 1	Alelo 2
rs78049492	2:48034356	G>T	c.*901G>T	/	0.136 A	0.864 C
rs71772028	2:47873730-47873733	-/TACA	c.*1046_*1049del4	/	/	/
rs55959776	2:48034279-48034323	Delección de 45 nucleótidos	c.*934_*978del45	/	/	/
rs55886808	2:48034346-48034347	-/TGTC	c.*910_*911ins4	/	/	/
rs55762324	2:48034365-48034366	-/JA	g.26856252_26856253insA	/	/	/
rs3136370	2:48034396-48034397	-/TT	g.26856283_26856284ins2	/	/	/
rs3136369	2:48034156	C/T	g.26856043T>C	/	0.006 C	0.994 T
rs2020906	2:48034084	A/T	c.*85T>A	/	0.021 A	0.979 T

Por otro lado, existen polimorfismos en algunos microRNA que podrían afectar a su unión con las secuencias no codificantes de los genes MLH1 y MSH2. En aquellos microRNA identificados para MSH6 no se han encontrado SNPs (Tabla 52).

Todos estos datos deben ser valorados con cautela, puesto que son aproximaciones preliminares basadas en la información que aportan las bases de datos. Es necesaria la comprobación tanto de la unión de estos microRNAs a los genes reparadores estudiados como la presencia del SNP puede afectar a la expresión del mRNA.

Tabla 52: SNP localizados en las secuencias en los microRNA que se unen a los genes MLH1, MSH2 y MSH6.

miRNA MLH1	SNP miRNA	Alteración	Cambio	Homocig N	Heterocig	Homoc M
hsa-miR-148b	rs12304902	C/G (ancestral C)	NT_029419.11:g.18588123C>G			
hsa-miR-148a	rs6968046	C/T (ancestral T)	NT_007819.16:g.25512006T>C			
hsa-miR-152	rs34201788	-/T				
hsa-miR-188-3p	NO	NO	NO			
hsa-miR-568	rs10713136	-/T				
hsa-miR-302b	rs11944440	C/T	NT_016354.18:g.37898338C>T			
hsa-miR-372	NO	NO	NO			
hsa-miR-302d	NO	NO	NO			
hsa-miR-520a-3p	NO	NO	NO			
hsa-miR-302c	rs1989927	A/G (ancestral A)	NT_016354.18:g.37897890G>A	0.50 AA	0.367 AG	0.583 GG
hsa-miR-520b	NO	NO	NO			
hsa-miR-520c-3p	NO	NO	NO			
hsa-miR-361-5p	rs2205706	C/T (ancestral C)	NM_000390.2:c.116+10497C>T	0.348 CC	0.217 CT	0.435 TT
hsa-miR-340	NO	NO	NO			
hsa-miR-568	rs10713136	-/T				
hsa-miR-374a	NO	NO	NO			
hsa-miR-208b	rs71735345	-/A	NT_026437.11:g.5817270delA			
hsa-miR-208b	rs35200394	-/G				
miRNA MSH2	SNP	Alteración	Cambio	Homocig N	Heterocig	Homocig M
hsa-mir-21	NO					
hsa-miR-9	NO					
hsa-miR-590-5p	NO					
hsa-miR-889	NO					
hsa-miR-559	NO					
hsa-miR-548b-5p	rs794251	A/G	NM_024581.4:c.160-2623G>A			
hsa-miR-548b-5p	rs794250	A/C	NM_024581.4:c.160-2606G>A???			
hsa-miR-548c-3p	NO					
hsa-miR-30c-2	NO					
hsa-miR-30c-1	NO					
hsa-miR-30a	NO					
hsa-miR-30b	NO					
hsa-miR-212	NO					
hsa-miR-132	NO					
hsa-miR-935	NO					
hsa-miR-499-3p	rs34313391	-/T				
hsa-miR-135a-1	rs755687	A/G				
hsa-miR-135a-2	rs12146836	C/T	NM_020140.2:c.456+13429T>A			
hsa-miR-135a-2	rs7298602	C/T	NM_020140.2:c.456+13431T>A			
hsa-miR-592	NO					
hsa-miR-545	rs249685	A/G	NM_005685.2:c.-7+523G>C			
miRNA MSH6	SNP	Alteración	Cambio	Homocig N	Heterocig	Homocig M
hsa-miR-409-3p	NO					
hsa-miR-561	NO					
hsa-miR-142-5p	NO					
hsa-miR-330-3p	NO					
hsa-miR-888	NO					
hsa-miR-1231	NO					
hsa-miR-632	NO					

Discusión

El síndrome de Lynch es una enfermedad que se transmite con un patrón de herencia autosómica dominante con penetrancia incompleta, producida por la presencia de mutaciones en la línea germinal de genes implicados en la ruta reparadora de errores de alineamiento durante la replicación del DNA (MMR), principalmente los genes MLH1, MSH2 y MSH6.

En este trabajo hemos estudiado los genes, los genes MLH1 y MSH2 en 153 probandos, los genes MLH1, MSH2 y MSH6 en 29 probandos y únicamente el gen MSH6 en 6 probandos con sospecha de síndrome de Lynch, seleccionados de acuerdo con los criterios establecidos por el Programa de Consejo Genético de la Junta de Castilla y León (PCGJCYL), que son menos estrictos que los empleados habitualmente (Criterios de Amsterdam y de Bethesda) en un intento por evitar la exclusión de familias portadoras de mutación en las que el estudio familiar es poco informativo. El gen MSH6 sólo se estudió en familias en las que o bien había tumores de endometrio o tenían una gran penetrancia habiendo descartado previamente la existencia de mutaciones en MLH1 o MSH2. De los 188 sujetos incluidos en nuestro estudio, 125 (66,49%) cumplían al menos uno de los criterios del PCGJCYL, mientras que 63 (33,51%) no cumplían los criterios o eran individuos sanos que se estudiaron por motivos compasivos al no existir familiares vivos portadores de enfermedad neoplásica (tabla 23, página 61).

El estudio de los pacientes de acuerdo con los criterios de Amsterdam o Bethesda mostró que 97 sujetos (51.59%) cumplían criterios de Amsterdam o Bethesda: 30 (30.93%) criterios de Amsterdam y 67 (69.07%) criterios de Bethesda. El resto de los probandos analizados eran sanos (el 20.74%, 39 de 188) o sin criterios de inclusión (27.65%, 52 de 188) (tabla 21, página 58)

El estudio molecular puso de manifiesto la existencia de mutaciones en 15 probandos. Doce mutaciones estaban presentes en familias que cumplían criterios de Amsterdam, Bethesda y de PCGJCYL. Si comparamos la efectividad en la detección de mutaciones, observamos que de 125 probandos que cumplían criterios del PCGJCYL, once eran portadores (8.8%); mientras que de 97 probandos que cumplían criterios internacionales once eran portadores de mutación (11.34%) (tabla 53).

Además, hemos encontrado mutaciones patogénicas en tres en familias sin criterios o en individuos sanos.

Tabla 53: probandos portadores que cumplen criterios internacionales versus probandos portadores de mutación que cumplen criterios de la Junta de Castilla y León.

Criterios	Nº probandos analizados	Nº probandos portadores	%
Criterios internacionales (Amsterdam y Bethesda)	97	11	11.34
PCGJCYL	125	11	8.8

Estos datos indican que los criterios establecidos por el PCGJCYL incorpora muchas más familias en el estudio lo que produce una disminución del porcentaje de pacientes portadores de mutación pero nos ha permitido diagnosticar tres familias portadoras de mutación que no cumplían criterios internacionales. Uno de estos probandos estaba afectado de cáncer de colon a los 51 años sin presentar antecedentes familiares y era portador de una mutación patogénica en el gen MLH1. Otro de los probandos era un individuo sano cuya familia no cumplía criterios. El tercer probando era un individuo sano, la familia cumplía los criterios pero todos los miembros afectados habían fallecido.

Por lo tanto, los criterios de Amsterdam y Bethesda, por sí mismos son lo suficientemente robustos como para identificar la mayoría de las familias con sospecha de síndrome de Lynch portadoras de mutación patogénica pero existe un pequeño porcentaje de casos portadores de mutación que no son detectados con estos criterios y que se benefician de criterios más relajados como los del PCGJCYL. Si hiciésemos una relación coste beneficio, parece evidente que el empleo de los criterios de Amsterdam o Bethesda son los más rentables desde esta perspectiva, pero desde un punto de vista ético nos plantea el problema de las tres familias portadoras de mutación que nunca hubiesen sido diagnosticadas.

La prevalencia de mutaciones en poblaciones analizadas en diferentes estudios varía del 86% en Finlandia (Nystrom-Lati y col., 1999) al 44% en familias que cumplían criterios de Amsterdam en España (Caldés., 2002). En nuestro estudio hemos observado que sólo 6 de 30 (20%) familias con criterios de Amsterdam incluidos en nuestro trabajo presentaban mutación patogénica. Esta discrepancia es difícil de explicar y podría deberse a una menor sensibilidad de los métodos de escrutinio empleados en nuestro trabajo o a una mayor proporción de casos de cáncer de colon familiar sin mutación en genes reparadores en la población controlada por el PCGJCYL.

La mayoría de los tumores asociados al síndrome de Lynch se diagnostican antes de los 45-50 años (la Chapelle y col., 2004; Lynch y col., 2009). En nuestro estudio la media de edad de diagnóstico del tumor en los probandos fue de 52 años y la mayor frecuencia de mutaciones se encontró en probandos con edades mayores de 50 años (tabla 54).

Tabla 54: Porcentaje de mutación según edad. ^A De todos los probandos de los cuales tenemos edad de diagnóstico del tumor.

Edad	Nº de probandos	Nº de probandos mutados	%
Menores de 50 años	76	3	3.95
Mayores de 50 años	79	9	10.13
Total ^A	155	12	/

Llama la atención la edad de diagnóstico de los tumores en los probandos que es superior a la de otras series, pensamos que esta observación puede reflejar deficiencias en los sistemas de

prevención de cáncer de colon y sugieren que sería conveniente mejorar los sistemas de escrutinio para hacer una detección precoz de este tipo de tumores.

En nuestro trabajo hemos identificado 61 mutaciones; el 21.31% eran mutaciones patogénicas asociadas con la enfermedad, el 59.02% polimorfismos poblacionales y el 19.67% variantes de significado desconocido. 11 de 61 (18.03%) mutaciones identificadas se describen por primera vez en este trabajo y de ellas 4 (36.36%) son variantes de significado desconocido. Estas mutaciones son en su mayoría mutaciones de cambio de aminoácido (*missense*) de las cuales no conocemos su efecto sobre la proteína y, por tanto, no podemos asociar la presencia de esta alteración con el desarrollo de la enfermedad.

El desconocimiento del efecto sobre la función y la estructura que provocan las variantes de significado desconocido impiden obtener un resultado informativo en el análisis genético. Puesto que los estudios funcionales resultan demasiado costosos en tiempo y esfuerzo; se suele realizar un estudio aproximativo en general no conclusivo. Los estudios de clasificación de las variantes de significado desconocido se centran, principalmente, en estudios de asociación de pacientes frente a controles, estudios de segregación familiar de la mutación, conservación entre especies, pérdida de heterocigosidad, estudios bioinformáticos, etc.

1. Estudio del gen MLH1

El estudio del gen MLH1 nos ha permitido caracterizar cuatro mutaciones patogénicas (tabla 24, página 62) de las que tres son mutaciones sin sentido (*nonsense*) que generan un codón de parada (p.R100X, p.R226X y p.R487X). La aparición de un codón de parada prematuro puede alterar la expresión y la función de la proteína dos maneras diferentes: a) El mRNA transcrito con un codón de parada prematuro es una molécula inestable y fácilmente degradable, de manera que el acortamiento de la vida media del mRNA mutante impediría la traducción de la proteína disminuyendo los niveles de la misma. En este sentido, Nomura y sus colaboradores describieron la degradación del mRNA de varias mutaciones *nonsense* del gen MLH1, entre ellas **p.R226X** a través del mecanismo de degradación de mutaciones nonsense (NMD, nonsense-mediated decay) (Nomura y col., 2000). b) En el caso de que el mRNA mutado sea estable, y eluda el mecanismo NMD, la proteína sintetizada a partir de él está truncada y puede carecer de dominios importantes para su función. No obstante, existen mutaciones *nonsense* en las que el codón de parada se encuentra en el extremo carboxilo terminal de la proteína, fuera de cualquier dominio funcional; en este caso, la pérdida de los últimos aminoácidos de la proteína no produciría ningún efecto negativo en la funcionalidad de la proteína y estas mutaciones no son consideradas patogénicas. Las mutaciones localizadas en la región central del gen MLH1 producen una proteína que carece tanto de los dominios funcionales de interacción con PMS2 ó MLH3 como de parte

del dominio ATPasa que le permite estabilizar la dimerización en los complejos MutL α y MutL β . LA mutación p.R100X produce una proteína que pierde los dominios ATP y los dominios localizados en el dominio carboxi-terminal. La mutación p.R487X produce una proteína truncada que carece del dominio de dimerización con PMS1 localizado en el extremo C-terminal de la proteína.

La cuarta mutación patogénica encontrada en nuestra serie es un cambio de sentido **p.T117M** que se localiza en el extremo amino-terminal de la proteína, dentro del dominio ATPasa, que se extiende desde, aproximadamente, desde el codon 30 hasta el 350 de la proteína y tiene una gran similitud con dominios de unión a ATP de otras ATPasas como la girasa B de *Escherichia coli* y la proteína HSP90 (Raschle y col., 2002). El dominio ATPasa permite la unión e hidrólisis de una molécula de ATP por parte de la subunidad MLH1 que facilita la estabilización de la interacción con PMS2 ó MLH3 dentro del heterodímero MutL α ó MutL β (Raschle y col., 2002). La secuencia aminoacídica que conforma el dominio ATPasa en esta proteína está muy conservado entre especies, por lo que un cambio de aminoácido puede afectar tanto a su estructura como a la función que realiza dentro de la proteína (Takahashi y col., 2007).

En el gen MLH1 hemos encontrado 10 polimorfismos entre los cuales se incluyen variantes intrónicas, mutaciones silenciosas y una mutación de cambio de aminoácido (tabla 25, página 63). Tanto las mutaciones intrónicas que se producen en sitios alejados de las secuencias canónicas de procesamiento como las mutaciones silenciosas en regiones codificantes que no cambian el aminoácido se han considerado clásicamente como polimorfismos al considerar que no modifican las proteínas codificadas por un gen. No obstante, hoy sabemos que algunas de estas secuencias pueden modular el procesamiento del RNA al estar localizadas en regiones reguladoras del procesamiento como los sitios ESE (*exonic splicing enhancer*) o ISE (*intronic splicing enhancer*). La regulación que ejercen las secuencias ESE se produce, entre otras, a través de las proteínas SR, que tienen la capacidad de reconocer y unirse a los sitios ESE y reclutar a los componentes de la maquinaria de maduración del RNA (Fairbrother y col., 2004). Las mutaciones que alteren alguna de estas secuencias pueden afectar a la unión de las proteínas SR y, consecuentemente, disminuir la eficacia en la maduración del RNA mensajero. La mayor densidad de los sitios ESE importantes se localizan tanto en los primeros 20-25 nucleótidos del extremo 5' como en los del extremo 3' (figura 54) (Fairbrother y col., 2004); en consecuencia, mutaciones localizadas en estas zonas tienen una mayor probabilidad de resultar dañinas para el procesamiento del RNA. Sin embargo, los SNPs aparecen distribuidos por los exones con menor frecuencia en las regiones próximas a las zonas de consenso de procesamiento del RNA mensajero y con una alta frecuencia en el interior en un intento de mantener la estabilidad del procesamiento del RNA (Fairbrother y col.; 2004).

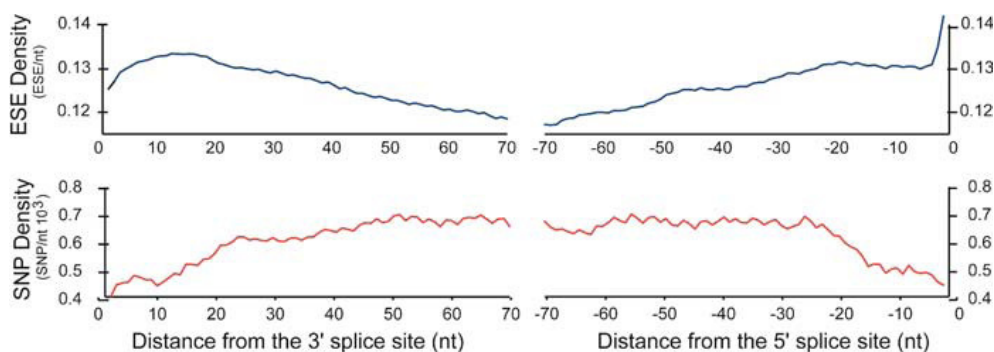


Figura 54: A) la gráfica muestra la densidad de los sitios ESE en función a su posición en el exón. B) la gráfica de la distribución de los SNPs a lo largo del exón.

El análisis bioinformático sobre el efecto de las mutaciones silenciosas que hemos encontrado en el gen MLH1 en nuestra serie que la mutación **c.702G>A** afecta a un sitio ESE. El cambio se localiza a 25 nucleótidos en el extremo 5' del exón 9, zona donde teóricamente empiezan a escasear los sitios ESE y se encuentra acompañado por otros tres sitios ESE que no se solapan con él (figura 2). Este sitio ESE se pierde cuando en esa posición aparece G en vez de A según la predicción del programa RESCUE ESE (Auclair y col., 2004); sin embargo, el programa ESE FINDER, que realiza predicciones más fiables y con mayor especificidad (Auclair y col., 2004) no encuentra alteraciones en el sitio ESE. En resumen, la presencia de este cambio de nucleótido no tiene una predicción teórica de peso que indique su efecto sobre el procesamiento del mRNA, siendo necesarios estudios sobre el RNA para comprobar que el procesamiento es normal.

En nuestro estudio hemos observado una mayor frecuencia del alelo mutado del polimorfismo **c.-93G>A** en población con cáncer de colon esporádico frente a población control y aún mayor con respecto a la población con diagnóstico de síndrome de Lynch, lo que sugiere que el alelo mutado confiere una predisposición a padecer cáncer de colon esporádico y no está asociado al síndrome de Lynch. Es bien conocido que muchos tumores colorectales presentan inestabilidad de microsatélites y pérdida de expresión de la proteína MLH1 como consecuencia del silenciamiento epigenético secundario a hipermetilación del promotor de este gen. El polimorfismo **c.-93G>A** se localiza en el promotor del gen MLH1 y la presencia del alelo mutado impide la correcta interacción de los factores de transcripción con el DNA y disminuye la expresión de MLH1, produciendo una disminución de la reparación post-replicativa del DNA (Mei y col., 2009), lo que explicaría su alta frecuencia en individuos con cáncer colorrectal esporádico e inestabilidad de microsatélites. Además, este polimorfismo se ha asociado con un aumento del riesgo de sufrir tumores de ovario, endometrio, linfoma Hodgkin y tumores colorrectales con inestabilidad de microsatélites y pérdida de expresión de MLH1 en individuos que no son portadores de mutaciones germinales en este locus (Nedja y col., 2009).

1.1. Variantes de significado desconocido en el gen MLH1

El estudio del gen MLH1 en pacientes diagnosticados de síndrome de Lynch nos permitió caracterizar seis variantes sin clasificar (tabla 28, página 66) que han sido descritas previamente por otros autores excepto la mutación c.23023C>T. No obstante, hasta el momento no se ha realizado la caracterización funcional de estas mutaciones. Nosotros hemos realizado un estudio *in silico* de las mismas como primera aproximación a la realización de estudios funcionales.

1.1.1. Variantes de significado desconocido localizadas en el dominio ATPasa

La variante de significado desconocido **p.G22A** se localiza en el extremo aminoterminal de la proteína MLH1, donde se localiza un dominio ATPasa que presenta muchas semejanzas con otras proteínas con función ATPasa como la DNA girasa B o HSP90 (Räschle y col., 2002). Este dominio permite la unión y la hidrólisis del ATP, que conlleva cambios conformacionales en la proteína que modifican la interacción con otros dímeros de MutL (como PMS2) o MutS (Arana y col., 2000). El dominio de unión con ATP tiene cuatro motivos y se localiza aproximadamente entre los residuos 30 y 182, siendo el residuo 34 el más importante en la hidrólisis del ATP (Räschle y col., 2002). La glicina en la posición 22 se encuentra incluida en una región muy conservada entre especies: VAGBIRRLDETVVNRIAAGEVIQPRANAIAKEMIENCLD, lo que pone de manifiesto su importancia en la función de la proteína, probablemente facilitando la estabilidad de la estructura terciaria del dominio ATPasa. Por otra parte, dado que la mutación c.65G>C se localiza en un exón, realizamos un estudio de predicción de sitios ESE, observando que la mutación genera un nuevo sitio de unión de la proteína SF/ASF (figura 30, página 69), lo que sugiere que la mutación podría alterar el procesamiento del RNA. El estudio *in silico* de predicción del efecto del cambio del codón sobre la proteína reveló que dos de los programas, SIFT y SNP3D, lo consideran no tolerable, mientras que el programa POLYPHEN lo considera benigno (tabla 30, página 70) El cambio de glicina por alanina produce la aparición de un resto hidrofóbico donde antes no lo había. A la vista de estos datos, y dado que la mutación no se encontró en un estudio poblacional en 100 individuos sanos, y que el paciente, con un claro fenotipo de síndrome de Lynch, no era portador de otra mutación concomitante en genes reparadores, podemos considerar que la mutación es probablemente patogénica, aunque son necesarios estudios funcionales para asignarla definitivamente a este grupo.

La mutación **p.R69K** está localizada también en el dominio ATPasa de la proteína MLH1, en un aminoácido muy conservado entre diferentes especies. El cambio de arginina, presente en las proteínas silvestres, por lisina, no parece tener ningún efecto negativo sobre la estructura terciaria puesto que ambos tienen una estructura y unas características químicas similares. El programa RESCUE-ESE

muestra que el cambio de adenina por guanina en la posición 206 (c.206G>A) añade dos sitios ESE potenciales al ya existente en el exón 2, generando un nuevo sitio de unión para la proteína SC35, por lo que esta mutación podría modificar el nivel del procesamiento del RNA. El estudio *in silico* realizado en este trabajo muestra que dos de los tres programas de predicción, *SIFT* y *SNP3D*, valoran el cambio como tolerable; mientras que el programa *POLYPHEN* lo considera como posiblemente patogénico (tabla 30, página 70). Otros grupos han realizado estudios funcionales de expresión transitoria de la proteína mutada en diferentes cepas de levadura, concluyendo que, si bien la proteína mutada no pierde totalmente su función, ésta aparece disminuida y, por tanto, existe un descenso en la capacidad reparadora de la célula (Ellison y col., 2001 y Takahashi y col., 2007). Más concretamente, Ellison y colaboradores, generaron en el laboratorio una proteína MLH1 híbrida entre la humana y la de levadura que complementa la deficiencia de reparación de una cepa de levadura. El híbrido MLH1 que posee la variante tiene una baja tasa de reparación, pero no se anula completamente. Por lo tanto, Ellison y colaboradores, clasifican a esta variable como un polimorfismo de eficiencia, es decir, un cambio en la proteína que sin anular la función, producen una variación en la eficiencia de su actividad (Ellison y col., 2001). El estudio poblacional de esta mutación puso de manifiesto que no se encuentra en la población sana y el paciente portador de la mutación no presentaba ninguna mutación patogénica en los genes reparadores estudiados, lo que sugiere que pueda tratarse de una mutación patogénica.

1.1.2. Variantes de significado desconocido localizadas en el dominio de dimerización

Hasta el momento se desconoce el significado de la mutación **p.K618A**. Esta mutación provoca el cambio de un aminoácido cargado por uno neutro en una zona conservada de la proteína (figura 32, página 70). El estudio de predicción de cambios en secuencias moduladoras del procesamiento del RNA (ESE) no mostró variaciones (tabla 29, página 68). Las predicciones *in silico* de patogenicidad muestran que todos los programas empleados en nuestro estudio la consideran probablemente patogénica (tabla 30, página 70). Los estudios funcionales llevados a cabo por otros grupos no son consistentes y, aunque estudios realizados en levaduras muestran una disminución en la eficiencia de la función de la proteína MLH1 (Takahashi y col. 2007), no se ha demostrado que la mutación afecte a la unión entre MLH1 y PMS2 (Perera y col., 2008). El estudio poblacional demostró que la mutación no está presente en población sana. Por otra parte, aunque Lui y colaboradores la han descrito asociada a una mutación intrónica que alteraba el procesamiento del RNA (Liu y col., 1999), nuestro paciente no presentaba otra mutación patogénica en genes reparadores del DNA. Finalmente, Castillejo y colaboradores, señalan que la variante no cosegrega con la enfermedad (Castillejo y col., 2011).

La variante de significado desconocido **p.P640L** supone un cambio de de una prolina por leucina en el dominio de dimerización de la proteína. Este residuo aparece muy conservado en todas las

especies analizadas. El programa ESE FINDER predice una alteración en la unión de la proteína SC35 durante la maduración del RNA mensajero, por lo que podría afectar al procesamiento; no obstante, un estudio llevado a cabo empleando minigenes, analizó el efecto *in vitro* de esta mutación sobre el procesamiento del RNA concluyendo que no modifica el procesamiento (Tournier y col., 2008). El análisis *in silico* de esta variante muestra que todos los programas de predicción utilizados en nuestro trabajo la consideran patogénica. El estudio poblacional mostró que no está presente en población sana y nuestros resultados muestran que el paciente portador de esta mutación no presenta otra mutación patogénica en los genes reparadores incluidos en nuestro estudio. Por todo ello consideramos que la mutación es probablemente patogénica aunque es necesario realizar estudios funcionales para confirmar nuestra hipótesis.

1.1.3. Variantes de significado desconocido en el extremo 3'UTR del gen MLH1

Las alteraciones en los extremos 3' UTR de los genes han sido consideradas de manera clásica como variantes neutras o polimorfismos al no localizarse dentro de la secuencia codificante o en las secuencias consenso implicadas en la maquinaria de maduración del RNA. Contrariamente a esta idea, investigaciones recientes demuestran que las secuencias UTR tienen su importancia en el mantenimiento de la estabilidad del mRNA y plantean una revisión sobre el efecto de estas mutaciones. En este sentido, la pérdida de tres nucleótidos CTT (**c.*32_34del**) en el extremo 3'UTR del gen MLH1 ha estado considerada como un polimorfismo durante mucho tiempo, en parte por presentarse en coexistencia con otras mutaciones consideradas patogénicas y, en parte, por su alta frecuencia de aparición en población sana. En 2008, Mao y colaboradores analizaron el efecto de esta alteración en un paciente con una recaída en leucemia y observaron un descenso en la expresión del mRNA. El mecanismo que desestabiliza al RNA transcrito todavía no es conocido. Sin embargo, el análisis de la estructura secundaria muestra la presencia de un gran bucle que, cuando se pierden los tres nucleótidos (CTT), reduce su tamaño y crea un bucle nuevo más pequeño. La diferencia de tamaño y el número de bucles podrían estar implicados en la estabilidad y/o la capacidad de unión del mRNA a otras proteínas (Mao y col., 2008).

En nuestro estudio la mutación c.*32_34del aparece asociada con una pérdida de los exones 4-6 en el gen MLH1. Aunque síndrome de Lynch es una enfermedad autosómica dominante, se han observado algunos casos de mutaciones bilalélicas en el gen MLH1 (Riccardone y col., 1999), constituyendo una entidad clínica denominada CoLoN, Síndrome de Lynch III o Síndrome de la deficiencia constitutiva del sistema reparador de errores (CR(M-D), que se caracteriza por la presencia de tumores en el colon, leucemias, linfomas y síntomas similares a la neurofibromatosis., (Bandipallin y col., 2005; Felton y col., 2007; Wimmer y Etzler, 2008). El paciente estudiado por nosotros portador de las dos

mutaciones no presenta ninguno de estos síntomas, lo que nos lleva a pensar que las dos mutaciones se encuentren en el mismo alelo, permaneciendo el segundo alelo activo (tabla 28, página 66).

En nuestro trabajo encontramos una gran variabilidad en el espectro de mutaciones patogénicas del gen MLH1 de las cuales dos han sido descritas previamente en población española. Los polimorfismos exónicos p.I209V, p.E234E y el polimorfismo intrónico c.1558+14G>A, así como la variante de significado desconocido p.K618A aparecen recurrentemente tanto en nuestra población de estudio como en otros estudios realizados en población española.

1.2. Análisis del gen MLH1 en Cáncer colorrectal esporádico con inmunohistoquímica negativa para MLH1

Como complemento a nuestro estudio hemos analizado el gen MLH1 en 12 pacientes con cáncer de colon esporádico cuyos tumores no expresaban la proteína MLH1, no encontrando mutaciones germinales en ninguno de los casos. La inestabilidad de microsatélites es una característica molecular importante de los tumores relacionados con el síndrome de Lynch; sin embargo, entre un 15 y un 40% de los tumores esporádicos muestran esta misma característica (Heinen y col., 2009). Además, una amplia proporción de tumores de cáncer colorrectal muestran ausencia de proteína MLH1 cuando se realiza la tinción del tejido tumoral con anticuerpos específicos para la proteína. En estos casos, el origen de la inestabilidad no está causada por mutaciones germinales en los genes reparadores, si no que es debida a mutaciones somáticas ocurridas en el tejido o por hipermetilación del promotor de MLH1.

Durante el análisis molecular del gen MLH1 en estos probandos encontramos en seis (el 50%) de ellos polimorfismo común, p.I219V. Esta mutación aparece frecuentemente descrita y está clasificada como polimorfismo; sin embargo, varios estudios de asociación indican que se asocia con un mayor riesgo de cáncer de colon, colitis ulcerosa o leucemias. Además, Campbell y colaboradores han observado una asociación significativa entre la presencia de la mutación, la dieta occidental y el riesgo de cáncer de colon esporádico (Campbell y col., 2008).

2. Estudio del gen MSH2

Hemos estudiado el gen MSH2 en 182 pacientes incluidos en el PCGJCYL y encontrado cuatro mutaciones patogénicas ya descritas previamente por otros grupos, de las que dos corresponden a mutaciones *nonsense* que incorporan un codón de parada y otras dos corresponden a pérdidas de nucleótidos que cambian la fase de lectura (*frameshift*) (tabla 31, página 71). La mayoría de las veces, el cambio en la fase de lectura se acompaña de la aparición prematura de un codón de parada y, al igual

que ocurre con las mutaciones *nonsense*, una señal de finalización de la traducción muy temprana producirá proteínas carentes de dominios funcionales e incapaces de ejercer su función. En todo caso, al igual que en las mutaciones *nonsense*, el efecto deletéreo de la mutación depende, en gran medida, de la localización del codón de parada prematuro.

Este estudio nos ha permitido caracterizar también 12 polimorfismos, dos de los cuales no han sido descritos anteriormente (tabla 33, página 73). Estos dos polimorfismos, c.645-39C>A y c.1277-12A>C, se encuentran alejados de las secuencias consenso de procesamiento y pensamos que no modifican el procesamiento del RNA. En el caso de las tres mutaciones silenciosas p.K113K, p.A328A y p.K579K, que ya habían sido descritas por otros grupos previamente, el análisis bioinformático con programas de predicción de efecto sobre secuencias ESE, muestra que todas las mutaciones modifican secuencias reguladoras de procesamiento del RNA (tabla 34, página 73), siendo necesarios estudios sobre RNA para determinar su posible efecto sobre la transcripción.

2.1. Variantes de significado desconocido en el gen MSH2

El estudio del gen MSH2 ha puesto de manifiesto la existencia de seis mutaciones de significado desconocido localizadas en diferentes regiones del gen, de las cuales tres no han sido descritas hasta el momento.

2.1.1. Variantes de significado localizadas en el dominio de unión al DNA.

La mutación **p.H46Q** se localiza en un dominio conservado de la proteína (Figura 38, página) en una zona cercana a los residuos 35 a 38, importantes en el reconocimiento del error por el heterodímero MutS. La mutación provoca un cambio no conservativo que podría alterar la estructura terciaria de la proteína como queda reflejado en el estudio *in silico*, en el que los programas POLYPHEN y SNP3D la consideran patogénica (tabla 37, página 80). El estudio de predicción de modificación de sitios ESE reveló que esta mutación afecta a los sitios de unión de la proteína SF2/ASF y SRp55 (Tabla 36, página 78). La mutación **p.N127S** se encuentra en una pequeña secuencia (KASPGNLSQFEDILFGN) conservada evolutivamente (figura 38, página 77). Aunque el cambio de aminoácido no es drástico, ya que se mantiene la polaridad, la diferencia de tamaño entre la asparagina y la serina sí podría tener un efecto sutil en la estructura terciaria de la proteína. Los análisis bioinformáticos realizados sugieren una posible patogenicidad (tabla 37, página 80). El estudio de su posible implicación en modificación del procesamiento del RNA, reveló que la mutación modifica un sitio de unión de la proteína SC35. Son necesarios estudios funcionales y de análisis de RNA para catalogar definitivamente estas mutaciones.

La mutación **p.E188Q** no ha sido descrita hasta el momento. La mutación se localiza en una región muy conservada filogenéticamente y el estudio poblacional mostró su ausencia en población control. Los análisis preliminares *in silico* muestran que es considerada patogénica por dos de los programas de predicción (tabla 37, página 80). Esta mutación cambia un aminoácido polar, alifático y con una carga negativa (ácido glutámico) por otro (glutamina) similar al anterior pero neutro dentro del dominio de unión al DNA de la proteína. Es posible que el cambio de carga en la secuencia implique un cambio en la eficiencia de interacción de este dominio con la molécula de DNA. La mutación c.562G>C (p.E188Q), además, provoca alteraciones en los sitios ESE: se añaden tres sitios ESE nuevos en una zona dónde no existían y se pierde un sitio SRp55 (figura 41, página 79).

La mutación **p.K249R**, provoca el cambio de una lisina por una arginina. Ambos son aminoácidos hidrofílicos cargados positivamente, de aproximadamente el mismo tamaño, por lo que la alteración sufrida por la proteína posiblemente sea nula, lo que concuerda con los datos que aportan los programas informáticos de predicción. La mutación se localiza en una zona de la proteína conservada entre especies que, aunque se localiza dentro del dominio de unión al DNA, no parece tener ninguna implicación en esta función (Obmolova y col., 2002). Los estudios bioinformáticos sobre su implicación en sitios ESE indican que esta mutación no se localiza en una secuencia ESE y por tanto no tiene ningún efecto sobre la maduración del RNA. Según estos datos, esta variante de significado desconocido podría ser considerada como polimorfismo, aunque se requieren estudios funcionales que analicen la eficiencia de reparación y de unión al DNA de esta variante.

Todas estas mutaciones se localizan en el dominio de unión a DNA dañado de la proteína MSH2. La capacidad de unión de los heterodímeros MutS α (MSH2-MSH6) y MutS β (MSH2-MSH3) al DNA desapareado puede ser medida por ensayos simples de co-precipitación entre el heterodímero y el DNA o con pruebas un poco más complejas de movilidad electroforética de DNA (*electrophoretic mobility shift assay* o EMSA).

2.1.2. Variantes de significado desconocido localizadas en el dominio de dimerización de la proteína MSH2

La variante de significado desconocido **p.Q377R** ha sido descrita previamente y el estudio bioinformático de predicción de estructura de la proteína sugiere que tiene un comportamiento benigno a pesar de estar próxima a una zona muy conservada de la proteína. El análisis de su posible implicación en sitios ESE muestra que altera un sitio de unión de las proteínas SF/ASF y SRp40, siendo necesarios más estudios para confirmar la benignidad de esta mutación.

La variante de significado desconocido **p.P616R** no ha sido descrita hasta el momento. Esta mutación se localiza en una secuencia de aproximadamente 10 aminoácidos (SNGAPPVPYVRPAILEK) muy conservada entre especies y con varios residuos aminoacídicos de importancia para la función. El estudio poblacional no permitió encontrarla en población sana. El cambio de una prolina por una arginina supone la adición de una carga positiva nueva a la secuencia proteína. Estos cambios son valorados como perjudiciales para la proteína por el programa informático PolyPhen, mientras que SIFT y SNPD3 lo consideran tolerable. El estudio de su posible implicación en la regulación de la maduración del RNA a través de sitios ESE mostró que la mutación altera un sitio reconocido por la proteína SRp55. No obstante, con los datos de los que disponemos no podemos clasificar la mutación de manera definitiva.

En población española se han identificado una gran variedad de mutaciones asociadas al desarrollo del síndrome de Lynch y, al igual que ocurren en MLH1, existe una gran variabilidad en su espectro mutacional. Dos de las mutaciones patogénicas que hemos encontradas en este trabajo aparecen descritas por Caldés en estudios realizados sobre población madrileña (Caldés y col., 2002). Ninguna de las variantes de significado desconocido localizadas en el gen MSH2 durante nuestro trabajo aparece por primera vez identificadas en población española.

3. Estudio del gen MSH6

El gen MSH6 se estudió en 35 pacientes encontrando 15 mutaciones, de las que una es una mutación patogénica **c.1458_1459dupT**, descrita por primera vez en este trabajo, que provoca un desfase de lectura y un codón de parada prematuro. Las mutaciones en el gen MSH6 suelen asociarse con características fenotípicas propias como una edad de aparición más tardía y una mayor incidencia de tumores de endometrio (Plaschke y col., 2004); la paciente portadora de la mutación aquí descrita presentaba cáncer de endometrio y cáncer de páncreas a los 66 años. La incidencia de mutaciones en MSH6 es inferior a la detectada en MLH1 y MSH2. No obstante, Sjursen y sus colaboradores han sugerido que la detección de mutaciones en el gen MSH6 puede estar infravalorada y que la selección de familias exclusivamente a través de los criterios de Amsterdam y Bethesda puede ser insuficiente y se pueden perder familias portadoras, recomendando realizar siempre un análisis de inestabilidad de microsatélites y/o inmunohistoquímica para no perder familias portadoras de mutación en MSH6 (Sjursen y col.; 2010).

Entre las otras 14 mutaciones hemos caracterizado 9 mutaciones intrónicas, una de ellas IVS6 - 16delT, no descrita hasta el momento (tabla 38, página 82). Estas mutaciones se encuentran alejadas de las zonas de procesamiento canónico del RNA y pensamos que no deben de influir en este proceso. Las 5 mutaciones silenciosas son consecuencia de transiciones en el DNA y el estudio bioinformático de su posible implicación en el procesamiento del RNA a través de sitios ESE mostró que tres de ellas, **c.186G>A**,

c.276A>G, y **c.540T>C** alteran las secuencias de varios sitios ESE y podrían afectar a la maduración del RNA mensajero (tabla 39, página 83). En los tres casos son necesarios estudios sobre RNA que corroboren estos resultados preliminares.

Hay pocos estudios realizados sobre mutaciones en el gen MSH6 en población española y en ellos se observa baja frecuencia de mutaciones patogénicas que suelen ser de nueva descripción. El gen MSH6, sin embargo, presenta una alta frecuencia de polimorfismos silenciosos e intrónicos descritos tanto en la población mundial como en la población española. En nuestro trabajo, seis de los catorce polimorfismos identificados son de nueva descripción en población española.

4. Estudio de grandes reordenamientos genómicos

El estudio de grandes reordenamientos en 75 probandos pertenecientes a familias incluidas en el PCGJCYL nos ha permitido caracterizar dos casos portadores de deleciones en el gen MLH1 y dos casos deleciones en MSH2. (tabla 40, página 85). Aunque trabajos previos han descrito una incidencia superior a la encontrada en nuestra serie (van der Klift y col., 2005; di Fiore y col., 2004; van der Klift y col., 2005), nuestros resultados son similares a los descritos en población española (Castellví-Bell y col., 2005). Esta amplia heterogeneidad en la incidencia de los grandes reordenamientos genómicos parece tener que ver con la alta heterogeneidad en los criterios usados para la selección de probandos y con los diferentes métodos empleados para su detección (di Fiore y col., 2004).

Los mecanismos que provocan los reordenamientos reorganizaciones parecen estar mediados por recombinaciones homólogas a través de secuencias Alu. En los estudios que se han realizado, el gen MSH2 es el que tiene mayor número de secuencias Alu y por eso se considera que es en el que aparecen con mayor frecuencia este tipo de alteraciones (van der Klift y col., 2005). Las dos familias portadoras de deleciones en el gen MSH2 tienen una gran diferencia en la manifestación de la enfermedad: así, el gran reordenamiento genómico identificado en el exón 1 gen MSH2 en nuestra serie se presenta en una familia con una alta agregación de tumores uroepiteliales. Este tipo de tumores extracolónicos son característicos en el síndrome de Lynch (Watson y col., 2008). Diversos estudios indican que el riesgo de cánceres urinarios está asociado a alteraciones en el gen MSH2 (van der Post y col., 2010). Por otro lado la familia portadora de una deleción en el exon 6 del gen MSH2 presenta un gran número de tumores colorrectales.

La técnica del MLPA es una herramienta muy útil para evaluar simultáneamente toda la región codificante del gen en busca de grandes pérdidas o amplificaciones. Debido a la complejidad en el desarrollo del ensayo de MLPA, es muy sensible al método de extracción, a la calidad y concentración del DNA y por tanto, su reproducibilidad es bastante baja. Dada la importancia de un buen diagnóstico, la

técnica de MLPA requiere ser confirmada. La PCR cuantitativa es una herramienta alternativa útil a la hora de confirmar, de manera específica, las alteraciones identificadas en el ensayo de MLPA debido a su sencillez, rapidez y especificidad. Con la ayuda de la PCR cuantitativa no sólo pudimos confirmar las tres alteraciones identificadas en este trabajo, sino que nos permitió valorar como normales siete probandos para los cuales el resultado del MLPA era dudoso (tabla 42, página 86).

5. Estudio bioinformático de miRNAs asociados a genes implicados en síndrome de Lynch.

La búsqueda bioinformática de microRNAs con capacidad de unión a alguno de los genes del sistema reparador de errores ha mostrado 42 miRNAs candidatos (tablas 46, 47, y 48, páginas 94 y 95). Entre los microRNAs identificados, encontramos 12 cuya expresión se describió alterada en tumores colorrectales (Slaby y col., 2009) (Tabla 55). La asociación de estos microRNAs con alguno de los genes reparadores del sistema MMR se ha descrito en muy pocos casos, siendo miR-21 el más descrito. miR-21 es un microRNA cuya función más estudiada y mejor descrita es la interacción con Bcl2; la sobre-expresión de este microRNA inhibe la función de Bcl2 asociándose a un aumento en la proliferación celular y el aumento de la vascularización (Schetter y col., 2007). Xi y colaboradores han identificado al gen MLH1 como una posible diana de este microRNA (Xi y col., 2008). miR-21, aparece, además, implicado en procesos de respuesta del tumor al tratamiento, habiéndose descrito que la sobre-expresión de miR-21 en líneas celulares de cáncer colorrectal produce resistencia al tratamiento con 5-fluorouracilo (5-FU) debido a que el microRNA regularía la expresión de MSH2 (Valeri y col., 2010).

Tabla 55: microRNAs con implicación en el cáncer.

miRNA	Gen diana	Función	Bibliografía
hsa-miR-21	Bcl2	La sobre-expresión de este microRNA inhibe la función de Bcl2 produciendo un aumento en la proliferación celular y el aumento de la vascularización.	Schetter y col., 2007
	MLH1	Asociado a proliferación celular y apoptosis en líneas celulares de cáncer de colon (HTC116)	Schetter y col., 2007
	MLH1	Aparece sobre-expresado en líneas celulares de cáncer de colon (HT29 y HTC116) cuando se tratan con 5-FU.	Schetter y col., 2007
hsa-miR-30a (familia)	MSH2, MLH1	Aparece sobre-expresado y asociado a proliferación celular y apoptosis	Schetter y col., 2007, y Xi y col., 2007
hsa-miR-372	MLH1, AKT1	Aparece sobre-expresado y asociado a proliferación celular y apoptosis	Schetter y col., 2007
	MLH1, MSH2	Aparece sobre-expresado en líneas celulares de cáncer de colon (HT29 y HTC116) cuando se tratan con 5-FU.	Schetter y col., 2007
hsa-miR-30c	MLH1	Aparece inhibido y asociado a proliferación celular y apoptosis	Schetter y col., 2007
hsa-miR-30e	MLH1	Aparece sobre-expresado en líneas celulares de cáncer de colon (HT29 y HTC116) cuando se tratan con 5-FU.	Schetter y col., 2007
hsa-miR-152	/	Aparece sobre-expresado en líneas celulares de cáncer de colon (HT29 y HTC116) cuando se tratan con 5-FU.	Schetter y col., 2007
hsa-miR-148a	Bcl2	Actúa directamente sobre Bcl2 e induce apoptosis en tumores colorrectales	Zhang y col., 2011
hsa-miR-148b	/	Inhibe la proliferación celular. Su expresión es baja en cáncer gástrico.	Zhang y col., 2011
hsa-miR-374a	/	Aparece sobre-expresado en cáncer colorrectal	

Otros 30 microRNAs con capacidad de interactuar con genes del sistema MMR no aparecen asociados al cáncer colorrectal en la literatura, ni se ha estudiado su expresión en este tipo de tumor. Son necesarios análisis de expresión y estudios funcionales que corroboren la presencia real de estos microRNAs en el tejido tumoral así como sus posibles funciones y, a su vez, descarten los falsos positivos. Esta lista de microRNAs abren la vía al estudio de sus funciones en busca de datos de los procesos tumorigénicos relacionados con el síndrome de Lynch, permitan la obtención de datos diagnósticos, pronósticos o de respuesta al tratamiento.

Cabe la posibilidad de que nuestro análisis bioinformático no haya detectado microRNAs que interactúan con los genes del sistema reparador, como es el caso del miR-155. Durante el desarrollo de este trabajo, Valeri y colaboradores, identificaron que miR-155 aparece sobre-expresado en tumores colorrectales con alta inestabilidad de microsatélites e interactúa con los genes MLH1, MSH2 y MSH6. miR-155 parece inhibir de manera directa la expresión de las proteínas MLH1, MSH2 y MSH6. La estabilidad de las proteínas implicadas en este proceso de debe a su capacidad de formar dímeros y por lo tanto la pérdida de la expresión de MLH1 o de MSH2 mediante de la regulación de miR-155 impediría la formación de estos heterodímeros y la pérdida de la capacidad de reparación del DNA (Valeri y col., 2010).

La presencia de mutaciones en las secuencia de unión de los microRNA o de la secuencia 3'UTR puede ser la causante de su desregulación y, en consecuencia, tener implicaciones en el desarrollo de cancer colorrectal o en la respuesta al tratamiento. En este trabajo hemos identificado 17 SNPs en los microRNAs y en los extremos 3' UTR de los genes reparadores. Así, hemos identificado un SNP (rs 2234891) en el extremo 3' no codificante del gen MLH1 que se sitúa en la zona de unión entre con miR-374a, localizado en el cromosoma X y del cual se ha descrito su sobre-expresión en cáncer colorrectal (Wang y col., 2010). El SNP consiste en la delección de cuatro nucleótidos (-/GATT) localizados en la secuencia diana del microRNA (Figura 55). Esta pequeña delección coincide con los tres últimos nucleótidos de la secuencia "semilla" (seed site) del microRNA. Esta secuencia que se sitúa del nucleótido 2 al 7 del extremo 5' del microRNA es la determinante para la especificidad y función del mismo (Bartel, 2009).

Extremo 3'UTR de MLH1:

```

5'GGTGTtaaATATGGTTATTTATGCACTGTGGGATGTGTTCTTCTCTCTGTATTCCGATACAAAGTGTGTATCAAAGTGTGATATACAAAGTG
STOP
TACCAACATAAGTGTGGTAGCACTTAAGACTTATACTGCCTTCTGATAGTATTCTTTATACAGTGGATTgattATAAA 3'
                                     rs 2234891
miR-374a: 3'GUGAAUAGUCCAACA—UAAUUAUU 5'
Nucleótido:                               7     2

```

Figura 55: alineamiento del extremo 3'UTR del gen MLH1 con miR-374a. En rojo aparece señalado el codon de parada, en azul se muestran las secuencias complementarias tanto del miR-374a como del extremo 3'UTR del gen. En minúsculas y subrayado aparece el SNP rs 223489. (Del extremo 3'UTR mostramos la secuencia genómica parcial).

En la actualidad los estudios sobre microRNA y sus alteraciones en cáncer son una nueva vía de estudio que amplía el conocimiento sobre los procesos de desarrollo de los tumores. Además, en el síndrome de Lynch la identificación de alteraciones en la regulación de la expresión de los genes implicados en el sistema reparador de errores post-replicativos resulta de gran valía para el diagnóstico de la enfermedad.

En resumen, en el presente trabajo hemos realizado el estudio los genes MLH1, MSH2 y MSH6 en busca de mutaciones en aquellos 188 pacientes con sospecha de síndrome de Lynch incluidos en el Programa de Consejo Genético de la Junta de Castilla y León. La comparación durante este trabajo de la efectividad entre los criterios utilizados establecidos por el PCGJCYL y los criterios internacionales de selección (Amsterdam y Bethesda), nos ha permitido comprobar que los criterios internacionales son lo suficientemente robustos por si mismos, pero que permiten que una baja proporción de familias portadoras sin criterios queden excluidas del estudio.

El pequeño porcentaje de pacientes que presentaban mutaciones patogénicas nos obligó a realizar un estudio en mayor profundidad de los efectos teóricos sobre la proteína de la alta proporción de variantes de significado desconocido que hemos identificado en este trabajo; así como realizar un estudio preliminar de los posibles microRNAs con capacidad de interacción sobre los genes MMR con el fin de aportar información útil que nos permitiera mejorar la efectividad del cribado.

Conclusiones

Conclusiones

1. Los criterios de Amsterdam y Bethesda son lo suficientemente robustos como para identificar la mayoría de las familias con síndrome de Lynch, pero existe un pequeño porcentaje que se benefician de criterios de inclusión más relajados como los del Programa de Consejo Genético de la Junta de Castilla y León.
2. El alelo G del polimorfismo c.-93A>G MLH1, es más frecuente en pacientes con cáncer de colon esporádico que en pacientes con síndrome de Lynch y en individuos sanos, lo que sugiere que esta variante genética que facilita la metilación de esta región promotora predispone al desarrollo de cáncer esporádico.
6. El análisis bioinformático realizado sobre miRNAs realizado en este trabajo, muestra la presencia un SNP en el gen MLH1 que altera la secuencia de unión del microRNA miR-374a. Estos datos teóricos abren una nueva vía de estudio y diagnóstico del síndrome de Lynch.

Bibliografía

- Aaltonen LA, Peltomäki P, Leach FS, Sistonen P, Pylkkänen L, Mecklin JP, Järvinen H, Powell SM, Jen J, Hamilton SR (1993) Clues to the pathogenesis of familial colorectal cancer. *Science*; 260(5109):812-6.
- Aaltonen LA., (2000) Hereditary intestinal cancer. *Semin Cancer Biol.*; 10(4): 289-98.
- Alani E. (1996) The *Saccharomyces cerevisiae* Msh2 and Msh6 proteins form a complex that specifically binds to duplex oligonucleotides containing mismatched DNA base pairs. *Mol. Cell. Biol.*; 17:2436-47.
- Ambros V., Bartel B., Bartel D. P., Burge C. B., Carrington J., Chen X., Dreyfuss G., Eddy S. R., Giffiths-Jones S., Marshall M., Matzke M., Ruvkun G. y Tuschl T. (2003) A uniform system for microRNA annotation. *RNA*; 9(3): 277-279.
- Apeessos A, Mihalatos M, Danielidis I, Kallimanis G, Agnantis NJ, Triantafillidis JK, Fountzilias G, Kosmidis PA, Razis E, Georgoulas VA, Nasioulas G. (2005) hMSH2 is the most commonly mutated MMR gene in a cohort of Greek HNPCC patients. *Br J Cancer.*;92(2):396-404.
- Auclair J., Buisine M.P., Navarro C., Ruano E., Montmain G., Desseigne F., Saurin J.C, Lasset C., Bonadona V., Giraud S., Puisieux A., Qing Wang. (2006) Systematic mRNA Analysis for the Effect of MLH1 and MSH2 Missense and Silent Mutations on Aberrant Splicing. *Human Mutation*; 27:145-154.
- Barnetson R.A., Cartwright N., van Vliet A., Haq N., Drew K., Farrington S., Williams N., Warner J., Campbell H., Porteous M.E., Dunlop M.G. (2008) Classification of Ambiguous Mutations in DNA Mismatch Repair Genes Identified in a Population-Based Study of Colorectal Cancer. *Human Mutation*; 29:367-374.
- Bartel D.P., (2009) MicroRNAs: Target Recognition and Regulatory Functions. *Cell*; 136:215-233.
- Bartel D.P.,(2004) MicroRNAs: Genomics, Biogenesis, Mechanism and Function. *Cell*, 116:281-297.
- Bellacosa A, Genuardi M, Anti M, Viel A, Ponz de Leon M. (1996) Hereditary nonpolyposis colorectal cancer: review of clinical, molecular genetics, and counseling aspects. *Am. J. Med. Genet.*; 24:353-364.

- Benatti P, Sassatelli R, Roncucci L, Pedroni M, Fante R, Di Gregorio C, Losi L, Gelmini R, Ponz de Leon M. (1993) Tumour spectrum in hereditary non-polyposis colorectal cancer (HNPCC) and in families with "suspected HNPCC". A population-based study in northern Italy. Colorectal Cancer Study Group. *Int. J. Cancer*; 28:371–377.
- Bisgaard ML, Fenger K, Vulgo S, Niebuhr E, Mohr J (1994) Familial adenomatous polyposis (FAP): frequency, penetrance and mutation rate. *Hum. Mutat.*; 3:121-125.
- Blasi M.F., Ventura I, Aquilina G., Degan P., Bertario L., Bassi C., Radice P., Bignami M. (2006) A Human Cell-Based Assay to Evaluate the Effects of Alterations in the MLH1 Mismatch Repair Gene. *Cancer. Res.*; 66:9036-9044.
- Boni V, Zarate R, Villa JC, Bandre's E, Gomez MA, Maiello E, Garcia-Foncillas J, Aranda E (2010) Role of primary miRNA polymorphic variants in metastatic colon cancer patients treated with 5-fluorouracil and irinotecan. *Pharmacogenomics J*, 2010: 1-8.
- Børresen AL, Lothe RA, Meling GI, Lystad S, Morrison P, Lipford J, Kane MF, Rognum TO, Kolodner RD. (1995) Somatic mutations in the hMSH2 gene in microsatellite unstable colorectal carcinomas. *Hum Mol Genet.*;4(11):2065-72.
- Budhu A., Ji J., Wang X.W., (2010) The clinical potential of microRNAs. *Journal of Hematology & Oncology*; 3:37-44
- Buerstedde JM, Alday P, Torhorst J, Weber W, Müller H, Scott R. (1995) Detection of new mutations in six out of 10 Swiss HNPCC families by genomic sequencing of the hMSH2 and hMLH1 genes. *J Med Genet.*;32:909-912.
- Bujalkova M, Zavodna K, Krivulcik T, Ilencikova D, Wolf B, Kovac M, Karner-Hanusch J, Heinimann K, Marra G, Jiricny J, Bartosova Z. (2008) Multiplex SNaPshot genotyping for detecting loss of heterozygosity in the mismatch-repair genes MLH1 and MSH2 in microsatellite-unstable tumors. *Clin Chem. Nov*; 54(11):1844-54.
- Caldés T, Godino J, Sanchez A, Corbacho C, De la Hoya M, Lopez Asenjo J, Saez C, Sanz J, Benito M, Ramon Y Cajal S, Diaz-Rubio E. (2004) Immunohistochemistry and microsatellite

- instability testing for selecting MLH1, MSH2 and MSH6 mutation carriers in hereditary non-polyposis colorectal cancer. *Oncol. Rep.*; 12:621-629.
- Caldés T., Godino J., De la Hoya M., García Carbonero I., Pérez Segura P., Eng Ch., Benito M., Díaz-Rubio E. (2002) Prevalence of germline mutations of MLH1 and MSH2 in hereditary nonpolyposis colorectal cancer families from Spain. *Int. J. Cancer*; 98: 774–779.
 - Campbell PT, Curtin K, Ulrich CM, Samowitz WS, Bigler J, Velicer CM, Caan B, Potter JD, Slattery ML. (2009) Mismatch repair polymorphisms and risk of colon cancer, tumour microsatellite instability and interactions with lifestyle factors. *Gut*.;58(5):661-7.
 - Capellá G, Pera G, Sala N, Agudo A, Rico F, Del Giudicce G, Plebani M, Palli D, Boeing H, Bueno-de-Mesquita HB, Carneiro F, Berrino F, Vineis P, Tumino R, Panico S, Berglund G, Simán H, Nyrén O, Hallmans G, Martínez C, Dorronsoro M, Barricarte A, Navarro C, Quirós JR, Allen N, Key T, Bingham S, Caldas C, Linseisen J, Nagel G, Overvad K, Tjønneland A, Boshuizen HC, Peeters PH, Numans ME, Clavel-Chapelon F, Trichopoulou A, Lund E, Jenab M, Kaaks R, Riboli E, González CA. (2008) DNA repair polymorphisms and the risk of stomach adenocarcinoma and severe chronic gastritis in the EPIC-EURGAST study. *Int J Epidemiol.* 37(6):1316-25.
 - Cartegni L., Wang J., Zhu Z., Zhang M. Q., Krainer A.R., (2003) ESE finder: a web resource to identify exonic splicing enhancers. *Nucleic Acids Res.*; 31:3568-3571.
 - Castellví-Bel S., Castells A., Strunk M., Ferrández A., Piazuelo E., Milà M., Piñol V., Rodríguez-Moranta F., Andreu M., Lanás A., Piqué J.M. (2005) Genomic rearrangements in MSH2 and MLH1 are rare mutational events in Spanih patients with hereditary nonpolyposis colorectal cancer. *Cancer Letters*; 225:93-98.
 - Castillejo A., Guarinos C., Martínez-Canto A., Barbera V.M., Egoavil C., Castillejo M.I., Pérez-Carbonell L., Sanches-Heras A.B., Segura A., Ochoa E., Lázaro R., Ruiz-Ponte C., Bufandas L., Andreus M., Castells A., Cariacedo A., Llor X., Clófent J., Alenda C., Paya A., Jover R., Soto J.L. (2011) Evidence for classification of c.1852_1853AA>GC in MLH1 as a neutral variant for Lynch syndrome. *BMC Medical Genetics*, 12:12.
 - Cederquist K, Emanuelsson M, Göransson I, Holinski-Feder E, Müller-Koch Y, Golovleva I, Grönberg H. (2004) Mutation analysis of the MLH1, MSH2 and MSH6 genes in patients with double primary cancers of the colorectum and the endometrium: a population-based study in northern Sweden. *Int J Cancer*.109(3):370-6.

- Chao A., Thun M.J., Connell C.J., McCullough M.L., Jacobs E.J., Flanders W.D., et al., (2005) Meat consumption and risk of colorectal cancer, *JAMA*; 293:172–182.
- Charames GS, Millar AL, Pal T, Narod S, Bapat B. (2000) Do MSH6 mutations contribute to double primary cancers of the colorectum and endometrium? *Hum Genet.*;107(6):623-9.
- Chen-Shtoyerman R., Theodor L., Harmati E., Friedman E., Dacka S., Kopelman Y., Sternberg A., Zarivach R., Bar-Meir S. and Fireman Z. (2003) Genetic Analysis of Familial Colorectal Cancer in Israeli Arabs. *Human Mutation. Mutation in Brief #594 Online.*
- Chow E, Macrae F. (2008) A review of juvenile polyposis syndrome. *J Gastroenterol Hepatol.*; 20:1634-1640.
- Cunningham JM, Kim CY, Christensen ER, Tester DJ, Parc Y, Burgart LJ, Halling KC, McDonnell SK, Schaid DJ, Walsh Vockley C, Kubly V, Nelson H, Michels VV, Thibodeau SN. (2001) The frequency of hereditary defective mismatch repair in a prospective series of unselected colorectal carcinomas. *Am J Hum Genet.*;69(4):780-90.
- Curia MC, Palmirotta R, Aceto G, Messerini L, Veri MC, Crognale S, Valanzano R, Ficari F, Fracasso P, Stigliano V, Tonelli F, Casale V, Guadagni F, Battista P, Mariani-Costantini R, Cama A. (1999) Unbalanced germ-line expression of hMLH1 and hMSH2 alleles in hereditary nonpolyposis colorectal cancer. *Cancer Res.*; 59:3570-3575.
- de la Rosa M, Fasano C, Panariello L, Scarano MI, Belli G, Iannelli A, Ciciliano F, Izzo P (2000) Evidence for a recessive inheritance of Turcot's syndrome caused by compound heterozygous mutations within the PMS2 gene. *Oncogene*; 19:1719-1723.
- Desai DC, Lockman JC, Chadwick RB, Gao X, Percesepe A, Evans DG, Miyaki M, Yuen ST, Radice P, Maher ER, Wright FA, de La Chapelle A. (2000) Recurrent germline mutation in MSH2 arises frequently de novo. *J. Med. Genet.*; 37:646-652.
- Devlin LA, Graham CA, Price JH, Morrison PJ. (2008) Germline MSH6 mutations are more prevalent in endometrial cancer patient cohorts than hereditary non polyposis colorectal cancer cohorts. *Ulster Med J.*; 77(1):25-30.
- Devlin LA, Price JH, Morrison PJ. (2005) Hereditary non-polyposis colon cancer. *Ulster Med J.*;74(1):14-21.

- Dherin C, Gueneau E, Francin M, Nunez M, Miron S, Liberti SE, Rasmussen LJ, Zinn-Justin S, Gilquin B, Charbonnier JB, Boiteux S. 2008 Characterization of a highly conserved binding site of Mlh1 required for exonuclease I-dependent mismatch repair. *Mol. Cell Biol.*; 29:907-918.
- Dherin C., Gueneau E., Francin M., Nunez M., Miron S., Liberti S.E., Rasmussen L.J., Zinn-Justin S., Gilquin B., Charbonnier J.-B., Boiteux S. (2009) Characterization of a highly conserved binding site of MLH1 required for Exonuclease I-Dependent Mismatch repair. *Mol. Cell. Biol.*; 29:907-918.
- Dong Y., Wu W.K.K., Wu C. W., Sung J. J. Y., Yu J., Ng S.S.M. (2011) MicroRNA dysregulation in colorectal cancer: a clinical perspective. *British Journal of Cancer* ; 104:893–898.
- Douglas J., Gruber S., Meister K.A., Bonner J., Watson P., Krush A. J., Lynch H.T. (2005) History and molecular genetics of syndrome in family G. *JAMA*, 294(17): 2195-2202.
- Dufner P, Marra G, Raschle M, Jiricny J. (2000) Mismatch recognition and DNA-dependent stimulation of the ATPase activity of hMutSalph α is abolished by a single mutation in the hMSH6 subunit. *J. Biol. Chem.*; 275:36550–36555.
- Botteri E., Iodice S., Bagnardi V., Raimondi S., Lowenfels A.B., Maisonneuve P.,(2008) Smoking and colorectal cancer: a meta-analysis, *JAMA*; 300:2765–2778.
- Paskett E.D., Reeves K.W., Pineau B., Albert P.S., Caan B., Hasson M., et al., (2005) The association between cigarette smoking and colorectal polyp recurrence (United States), *Cancer Causes Control* ;6:1021–1033.
- Edelbrock MA, Kaliyaperumal S, Williams KJ. (2009) DNA mismatch repair efficiency and fidelity are elevated during DNA synthesis in human cells. *Mutat. Res.*; 9:59-66.
- Ellison A.R., Lofing J., Bitter G.A. (2004) Human MutL homolog (MLH1) function in DNA mismatch repair: a prospective screen for missense mutations in the ATPase domain *Nucleic Acids Research*; 32:5321–5338.
- Faber C, Kirchner T, Hlubek F (2009) The impact of microRNAs on colorectal cancer. *Virchows Arch*; 454:359-367.

- Fearnhead N.S., Wilding J.L., Winney B., Tonks S., Bartlett S., Bicknell D.C., Tomlinson I.P.M., McC. Mortensen N.J., and Bodmer W.F. (2004) Multiple rare variants in different genes account for multifactorial inherited susceptibility to colorectal adenomas. *PNAS*; 101: 15992–15997.
- Fidalgo P., Almeida M.R., West S., Gaspar C., Maia L., Wijnen L., Albuquerque C., Curtis A., Cravo M, Fodde R., Leitaó C.N., Burn J. (2000) Detection of mutations in mismatch repair genes in Portuguese families with hereditary non-polyposis colorectal cancer (HNPCC) by a multi-method approach. *European Journal of Human Genetics*; 8:49–53.
- Fredriksson H, Ikonen T, Autio V, Matikainen MP, Helin HJ, Tammela TL, Koivisto PA, Schleutker J. (2006) Identification of germline MLH1 alterations in familial prostate cancer. *Eur. J. Cancer*; 42:2802-2806.
- Fukui K. (2010) DNA mismatch Repair in Eukaryotes and Bacteria. *J Nucleic Acids.*; 2010: 1-16.
- Furihata M, Takeuchi T, Ohtsuki Y, Terao N, Kuwahara M, Shuin T. (2001) Genetic analysis of hMLH1 in transitional cell carcinoma of the urinary tract: promoter methylation or mutation. *J. Urol.*; 165:1760-1764.
- Ghimenti C., Tannergård P., Wahlberg S., Liu T., Giulianotti P.G, Mosca F., Fornaciari G., Bevilacqua G., Lindblom A., Caligo M.A. (1999) Microsatellite instability and mismatch repair gene inactivation in sporadic pancreatic and colon tumours. *British Journal of Cancer*; 80:11–16.
- Gómez A, Salguero G, García H, Aristizábal F, Gutiérrez O, Angel LA, Padrón J, Martínez C, Martínez H, Malaver O, Barvo R, Giraldo A. (2005) Detection mutations in the DNA mismatch repair genes of hMLH1 and hMSH2 genes in Colombian families with suspicion of hereditary non-polyposis colorectal carcinoma (Lynch syndrome) *Biomedica.*; 25:315-324.
- Grabowski M., Mueller-Koch Y., Grabson-Frodl E., Koehler U., Keller G., Vogelsang H., Dietmaier W., Kopp R., Siebers U., Schmitt W., Neitzel B., Gruber M., Doerner C., Kerker B., Ruemele P., Henke G., Holinski-Feder E. (2005) Deletions account for 17% of pathogenic germline alterations in MLH1 and MSH2 in Hereditary nonpolyposis Colorectal Cancer (HNPCC) families. *Genetic Testing*; 9:138-146.
- Gradia S, Subramanian D, Wilson T, Acharya S, Makhov A. (1999) hMSH2-hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA. *Mol. Cell*; 3:255–261.

- Gryfe R. (2009) Inherited Colorectal Cancer Syndromes. Clinics in colon and rectal cancer surgery; 22:4.
- Hampel H., Frankel W., Martin E., Arnold M., Khanduja K., Kuebler P., Nakagawa H., Sotamaa K., Prior T, Westman J., Panescu J., Fix D., Lockman J., Comeras I., de la Chapelle A (2005) Screening for the Lynch Syndrome (Hereditary Nonpolyposis Colorectal Cancer) N. Engl. J. Med.; 352:1851-1860.
- Hegde M., Blazo M., Chong B., Prior T., Richards C. (2006) Assay Validation for Identification of Hereditary Nonpolyposis Colon Cancer-Causing Mutations in Mismatch Repair Genes MLH1, MSH2, and MSH6. Journal of Molecular Diagnostics; 7:525-534.
- Hemminki A, Peltomaki P, Mecklin JP, Jarvinen H, Salovaara R, Nystrom-Lathi M, la Chapelle A, Aaltonen L. (1994) Loss of wild type MLH1 gene is a feature of hereditary nonpolyposis colorectal cancer. Nature Genetics; 8:405-410.
- Hienonen T, Laiho P, Salovaara R, Mecklin JP, Jarvinen H, Sistonen P, Peltomaki P, Lehtonen R, Nupponen NN, Launonen V, Karthu A, Aaltonen LA. (2003) Little evidence for involvement of MLH3 in colorectal cancer predisposition. Int. J. Cancer; 106:292-296.
- Hizawa K, Iida M, Yao T, Aoyagy K, Fujishima M. (1997) Juvenile polyposis of the stomach: clinicopathological features and its malignant potential. Journal of Clinical Pathology; 50:771-774.
- Hsieh P., (2001) Molecular mechanisms of DNA mismatch repair. Mutat Res.; 486(2):71-87.
- Hudler P, Vouk K, Liovic M, Repse S, Juvan R, Komel R. (2004) Mutations in the hMLH1 gene in Slovenian patients with gastric carcinoma. Clin Genet.; 65:405-411.
- Iyer RR, Pluciennick A, Burdett V, Modrich P. (2006) DNA Mismatch Repair: Functions and Mechanisms. Chem. Rev.; 106:302-323.
- Jäger A.C., Rasmussen M., Bisgaard H.C., Singh K.K., Nielsen F.C., Rasmussen L.J. (2001) HNPCC mutations in the human DNA mismatch repair gene hMLH1 influence assembly of hMutLa and hMLH1± hEXO1 complexes. Oncogene; 20:3590-3595.
- Jemal A., Siegel R., Xu J., Ward E. (2010) Cancer Statistics. Cancer J Clin; 60:277-300.

- Tamakoshi K., Wakai K., Kojima M., Watanabe Y., Hayakawa N., Toyoshima H., et al., (2004) A prospective study of body size and colon cancer mortality in Japan: The JACC Study, *Int. J. Obes. Relat. Metab. isord.*; 28:551–558.
- Khan N., Afaq F., Mukhtar H. (2009) Lifestyle as risk factor for cancer: Evidence for human studies. *Cancer Letters*; 293:133-143.
- Klaus K, Herfarth F, Ogunbiyi OA, Moley JF, Kodner IJ, Wells SA Jr. (1998) Four new mutations in the DNA mismatch repair gene MLH1 in colorectal cancers with microsatellite instability. *Hum Mutat.*; 12:73.
- Knudsen A, Bisgaard M, Vulgo S. (2003) Attenuated familial adenomatous polyposis (AFAP). A review of the literature. *Familial Cancer*; 2:43–55.
- Kolodner RD, Tytell JD, Schmeits JL, Kane MF, Gupta RD, Weger J, Wahlberg S, Fox EA, Peel D, Ziogas A, Garber JE, Syngal S, Anton-Culver H, Li FP. (1999) Germ-line msh6 mutations in colorectal cancer families. *Cancer Res.*;59(20):5068-74.
- Kunz C, Saito Y, Schär P. (2009) DNA Repair in mammalian cells: Mismatched repair: variations on a theme. *Cell. Mol. Life Sci*; 66:1021-1038.
- Thygesen L.C., Gronbaek M., Johansen C., Fuchs C.S., Willett W.C., Giovannucci E., (2008) Prospective weight change and colon cancer risk in male US health professionals, *Int. J. Cancer .*; 123:1160–1165.
- la Chapelle A., (2004) Genetic predisposition to Colorectal Cancer. *Nature Reviews*; 4:769-780.
- Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T (2001) Identification of novel genes coding for small expressed RNAs. *Science*; 294(5543):853-8.
- Lagos-Quintana M., Rauhut R., Lendeckel W., Tuschl T. (2001) Identification of novel genes coding for small expressed RNAs. *Science*; 294(5543):853-8.
- Lamberti C, Kruse R, Ruelfs C, Caspari R, Wang Y, Jungck M, Mathiak M, Malayeri HR, Friedl W, Sauerbruch T, Propping P. (1999) Microsatellite instability-a useful diagnostic tool to select patients at high risk for hereditary non-polyposis colorectal cancer: a study in different groups of patients with colorectal cancer. *Gut.*; 44:839-843.

- Lamers MH, Perrakis A, Enzlin JH, Winterwerp HH, de Wind N, Sixma TK. (2000) The crystal structure of DNA mismatch repair protein MutS binding to a G x T mismatch. *Nature*; 407:711-717.
- Landi D, Gemignani F, Naccarati A, Pardini B, Vodicka P, Vodickova L, Novotny J, Forsti A, Hemminki K, Canzian F, Landi S (2008) Polymorphisms within micro-RNA-binding sites and risk of sporadic colorectal cancer. *Carcinogenesis*; 29: 579–584.
- Lanza G., (2007) mRNA/microRNA gene expression profile in microsatellite unstable colorectal cancer. *Mol. Cancer.*; 6:54–65.
- Leach F.S., Nicolaidis N.C., Papadopoulos N., Liu B., Jen J., Parsons R., Peltomiiki P., Sistonen P., Aaltonen L.A., Nystriim-Lahti M., Guan X.-Y., Zhang Ji, Meltzer P.S., J-F Yu, Kao F.-T., Chen D.-J., Cerosaletti K.M., Fournier K., Todd S., Lewis T., Leach R., Naylor S.L., Weissenbach J., Mecklin J-P, Jlrvinen H., Petersen G.M., Hamilton S., Green J., Jass J., Watson P., Lynch H.T., Trent J.M., de la Chapelle A., Kinzler K.W., Vogelstein B. (1993) Mutations of a mutS Homolog in Hereditary Nonpolyposis Colorectal Cancer. *Cell*; 75:1215-1225.
- Lee HC, Kim JG, Chae YS, Sohn SK, Kang BW, Moon JH, Jeon SW, Lee MH, Lim KH, Park JY, Choi GS, Jun SH (2010) Prognostic impact of microRNA-related gene polymorphisms on survival of patients with colorectal cancer. *J. Cancer Res. Clin. Oncol.*; 136: 1073–1078.
- Lee SC, Guo JY, Lim R, Soo R, Koay E, Salto-Tellez M, Leong A, Goh BC. (2005) Clinical and molecular characteristics of hereditary non-polyposis colorectal cancer families in Southeast Asia. *Clin. Genet.*; 68:137-145.
- Lee SC, Guo JY, Lim R, Soo R, Koay E, Salto-Tellez M, Leong A, Goh BC. (2005) Clinical and molecular characteristics of hereditary non-polyposis colorectal cancer families in Southeast Asia. *Clin Genet.*;68(2):137-45.
- Lee, R.C., Feinbaum, R.L., and Ambros, V. 1993. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*; 75: 843–854.
- Leung S.Y., Chan T.L., Chung L.P., Chan A., Fan Y.W., Hung K.N., Kwong W.K., Ho J.W.C Yuen S.T. (1998) Microsatellite Instability and Mutation of DNA Mismatch Repair Genes in Gliomas. *American Journal of Pathology*; 153:1181-1188.

- Levene S., Scott G., Price P., Sanderson J., Evans H., Taylor C., Bass S., Lewis C. and Hodgson. (2003) Does the occurrence of certain rare cancers indicate an inherited cancer susceptibility? *Familial Cancer*; 2:15–25.
- Lindblom A, Tannergard P, Werelius B, Nordenskjold M (1993) Genetic mapping of a second locus predisposing to hereditary non-polyposis colon cancer. *Nat. Genet.*; 5:279-282.
- Liu B, Parsons R, Papadopoulos N, Nicolaides NC, Lynch HT, Watson P, Jass JR, Dunlop M, Wyllie A, Peltomäki P, de la Chapelle A, Hamilton SR, Vogelstein B, Kinzler KW. (1997) Analysis of mismatch repair genes in hereditary non-polyposis colorectal cancer patients. *Nat. Med.*; 2:169-174.
- Liu B., Nicolaides C., Markowitz S., Willson J., Parsons R.E., Jen J., Papadopoulos N., Peltomaki P., de la Chapelle A., Hamilton S.R., Kinzler K.W., Vogelstein B. (1995) Mismatch repair defects in sporadic colorectal cancers with microsatellite instability. *Natures Genetics*; 9:48-55.
- Liu B., Parsons R.E., Hamilton S.R., Petersen G.M., Lynch H.T., Watson P., Markowitz S., Willson J.K.V, Green J, de la Chapelle A., Kinzler K.W. and Vogelstein B. (1994) HSMH2 Mutations in Hereditary Nonpolyposis Colorectal Cancer Kindreds. *Cancer Research*; 54:4590-4594.
- Liu M, Chen H., (2010) The role of microRNAs in colorectal cancer. *J. Genet. Genomics*; 37:347–358.
- López-Abente. (2004) “La situación del cáncer en España”. *An. Sist. Sanit. Navar.*; 2:165-173.
- Lucci-Cordisco E, Rovella V, Carrara S, Percesepe A, Pedroni M, Bellacosa A, Caluseriu O, Forasarig M, Anti M, Neri G, Ponz de Leon M, Viel A, Genuardi M. (2001) Mutations of the 'minor' mismatch repair gene MSH6 in typical and atypical hereditary nonpolyposis colorectal cancer. *Fam Cancer.*;1(2):93-9.
- Lynch H.T., Smyrk T., McGinn T., Lanspa S., Cavalieri J., Lynch J., Slominski-Castor S., Cayouette M., Priluck I., Luce M. (1995) Attenuated familial adenomatous polyposis (AFAP) a phenotypically and genotypically distinctive variant of FAP. *Cancer*, 76, 12: 2427-2433.
- Lynch HT, Krush AJ (1971) The cancer family syndrome and cancer control. *Surg. Gynecol. Obstet.*; 132:247–250.

- Lynch HT, Lynch PM, Lanspa SJ, Snyder CL, Lynch JF, and Boland CR (2009) Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications. *Clin. Genet.*; 76(1): 1–18.
- Akhter M., Kuriyama S., Nakaya N., Shimazu T., Ohmori K., Nishino Y., et al., (2007) Alcohol consumption is associated with an increased risk of distal colon and rectal cancer in Japanese men: the Miyagi cohort study, *Eur. J. Cancer*; 43:383–390.
- Sarebo M., Skjelbred C.F., Breistein R., Lothe I.M., Hagen P.C., Bock G., et al., (2006) APC mutations and the risk of developing sporadic colorectal adenomas and carcinomas, *BMC Cancer*; 6:71.
- Woods M.O., Williams P., Careen A., Edwards L., Bartlett S., McLaughlin J., Younghusband H. B. (2007) A new variant database for mismatch repair genes associated with Lynch syndrome. *Hum. Mut.*; 28:669-673.
- Malkov VA, Biswas I, Camerini-Otero RD, Hsieh P. (1997) Photocross-linking of the NH₂-terminal region of Taq MutS protein to the major groove of a heteroduplex DNA. *J. Biol. Chem.*; 19:23811–23817.
- Mangold E., Pagenstecher C., Friedl W., Mathiak M., Buettner R., Engel C., Loeffler M., Holinski-Feder E., Müller-Koch M., Keller G., K. Schackert H., Krüger S., Goecke T., Moeslein G., Kloor M., Gebert J., Kunstmann E, Schulmann K, Rüschoff J. and Propping P. and the German HNPCC Consortium. (2005) Spectrum and frequencies of mutations in MSH2 and MLH1 identified in 1,721 german families suspected of hereditary nonpolyposis colorectal cancer. *Int. J. Cancer*; 116:692–702.
- Mao G., Pan X., Gu L. (2008). Evidence That a Mutation in the MLH1 3'-Untranslated Region Confers a Mutator phenotype and Mismatch Repair Deficiency in Patients with Relapsed Leukemia. *The Journal of Biological Chemistry*; 283:3211–3216.
- Marra G, Boland CR. (1995) Hereditary nonpolyposis colorectal cancer: the syndrome, the genes, and historical perspectives. *J Natl Cancer Inst*; 87:1114-1125.
- Marsischky GT, Kolodner RD. (1999) Biochemical characterization of the interaction between the *Saccharomyces cerevisiae* MSH2-MSH6 complex and mispaired bases in DNA. *J. Biol. Chem.*; 274:26668–26682.

- Marti T.M., Kunz C., Fleck O., (2002) DNA Mismatch Repair and Mutation Avoidance Pathways. *Journal of Cellular Physiology*; 191:28-48.
- Mei M, Liu D, Dong S, Ingvarsson S, Goodfellow PJ, Chen H. (2010) The MLH1 -93 promoter variant influences gene expression. *Cancer Epidemiol.*;34(1):93-5. ARCH ARTICLE Open Access
- Mitchell R. J., Farrington S. M., Dunlop M. G., and Campbell H. (2002) Mismatch Repair Genes hMLH1 and hMSH2 and Colorectal Cancer: A HuGE Review. *Am. J. Epidemiol.*; 156:885–902.
- Mohd AB, Palama B, Nelson SE, Tomer G, Nguyen M, Huo X, Buermeyer AB. (2006) Truncation of the C-terminus of human MLH1 blocks intracellular stabilization of PMS2 and disrupts DNA mismatch repair. *DNA Repair. Amst.*; 5:347-361.
- Mongiat-Artus P, Miquel C, Fléjou JF, Coulet F, Verine J, Buhard O, Soliman H, Teillac P, Praz F. (2006) Spectrum of molecular alterations in colorectal, upper urinary tract, endocervical, and renal carcinomas arising in a patient with hereditary non-polyposis colorectal cancer. *Virchows Arch.*;449(2):238-43.
- Moskal A., Norat T., Ferrari P., Riboli E. (2007) Alcohol intake and colorectal cancer risk: a dose–response meta-analysis of published cohort studies, *Int. J. Cancer*; 120:664–671.
- Moslein G, Tester DJ, Lindor NM, Honchel R, Cunningham JM, French AJ, Halling KC, Schwab M, Goretzki P, Thibodeau SN. (1996) Microsatellite instability and mutation analysis of hMSH2 and hMLH1 in patients with sporadic, familial and hereditary colorectal cancer. *Hum. Mol. Genet.*; 5:1245-1252.
- Murata H, Khattar NH, Kang Y, Gu L, Li GM. (2002) Genetic and epigenetic modification of mismatch repair genes hMSH2 and hMLH1 in sporadic breast cancer with microsatellite instability. *Oncogene*; 21:5696-5703.
- Gooderham N.J., Murray S., Lynch A.M., Yadollahi-Farsani M., Zhao K., Rich K., et al., (1996) Assessing human risk to heterocyclic amines, *Mutat. Res.*; 376:53–60.
- Bingham S.A., Pignatelli B., Pollock J.R., Ellul A., Malaveille C., Gross G., et al., (1996) Does increased endogenous formation of N-nitroso compounds in the human colon explain the association between red meat and colon cancer?, *Carcinogenesis*; 17:515–523.

- Nagy R., Sweet K., Eng Ch. (2004) Highly penetrant hereditary cancer syndromes. *Oncogene*; 23:6445-6470.
- Nakagawa H, Lockman JC, Frankel WL, Hampel H, Steenblock K, Burgart LJ, Thibodeau SN, de la Chapelle A (2004) Mismatch repair gene PMS2: disease-causing germline mutations are frequent in patients whose tumours stain negative for PMS2 protein, but paralogous genes obscure mutation detection and interpretation. *Cancer Res.*; 64:4721-4727.
- Nakken S., Alseth I., Rognes T. (2007) Computational prediction of the effects of non-synonymous single nucleotide polymorphisms in human DNA repair genes. *Neuroscience*; 145(5): 1273-9.
- Ng PC, Henikoff S. (2001) Predicting deleterious amino acid substitutions. *Genome Res.* ;11:863-74.
- Nilbert M, Wikman FP, Hansen TV, Krarup HB, Orntoft TF, Nielsen FC, Sunde L, Gerdes AM, Cruger D, Timshel S, Bisgaard ML, Bernstein I, Okkels H. (2009) Major contribution from recurrent alterations and MSH6 mutations in the Danish Lynch syndrome population. *Fam Cancer.*;8(1):75-83.
- Nomura S, Sugano K, Kashiwabara H, Taniguchi T, Fukayama N, Fujita S, Akasu T, Moriya Y, Ohhigashi S, Kakizoe T, Sekiya T. (2000) Enhanced detection of deleterious and other germline mutations of hMSH2 and hMLH1 in Japanese hereditary nonpolyposis colorectal cancer kindreds. *Biochem. Biophys. Res. Commun*; 271:120-129.
- Nyström-Lathi M., Ying Wu, Moisio A.L., Hofstra R.M.W., Ostinga J., Mecklin J.-P., Järvinen H., Leisti J., Buys C.H.C.M., de la Chapelle A., Peltomaki P. (1996) DNA mismatch repair gene mutations in 55 kindreds with verified or putative hereditary non-polyposis colorectal cancer. *Human Molecular Genetics*; 5; 6:763-769.
- Obmolova G, Ban C, Hsieh P, Yang W. (2000) Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature.*; 407:703-710.
- Ollila S., Bebek D.D., Jiricny J., Nyström M. (2008) Mechanisms of Pathogenicity in Human MSH2 Missense Mutants. *Hum. Mut.*; 29(11):1355-1363.

- Orita M., Iwahana H., Kanazawa H.; Hayashi K., Sekiya T. Proc. (1989) Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms (mobility shift of separated strands/point mutation/riiction franment length polymorphism) Natd. Acad. Sci. USA; 86:2766-2770.
- Otway R, Tetlow N, Hornby J, Kohonen-Corish M. (2000) Evaluation of enzymatic mutation detectiontrade mark in hereditary nonpolyposis colorectal cancer. Hum. Mutat.; 16:61-67.
- Ou J, Niessen RC, Lutzen A, Sijmons RH, Kleibeuker JH, de Wind N, Rasmussen LJ, Hofstra RM. 2007. Functional analysis helps to clarify the clinical importance of unclassified variants in DNA mismatch repair genes. Hum. Mutat.; 28:1047–1054.
- Overbeek LI, Kets CM, Hebeda KM, Bodmer D, van der Looij E, Willems R, Goossens M, Arts N, Brunner HG, van Krieken JH, Hoogerbrugge N, Ligtenberg MJ. (2007) Patients with an unexplained microsatellite instable tumour have a low risk of familial cancer. Br J Cancer.;96(10):1605-12.
- Liang P.S., Chen T.Y., Giovannucci E., (2009) Cigarette smoking and colorectal cancer incidence and mortality: systematic review and meta-analysis, Int. J. Cancer; 124:2406–2415.
- Palicio M, Balmaña J, González S, Blanco I, Marcuello E, Peinado MA, Julià G, Germà JR, López López JJ, Brunet J, Capellà G. (2002) Mismatch repair gene analysis in Catalanian families with colorectal cancer. J Med Genet.;39(6):E29.
- Papadopoulos N, Nicolaidis NC, Wei YF, Ruben SM, Carter KC, Rosen CA, Fleischmann RD, Fraser CM, Adams MD (1994) Mutation of a mutL homolog in hereditary colon cancer. Science; 263:1625-1629.
- Park SJ, Lee KA, Park TS, Kim NK, Song J, Kim BY, Choi JR. (2008) A novel missense MSH2 gene mutation in a patient of a Korean family with hereditary nonpolyposis colorectal cancer. Cancer Genet Cytogenet.;182(2):136-9.
- Peel DJ, Ziogas A, Fox EA, Gildea M, Laham B, Clements E, Kolodner RD, Anton-Culver H. (2000) Characterization of hereditary nonpolyposis colorectal cancer families from a population-based series of cases. J. Natl. Cancer. Inst.; 92:1517-1522.

- Peltomäki P, Vasen HF, (1997) Mutations predisposing to hereditary nonpolyposis colorectal cancer: database and results of the collaborative study. The International Collaborative Group on Hereditary Nonpolyposis Colorectal Cancer. *Gastroenterology*; 113:1146-1158.
- Perez-Cabornero L, Velasco E, Infante M, Sanz D, Lastra E, Hernández L, Miner C, Duran M. (2009) A new strategy to screen MMR genes in Lynch Syndrome: HA-CAE, MLPA and RT-PCR. *Eur J Cancer*.;45(8):1485-93.
- Peterlongo P, Nafa K, Lerman GS, Glogowski E, Shia J, Ye TZ, Markowitz AJ, Guillem JG, Kolachana P, Boyd JA, Offit K, Ellis NA. (2003) MSH6 germline mutations are rare in colorectal cancer families. *Int J Cancer*.;107(4):571-9.
- Piñol V., Andreu M., Castells A., Payá A., Bessa X., Jover R., (2004) Frequency of non-polyposis colorectal cancer and othe colorectal cancer familial forms in Spain: a multicentre, prospective, nationwide study. *Eur. J. Gastroenterol. Heptol.*; 16:39-45.
- Piñol V., Castells A., Andreu M. (2005) Accuracy of Revised Bethesda Guidelines, Microsatellite Instability, and Immunohistochemistry for the Identification of Patients With Hereditary Nonpolyposis Colorectal Cancer. *JAMA*; 293:1986-1994.
- Plaschke J, Kruppa C, Tischler R, Bocker T, Pistorius S, Dralle H, Rüschoff J, Saeger HD, Fishel R, Schackert HK. (2000) Sequence analysis of the mismatch repair gene hMSH6 in the germline of patients with familial and sporadic colorectal cancer. *Int J Cancer*.;85(5):606-13.
- Plasilova M., Zhang J, Okhowat R., Marra G., Mettler M., Mueller H., Heinimann K. (2006) A De Novo MLH1 Germ Line Mutation in a 31-Year-Old Colorectal Cancer Patient. *Genes, Chromosomes & Cancer*; 45:1106–1110.
- Plotz G., Welsch C., Giron-Monzon L., Friedhoff P., Albrecht M., Piiper A., Biondi R.M., Lengauer T, Zeuzem S., Raedle J. (2006) Mutations in the MutSa interaction interface of MLH1 can abolish DNA mismatch repair. *Nucleic Acids Research*; 22:6574–6586.
- Ponz de Leon M., Benatti P., Di Gregorio C., Pedroni M., Losi L., Genuardi M., Viel A., Fornasarig M., Lucci-Cordisco E., Anti M., Ponti G., Borghi F., Lamberti I. and Roncucci L. (2004). Genetic testing among high-risk individuals in families with hereditary nonpolyposis colorectal cancer. *British Journal of Cancer*; 90:882-887.

- Räschle M, Marra G, Nyström-Lathi M, Schär P, Jiricny J, (1999) Identification of hMutLb, a Heterodimer of hMLH1 and hPMS1. *The Journal of Biological*; 274:32368-32375.
- Ravi R. Iyer, Anna Pluciennik, Vickers Burdett, and Paul L. Modrich. (2006) DNA Mismatch Repair: Functions and Mechanisms. *Chem. Rev.*; 106:302-323.
- Robinson K., Tao L., Vandrovцова J. , Halvarsson B., Clendenning M., Frebourg T., Papadopoulos N., Kinzler K.W., Vogelstein B., Peltomäki P., Kolodner R.D., Nilbert M., Lindblom A. (2007) Lynch Syndrome (Hereditary Nonpolyposis Colorectal Cancer) diagnostics. *J. Natl. Cancer Inst.*; 99:291-299.
- Rodriguez A., Griffiths-Jones S., Ashurst J.L., Bradley A. (2004) Identification of Mammalian microRNA Host Genes and Transcription Units. *Genome Res*; 14(10a): 1902–1910.
- Rodriguez-Bigas MA, Boland CR, Hamilton SR, Henson DE, Jass JR, Khan PM, Lynch H, Perucho M, Smyrk T, Sobin L, Srivastava S. (1997) A National Cancer Institute Workshop on Hereditary Nonpolyposis Colorectal Cancer Syndrome: meeting highlights and Bethesda guidelines. *J. Natl. Cancer Inst.*; 89:1758-1762.
- Rubio-Del-Campo A, Salinas-Sánchez AS, Sánchez-Sánchez F, Giménez-Bachs JM, Donate-Moreno MJ, Pastor-Navarro H, Carrión-López P, Escribano J. (2008) Implications of mismatch repair genes hMLH1 and hMSH2 in patients with sporadic renal cell carcinoma. *BJU Int.*;102(4):504-9.
- Rustgy A.K., (2007) The genetics of hereditary colorectal cancer. *Gene Dev.*; 21:2525-2538.
- Larsson S.C., Bergkvist L., Wolk A., (2005) High-fat dairy food and conjugated linoleic acid intakes in relation to colorectal cancer incidence in the Swedish mammography cohort, *Am. J. Clin. Nutr.* 82:894–900.
- Sachio Nomura, Kokichi Sugano, Hidefumi Kashiwabara, Takahiro Taniguchi, Noriko Fukayama, Shin Fujita, Takayuki Akasu, Yoshihiro Moriya, Seiji Ohhigashi, Tadao Kakizoe, Takao Sekiya. (2000) Enhanced Detection of deleterious and Other Germline Mutations of hMSH2 and hMLH1 in Japanese Hereditary Nonpolyposis Colorectal Cancer Kindreds. *Biochemical and Biophysical Research Communications*; 271:120–129.

- Gille J.J., Hogervorst F.B., Pals G., Wijnen J.T., van Schooten R.J., Dommering C.J., Meijer G.A., Craanen M.E., Nederlof P.M., de Jong D., McElgunn C.J., Schouten J.P., Menko F.H. (2002) Genomic deletions of MSH2 and MLH1 in colorectal cancer families detected by a novel mutation detection approach. *Br J Cancer*.;87(8):892-7.
- Samowitz WS, Curtin K, Lin HH, Robertson MA, Schaffer D, Nichols M, Gruenthal K, Leppert MF, Slattery ML. (2001) The colon cancer burden of genetically defined hereditary nonpolyposis colon cancer. *Gastroenterology*; 121:1005-1008.
- Sanchez de Abajo A, de la Hoya M, van Puijenbroek M, Godino J, Díaz-Rubio E, Morreau H, Caldes T. (2006) Dual role of LOH at MMR loci in hereditary non-polyposis colorectal cancer? *Oncogene*;25(14):2124-30.
- Sánchez-Céspedes M. (2007) A role for LKB1 gene in human cancer beyond Peutz-Jeghers syndrome. *Oncogene*; 26(57):7825-32.
- Shcherbakova P.V., Kunkel T.A. (1999) Mutator Phenotypes Conferred by MLH1 Overexpression and by Heterozygosity for mlh1 Mutations. *Molecular and cellular Biology*; 19:3177–3183.
- Shin KH, Han HJ, Park JG. (1998) Growth suppression mediated by transfection of wild-type hMLH1 in human cancer cells expressing endogenous truncated hMLH1 protein. *Int. J. Oncol.*; 12:609-615.
- Sieber MO, Lipton S, Crabtree M, Heinemann K, Fidalgo P, Philips R, Bisgaard ML, Orntoft TO, Aaltonen L, Hodgson S, Thomas Huw, Tomlinson Ian. (2003) Multiple Colorectal Adenomas, Classic Adenomatous Polyposis, and Germ-Line Mutations in MYH. *N. Engl. J. Med.*; 348:791-799.
- Slaby O., Svodoba M., Michalek J., Vyzula R. (2009) MicroRNAs in colorectal cancer: translation of molecular biology into clinical application. *Molecular Cancer*, 8:102-115.
- Smith P.J., Zhang C., Wan J. et al. (2006) An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers. *Hum. Mol. Genet.* 115(16):2490-508.
- Smith TR, Levine EA, Freimanis RI, Akman SA, Allen GO, Hoang KN, Liu-Mares W, Hu JJ. (2008) Polygenic model of DNA repair genetic polymorphisms in human breast cancer risk. *Carcinogenesis*.;29(11):2132-8.

- Song H, Ramus SJ, Quaye L, DiCioccio RA, Tyrer J, Lomas E, Shadforth D, Hogdall E, Hogdall C, McGuire V, Whittemore AS, Easton DF, Ponder BA, Kjaer SK, Pharoah PD, Gayther SA. (2006) Common variants in mismatch repair genes and risk of invasive ovarian cancer. *Carcinogenesis*;27(11):2235-42.
- Soravia S, Berk T, Madlensky L, Mitri A, Cheng H, Gallinger S, Cohen Z, Bapat Bharati. (1998) Genotype-Phenotype Correlations in Attenuated Adenomatous Polyposis Coli. *Am. J. Hum. Genet.*; 62:1290–1301.
- Stitzel N.O., Tseng Y.Y., Pervouchine D. et al., (2003) Structural localization of disease-associated single-nucleotide polymorphism. *J. Mol. Bio.* 327:1021-1030.
- Sutter C, Dallenbach-Hellweg G, Schmidt D, Baehring J, Bielau S, von Knebel Doeberitz M, Gebert J. (2004) Molecular analysis of endometrial hyperplasia in HNPCC-suspicious patients may predict progression to endometrial carcinoma. *Int J Gynecol Pathol.*;23(1):18-25.
- T. Mizoue, M. Inoue, K. Wakai, C. Nagata, T. Shimazu, I. Tsuji, et al., Alcohol drinking and colorectal cancer in Japanese: a pooled analysis of results from five cohort studies, *Am. J. Epidemiol.*; 167:1397–1406.
- Takahashi M, Shimodaira H, Andreutti C, Iggo R., Kolodner R.D. & Ishioka C. (2007) Functional analysis of human MLH1 variants using yeast and in vitro mismatch repair assays. *Cancer Research*; 67:4595-4604.
- Tinat J., Baert-Desurmont S., Latouche J. B., Vasseur S., Martin C., Bouvignies E., Frébourg T. (2008) The three nucleotide deletion within the 30untranslated region of MLH1 resulting in gene expression reduction is not a causal alteration in Lynch syndrome. *Familial Cancer*; 7:339–340.
- Tournier I, Vezain M, Martins A, Charbonnier F, Baert-Desurmont S, Olschwang S, Wang Q, Buisine MP, Soret J, Tazi J, Frébourg T, Tosi M. (2008) A large fraction of unclassified variants of the mismatch repair genes MLH1 and MSH2 is associated with splicing defects. *Hum Mutat.*;29(12):1412-24.
- Tulupova E, Kumar R, Hanova M, Slyskova J, Pardini B, Polakova V, Naccarati A, Vodickova L, Novotny J, Halamkova J, Hemminki K, Vodicka P. (2008) Do polymorphisms and haplotypes of mismatch repair genes modulate risk of sporadic colorectal cancer? *Mutat Res.*;648(1-2):40-5.

- Umar A. (2004) Lynch syndrome (HNPCC) and microsatellite instability. *Dis. Markers*; 20:179-180.
- Vahteristo P, Ojala S, Tamminen A, Tommiska J, Sammalkorpi H, Kiuru-Kuhlefelt S, Eerola H, Aaltonen LA, Aittomäki K, Nevanlinna H. (2005) No MSH6 germline mutations in breast cancer families with colorectal and/or endometrial cancer. *J Med Genet.*;42(4):e22.
- Valeri N., Gasparini P., Fabbri M., Braconi C., Veronesse A., Lovat F., Adair B., Vannini I., Fanini F, Bottoni A., Costinean S., Sandhu G., Nuovo G.J., Alder H., Gafa R., Calore F., Ferracin M., Lanza G., Volinia S., Negrini M., McIlhatton M.A., Amadori D., Fischel R., Croce C.M. (2010) Modulation of mismatch repair and genomic stability by miR-155. *PNAS*; 107 (15):6982-6987.
- Varesco L. (2004) Familial adenomatous polyposis: genetics and epidemiology. *Tech. Coloproctol*; 8:S305–S308.
- Vasen HF et al. (1996) Cancer risk in families with hereditary nonpolyposis colorectal cancer diagnosed by mutation analysis. *Gastroenterology*; 110:1020-1027.
- Vasen HF, Mecklin JP, Khan PM. (1991) The International Collaborative Group on Hereditary Nonpolyposis Colorectal Cancer. *Dis. Colon Rectum*; 34:424–425.
- Vasen HF, Watson P, Mecklin JP, et al. (1999) New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative Group on HNPCC. *Gastroenterology*; 116:1453–1456.
- Verma L, Kane MF, Brassett C, Schmeits J, Evans DG, Kolodner RD, Maher ER. (1999) Mononucleotide microsatellite instability and germline MSH6 mutation analysis in early onset colorectal cancer. *J Med Genet.*;36(9):678-82.
- Wagner A, Hendricks Y, Meijers-Heijboer EJ, de Leeuw WJ, Morreau H, Hofstra R, Tops C, Bik E, Brocker-Vriends AH, van der Meer C, Lindhout D, Vasen HF, Breuning MH, Cornelisse CJ, van Krimpen C, Niermeijer MF, Zwinderman AH, Wijnen J, Fodde R (2001) Atypical HNPCC owing to MSH6 germline mutations: analysis of a large Dutch pedigree. *J. Med. Genet.*; 38:318-322.
- Wagner A., Barrows A., Wijnen J., van der Klift H., Franken P., Verkuijlen P., Nakagawa H., Geugien M., Jaghmohan-Changur S., Breukel C., Meijers-Heijboer H., Morreau H., van

- Puijtenbroek M., Burn J., Coronel S., Kinarski Y., Okimoto R., Watson P, Lynch J.F., de la Chapelle A., Lynch H.T., and Fodde R. (2003) Molecular Analysis of Hereditary Nonpolyposis Colorectal Cancer in the United States: High Mutation Detection Rate among Clinically Selected Families and Characterization of an American Founder Genomic Deletion of the MSH2 Gene. *Am. J. Hum. Genet.*; 72:1088–1100.
- Wahlberg SS, Nyström-Lahti M, Kane MF, Kolodner RD, Peltomäki P, (1997) Lindblom A. Low frequency of hMSH2 mutations in Swedish HNPCC families. *Int J Cancer.*;74(1):134-7.
 - Wang Q., Lasset C., Desseigne F., Saurin J.C., Maugard C., Navarro C., Ruano E., Descos L., Trillet-Lenoir V., Bosset J.F., Puisieux A. (1999) Prevalence of germline mutations of hMLH1, hMSH2, hPMS1, hPMS2, and hMSH6 genes in 75 French kindreds with nonpolyposis colorectal cancer. *Hum. Genet.*; 105 :79–85.
 - Warren JJ, Pohlhaus TJ, Changela A, Iyer RR, Modrich PL, Beese LS. (2007) Structure of the human MutSalpha DNA lesion recognition complex. *Mol. Cell.*; 26:579–592.
 - Warthin A. (1931) Heredity of carcinoma in man. *Ann. Int. Med.*; 4: 681-696.
 - Watson P, Vasen HF, Mecklin JP, Bernstein I, Aarnio M, Järvinen HJ, Myrhøj T, Sunde L, Wijnen JT, Lynch HT. (2008) The risk of extra-colonic, extra-endometrial cancer in the Lynch syndrome. *Int J Cancer.* 123(2):444-9.
 - Weber T, Conlon W, Petrelli NJ, Rodríguez-Bigas M, Keitz B, Pazik J, Farrell C, O'Malley L, Oshalim M, Abdo M, Anderson G, Stoler D, Yandell D. (1997) Genomic DNA-based hMSH2 and hMLH1 Mutation Screening in 32 Eastern United States Hereditary Nonpolyposis Colorectal Cancer Pedigrees. *Cancer Research*; 57:3798-3803.
 - Whitelaw SC et al. (1997) Clinical and molecular features of the hereditary mixed polyposis syndrome. *Gastroenterology*; 112:327–334.
 - Wijnen J, Fodde R, Khan PM. (1994) DGGE polymorphism in intron 10 of MSH2, the HNPCC gene. *Hum Mol Genet.*;3(12):2268.
 - Wijnen JT, Vasen HF, Khan PM, Zwinderman AH, van der Klift H, Mulder A, Tops C, Møller P, Fodde R. (1998) Clinical findings with implications for genetic testing in families with clustering of colorectal cancer. *N Engl J Med.*; 339(8):511-8.

- Wojcik SE, Rossi S, Shimizu M, Nicoloso MS, Cimmino A, Alder H, Herlea V, Rassenti LZ, Rai KR, Kipps TJ, Keating MJ, Croce CM, Calin GA (2010) Non-coding RNA sequence variations in human chronic lymphocytic leukemia and colorectal cancer. *Carcinogenesis*; 31: 208–215.
- Woods MO, Hyde AJ, Curtis FK, Stuckless S, Green JS, Pollett AF, Robb JD, Green RC, Croitoru ME, Careen A, Chaulk JA, Jegathesan J, McLaughlin JR, Gallinger SS, Younghusband HB, Bapat BV, Parfrey PS. (2005) High frequency of hereditary colorectal cancer in Newfoundland likely involves novel susceptibility genes. *Clin Cancer Res.*; 11:6853-6861.
- Y.M. Coyle, Lifestyle, genes, and cancer, *Meth. Mol. Biol.* 472 (2009) 25–56.
- Yan SY, Zhou XY, Du X, Zhang TM, Lu YM, Cai SJ, Xu XL, Yu BH, Zhou HH, Shi DR. (2007) Three novel missense germline mutations in different exons of MSH6 gene in Chinese hereditary non-polyposis colorectal cancer families. *World J Gastroenterol.*;13(37):5021-4.
- Yap HL, Chieng WS, Lim JR, Lim RS, Soo R, Guo J, Lee SC. (2009) Recurring MLH1 deleterious mutations in unrelated Chinese Lynch syndrome families in Singapore. *Fam Cancer.*;8(2):85-94.
- Yu Z, Li Z, Jolicoeur N, Zhang L, Fortin Y, Wang E, Wu M, Shen SH (2007) Aberrant allele frequencies of the SNPs located in microRNA target sites are potentially associated with human cancers. *Nucleic Acids Res*; 35:4535– 4541.
- Yue P., Melamoult E., Moulton J., (2006) SNPs3D: candidate gene and SNP selection for associations studies. *BMC Bioinformatics*; 7:166. (a)
- Yue P., Moulton J., (2006) Identification and analysis of deleterious human SNPs, *J. Mol., Bio.*; 356(5): 1263-74. (b)
- Zhou XP, Waite KA, Pilarski R, Hampel H, Fernandez MJ, Bos C, Dasouki M, Feldman GL, Greenberg LA, Ivanovich J, Matloff E, Patterson A, Pierpont ME, Russo D, Nassif NT, Eng C (2003). "Germline PTEN promoter mutations and deletions in Cowden/Bannayan-Riley-Ruvalcaba syndrome result in aberrant PTEN protein and dysregulation of the phosphoinositol-3-kinase/Akt pathway". *Am. J. Hum. Genet.*; 73: 404–411.

