



*“DETECCIÓN DE MUTACIONES DEL GEN
TP53 EN LEUCEMIA LINFÁTICA CRÓNICA
A PARTIR DE LA EXTRACCIÓN DE DNA DE
MUESTRAS DE TEJIDO PARAFINADO Y
SANGRE PERIFÉRICA”*

Trabajo de Fin de Grado

Autor: Francisco González Sammarco

Facultad de Medicina | Universidad de La Laguna | 2019

“Detección de mutaciones del gen *TP53* en leucemia linfática crónica a partir de la extracción de DNA de muestras de tejido parafinado y sangre periférica”

Francisco González Sammarco¹, Eduardo Salido Ruiz².

1. Alumno de 6º de medicina, Universidad de La Laguna, Tenerife, España. Contacto: alu0100818659@ull.edu.es.
2. Tutor de TFG, MD, PhD, Departamento de Anatomía Patológica, Hospital Universitario de Canarias, La Laguna, Tenerife, España. Contacto: edsalido@gmail.com.

Resumen

La leucemia linfática crónica (CLL en adelante), es el tipo de leucemia más común en occidente en la población adulta. El desarrollo de la enfermedad es motivado por alteraciones genómicas específicas que afectan fundamentalmente a la apoptosis de las células B.

En el universo que conforma la elevadísima complejidad genómica de la enfermedad, destacan las deleciones y las trisomías. Dentro de las deleciones, es fundamental resaltar la que ocurre a nivel del brazo corto del cromosoma 17, donde se halla el gen *TP53*, pues se conoce que tiene un papel importante a nivel pronóstico y terapéutico.

En este proyecto, se estudiarán las diferentes mutaciones del gen *TP53*, así como su patrón de frecuencia en pacientes diagnosticados de CLL a partir de muestras histológicas parafinadas y sangre periférica haciendo uso de técnicas de secuenciación molecular de nueva generación.

Abstract

Chronic lymphocytic leukemia (CLL) is the most common type of leukemia in the adult population in the West. The development of the disease is motivated by specific genomic alterations that fundamentally affect the apoptosis of B cells.

In the extremely high genomic complexity of the disease, deletions and trisomies play an important role. Among the deletions, it is essential to highlight the one that occurs in the short arm of chromosome 17, where the *TP53* gene is found. It is known that people with *TP53* mutated gene have a poor prognosis and less response to the treatment.

In this project, we will study the different mutations of the *TP53* gene and also its frequency pattern in patients diagnosed with CLL from paraffinated histological and blood samples using molecular next generation sequencing techniques.

Palabras clave

Leucemia linfática crónica (CLL), *TP53*, secuenciación de nueva generación (NGS).

Introducción

El gen *TP53* se aloja en el cromosoma 17p13.1, concretamente entre las regiones 17: 7.661.779-7.687.550 según “Ensembl” para la versión humana GRCh38.p12. Este gen tiene 11 exones y codifica 29 transcritos diferentes gracias a los mecanismos de splicing alternativo. Un mRNA de aproximadamente 2,8 kB es la molécula cuya traducción da lugar a la proteína *p53*. El péptido cuenta con 393 aminoácidos, 53 kDa de peso molecular y una estructura compleja^{1,2}.

En condiciones normales, la proteína *p53* actúa a distintos niveles en las células: 1) regula el ciclo celular, mediando el paso de G1 a S; En caso de existir alteraciones en el DNA, *p53* detiene el ciclo para que actúen los mecanismos reparadores. 2) regula los genes diana que inducen la apoptosis. 3) participa en la diferenciación celular¹. En células estables se mantiene inactivo a través de la unión con la ubiquitina ligasa *MDM2*, y es degradada por la ubiquitina *p53*; En caso de exposición tóxica celular, los niveles de *p53* se elevan, iniciando la muerte celular programada.

La actividad de este gen ante situaciones tóxicas para las células se pierde en una gran variedad de neoplasias malignas que afectan al ser humano, ya sea por alteración del propio gen o de las señales que actúan regulándolo².

A lo largo del gen *TP53* han sido identificadas en torno a 29.000 mutaciones somáticas distintas para todos los tipos de tumores. En pacientes con CLL, las mutaciones somáticas descritas hasta la actualidad representan un total de 188 y las de línea germinal un total de 3. Aproximadamente el 90% de las mutaciones tanto somáticas como de línea germinal reportadas se localizan entre el exón 5 y el exón 8. Es por ello, que los exones 4-10 sean objeto de estudio en la mayoría de proyectos experimentales. El exón donde ha sido descrito un mayor número de mutaciones somáticas es el 7. A su vez, dentro de este, se localizan más frecuentemente en el codón 248³. Otros codones frecuentemente afectados son los 175, 179, 209 y 273.

En los exones del 1-4 y del 9-10 se han descrito mayoritariamente otro tipo de mutaciones. A dicho nivel ocurren principalmente mutaciones con desplazamiento del marco de lectura (frameshit), llegando a representar el 53% de las mutaciones en dichos exones. Otra menos frecuente es la mutación sin sentido (nonsense), que en dichas regiones llega a suponer el 9% del total.

En general, las mutaciones somáticas que afectan al gen *TP53* en la CLL son mayoritariamente las mutaciones por cambio de sentido (missense), representando un 74% del total. Esto significa que por un cambio en una de las bases del DNA se incorpora un aminoácido distinto al de la variante normal durante la síntesis de la proteína, alterando la estructura de la misma. Otras menos frecuentes son: las deleciones, las cuales ocupan un 16% del total y las mutaciones sin sentido (nonsense), que representan un 4% del total. En ellas, lo que ocurre es la aparición de un codón de parada de manera prematura. Finalmente, un 3% del total lo ocupan las inserciones.

A nivel de nucleótidos, las tres sustituciones más frecuentes en cuanto a porcentajes son: 1) transición G:C>A:T fuera de la región CpG 13%; 2) G:C>A:T en región CpG 20%; 3) G:C>C:G 14%^{3,4}. Otras sustituciones frecuentes son transición la A:T>G:C 13%; la transversión A:T>C:G 7% y la transversión A:T>T:A 6%.

La leucemia linfática crónica es un tipo de neoplasia linfoproliferativa crónica de expresión leucémica. La enfermedad está clasificada en base a su origen a partir de células B. La leucemia linfática crónica también puede ser hallada en la literatura como leucemia linfocítica crónica o CLL.

Constituye la forma más frecuente de leucemia en países occidentales⁵. La enfermedad tiene una incidencia de 4,2 casos por cada 100.000 habitantes por año. Ocurre normalmente en adultos y ancianos, siendo la edad media de presentación los 60 años⁶. Existe cierta predominancia en hombres, en donde es 1,7-2:1 veces más frecuente según las diferentes series⁵.

A diferencia de otras neoplasias linfoides, las translocaciones cromosómicas son raras en la CLL. Las alteraciones genéticas más representativas de la enfermedad son la pérdida o adición de material cromosómico en forma de: 1) Deleciones (13q, 11q23 y 17p) y 2) Trisomías (12q)⁷. A estas alteraciones del material genético se añaden otras mutaciones subyacentes, responsables de la heterogeneidad de la enfermedad en términos de respuesta al tratamiento y supervivencia⁸.

La deleción del brazo largo del cromosoma 13 (del(13q14)) representa la aberración cromosómica más frecuente y representativa de la enfermedad, llegando a estar presente en torno al 55% de los pacientes diagnosticados de CLL⁵. La mutación aislada de esta región se ha descrito como un factor de buen pronóstico.

La mutación por pérdida cromosómica del brazo largo del cromosoma 11 (del(11q23)) está presente en torno al 25% de los pacientes con enfermedad avanzada y en un 10% en enfermos con estadios precoces⁵. Esta región alberga al gen *ATM*, encargado de la reparación del ácido desoxirribonucleico (DNA) y de regular la división celular. Como consecuencia de alteraciones a nivel de esta región, los pacientes presentan de manera muy frecuente linfadenopatías grandes, rápida progresión de la enfermedad y una supervivencia global inferior.

Otro factor determinante de esta región es su implicación indirecta sobre p53. Las alteraciones a nivel del cromosoma 11q, que codifica microRNA-15 y microRNA-16,

participan de manera notable en la patogénesis de la CLL. Se conoce que existe un circuito directo entre microRNA-15/16 y TP53⁹ el cual puede verse alterado en la CLL.

La delección del brazo corto del cromosoma 17 (del(17p) ocurre en un 5-8% de los pacientes⁵. Esta mutación incluye prácticamente de manera constante la banda 17p13, donde se aloja el gen TP53, el gen supresor tumoral por excelencia. Los enfermos con esta mutación muestran una importante resistencia al tratamiento combinado de inmunquimioterapia.

Los enfermos con CLL en los que se han demostrado alteraciones aisladas del gen *TP53* o de todo cromosoma 17p, muestran un peor pronóstico independientemente del estado mutacional de *IGHV*⁵. El impacto del estado mutacional de la región que aloja al gen *TP53* es tal, que en determinados países se adoptan las técnicas de hibridación fluorescente in situ (FISH) para detectar mutaciones del cromosoma 17p13 con el fin de evitar tratamientos que resultan inefectivos en pacientes con CLL y dicha anomalía^{5,10}.

La alteración del gen *TP53* de manera aislada es también notoriamente trascendental para los enfermos, ya que se trata de un factor pronóstico relevante y ayuda a establecer planes de tratamiento¹¹. Para pacientes diagnosticados de CLL, se está enfatizando el rol de las mutaciones aisladas del gen *TP53*. En el momento del diagnóstico, apenas el 10% tiene la mutación presente. Sin embargo, se ha demostrado que a medida que avanza la enfermedad dicho porcentaje se ve aumentado de manera significativa¹².

Dada la alta prevalencia de las mutaciones del gen *TP53* en pacientes con CLL evolucionada y la repercusión que dicha alteración tiene sobre el manejo terapéutico y la supervivencia, diversos estudios proponen la creación nuevas escalas pronósticas que tengan en cuenta tanto las delecciones de 17p13 como las mutaciones aisladas de *TP53*¹¹. Sin embargo, en la práctica es más complicado pues mientras la delección del cromosoma 17p13 puede ser demostrada haciendo uso de FISH, las mutaciones aisladas del gen *TP53* deben demostrarse mediante secuenciación de nueva generación, técnica que será desarrollada en este estudio.

Más allá de si se trata una delección completa de la región 17p o si ocurre de manera aislada a nivel de *TP53*, existen otros mecanismos que pueden alterar la proteína que transcribe esta región: la proteína p53. Esta proteína se encuentra ampliamente marcada por reguladores distintos al gen *TP53*, tales como: microRNA 15/16, *MDM2*, *ATM*, y microRNA 34. Todos estos factores actúan mediando la expresión de *p53*. De esta manera, alteraciones en otros cromosomas repercuten directamente en la transcripción de dichas moléculas e indirectamente sobre la acción de p53.

Por último, se ha descrito que la trisomía del brazo largo del cromosoma 12 ocurre en un 20% de pacientes con CLL⁵. Las alteraciones a este nivel se correlacionan con sobreexpresión de *MDM2*, una proteína que autorregula de manera negativa la expresión de p53. La alteración del cromosoma 12 ha sido correlacionada con histología atípica y evolución clínica agresiva.

Existen otras mutaciones que han sido asociadas recientemente a la CLL tales como la del receptor *NOTCH1* así como *MYD88*, *SF3B1*, *FBXW7*, *POT1*, *CHD2*, *RPS15*, *IKZF3*, *ZNF292*, *ZMYM3*, *ARID1A* y *PTPN11*^{4,5,13,14}.

Tal como ha sido descrito, no cabe duda en destacar que todas las alteraciones que ocurren a distinto nivel cromosómico en la CLL afectan de manera directa e indirecta a la expresión de *TP53*. Las alteraciones a nivel del cromosoma 11, 12, 13 son las más frecuentes en la enfermedad, y están asociadas sobre todo a mutaciones a nivel de proteínas reguladoras de p53 como: *MDM2* y *ATM*. Estas alteraciones, a pesar de ocurrir en cromosomas diferentes del 17p13 (donde se aloja el gen *TP53*) hacen que, por la existencia de una compleja red de regulaciones positivas y negativas entre proteínas, la proteína p53 se vea alterada.

La CLL tiene un inmunofenotipo característico. Las células expresan marcadores B tales como CD19, CD20 tenue, CD23 así como CD5 siendo estas últimas una combinación propia de la CLL. El antígeno CD23 es realmente útil en el diagnóstico diferencial de CLL con linfoma de células del manto, enfermedad en la que dicho marcador es negativo^{12,15}. Otros antígenos de linfocito B positivos en la enfermedad son CD79a y pax-5. Es importante resaltar que en la enfermedad es frecuente una baja expresión de Inmunoglobulinas (Ig) de superficie CD20 y CD79b frente a aquellos

encontrados en células B normales. Cada clona de células leucémicas expresa exclusivamente una Ig de cadena ligera, o la kappa o lambda. Estos marcadores son útiles para el diagnóstico si se hace uso de la escala de Matutes.

El estudio por citometría de flujo muestra una expresión débil de Ig de superficie (IgM o IgM + IgG o, pocas veces IgG) con restricción de cadena ligera kappa o lambda. Hay expresión de CD19, CD20 (débil), CD79a, CD5, CD23, CD43, CD200 con CD10 negativo y FMC7 normalmente negativo o débilmente positivo. Se ha demostrado que los pacientes con ZAP70 positivo en > 30% de las células mediante citometría de flujo tienen peor pronóstico que ZAP70 negativo¹⁵.

Algunos casos de CLL tienen un inmunofenotipo extraño bien sea por CD5 negativo, CD23 negativo, por FMC7 positivo o CD79b positivo. En estos casos es muy importante realizar un diagnóstico diferencial con otros tipos de leucemias de células B, tales como linfoma de la zona marginal.

El recuento leucocitario encontrado en los diferentes pacientes con CLL es notoriamente variable. Pueden existir recuentos de más de 200.000 leucocitos por milímetro cúbico. Sin embargo, estos han sido encontrados en pacientes con grandes masas tumorales. También existen pacientes asintomáticos que tienen en sangre periférica valores de linfocitos B monoclonales demasiado bajos como para recibir el diagnóstico de CLL, pero con las mismas mutaciones a nivel cromosómico que se describen en la enfermedad. De ellos, solo el 1% evolucionará a CLL sintomática cada año. Clásicamente una cifra de más de 5.000 elementos linfoides B por milímetro cúbico durante un periodo de al menos tres meses suele ser identificativo de CLL. Para su confirmación, es importante demostrar la naturaleza clonal de las células mediante el uso de citometría de flujo.

El estudio histológico de la enfermedad muestra generalmente un borramiento difuso de los ganglios linfáticos por un infiltrado de linfocitos predominantemente pequeños, de 6-12 micrómetros de diámetro con núcleos redondeados o ligeramente irregulares, cromatina condensada y citoplasma escaso⁷. A menudo, los linfocitos se reúnen en agregados laxos denominados centros de proliferación, en cuyo interior hay

células mitóticamente activas. Estos centros de proliferación son prácticamente patognomónicos de CLL.

Al estudiar la sangre periférica, se puede evidenciar que contiene un número mayor de linfocitos redondeados pequeños con citoplasma escaso. Un hallazgo característico al extender las células son las sombras de Gumprecht, que refleja la fragilidad de los linfocitos.

Por último, destacar que la médula ósea casi siempre está afectada por infiltrados intersticiales o agregados de células tumorales. Los infiltrados también se ven prácticamente siempre en la pulpa esplénica blanca y roja y en los tractos portales hepáticos⁶.

La CLL es diagnosticada mayoritariamente en pacientes asintomáticos. Cuando la clínica se hace notable incluye síntomas inespecíficos como fatigabilidad, pérdida de peso y anorexia^{5,7}. En un 50-60% de los pacientes sintomáticos ocurre una linfadenopatía generalizada acompañada de hepatoesplenomegalia. En un gran conjunto de pacientes, la clínica deriva de la infiltración progresiva de la médula, ganglios y órganos linfoides por las clonas malignas. Así, progresivamente irá padeciendo consecuencias clínicas típicas de la anemia, la neutropenia y finalmente la trombopenia. Dentro de ellas, se engloban consecuencias tales como el cansancio, la tendencia a las infecciones, especialmente por bacterias y al sangrado respectivamente.

La evolución y el pronóstico de la enfermedad son extremadamente variables y dependen fundamentalmente de la fase clínica en la que se encuentra la enfermedad. La mediana de supervivencia global es de 4-6 años, pero es mayor de 10 años en sujetos con cargas tumorales mínimas en el momento del diagnóstico. Otras variables que se correlacionan con un peor evolución son: 1) la presencia de deleciones 11q y 17p, siendo las que afectan al 17p las mutaciones objeto de este estudio; 2) la ausencia de hipermutación somática; 3) la expresión de ZAP-70, una proteína que aumenta las señales inducidas por el receptor de Ig; y 4) la presencia de mutaciones de *NOTCH1*⁵.

La tendencia de la CLL a transformarse en un tumor agresivo es otro factor pronóstico importante. La forma más frecuente de evolución es a un linfoma difuso de

linfocitos B grandes. Sin embargo, puede progresar a un linfoma de Hodgkin clásico, a un linfoma periférico de linfocitos T y excepcionalmente a una leucemia/linfoma linfoblástico B. Todas estas formas de progresión de enfermedad se denominan síndrome de Richter (que ocurre en torno al 10-15% de los pacientes) y suele tener una mediana de presentación de 2-3 años desde el diagnóstico de CLL, sobre todo en adultos en torno a 70 años de edad⁷. Esta transformación suele estar anunciada por el desarrollo de una masa de crecimiento acelerado en un ganglio linfático o en el bazo. Probablemente se deba a la adquisición de mutaciones adicionales, aún no conocidas, que aumentan el crecimiento.

En la clínica, para realizar una aproximación pronóstica universal se hace uso de las escalas de Rai y Binet. La escala pronóstica de Rai modificada define como paciente de bajo riesgo (Rai 0) a aquel que tiene linfocitosis con células leucémicas (clonales) en sangre periférica o médula ósea (siendo el recuento de células linfocíticas mayor del 30%). Aquellos con linfocitosis, adenomegalias, esplenomegalia y/o hepatomegalia son considerados estadios intermedios (Rai I-II). Finalmente, la enfermedad de alto riesgo está definida en pacientes con anemia y/o trombocitopenia secundaria a la evolución de la enfermedad (Rai III y IV respectivamente).

El sistema de estadificación de Binet se ampara en el número de áreas afectas. Son definidas por la presencia de adenomegalias mayores de 1 centímetro de diámetro u organomegalias así como si existe o no anemia y/o trombocitopenia.

Los pacientes sintomáticos por lo general reciben tratamiento quimioterápico e inmunoterapia con anticuerpos contra proteínas presentes en la superficie de las células de CLL, especialmente CD20. El trasplante de células madre hematopoyéticas se ofrece a los pacientes relativamente jóvenes. El tratamiento nuevo más esperanzador se basa en los inhibidores de BTK, una quinasa presente en la transducción de señales de crecimiento y supervivencia que ocurre en las células tumorales a partir de señales generadas por el receptor de linfocitos B.

Es importante conocer que la NGS constituye una técnica que identifica diferencias en las secuencias de DNA o ácido ribonucleico (RNA) complementario de un paciente afecto de CLL al compararlo con uno de referencia. Además de la detección de

nuevas mutaciones, la NGS se puede utilizar para establecer el tamaño del clon neoplásico y trazar el modelo de adquisición de eventos que conllevan progresión de enfermedad¹⁶.

De cara al futuro, existe un gran potencial en el uso de la NGS. Ejemplo de ello son enfermedades como la CLL, donde además de conocer la fisiopatología neoplásica, permite predecir de manera precoz una situación potencial de mayor riesgo de evolución a otras formas más agresivas de la enfermedad como puede ser el síndrome de Richter, y poder atajarlo con terapia dirigida a una clona específica^{16,17}.

Las tecnologías de secuenciación nueva generación han revolucionado el diagnóstico genético de la CLL⁸. Por ello, en este estudio experimental se obtendrá DNA a partir de muestras de tejido parafinado y sangre periférica y se estudiarán las mutaciones que afecten al gen *TP53*.

Por todas estas características, la NGS es un procedimiento idóneo para abordar la complejidad mutacional de *TP53*, ya que permite la detección de alteraciones de tipo somático presentes en bajos porcentajes en las muestras gracias al multiplexado. Otra característica por la que destaca esta nueva técnica es porque permite el estudio de un gran número de muestras en menor tiempo y coste^{16,17}. Es por esto último y porque el panorama genómico de la CLL es de notoria importancia pronóstica, que las técnicas de NGS comenzarán a desplazar los métodos antiguos de estudio genómico como las técnicas citogenéticas moleculares o la secuenciación de primera generación^{8,17}.

Objetivos

En este estudio experimental se extraerá, purificará y amplificará el material genético de muestras guardadas en archivo en tejido parafinado y muestras frescas de sangre periférica de pacientes diagnosticados de leucemia linfática crónica para analizar y cuantificar mutaciones en la secuencia del gen *TP53* entre los exones 4 y 10 mediante el uso de técnicas de secuenciación de nueva generación.

Material y métodos

Se seleccionaron pacientes diagnosticados de leucemia linfática crónica entre febrero de 2016 y junio de 2018 mediante el buscador Snomed OpenLab haciendo uso del protocolo de búsqueda bajo el código M 9823/3. Se obtuvieron un total de 60 pacientes diagnosticados de CLL con biopsias recogidas de médula ósea y ganglios linfáticos.

Se seleccionó dentro del total de pacientes a aquellos que no había sido por punción de médula ósea dado que las técnicas para la desmineralización alteran el DNA. De esta forma, se completó una n =16 de los cuales n=11 de ganglios linfáticos, n=1 de cavum, n=1 de pleura, n=1 de órbita, n=1 de mama y n=1 de hígado.

Dado los resultados posteriormente descritos, se amplió la búsqueda hasta febrero de 2019 y se seleccionaron 5 muestras recogidas de sangre periférica de pacientes diagnosticados de CLL. De esta manera, el tamaño muestral final refleja un total de n=21 pacientes.

Con los códigos de muestra asociado a cada paciente, se buscó en archivo las muestras parafinadas y se seleccionaron los cortes con mayor cantidad de tejido para la extracción del DNA mediante la técnica protocolizada de laboratorio denominada purificación. Se trata de un proceso mediante el cual se extrae DNA a partir de muestras de tejido parafinado haciendo uso de técnicas físico-químicas. Para ello, se hace uso del kit de purificación “Easy-DNA™ gDNA Purification Kit de Invitrogen”. El mismo proceso fue aplicado posteriormente al DNA de las nuevas muestras de sangre periférica seleccionadas.

El proceso por el cual a partir de una muestra de archivo en parafina se obtiene DNA se divide en dos fases claramente definidas. En primer lugar, la denominada lisis celular. En esta primera fase se debe romper la membrana celular y los componentes del citoplasma, como proteínas, las cuales se deben desnaturalizar. En segundo lugar, se busca obtener el DNA ya separado del resto de elementos celulares. El conjunto del proceso descrito es denominado purificación, cuyo objetivo final es conseguir DNA de buena calidad, y además una cantidad suficiente para el estudio.

La primera fase, la lisis celular, explica que antes de purificar el DNA se deben eliminar otros elementos celulares. En primer lugar, hay que fundir la parafina de la muestra durante 1 hora. Posteriormente hay que desparafinar e hidratar con Xilol durante 5 minutos dos veces, y realizar la siguiente secuencia: Etanol 100° durante 5 minutos, Etanol 96° durante 5 minutos, Etanol 70° durante 5 minutos y finalmente lavar con agua destilada. Se debe delimitar la zona de tumor en microscopio y marcar el área tumoral con rotulador permanente. Con una punta desechable se descarta la zona no tumoral y se recoge en tubo de 1,5 ml el material tumoral pipeteando con TENS, hasta 500µl.

Para comprender lo que ocurre en la fase de lisis celular es necesario interiorizar el efecto de cada proceso químico sobre el DNA. En esta fase se hace uso de dos sustancias, el TENS y la proteinasa K. El primero de ellos está compuesto por cuatro componentes: 1) TRIS HCL: un tamponador que asegura que el pH sea el adecuado; 2) EDTA : es un agente quelante que inhibe a las nucleasas de DNA uniéndose a iones metálicos (Mg^{2+}); 3) NaCl, que asegura la osmolaridad; 4) SDS: actúa como detergente, rompiendo la membrana lipídica. En segundo lugar se añade la proteinasa K, una exo y endoproteasa que se encarga de degradar proteínas y nucleasas.

La segunda fase comienza tras una hora en la que se calientan las muestras a 90°C. Se añade Fenol-Cloroformo 500µl, y se mezcla bien hasta que ambas fases sean homogéneas. Posteriormente, se centrifuga a máximas r.p.m. durante 10 minutos. Acto seguido se pasa a fase acuosa a un tubo de 2ml, con precaución para no derramar la interfase y se añade 50 µl de AcNa y 1500 µl de Etanol abs. Tras ello, se deja 1 hora a -20°C.

Posteriormente, se centrifuga a máximas r.p.m. 15 minutos y se tira el sobrenadante, se añaden 500 µl de EtOH 70° y se centrifuga durante 5 minutos. Se descarta el sobrenadante, se drenan los tubos en papel, se quitan las gotas de etanol sobre de la pared del tubo y se deja secar 5-10 min. Tras ello, se suspende el fondo de la muestra en 50ul de TE.

Finalmente, hay que cuantificar la cantidad de DNA ya purificado que hay incluido en cada tubo, que se corresponde a cada muestra de los diferentes pacientes. Para

ello se hace uso del programa NanoDrop. En el total de las 21 muestras, se ha estudiado la cantidad de DNA expresado en ng/ μ l, cuyos valores se ven expresados en el *anexo 1*.

Una vez realizada la purificación del DNA, se entra en la segunda fase del experimento: antes de comenzar, es necesario equiparar mediante la adición de DNA SD de manera que se disuelva los nanogramos de DNA a un rango deseado y de esta forma poder trabajar con las mismas concentraciones de DNA en todas las muestras. En el caso de este ensayo, se utilizó como valor objetivo una concentración de DNA de entre 80 y 150 nanogramos por microlitro. Dichos valores pueden ser comprobados también en el *anexo 1*.

Para que el estudio de las mutaciones de un exón sea posible, se precisa amplificar las regiones que serán analizadas. Para ello, se procederá a la realización de la reacción en cadena de la polimerasa (PCR) de las muestras cuyas concentraciones han sido equiparadas anteriormente. El objetivo final de este procedimiento es obtener un gran número de copias de DNA a partir del fragmento extraído de cada muestra de los diferentes pacientes. El material de DNA purificado y diluido de las muestras de los 21 pacientes diagnosticados de CLL se amplificará gracias al uso de cebadores específicamente diseñados para las regiones diana, es decir, aquellas donde se buscará cuantificar las mutaciones.

Para el diseño de los cebadores se hace uso de dos programas de bioinformática: “Ensembl” y “Primer 3”^{2,18}. En primer lugar, se extrae la secuencia de intrones y exones del gen *TP53* de Ensembl. Dicha página arroja diversos transcritos para el mismo gen. En este estudio se trabajó con el *TP53_201* ya que mostraba buena calidad de información extraídos a partir de diferentes nodos de bioinformática fiables.

Sobre dicho transcrito se trabaja y se decide, tras la revisión de la bibliografía, cuáles son las regiones del gen con mayor acúmulo cuantitativo de mutaciones de *TP53* en pacientes diagnosticados de CLL. Para este estudio se ha decidido cuantificar mutaciones en los exones 4-10 del gen porque, tal como se redactó en la introducción, el 90% de mutaciones de *TP53* descritas en pacientes con CLL se hallan en dichas regiones³. Los tipos de mutaciones y la región en la que se encuentran, podrán ser cuantificadas y analizadas más adelante en los resultados del experimento.

En segundo lugar, tras haber decidido las regiones a estudiar (exón 4-10) se selecciona la longitud (definida en pares de bases) deseada que se va a flanquear entre cada pareja de cebadores. Para ello, se establecieron los límites entre 170 y 270 kB para cada exón y sus intrones colindantes. Posteriormente, se selecciona dicha región y haciendo uso de “Primer 3” se obtienen los cebadores específicos para la secuencia deseada. Los cebadores que se han diseñado así como la región que flanquean se puede ver en el *anexo 2*.

Una vez diseñados, se envía el archivo a la empresa Suiza “Mycrosynth AC” la cual sintetizará los cebadores específicos (tanto el forward como el reverse) con los que se trabajará la amplificación de las regiones diseñadas.

Cuando se reciben las parejas de cebadores ya sintetizados, se procede a iniciar la PCR. El protocolo de PCR comienza por la mezcla de 15 microlitros de MasterMix, 1 microlitro de DNA de muestra parafinada y 0,5-1 microlitro del cebador. Esta mezcla irá en cada uno de los tubos que se desea amplificar.

Finalmente, se sella la parte superior de los tubos con cinta y se inicia la PCR con el programa hot55_30 o 58HOT. Como se demostrará más adelante, el primer protocolo de PCR arrojó amplificaciones débiles de baja calidad, por lo que tras diversas pruebas se optó por el programa 58HOT. El nuevo proceso de PCR consiste en 30 ciclos en los que se expone el material a diversas temperaturas, aquellas que son óptimas para que exista hibridación y por tanto amplificación. Concretamente, el protocolo seguía los siguientes pasos: en primer lugar, se eleva la temperatura a 95° C durante 15 minutos. Posteriormente ocurren tres pasos cortos de 20 segundos, todos de diferentes temperaturas: 95° C, 58° C y 72° C respectivamente. Al finalizar, se volverá al paso número 2 (temperatura de 95° C durante 20 segundos). Estos tres pasos se repetirán 29 veces más. Finalmente, tras los 30 ciclos nombrados, se mantendrá durante 5 minutos una temperatura de 72° C y, tras este, se procede a un último paso que consiste en el enfriamiento a 18° C.

Al finalizar la PCR se comprueba si la amplificación ha sido correcta mediante una electroforesis. Para ello, se prepara un gel con agarosa al 2,5%. Para la preparación

del gel, se añade 40 mililitros de tampón y 0,8 gramos de agarosa. Se debe llevar a ebullición y agitar hasta que la agarosa entre en solución y añadir finalmente 15 microlitros bromuro de etidio. Posteriormente, se vierte el líquido en una superficie plana y se deja reposar durante 20 minutos con los moldes de los pocillos introducidos. Tras 20 minutos se obtiene un gel, que contará con diferentes pocillos en cuyo interior se añadirá: en el primero de la fila, un DNA control denominado “Puc Hae III”; en los siguientes pocillos se irá añadiendo 3 microlitros de los productos obtenidos de la PCR. El tiempo de electroforesis empleado son 20 minutos a 120 voltios.

Para comprobar los resultados arrojados tras la primera purificación y PCR se introduce el gel resultante de la electroforesis en la máquina “Gel Doc EZ Imager de Bio-Rad” y gracias al programa “Image Lab” se obtiene una imagen en negativo del gel, donde están representadas las bandas de amplificación de los diferentes cebadores con el DNA a estudio.

Tras seleccionar las muestras que han amplificado de manera óptima se procederá a la tercera fase, aquella que permitirá crear las librerías para enviar a secuenciar el material a estudio. A partir de las 21 muestras iniciales, se ha seleccionado un total de 12, correspondientes con aquellas cuya nomenclatura es 526, 528, 539, 545, 546, L6, L9, L10, L11, L12, L13 y L16.

Las librerías son el resultado del conjunto de una serie de fragmentos de ADN (en este estudio, el material genético resultante de la amplificación) a cuyos extremos se une una pareja de oligonucleótidos, llamados adaptadores. Para poder crear las librerías de las muestras que serán posteriormente secuenciadas es necesario seguir los siguientes pasos: 1) Purificación del material; 2) Cuantificación de la concentración de DNA de cada exón tras la purificación; 3) Estudio de los geles de electroforesis y concentraciones de DNA de cada tubo; 4) Unión de los diferentes exones en un mismo tubo; 5) Preparación librerías mediante la adición de las colas 3’ y 5’ para la NGS ; 6) Amplificación por PCR; 7) Electroforesis digital de los resultados con “Agilent D1000 ScreenTape System Quick Guide”; 8) Enviar las librerías listas para la secuenciación de nueva generación.

En primer lugar, se ha de volver a purificar cada exón de las diferentes muestras para eliminar los dímeros de cebadores que han quedado de la primera PCR. Para ello, se hace uso del protocolo de purificación de SequalPrep, de la compañía Invitrogen.

En una plantilla de 96 pocillos (12 muestras de diferentes pacientes x 8 exones a estudiar de cada uno) se introduce todo el material de PCR de cada muestra, cuya cantidad resultante fue de 13µl por cada pocillo. Posteriormente, en cada uno de los pocillos se pipetea la misma cantidad (13µl) de “SequalPrep™ Binding Buffer” y se mezcla bien pipeteando arriba y abajo. Una vez realizado, se deja incubando durante una hora a temperatura ambiente. Tras ello, se pasa a la fase de lavado. Se aspira el contenido de cada tubo procurando no arañar las paredes y se descarta. Luego, se añaden 50 µl de “SequalPrep™ Normalization Wash Buffer” y se mezcla pipeteando arriba y abajo. Acto seguido, se aspira completamente el contenido, procurando que el interior de cada tubo esté lo más seco posible. Finalmente, se pasa a la fase de elución. La fase comienza por añadir 20 µl de “SequalPrep™ Normalization Elution Buffer” en cada tubo de la plantilla. Se mezcla bien pipeteando arriba y abajo y se deja incubando durante 5 minutos. Mientras transcurre dicho tiempo, se prepara una nueva plantilla de 96 pocillos donde transferir todo el contenido final de la fase de purificación. Una vez finalizado, el producto puede ser guardado a 4°C para continuar con el paso 2) Cuantificar la concentración de DNA de cada tubo tras la purificación.

Para normalizar la concentración de DNA entre las diferentes muestras es principalmente necesario cuantificarlo mediante un método fluorimétrico, como Qubit. Para ello, Con el objetivo de optimizar los recursos se han cuantificado dos exones por muestra. En base a las recomendaciones, y con la finalidad de minimizar errores, se realizaron dos mediciones para cada muestra, los resultados de esta medición se pueden ver en el *anexo 3*.

Teniendo en cuenta la distribución de las diferentes concentraciones de DNA de cada muestra y los geles de la primera electroforesis, se inicia el paso 3) Estudio de los geles de electroforesis y concentraciones de DNA de cada tubo y 4) unión de todos los exones de un mismo paciente en un tubo. En estos pasos se decidirá, en primer lugar la cantidad (en µl) de los diferentes exones a estudio que se mezclara en un mismo tubo. Finalmente, tras el estudio concienzudo de los geles, se decidió pipetear 5 µl de los exones

4a, 5, 7 y 8; 7 μ l del exón 6 y 10 μ l de los exones 4b, 9 y 10. La decisión se basó en el grado de amplificación de cada exón visto en los geles de electroforesis, siendo muy buena, buena y regular respectivamente para cada uno de los tres grupos.

En segundo lugar, se pasa a la preparación de las librerías, la final de la fase de experimentación. La preparación de las librerías por parte del investigador es el primer paso para la posterior secuenciación masiva. En esta fase, se reciben los códigos de barras específicos para NGS que son sintetizados por cada casa comercial, en el caso de nuestro experimento, los adaptadores los suministra Illumina, y tienen en torno a 8-12 pares de bases cada uno. En una plantilla como representa el *anexo 4* se preparan las librerías para la segunda PCR. En cada tubo se añaden: 20 μ l de MasterMix y 0,5 μ l de la mezcla de códigos de barra (forward y reverse) y finalmente 5 μ l de DNA de cada muestra (que contiene todos los exones a estudiar).

El programa de PCR elegido se nombró 55_65. Durante un tiempo de 60 minutos realizó el total de la PCR que consta de: cuatro primeros pasos que se repiten tres veces, para posteriormente finalizar con dos pasos de 20 segundos cada uno. Los primeros cuatro pasos consisten en: calentar a 95° C durante 15 minutos, una vez finalizado, una transición de 20 segundos hasta 55° C que se mantiene durante 20 segundos. Tras este, se aumenta a 72° C durante 20 segundos y se pasa al paso número uno. Finalmente, tras tres ciclos se avanza a elevar la temperatura a 95° C durante 20 segundos y acaba el programa descendiendo a 65° C durante 20 segundos.

Para el análisis de los productos de PCR se preparó un gel de agarosa al 2,5% y se realizó un programa de electroforesis a 120 V durante 20 minutos. La imagen del gel se obtuvo, tal como en la primera PCR, mediante el “Gel Doc EZ Imager de BioRad”.

Tras comprobar en el gel de electroforesis que la PCR ha sido efectiva, se vuelve a purificar el DNA resultante antes de enviar definitivamente las muestras a secuenciar. A partir de los 20 μ l resultantes en cada uno de los tubos se inicia una serie de procesos haciendo uso de imanes y sustancias de lavado, obteniendo finalmente un producto purificado.

Una vez se purifica el DNA, se realiza una electroforesis digital, haciendo uso de la máquina “Agilent D1000 ScreenTape System”. Esta es capaz de realizar el análisis de cualquier DNA que tiene 35 y 1000 pares de bases de longitud, y lo hace con una sensibilidad muy alta, pues es capaz de detectar a partir de 0,1 ng/μl de concentración de DNA. Previamente al inicio, se debe preparar cada muestra de la siguiente manera: durante 30 minutos se lleva a temperatura ambiente cada muestra y se mezcla con vortex. Se lleva a otro tubo 3 μl de “D1000 Sample Buffer” y 1 μl de DNA. Se mezcla bien a 2000 r.p.m. durante un minuto y se carga en la máquina “2200 TapeStation”, donde durante aproximadamente 20 minutos se efectuará la electroforesis.

Cuando las librerías están preparadas, son enviadas al instituto de tecnologías renovables donde serán secuenciadas por técnicas de nueva generación, donde se realizarán los siguientes tres pasos: 1) Generación de grupos, 2) Secuenciación y 3) Análisis de los datos.

El fundamento por el cual mediante la secuenciación masiva se obtienen datos sobre las regiones deseadas a estudiar es gracias a una polimerasa que cataliza la incorporación de desoxirribonucleótidos trifosfato al DNA durante diversos ciclos de síntesis. En cada ciclo, en el momento de la incorporación, los diferentes nucleótidos son identificados por la fluorescencia que emiten, por lo que se obtiene una secuencia de nucleótidos basándose en dicho método. Gracias a las NGS, se puede realizar este proceso de manera paralela en millones de fragmentos de DNA. Una vez finaliza la secuenciación, las diferentes lecturas son alineadas y analizadas en busca de alteraciones del DNA tales como polimorfismos de un nucleótido, inserciones o deleciones, siendo el estudio y cuantificación de todos ellos el objetivo de este trabajo experimental.

Resultados:

La primera prueba, reflejada en la *figura 1* se realizó con DNA de alta calidad, para comprobar que el procedimiento de amplificación sea óptimo. Los resultados arrojados de la primera prueba fueron positivos, confirmando una amplificación del DNA del tamaño esperado con cada pareja de cebadores. Después se decidió repetir el proceso con DNA purificado de bloques de parafina de los diferentes casos de LLC seleccionados para el estudio. Los geles de las *figuras 2 y 3* muestran la variabilidad de la calidad de las

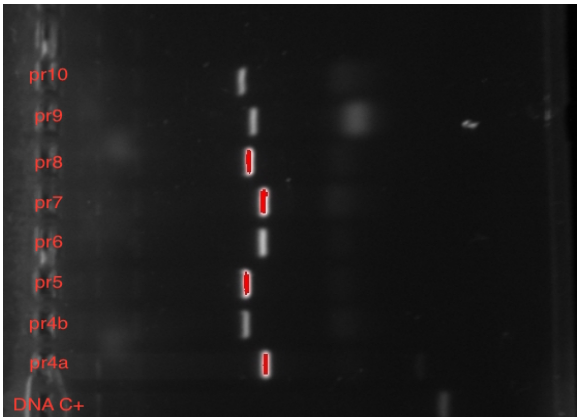
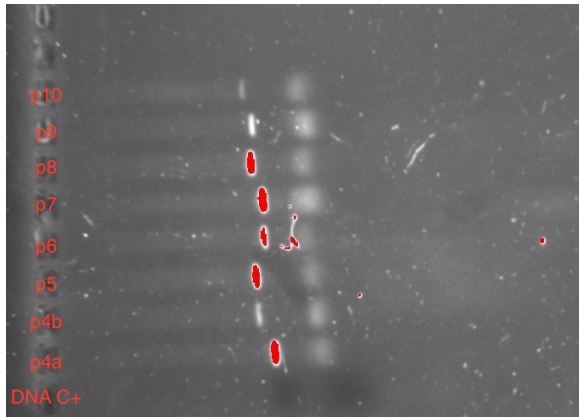
amplificaciones conseguidas, a pesar de haber mantenido las condiciones de amplificación sin cambios con respecto al control que se usó para poner a punto la técnica.

Para comprobar que no se trataba de un error de técnica se decidió probar una nueva amplificación con DNA control positivo, *figura 4*, obteniendo una amplificación correcta del DNA para cada pareja de cebadores.

Al no tratarse de un error de procedimiento, se optó por una nueva purificación del DNA de las muestras. Las seleccionadas fueron las muestras que previamente habían arrojado amplificaciones débiles y, dentro de ellas, las que mayor concentración de DNA se extrajo a partir de la primera purificación del DNA de tejido parafinado. Finalmente se optó por volver a purificar L9, L12 y L13. Tras el proceso de purificación y la nueva PCR se obtuvieron los resultados reflejados en la *figura 5*. Como se puede comprobar, la nueva purificación no resultó efectiva y en el gel se obtuvieron amplificaciones muy débiles y dímeros de cebadores.

Por tanto, tras demostrar que el problema no se encontraba ni en la purificación ni en el procedimiento, se plantearon dos nuevas incógnitas: será problema de la calidad del DNA de las muestras o será problema de las temperaturas y el número de ciclos de PCR elegido. Para resolver ambas, se realizaron dos nuevos experimentos: en primer lugar, se seleccionó DNA control de alta calidad y se amplificó mediante PCR a diferentes programas: 1) programa a 58° C; 2) programa a 54° C; y 3) programa a 52° C. Los resultados obtenidos en esta primera prueba fueron llamativos; como muestra la *figura 6*, la amplificación del DNA con el programa a 58° C resultó más eficaz que las otras dos.

Como consecuencia de la prueba anteriormente descrita, se optó por seguir el experimento con ese protocolo de PCR. Sin embargo, la segunda incógnita seguía sin resolverse, por lo que se decidió realizar un experimento con DNA extraído de una muestra de plasma de un paciente con CLL para ver si la calidad de este es superior a las que son extraídas de tejido parafinado guardado en archivo. La *figura 7* muestra una amplificación óptima según lo esperado, por lo que se puede concluir que resultó correcto y positivo para el curso del experimento tanto el cambio de programa de PCR como la mayor calidad de DNA extraído de la muestra de sangre.

<p><i>equiparadas junto a los cebadores de los exones 4a, 4b y 5.</i></p>	
	
<p><i>Figura 7: Fotografía del gel de electroforesis. DNA obtenido de sangre periférica junto con 8 cebadores de los exones para TP53 y PCR a 58°C.</i></p>	<p><i>Figura 8: Fotografía del gel de electroforesis. DNA de L12 junto con 8 cebadores de los exones para TP53 y PCR a 58°C.</i></p>

Para avanzar con el proyecto, decidimos abandonar las muestras con nomenclatura: L1, L2, L3, L4, L5, L7, L14 y L15 puesto que repetidos intentos de amplificación resultaron en bandas débiles, muchos dímeros de cebadores e inconsistencias entre experimentos. Intentos de variar la cantidad de DNA inicial que utilizamos para la amplificación (entre 50 y 500 ng/μL por tubo) no resultaron eficaces para solucionar las inconsistencias en las amplificaciones.

Los resultados obtenidos por las muestras: L6, L9, L10, L11, L12, L13, L16 y el DNA purificado a partir de muestras de sangre periférica guardada en archivo: 526, 528, 539, 545 y 546, fueron más consistentes por lo que el experimento continuó con todas estas muestras. Tras las PCR para cada una de las muestras, se obtuvieron amplificaciones del tamaño esperado para cada pareja de cebadores, similares a lo que muestra la *figura 8*.

Una vez se seleccionaron las muestras para continuar con el experimento, se volvieron a purificar y a cuantificar, tal como se describió en métodos. Los resultados arrojados tras la segunda cuantificación se pueden analizar en el *anexo 3*. En base a estos y a los geles resultantes de la primera PCR se decidió qué cantidad de material de cada

exón iba a ser pipeteado en el tubo final, el cual contendría todos los exones (que se han decidido estudiar en este experimento) de un mismo paciente.

Tras seguir el protocolo de Illumina para realizar la preparación de librerías, se realizó la segunda PCR. Los resultados arrojados se pueden ver en la *figura 9*. Se concluye que tanto la purificación como la amplificación fueron correctas, confirmándose una amplificación uniforme, consistente y del tamaño deseado para cada una de las muestras, así como la ausencia de dímeros de cebadores.

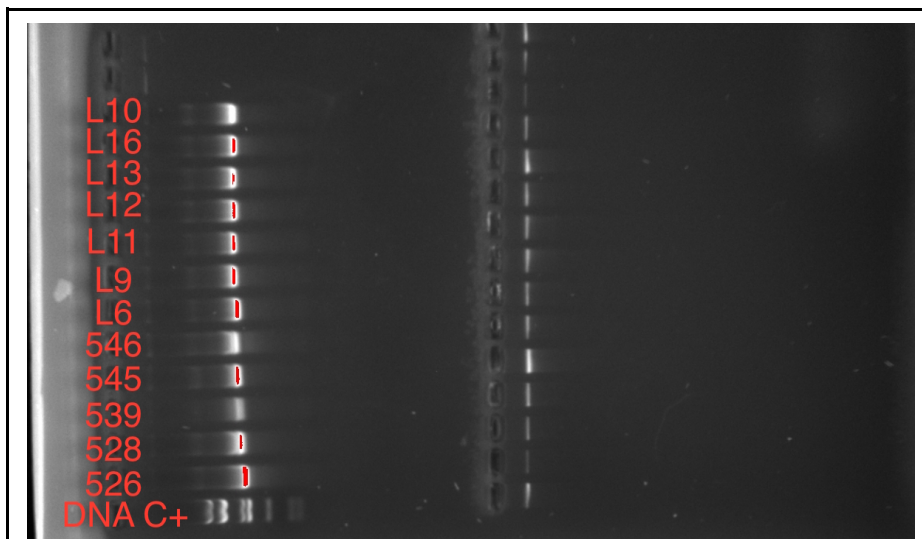


Figura 9: Fotografía del gel de electroforesis. Segunda PCR tras purificación.

A partir de los resultados positivos de la anterior electroforesis, se realizó una purificación final, con el objetivo de eliminar los dímeros de cebadores. Tal como se describió en métodos, tras dicha purificación, se cuantificaron los resultados haciendo uso de la electroforesis digital.

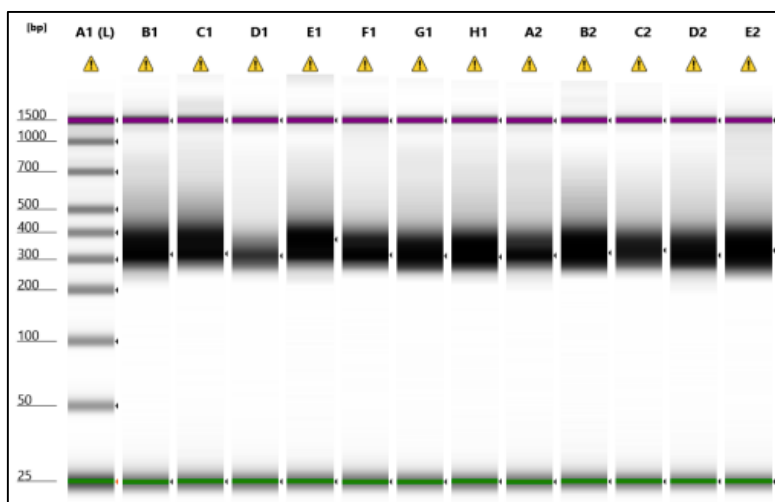


Figura 10: Electroforesis digital. Amplificación para todas las muestras en torno a 320 pares de bases (bp).

Las gráficas obtenidas, resultaron similares para todas las muestras de pacientes, teniendo la curva una disposición similar con un pico medio en torno a las 325 pares de bases, tal como se puede observar en la figura 11.

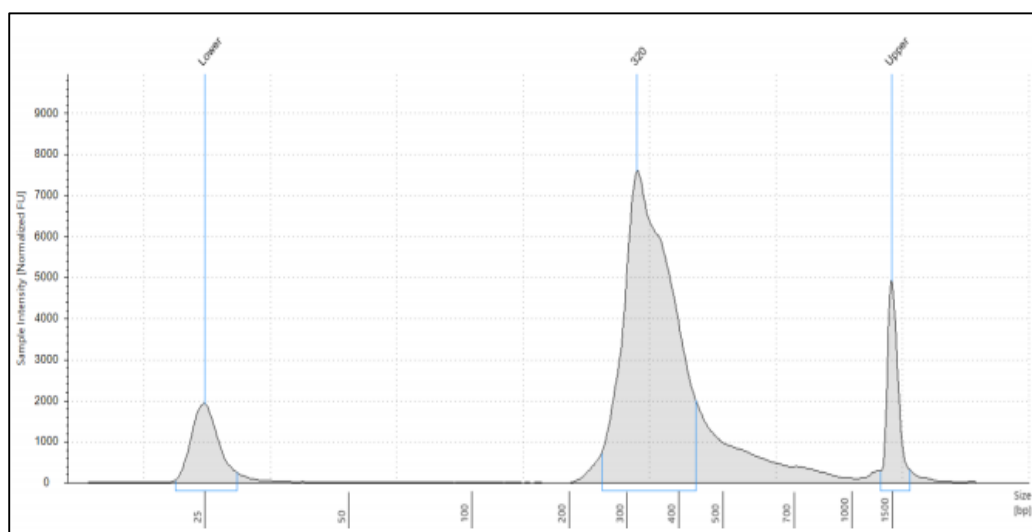


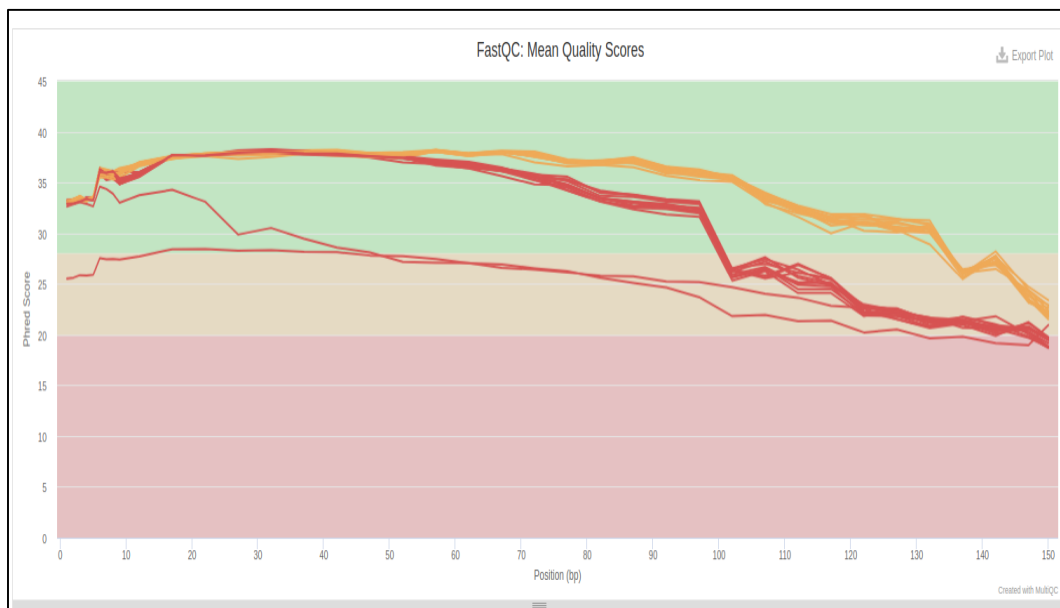
Figura 11: Gráfico de la muestra con nomenclatura 526. Pico en 346 bp. Todos los otros gráficos resultantes guardan curvas similares a la expuesta.

Finalmente, cuando se consiguió que las amplificaciones hayan resultado homogéneas en cuanto a longitud en pares de bases, se enviaron las muestras al Instituto de Tecnologías Renovables para que sean secuenciadas. Junto a las muestras, se les envía un archivo de texto plano en el cual se delimita dónde comienza y acaba cada secuencia

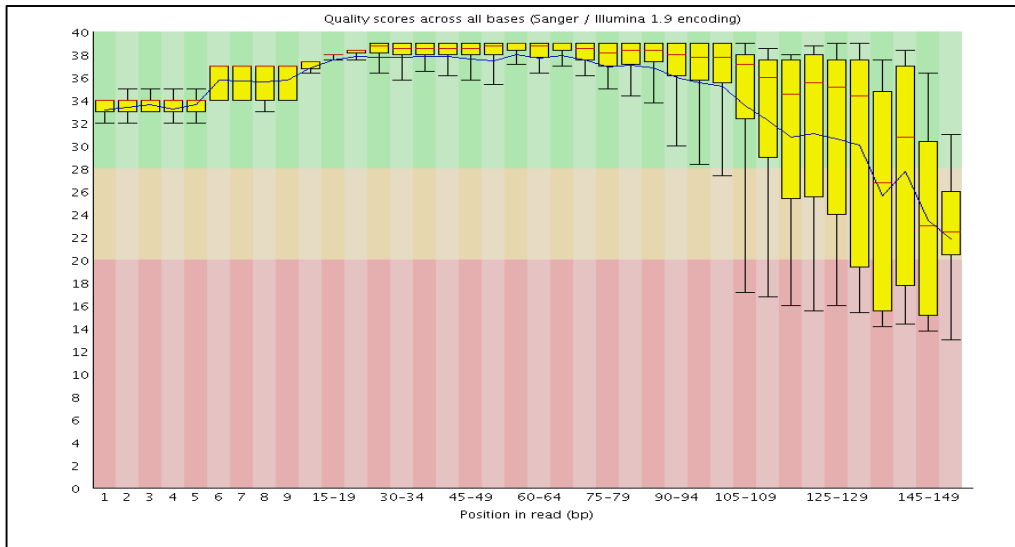
intrón-exón-intrón a estudio en base a su posición en el genoma humano. Esto es llamado “Bedfile”, y se realiza con la herramienta bioinformática “Ensembl”². Los resultados pueden verse en el *anexo 6*.

Una circunstancia relevante en la siguiente fase del proyecto ha sido el desacople de tiempos en la realización del mismo. Conseguimos financiar el alto coste de los reactivos de NGS para los meses de verano del 2018, pero no fue hasta marzo de 2019 cuando tuvimos las librerías listas para secuenciar. Consecuentemente, tuvimos que usar kits de secuenciación que ya habían caducado. Es conocido, que con el tiempo los reactivos de secuenciación NGS pierden eficacia y se consiguen carreras más cortas. No obstante, no teníamos más alternativa que seguir adelante porque la inversión económica del proyecto (2.000 €) ya estaba ejecutada.

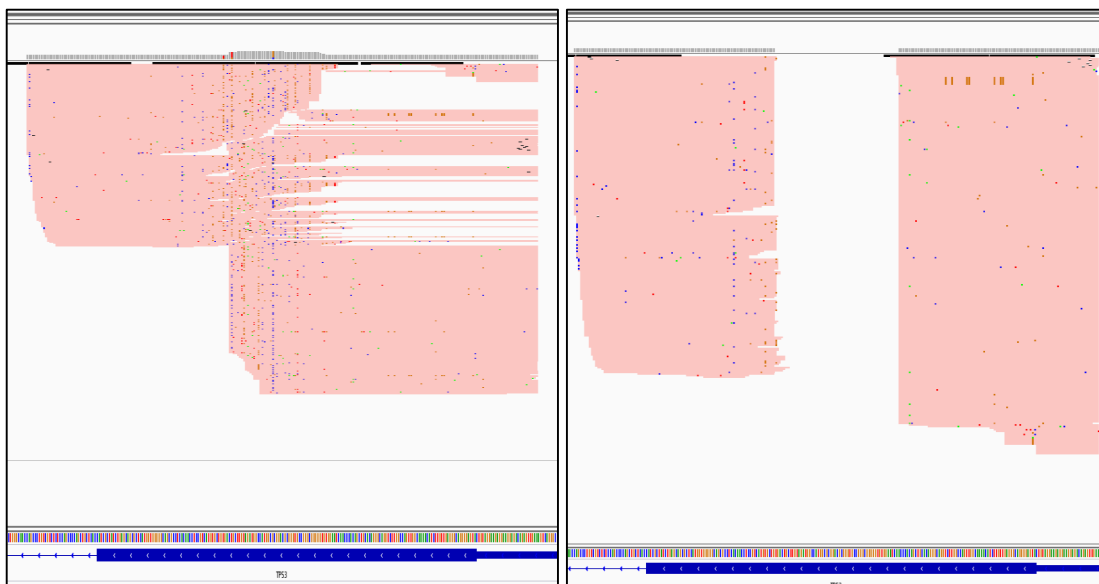
El gráfico inferior muestra las calidades medias por base en las lecturas generadas. Se aprecian caídas en la calidad aproximadamente a partir de la base 95.



Al analizar con más detalle, tal y como se muestra en la figura inferior (se muestra la lectura en sentido forward para una de las muestras; empeora incluso si miramos la lectura en sentido reverse), las calidades por base a partir de la base 90-95 caen drásticamente. Esto no es normal en Illumina, lo que achacamos que el kit está caducado. No afecta a distancias cortas, pero empieza a notarse a distancias más largas.

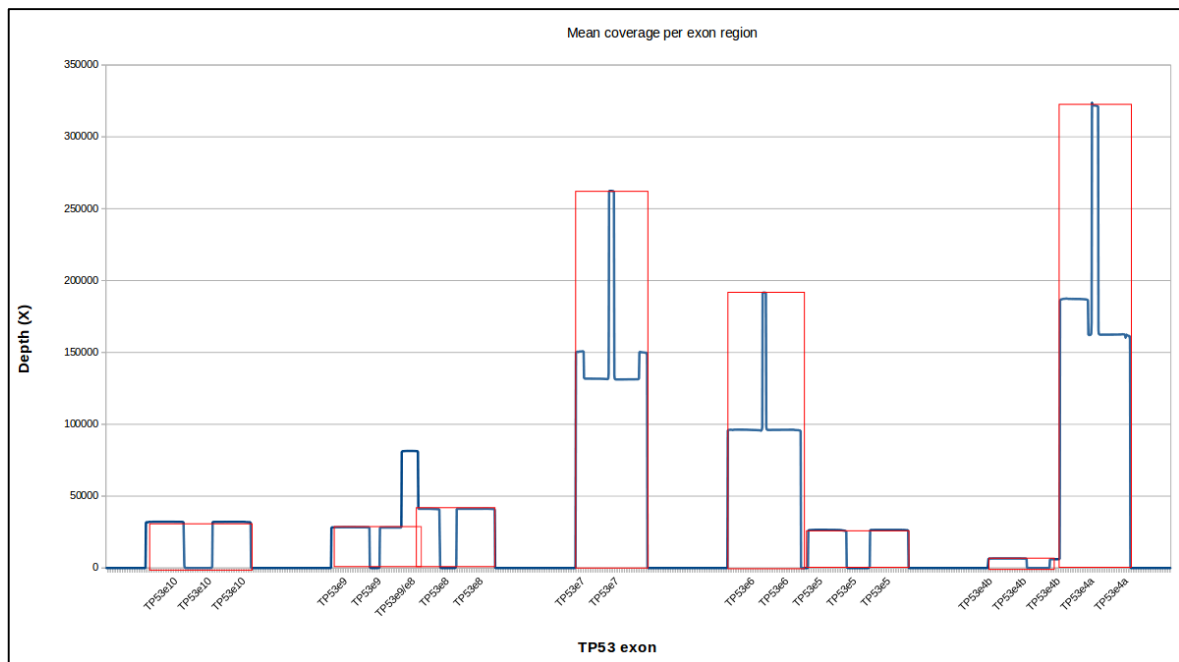


Abajo se muestran las lecturas alineadas al fragmento del exón 5 de *TP53* de una de las muestras. Datos crudos en la izquierda, y datos después de recortar a partir de base 95 en la derecha. Tras el recorte, se aprecia que las regiones centrales de los exones no resultan cubiertas.



Si representamos la cobertura de cada uno de los exones en la reacción de secuenciación, destaca una gran heterogeneidad. El "skyline" azul es el obtenido, como consecuencia de la mala calidad de lectura y mapeado. El "skyline" rojo es el teórico, que

nos da una idea de la cobertura a esperar con el diseño y con las concentraciones de carga actuales.



Hay ciertos exones (4a, 6 y 7) que tienen mucha más cobertura que los demás. En la gráfica anterior se puede comprobar como dichos exones obtienen un gran número de copias en la secuenciación frente a la poca amplificación que ocurre en el resto de exones. Para solucionar la heterogeneidad que se mostró en la secuenciación existente entre los diferentes exones se necesita de mejoras en la normalización/pooling.

Con estos resultados nos planteamos probar un diseño diferente a continuación, pero ya no tendremos tiempo de incluirlo en el presente trabajo de fin de grado. Para futuros proyectos, vamos a cambiar a un genotipado basado en "sondas invertidas" (MIPs: molecular inversion probes) que han demostrado conseguir mayor homogeneidad en la cobertura de las zonas de interés¹⁹.

Discusión

La bibliografía consultada permite concluir que determinación de mutaciones en el gen TP53 es beneficiosa para el mejor manejo de los pacientes con LLC. Por tanto, tiene sentido que los departamentos diagnósticos de hospitales como el nuestro se

preparen para ser capaces de determinar estas mutaciones de manera eficiente. Puesto que la secuenciación actualmente a nuestro alcance (tipo Sanger) no es eficiente para el rastreo de mutaciones somáticas en múltiples exones, se hace necesario adquirir la capacidad de realizar secuenciación de nueva generación (NGS). Este reto tiene dificultades que hemos comprobado durante la ejecución del proyecto:

1) Las muestras de archivo disponibles (tejido fijado en formaldehído e incluido en parafina) proporcionan DNA de baja calidad, que dificulta la amplificación homogénea de todos los exones de manera reproducible y con rendimientos similares entre distintos pacientes. Este hecho hace muy complicado conseguir librerías de fragmentos que tengan coberturas buenas a todo lo ancho de la zona a secuenciar y que dejen a todos los pacientes representados por igual en reacciones de secuenciación multiplexadas. Mejores resultados se obtienen con DNA de sangre periférica, pero esto limita la detección de mutaciones a circunstancias con buena representación neoplásica en sangre periférica.

2) El tamaño de los amplicones que se puede secuenciar de manera fiable debe recortarse por debajo de unos 150 bp ya que el rendimiento de los kits de secuenciación de 300 bp (150x2) cae drásticamente con el tiempo de almacenado del kit. También es más complicada la amplificación homogénea de fragmentos próximos a 300 bp a partir de DNA obtenido de bloques de tejido en parafina.

3) El procedimiento de doble amplificación por PCR (nested PCR) para preparar la librería tiene el principal inconveniente en la dificultad de conseguir amplificaciones homogéneas que permitan una cobertura uniforme de la zona a secuenciar, retrasando el tiempo hasta la creación de la librería y ampliando la posibilidad de que el kit pierda eficacia y los resultados de la secuenciación no sean valorables.

Este proyecto se continuará con la implementación de un diseño tipo MIPs, del que esperamos conseguir librerías idóneas para el genotipado de mutaciones en *TP53* en muestras de tejido incluido en parafina.

Conclusiones

- El estado mutacional del gen *TP53* ha mostrado ser determinante para los pacientes diagnosticados de leucemia linfática crónica en cuanto a respuesta al tratamiento, evolución de la enfermedad y pronóstico.
- La inclusión de pruebas diagnósticas de nueva generación en laboratorios como el Departamento de Anatomía Patológica del HUC puede ser útil para predecir la evolución de las enfermedades oncohematológicas.
- Se debe limitar la longitud de los amplicones para obtener una secuenciación satisfactoria y fiable.
- Para lograr demostrar la hipótesis del estudio, se requiere continuar con la experimentación a partir de los datos recogidos en este proyecto, principalmente mediante la implementación de un diseño tipo MIPs.

¿Qué he aprendido de este TFG?

Gracias a la implicación de mi tutor y a la posibilidad de acceder a la experimentación, haciendo uso de procedimientos y protocolos de trabajo con material genético he podido comprobar la alta complejidad y el alto grado de esfuerzo necesario para llevar a cabo una investigación desde el inicio. El hecho de poder comenzar con una hipótesis e ir a buscar su justificación mediante la experimentación hace que incontables horas de trabajo sean requeridas.

A pesar de no obtener correctos resultados, existe una línea de investigación ya abierta gracias a este TFG que seguirá buscando su camino hasta poder encontrar una resolución final satisfactoria y demostrar la eficiencia de la secuenciación en la práctica clínica diaria.

Bibliografía

1. M. López, M. Anzola, N. Cuevas-Salazar, J.M. Aguirre, M. Martínez de Pancorbo. p53, un gen supresor tumoral. *Gac Med Bilbao*. 2001; 98: 21-27.
2. P.J. Kersey, J.E. Allen, A. Allot, M. Barba, S. Boddu, B.J. Bolt, et al. *Ensembl Genomes 2018: an integrated omics infrastructure for non-vertebrate species* *Nucleic Acids Research* 2018; 46(1): 802–808.

3. Bouaoun L., Sonkin D., Ardin M., Hollstein M., Byrnes G., Zavadil J., et al. TP53 Variations in Human Cancers: New Lessons from the IARC TP53 Database and Genomics Data. *Human Mutation*. 2016; 37(9):865-76.
4. Puente X., Beà S., Valdés-Mas R., Villamor N., Gutiérrez-Abril J., Martín-Subero, et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature*. 2015; 526(7574):519-24.
5. Hallek M. Chronic lymphocytic leukemia: 2017 update on diagnosis, risk stratification, and treatment. *Am J Hematol*. 2017; 92 (9):946–965.
6. Medeiros, Miranda, Wang, Vega, Muzzafar, Yin, Bueso-Ramos, Lin. *Diagnóstico en patología: Ganglios linfáticos y bazo con linfomas extraganglionares*. 1ª ed. Texas: Marban; 2014.
7. Vinay Kumar, Abul K. Abbas, Jon.C. Aster. *Patología estructural y funcional*. 9ª ed. Madrid: Elsevier; 2015.
8. Ana E. Rodríguez-Vicente, Vasilis Bikos, María Hernández-Sánchez, Jitka Malcikova, Jesús-María Hernández-Rivas, Sarka Pospisilova. Next-generation sequencing in chronic lymphocytic leukemia: recent findings and new horizons. *Oncotarget*. 2017; 8(41): 71234–71248.
9. Cheng Wang, Xin Wang. The role of TP53 network in the pathogenesis of chronic lymphocytic leukemia. *Int J Clin Exp Pathol*. 2013; 6(7):1223-1229.
10. Edelman J., Gribben J.G. Managing Patients With TP53-Deficient Chronic Lymphocytic Leukemia. *Journal of Oncology Practice*. 2017; 13(6):371-377.
11. Buccheri, Barreto W.G., Fogliatto L.M., Capra M., Marchiani M., Rocha V. Prognostic and therapeutic stratification in CLL: focus on 17p deletion and p53 mutation. *Annals of Hematology*. 2018; 97(12): 2269-2278.
12. Dennis H. Wright, Bruce J. Addis. *Diagnostic Lymph node Pathology*. 2ª ed. EEUU: Hoder Arnold; 2011.
13. Puente X., Pinyol M., Quesada V., Conde L., Ordóñez G.R., Villamor N., et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature*. 2011; 475(7354): 101-5.
14. Valeria Spina, Davide Rossi. Overview of non-coding mutations in chronic lymphocytic leukemia *Molecular Oncology*. 2019; 13(1): 99–106.
15. Steven H. Swerdlow, Elias Campo, Nancy Lee Harris. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*. 4ª ed. Francia: International Agency for Research of Cancer; 2008.

16. Zoni Matta A.C., Martínez López, J. Eficacia de la secuenciación masiva para el diagnóstico y seguimiento de enfermedades oncohematológicas. Madrid: Ministerio de Sanidad, Servicios Sociales e Igualdad. Unidad de Evaluación de Tecnologías Sanitarias de la Comunidad de Madrid. 2015. 680-17-030-8.
17. Landau D.A., Carter S.L., Stojanov P., McKenna A., Stevenson K., Lawrence M.S., et al. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell*. 2013; 152(4): 714-26.
18. Koressaar T, Lepamets M, Kaplinski L, Raime K, Andreson R, Remm M. Primer3_masker: integrating masking of template sequence with primer design software. *Bioinformatics*. 2018; 34(11): 1937-1938.
19. Cantsilieris S, Stessman HA, Shendure J, Eichler EE. Targeted Capture and High-Throughput Sequencing Using Molecular Inversion Probes (MIPs). *Methods Mol Biol*. 2017;1492: 95-106.

Anexos

Anexo 1:

Nomenclatura que se ha utilizado durante el proyecto para cada una de las muestras de los diferentes pacientes diagnosticados de CLL. Cálculo de la concentración de DNA presente en cada una de las muestras purificadas mediante NanoDrop¹. Equiparación de las concentraciones de DNA para obtener resultados similares de amplificación (quinta columna). Se volvió a purificar las muestras L9, L12, L13 y se equipararon sus concentraciones (sexta y séptima columna).

N°LLC/año	Localización	Nomenclatura	[DNA] (ng/μl) ¹	[DNA] diluido (ng/μl)	[DNA] repurificado (ng/μl)	[DNA] repurificado y diluido (ng/μl)
276/16	Ganglio	L1	85	85		
908/16	Cavum	L2	152	116		
15333/16	Mama	L3	200	105		
34410/16	Hígado	L4	99	118		

36515/16	Ganglio	L5	549	123		
36531/16	Ganglio	L6	1440	117		
37852/16	Órbita	L7	457	107		
276/16	Ganglio	L8	67	103		
1076/17	Ganglio	L9	3087	126	590	89
21909/17	Pleura	L10	2180	137		
24737/17	Ganglio	L11	120	92		
29064/17	Ganglio	L12	2947	145	447	148
6089/18	Ganglio	L13	5291	92	333	142
7584/18	Ganglio	L14	4673	107		
9356/18	Ganglio	L15	192	112		
12990/18	Ganglio	L16	108	127		
526/19	Plasma	526	150			
528/19	Plasma	528	167			
539/19	Plasma	539	173			
545/19	Plasma	545	145			
546/19	Plasma	546	138			

Anexo 2:

Región que flaquea cada cebador para cada uno de los exones a estudio. Todos los cebadores comienzan en el intrón previo al exón a secuenciar y termina en el intrón posterior, a excepción del exón 4 que por su longitud, ha sido dividido en dos. La región destacada en negrita se corresponde con el cebador. Las letras minúsculas significa que dichos nucleótidos forman parte del intrón. Las letras mayúsculas representan el exón.

Primer 4aF	5'-CTA CAC GAC GCT CTT CCG ATC TGA CTG CTC TTT TCA CCC ATC T-3'
Región del exón 4a:	5'... tgactgctcttttcacccatc taacagTCCCCCTTGCCGTCCCAAGCAATGGATGATTTGATGCTGTCCCCGGACGATATTGAACAATGGTTCACCTGAAGACCCAGGTCCAGATGAAGCTCCCAGAATGCCAGAGGCTGCTCCCCCGTGGCCCC TGACCAGCAGCTCCTACACC ...3'
Primer 4aR	5'-CAG ACG TGT GCT CTT CCG ATC TGG TGT AGG AGC TGC TGG TG-3'

Primer 4bF	5'-CTA CAC GAC GCT CTT CCG ATC TGT CCA GAT GAA GCT CCC AGA-3'
------------	---

Región del exón 4b	5'... CCAGATGAAGCTCCCAGA ATGCCAGAGGCTGCTCCCCCGTGGCCCTGCACCAGCAGCTCCTA CACCGGGCGGCCCTGCACCAGCCCCCTCCTGGCCCTGTCATCTTCTGTCCCTTCCCAGAAAACCTA CCAGGGCAGCTACGGTTTCCGTCTGGGCTTCTTGCATTCTGGGACAGCCAAGTCTGTGACTTGCACG gtcagttgcctgaggggctggettcca atgagacttcaatgcctg ...3'
Primer 4bR	5'-CAG ACG TGT GCT CTT CCG ATC TGC CAG GCA TTG AAG TCT CAT-3'

Primer 5F	5'-CTA CAC GAC GCT CTT CCG ATC TTT CAA CTC TGT CTC CTT CCT-3'
Región del exón 5	5'... tttcaactctgtctctctctct cttctctacagTACTCCCCTGCCCTCAACAAGATGTTTTGCCAA CTGGCCAAGACCTGCCCTGTGCAGCTGTGGGTTGATTCCACACCCCGCCCGGCACCCGCGTCCGCG CCATGGCCATCTACAAGCAGTCACAGCACATGACGGAGGTTGTGAGGCGCTGCCCCACCATGAGCG CTGCTCAGATAGCGATGgtgagcagctgggg ctggagagacgacagggctg ...3'
Primer 5R	5'-CAG ACG TGT GCT CTT CCG ATC TCA GCC CTG TCG TCT CTC CAG-3'

Primer 6F	5'-CTA CAC GAC GCT CTT CCG ATC TGC CTC TGA TTC CTC ACT GAT-3'
Región del exón 6	5'... ctctgattcctcactgatt gctcttagGTCTGGCCCTCCTCAGCATCTTATCCGAGTGAAGG AAATTTGCGTGTGGAGTATTTGGATGACAGAAACACTTTTCGACATAGTGTGGTGGTCCCTATGAG CCGCCTGAGgtctctggtttgcaactgggg tctctgggaggggggtaa ...3'
Primer 6R	5'-CAG ACG TGT GCT CTT CCG ATC TTT AAC CCC TCC TCC CAG AGA-3'

Primer 7F	5'-CTA CAC GAC GCT CTT CCG ATC TAG GCG CAC TGG CCT CAT CTT-3'
Región del exón 7	5'... aggggcactggcctcactctt gggctgtgttatctctctagGTTGGCTCTGACTGTACCACCATC CACTACAACACTACATGTGTAACAGTTCTGCATGGGCGGCATGAACCGAGGCCATCCTCACCATCA TCACACTGGAAGACTCCAGgtcaggag gccacttgccaccctgcaca ...3'
Primer 7R	5'-CAG ACG TGT GCT CTT CCG ATC TTG TGC AGG GTG GCA AGT GGC-3'

Primer 8F	5'-CTA CAC GAC GCT CTT CCG ATC TTT CCT TAC TGC CTC TTG CTT-3'
Región del exón 8	5'... tttcttactgctcttctgctt ctcttttctctatctctgagtagTGGTAATCTACTGGGACGGAACA GCTTTGAGGTGCGTGTGTTGTCCTGTCTGGGAGAGACCGGCGCACAGAGGAAGAATCTCCGCAAG AAAGGGGAGCCTCACCACGAGCTGCCCCAGGGACACTAAGCGAGgtaagcaagcaggacaagaagc ggtggaggagaca ccaaggggtgcagttatgcct ...3'
Primer 8R	5'-CAG ACG TGT GCT CTT CCG ATC TAG GCA TAA CTG CAC CCT TGG-3'

Primer 9F	5'-CTA CAC GAC GCT CTT CCG ATC TGA CAA GAA GCG GTG GAG-3'
Región del exón 9	5'... gacaagaagcgggtggagg agaccaaggggtgcagttatgcctcagattcacttttatacctttcc ttgcctctttctctagCACTGCCCAACAACACCAGCTCCTCTCCCAGCCAAAGAAGAAACCACTGGAT GGAGAATATTTACCCCTTCAGgtactaagtcttgggacctcttatcaagtggaaagtttcc agtctaa cactcaaaatgccg ...3'
Primer 9R	5'-CAG ACG TGT GCT CTT CCG ATC TCG GCA TTT TGA GTG TTA GAC-3'

Primer 10F	5'-CTA CAC GAC GCT CTT CCG ATC TCA ATT GTA ACT TGA ACC ATC-3'
Región del exón 10	5'... caattgtaacttgaaccatc ttttaactcaggtactgtgtatatacttacttctccccctcctctgttgctgcagATCCGTGGGCGTGAGCGCTTCGAGATGTTCCGAGAGCTGAATGAGGCCTTGGAACTCAAGGATGCCAGGCTGGGAAGGAGCCAGGGGGGAGCAGGGCTCACTCCAGgtgagtgcacctcagccccttctggccctactcccctgccttcttaggttgaaagcc ataggattccattctcatcc ...3'
Primer 10R	5'-CAG ACG TGT GCT CTT CCG ATC TGG ATG AGA ATG GAA TCC TAT-3'

Anexo 3:

Resultados de medición mediante un método fluorimétrico Qubit. Siguiendo el protocolo, se realizaron dos determinaciones para cada muestra.

Muestra	Nomenclatura	1º medición [DNA] (ng/µl)	2º medición [DNA] (ng/µl)
526	D1	0,09	0,1
528	H2	0,19	0,22
539	B3	0,05	0,09
545	E4	0,09	0,12
545	H4	0,06	0,06
546	A5	0,18	0,28
546	E5	0,03	0,12
546	G5	0,13	0,17
L6	E6	0,04	0,21
L6	G6	0,08	0,12
L10	A12	0,09	0,06
L10	B12	0,09	0,14
L11	G8	0,05	0,06
L12	G9	0,2	0,22
L12	A9	0,2	0,25
L13	C10	0,15	0,17
L13	G10	0,19	0,27
L16	G11	0,09	0,13
L16	E11	0,05	0,11

Anexo 4:

Protocolo de preparación de librerías para NGS:

bc1F	bc2F	bc3F	
-------------	-------------	-------------	--

Tubo 1: bc1F + bc10R + 526	Tubo 2: bc2F + bc10R + 528	Tubo 3: bc3F + bc10R + 539	bc10R
Tubo 4: bc1F + bc12R + 545	Tubo 5: bc2F + bc12R + 546	Tubo 6: bc3F + bc12R + L6	bc12R
Tubo 7: bc1F + bc13R + L9	Tubo 8: bc2F + bc13R + L11	Tubo 9: bc3F + bc13R + L12	bc13R
Tubo 10: bc1F + bc14R + L13	Tubo 11: bc2F + bc14R + L16	Tubo 12: bc3F + bc14R + L10	bc14R

Correspondencia entre Tubos y nomenclatura de muestra

N° tubo	Tubo 1	Tubo 2	Tubo 3	Tubo 4	Tubo 5	Tubo 6	Tubo 7	Tubo 8	Tubo 9	Tubo 10	Tubo 11	Tubo 12
Nomenclatura	526	528	539	546	545	L6	L9	L11	L12	L13	L16	L10

Anexo 5:

Resultados de la electroforesis digital

Pared de la electroforesis	Nomenclatura	Concentración DNA (ng/μl)	Pico en bp
A1	DNA positivo	23,9	-
B1	526	45,2	320
C1	528	22	322
D1	539	20,7	314
E1	546	37,6	374
F1	545	41,8	316
G1	L6	47,7	315
H1	L9	32,7	310
A2	L11	35,6	326
B2	L12	61,4	325
C2	L13	37,3	335
D2	L16	47,6	316
E2	L10	62,7	334

Anexo 6:

Bedfile para la cuarta fase de la NGS: alineación.

Cromosoma	Inicio de la región en el genoma	Fin de la región en el genoma	Primers para gen-región
17	7.676.298	7.676.123	TP53-4a
17	7.676.192	7.675.949	TP53-4b
17	7.675.267	7.675.019	TP53-5
17	7.674.998	7.674.820	TP53-6
17	7.674.330	7.674.154	TP53-7
17	7.673.879	7.673.648	TP53-8
17	7.673.688	7.673.474	TP53-9
17	7.670.790	7.670.531	TP53-10
17	7.670.790	7.670.531	TP53-10