

Trabajo fin de grado

PUESTA A PUNTO DE LA TÉCNICA “DROPLET” DIGITAL PCR PARA EL ANÁLISIS DE VARIACIONES EN EL NÚMERO DE COPIAS EN TUMORES SÓLIDOS

Universitat Politècnica de València
Escuela Técnica Superior de Ingeniería Agronómica y del Medio Natural
(ETSIAMN)

Grado en Biotecnología - Curso 2018/2019

Presentado por: **Almudena Chicote González**

Tutor UPV: **Consuelo Sabater Marco**

Tutora empresa: **Dr. Mainer Ibarrola Villava**

Tipo de licencia: **Creative Commons**

Valencia, 20 de Junio de 2019

Título

Puesta a punto de la técnica “droplet” digital PCR para el análisis de variaciones en el número de copias en tumores sólidos.

Resumen

El cáncer es un problema sanitario de primer orden y la segunda causa de muerte en los países desarrollados por detrás de las enfermedades cardiovasculares. Según la OMS, los casos de cáncer han crecido un 28% en los últimos 6 años. En España, en 2018, se ha registrado una incidencia de más de 270.000 casos y una mortalidad de más 110.000, siendo los tipos tumorales más comunes el cáncer de pulmón, mama, colorrectal, próstata y estómago. La caracterización molecular del tejido tumoral basada en la búsqueda de mutaciones somáticas, variaciones en el número de copias y otras alteraciones características de cada tipo tumoral, se ha convertido en una práctica habitual en oncología clínica. La identificación de estas alteraciones ha permitido la aparición de nuevos biomarcadores para el diagnóstico, pronóstico y el tratamiento de la enfermedad. Estas determinaciones se realizan de forma rutinaria en muestras parafinadas de biopsias de tejido tumoral. Sin embargo, la sangre de pacientes con cáncer contiene distintos tipos de materiales que son susceptibles de un análisis molecular detallado: siendo el ADN tumoral circulante libre extracelular (ADNct) el más estudiado. El estudio de estos materiales obtenidos del plasma se conoce como biopsias líquidas y permite obtener información molecular y diagnóstica de relevancia similar a la obtenida mediante biopsias invasivas convencionales. El ADNct deriva de tumores primarios, lesiones metastásicas y de células tumorales circulantes (CTC) lisadas y su aislamiento es mucho más simple que las CTCs. La aplicación de estas técnicas no invasivas supone una revolución en el diagnóstico actual, pronóstico y seguimiento de la enfermedad, beneficiándose tanto pacientes en estadios iniciales como avanzados. En la actualidad se han desarrollado múltiples métodos que nos permiten detectar y evaluar el ADNct, como la secuenciación de nueva generación y varias técnicas de PCR incluida la PCR digital (dPCR). Las técnicas de dPCR son más precisas y la sensibilidad es cada vez mayor llegando a detectar porcentajes de mutación menores de 0,01%. La tecnología “droplet” dPCR (ddPCR) permite el procesamiento de aproximadamente 2 millones de reacciones de PCR utilizando ensayos TaqMan convencionales. Para ello se encapsulan en gotitas de aceite tanto los reactivos de PCR como la muestra de ADN, y dentro de cada una de ellas se da la reacción de PCR. Al final, mediante un citómetro de flujo automatizado se procede a la lectura de la fluorescencia de cada gotita. Las variaciones en el número de copias (CNV) son deleciones o amplificaciones de segmentos del genoma que pueden tener un tamaño de cien hasta varios millones de pares de bases. En los últimos años se ha demostrado la implicación de las CNV en muchas enfermedades, entre ellas el cáncer, donde además, recientemente se ha visto su importancia en la respuesta a determinados tratamientos. El objetivo de este trabajo es la puesta a punto de la determinación de CNV para distintos genes diana en muestras de diferentes tumores sólidos, tanto en muestras parafinadas como en muestras de biopsias líquidas.

Palabras Clave

Biopsia líquida - Droplet digital PCR (ddPCR) - Variación en el número de copias (CNV)
- Biomarcadores - Tumores sólidos

Title

Tuning of the "droplet" digital PCR technique for determination of copy number variation in samples of different solid tumors.

Abstract

Cancer is a first order health problem and the second cause of death in the developed countries, behind cardiovascular diseases. According to the WHO, cancer cases have grown 28% in the last 6 years. In Spain, in 2018 there has been an incidence of more than 270,000 cases and a mortality of more than 110,000. The most common are lung, breast, colorectal, prostate and stomach cancer. The molecular characterization of tumor tissue, which consists of finding somatic mutations, variations in the number of copies and other characteristic alterations of each tumor type, has become a common practice in clinical oncology. The identification of these alterations has allowed the appearance of new biomarkers for the diagnosis, prognosis and treatment of the disease. These determinations are carried out routinely in paraffin samples of tumor tissue biopsies. However, the patients blood contains different types of materials that are amenable to detailed molecular analysis, being the most studied the extracellular free circulating tumor DNA (ctDNA). The study of these materials obtained from plasma is known as liquid biopsies and it allows to obtain relevant molecular and diagnostic information, similar to that obtained by conventional invasive biopsies. The ctDNA is derived from primary tumors, metastatic lesions and lysed circulating tumor cells (CTC) and its isolation is much easy than CTCs. The application of these non-invasive techniques represents a revolution in the current diagnosis, in the prognosis and follow-up of the disease, benefiting both patients in the initial and advanced stages. Nowadays, multiple methods have been developed that allow us to detect and evaluate the ctDNA, such as next generation sequencing and several PCR techniques including digital PCR (dPCR). The dPCR techniques are more precise and their sensitivity is increasing, reaching detection of mutation percentages lower than 0.01%. Droplet dPCR (ddPCR) technology allows the processing of approximately 2 million PCR reactions using conventional TaqMan assays. For this purpose, both the reagents and the DNA sample are encapsulated in oil droplets, within each of which the PCR reaction occurs. At the end, an automated flow cytometer proceeds to read the fluorescence of each droplet. The copy number variation (CNV) are deletions or amplifications of genome segments that can have a size of one hundred to several million base pairs. In recent years, the implication of CNVs has been demonstrated in many diseases, including cancer, where in addition, its importance in the response to certain treatments has recently been seen. The objective of this work is to tuning the ddPCR technique for CNV determination in different target genes in samples of different solid tumors, both in paraffin samples and in samples of liquid biopsies.

Key Words

Liquid biopsies – Droplet digital PCR (ddPCR) – Copy number variation (CNV) – Biomarkers – Solid tumor

Autora del TFG: **Dña. Almudena Chicote González**
Valencia, Junio de 2019.

Tutor Académico: **Prof. Dña Consuelo Sabater Marco**

Cotutor: **Dña. Mainer Ibarrola Villava**

Tipo de licencia: **Creative Commons**

Índice

1. Introducción	1
1.1. Cáncer	1
1.2. Biopsia Líquida	3
1.3. Importancia del análisis de copy number variation o CNV	5
1.4. Técnicas de análisis	6
1.4.1. Droplet digital PCR (ddPCR)	6
2. Objetivos	7
3. Materiales y métodos	7
3.1. Materiales	7
3.1.1. Muestras	7
3.1.2. Equipo para la ddPCR y Sondas CNV	8
3.2. Métodos	10
3.2.1. Extracción de material genómico	10
3.2.1.1. Extracción de ADN a partir de tejido tumoral conservado en parafina	10
3.2.1.2. Extracción de ADN genómico a partir de muestras de sangre	11
3.2.1.3. Extracción de ADN circulante a partir de muestras de plasma	11
3.2.1.4. Extracción de ADN genómico a partir de líneas celulares	12
3.2.2. Cuantificación del ADN mediante Nanodrop y Qubit	12
3.2.3. Digestión del ADN genómico con enzimas de restricción	13
3.2.4. Reacción ddPCR	14
4. Resultados y discusión	17
4.1. Digestión del ADN	17
4.2. Determinación de la cantidad de ADN inicial	19
4.2.1. Cantidad adecuada de ADN inicial procedente de ADN genómico	19
4.2.2. Cantidad adecuada de ADN inicial procedente de tejido tumoral parafinado	20
4.2.3. Cantidad adecuada de ADN inicial procedente de plasma	22
4.3. Establecimiento del threshold	22
4.4. Modificación del programa ddPCR	24
4.4.1. Modificación del número de ciclos	25
4.4.2. Modificación del tiempo de annealing	25
4.5. Variación en el número de copias (CNV)	26
4.5.1. Evaluación del número de copias de los distintos genes de interés en pacientes con tumores sólidos	26
4.5.2. Correlación entre CNV en muestras de plasma y CNV en muestras de ADN genómico	29
4.5.3. Correlación entre CNV de muestras de plasma y CNV de muestras de ADN procedente de parafinas	32
5. Conclusiones	33
6. Bibliografía	33

Índice de figuras

1.	Tasas estimadas de mortalidad estandarizadas por edad en 2018, ambos sexos, todas las edades (GLOBOCAN 2018)	1
2.	Esquema de la monitorización de un paciente mediante la medición del ctDNA (Corcoran and Chabner, 2019)	4
3.	Equipo ddPCR. De izquierda a derecha; QX200 AutoDG™, termociclador C1000 Touch y lector QX200	8
4.	Cartridges DG32™, reactivos ddPCR y aceites del generador de ‘droplets’ y del lector (BIORAD 2019)	9
5.	Blast2Seq entre los exones de <i>GSTT1</i> y la sonda de CNV de <i>GSTT1</i> (NCBI, 2019)	13
6.	Montaje del AutoDG™ QX200 instrumento para la generación automática de gotitas, a la izquierda cómo se muestra en la pantalla del robot y a la derecha cómo se ve cuando está montado	14
7.	Gráfico de barras en el que se representa el número de eventos totales por muestra	15
8.	Gráfico 1D; en azul droplets positivas para el fluoróforo FAM, en verde droplets positivas para el fluoróforo HEX y en negro las droplets dobles negativas. (QUANTASOFT)	16
9.	Gráfico 2D; en azul droplets positivas para el fluoróforo FAM, en verde droplets positivas para el fluoróforo HEX, en naranja las dobles positivas y en negro las droplets dobles negativas (QUANTASOFT)	16
10.	Gráfico 1D ensayo CNV de <i>GSTT1</i> siendo las muestras en la posición A1 y B1 ADN genómico sin digerir, y muestras digeridas en C1-G1.	17
11.	Gráfico 2D del ensayo CNV de <i>ERBB2</i> , resultado fallido para el ADN de la línea celular NCI-N87 por la utilización de una enzima de digestión incorrecta	18
12.	Gráficos 1D, a la izquierda ensayo de CNV de <i>PIK3CA</i> con ADN de tejido parafinado digerido, a la derecha ensayo de CNV de <i>PIK3CA</i> con las mismas muestras sin digerir	19
13.	Gráfico 1D de ensayo de CNV de <i>GSTT1</i> en gradiente de cantidades de ADN genómico	20
14.	Gráfico de número de copias del ensayo para el gen <i>GSTT1</i> con el gradiente de concentración de ADN genómico total	20
15.	Gráfico 1D del ensayo de CNV para el gen <i>ERBB2</i> con el gradiente de concentración de ADN total procedente de tumores incluidos en parafina.	21
16.	Gráfico de número de copias del ensayo de CNV para el gen <i>ERBB2</i> con el gradiente de concentración de ADN total procedente de tumores incluidos en parafina.	21
17.	Gráfico de número de copias del gen <i>GSTT1</i> en muestras de cfDNA procedente de plasma	22
18.	Gráfico 2D ejemplo de lluvia generada principalmente cuando el material genético utilizado proviene de tejido parafinado	23
19.	Tres threshold diferentes para el mismo ensayo con ADN genómico	23
20.	Tres threshold diferentes para el mismo ensayo con ADN procedente de parafinas	24
21.	Izquierda; gráfico 2D utilizando la sonda para el gen <i>AGO1</i> , derecha; gráfico 2D utilizando la sonda para el gen <i>TERT</i>	24
22.	A la izquierda; gráfico 1D resultado de ddPCR realizada con 40 ciclos. A la derecha; gráfico 2D resultado de ddPCR realizada con 50 ciclos	25

23.	A la izquierda; gráfico 2D resultado de ddPCR realizada con un minuto y medio de tiempo de annealing. A la derecha; gráfico 2D resultado de ddPCR realizada con un minuto de tiempo de annealing.	26
24.	A la izquierda; gráfico 2D resultado de ddPCR realizada con dos minutos de tiempo de annealing. A la derecha; gráfico 2D resultado de ddPCR realizada con un minuto de tiempo de annealing.	26
25.	Número de copias del gen <i>GSTT1</i> en muestras de ADN genómico de pacientes con melanoma maligno	27
26.	Gráficos 2D, para 0, 1 y 2 copias del gen <i>GSTT1</i>	27
27.	Número de copias del gen <i>ERBB2</i> en pacientes con cáncer colorrectal y en pacientes con cáncer gástrico.	28
28.	Número de copias del gen <i>PIK3CA</i> en pacientes con cáncer de cabeza y cuello	29
29.	Número de copias del gen <i>PIK3CA</i> en pacientes con cáncer gástrico	29
30.	Correlación lineal entre el número de copias del gen <i>GSTT1</i> en muestras de ctDNA de plasma y en muestras de ADN genómico	30
31.	Gráfico de caja y bigotes en el que se representan los positivos para FAM tanto en muestras de ctDNA de plasma cómo en muestras de ADN genómico	31
32.	Gráfico de caja y bigotes en el que se representan los positivos para HEX tanto en muestras de ctDNA de plasma cómo en muestras de ADN genómico	31
33.	Número de copias del gen <i>PIK3CA</i> en muestras de ctDNA de plasma frente a número de copias de muestras de tejido parafinado	32

Índice de tablas

1.	Sistema de estadificación TNM	2
2.	Ventajas de la ddPCR frente a otras tecnologías	6
3.	Sondas CNV disponibles en el laboratorio, fluoróforo, quencher, longitud, ensayos de referencia y enzima de restricción recomendada para cada una.	9
4.	Programa de PCR para CNV recomendado por Bio-Rad	15
5.	Archivo generado por Quantasoft, con datos del ensayo CNV de <i>GSTT1</i> .	18
6.	Número de copias del gen <i>GSTT1</i> por muestra de ADN genómico según los diferentes thresholds.	23
7.	Número de copias del gen <i>ERBB2</i> por muestra de ADN procedente de tejido parafinado con los diferentes thresholds	24
8.	Número de copias del gen <i>GSTT1</i> por muestra en ADN genómico y ctDNA	30
9.	Número de copias del gen <i>GSTT1</i> por muestra tanto en ctDNA de plasma cómo en ADN procedente de tejido parafinado	32

Abreviaturas

ADN: Ácido desoxirribonucleico
AGO1: Argonauta RISC Componente 1
AJCC: American Joint Committee on Cancer
AP3B1: Adaptor Related Protein Complex 3 Subunit Beta 1
AR: Receptor de andrógenos
ARNm: Ácido ribonucleico mensajero
ATP: Adenosín trifosfato
BSA: Albúmina de suero bovino
CDKN2A: Inhibidor 2A de quinasa dependiente de ciclinas
cDNA: Ácido desoxirribonucleico complementario
cfDNA: ADN libre circulante
Ch1: Canal 1
Ch2: Canal 2
CNV: Variación en el número de copias
Ct: Ciclo umbral
CTC: Células tumorales circulantes
ctDNA: ADN circulante tumoral
ddPCR: Droplet digital polymerase chain reaction
dUTP: 2'-Deoxyuridina, 5'-Trifosfato
EDTA: Ácido etilendiaminotetraacético
EGFR: Receptor del factor de crecimiento epidérmico
ERBB2: Erb-B2 Receptor Tirosina Kinasa 2
FAM: Carboxyfluoresceína
FFPE: Formalin-fixed paraffin-embedded
GSTM1: Glutathion S-Transferasa Mu 1
GSTT1: Glutathion S-Transferasa Theta 1
HEX: Hexacloro-Fluoresceína
HPLC: High performance liquid chromatography
KRAS: KRAS proto-oncogen
LB: Lysogeny broth
MicroRNA: Micro ácido nucleico
MM: Melanoma Maligno
PIK3: Fosfoinositol 3-quinasa
PIK3CA: Fosfatidilinositol-4,5-bisfosfato 3-quinasa subunidad catalítica alfa
PtdIns: Fosfatidilinositol
qPCR: PCR cuantitativa
R: Coeficiente de Pearson
RCB: Residual cancer burden
ROS: Especies reactivas de oxígeno
RPMI: Roswell Park Memorial institute medium
RPP30: Ribonucleasa P/MRP Subunidad P30
SDS: Dodecilsulfato sódico
SNP: Single Nucleotide Polymorphism
TERT: Telomerasa transcriptasa inversa
UICC: The Union for International Cancer Control
UV: Ultravioleta

1. Introducción

1.1. Cáncer

Se denomina cáncer a las enfermedades en las que hay células anormales que proliferan descontroladamente y pueden invadir otros tejidos (NATIONAL CANCER INSTITUTE, 2019). Weinberg y Hanahan (2000) describieron las características principales de las células cancerosas: autosuficiencia para la proliferación, falta de sensibilidad a las señales inhibitorias del crecimiento, evasión de la apoptosis, potencial de replicación ilimitado, angiogénesis sostenida y capacidad de invadir tejidos y producir metástasis.

La incidencia y mortalidad del cáncer aumenta cada año a nivel mundial, siendo esta enfermedad la segunda causa de muerte del mundo. Sin embargo, cuando hablamos de cáncer no se trata de una sola enfermedad, sino que son más de 277 enfermedades distintas dependiendo del tejido en que se originen (Hassanpour y Dehghani, 2017). En el año 2018 el cáncer fue la causa de 9,6 millones de muertes. En ambos sexos combinados, el cáncer de pulmón es el más común (11.6 % del total de casos en el mundo) y el que más muertes provoca (18,4 % del total de muertes por cáncer), seguido de cerca por el cáncer de mama femenino (11.6 %), cáncer de próstata (7.1 %) y cáncer colorrectal (6.1 %) por incidencia y cáncer de estómago (8.2 %) y cáncer de hígado (8.2 %) por mortalidad. Sin embargo, el tipo de cáncer más frecuente en cada país varía sustancialmente, debido a los diferentes estilos de vida (factores ambientales) así como a diferencias genéticas (Figura 1). (Bray et al., 2018)

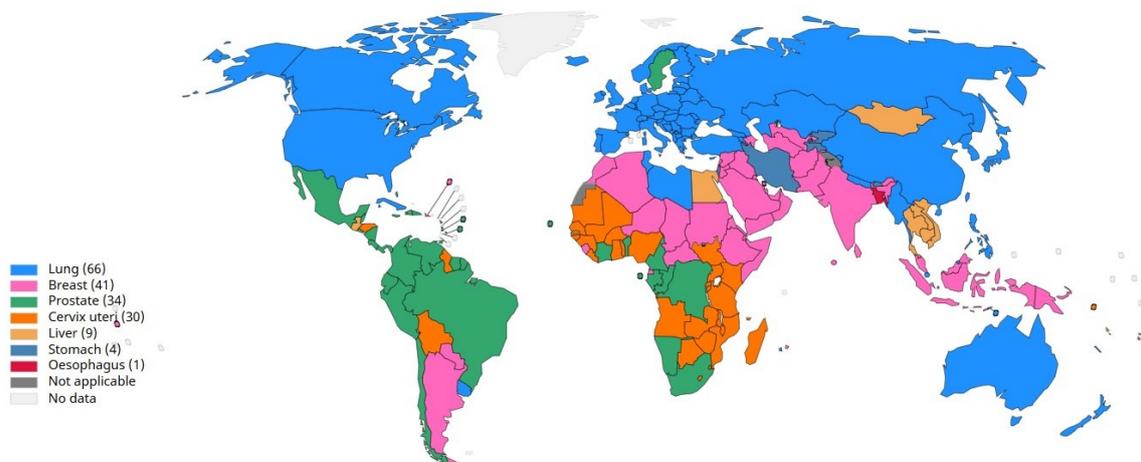


Figura 1: Tasas estimadas de mortalidad estandarizadas por edad en 2018, ambos sexos, todas las edades (GLOBOCAN 2018)

El concepto de biología molecular del cáncer surge en 1975 con el término proto-oncogen, descrito en el laboratorio de Harold Varmus y J.Michael Bishop en California; según los autores un proto-oncogen podía iniciar un cáncer si se alteraba. Un proto-oncogen codifica factores de transcripción que en definitiva controlan la expresión de otros genes relacionados con la división celular y el control del ciclo celular, si estos genes se alteran, su actividad puede verse afectada y así contribuyen al desarrollo del cáncer. Es a partir de este momento cuando se comienzan a investigar los mecanismos moleculares que conducen a la formación del cáncer (Weinberg, 2014).

Según el origen celular de los tumores estos adquieren una denominación u otra. Aquellos

que tienen su origen en las células epiteliales se denominan carcinomas, mientras que si se originan en células del tejido conectivo (cartílago, hueso, músculo o vasos sanguíneos) se denominan sarcomas. En cambio, se denominará leucemia si se origina en las células de la médula ósea, linfoma o mieloma si su origen son los ganglios linfáticos o en las células plasmáticas respectivamente y tumores del sistema nervioso si se originan en las células del neuroectodermo. (CANCER RESEARCH UK, 2019)

Es importante para el pronóstico y la elección del tratamiento conocer el tamaño del tumor, si se ha propagado y hacia dónde. Esta información nos la da el estadio del tumor. El sistema de estadificación utilizado es el TNM, a excepción de los tumores cerebrales o espinales y los cánceres hematológicos. (NATIONAL CANCER INSTITUTE, 2019) Este sistema lo desarrolló el American Joint Committee on Cancer (AJCC) junto con el TNM Committee of the International Union Against Cancer (UICC). La T hace referencia al tamaño del tumor primario, la N indica la migración del tumor a los ganglios linfáticos, mientras que la M hace referencia a la metástasis. Detrás de cada letra habrá un número (Tabla 1); cuanto mayor es el número mayor será el tumor, más ganglios linfáticos se habrán visto afectados y el cáncer se habrá diseminado a distintas partes del organismo. (Amin, 2017)

Tabla 1: Sistema de estadificación TNM

	Índices Numéricos	Descripción
T	TX	No puede medirse un tumor primario
	T0	No puede encontrarse un tumor primario
	T1, T2, T3 y T4	Cuanto más grande es el número después de la T, mayor es el tumor o más se ha extendido a los tejidos cercanos
N	NX	No puede medirse el cáncer en los ganglios linfáticos cercanos
	N0	No hay cáncer en los ganglios linfáticos cercanos
	N1, N2, N3 y N4	Cuanto más grande es el número después de la N, más son los ganglios linfáticos afectados por el cáncer
M	MX	No puede medirse la metástasis
	M0	El cáncer no se ha diseminado a otras partes del cuerpo
	M1	El cáncer se ha diseminado a otras partes del cuerpo

Cuándo un paciente presenta síntomas que pueden indicar la presencia de un tumor, lo primero que se debe hacer es descartar que no se deben a otras patologías, para ello el doctor pide el historial clínico familiar y realiza una exploración física del paciente. Tras ello, se realiza un análisis de sangre y u orina, para comprobar si hay sustancias que puedan indicar la presencia de un tumor. Para obtener un diagnóstico más preciso será necesario llevar a cabo otros procedimientos cómo los estudios de imágenes tomadas con tomografía computarizada, exploración nuclear, ecografía, resonancia magnética, tomografía por emisión de protones o rayos X. Sin embargo, en la mayoría de casos será necesario obtener una biopsia, este tejido lo analizará el patólogo bajo el microscopio para verificar la presencia de células cancerosas (NATIONAL CANCER INSTITUTE, 2019).

Cuando el cáncer se ha diagnosticado se procede a la cirugía dónde tratan de eliminar todo

el tumor, pero cómo no se puede estar totalmente seguro de que no hayan quedado restos del tumor se administra quimioterapia a la mayoría de los pacientes, también radioterapia en algunas ocasiones. Hoy en día existen alternativas que pueden introducirse solas o junto con quimioterapia, cómo la inmunoterapia, las terapias dirigidas con anticuerpos humanizados, la hormonoterapia o el trasplante de células madre. (Arruebo et *al.*, 2011) El objetivo de la investigación en el tratamiento del cáncer es llegar a la medicina personalizada en la que cada paciente reciba una terapia de acuerdo con las alteraciones genéticas del tumor.

El crecimiento celular y la proliferación tienen un papel importante en la progresión del cáncer. En el tumor se genera un ambiente de hipoxia que favorece la supervivencia y la proliferación de las células madre tumorales, esto puede agravarse en los tumores hormono-dependientes por la sobreproducción de ciertas hormonas. Además, la metástasis y el crecimiento son el resultado de la angiogénesis, de la transición epitelial a mesenquimal y del desencadenamiento de la autofagia entre otros procesos. Por tanto, hay muchos tratamientos destinados a inhibir estas vías de señalización para interferir en la proliferación del tumor y detenerla. (Feitelson et *al.*, 2015) Cetuximab es un anticuerpo monoclonal utilizado para tratar a pacientes con cáncer colorrectal metastásico, cuyos tumores sobreexpresan *EGFR* y no presentan mutaciones en el gen *KRAS*, normalmente en combinación con quimioterapia. Este anticuerpo actúa bloqueando los receptores EGFR y provocando que no se genere la transducción de la señal. (Reynolds y Wagstaff, 2004)

La biopsia convencional consiste en la extracción de una pequeña parte del tejido, y se puede llevar a cabo en cualquier órgano o parte del cuerpo. Algunas requieren la extirpación del tejido en un quirófano, cómo la biopsia excisional, la incisional o la biopsia endoscópica. No obstante, existen biopsias no tan invasivas cómo la que se realiza mediante aspiración con aguja fina. (AMERICAN CANCER SOCIETY, 2019) Estos procedimientos tienen ciertas limitaciones: son técnicas invasivas, no aportan información a tiempo real por lo que no permiten una monitorización del paciente, en ocasiones se pierde la heterogeneidad del tumor y pueden ser costosas. (LABCE, 2019) El tejido biopsiado normalmente es parafinado, pero este tipo de material causa la fragmentación de los ácidos nucleicos. Cabe destacar, que no existe ningún pretratamiento para mejorar la calidad de los ácidos nucleicos (Lenze et *al.*, 2012). Estas limitaciones podrían solventarse con el uso de biopsia líquida.

1.2. Biopsia Líquida

La biopsia líquida consiste en la detección y el aislamiento de células y moléculas que se liberan de la masa tumoral, cómo son las células tumorales circulantes (CTCs), los ácidos nucleicos circulantes (cfDNA, por sus siglas en inglés cell free DNA), y pequeñas vesículas llamadas exosomas, en distintos fluidos corporales. (Palmirotta et *al.*, 2018) La biopsia líquida se está convirtiendo en una herramienta muy interesante para identificar variaciones moleculares en el tumor, así como para conocer la heterogeneidad, detectar recaídas y monitorizar la evolución de la enfermedad. El plasma ha sido ampliamente estudiado como biopsia líquida, sin embargo, también podemos analizar variaciones moleculares en ADN circulante obtenido de la orina, el líquido cefalorraquídeo, la saliva, heces, el fluido pleural y los ascitos. (Corcoran y Chabner, 2019)

El ADN libre circulante o cfDNA fue descrito por Mandel y Metais (1948), se trata

de ADN fragmentado que se libera a la sangre y proviene de la apoptosis o la necrosis celular. Siempre se encuentra en fragmentos de entre 150 y 200 pb. La vida media en sangre de las moléculas de cfDNA es de aproximadamente una hora o menos. El cfDNA se encuentra en el plasma de todas las personas (10-15 ng/mL) y puede aumentar si hay daños como la inflamación entre otros. Sin embargo, hace 40 años se descubrió que los pacientes con cáncer tenían mayor nivel de cfDNA que es liberado por las células tumorales, a este tipo se le denominaría ctDNA (circulating tumor DNA) y constituye una pequeña porción de su cfDNA (Corcoran y Chabner, 2019). Se han encontrado variaciones entre personas con el mismo tipo de cáncer, es decir, existe una diferencia de muerte celular entre individuos, además presentan diferencias en la detectabilidad de ese ctDNA. En consecuencia su detección y análisis es un reto (Stewart et al., 2018). Las alteraciones más buscadas en este ctDNA son las mutaciones en un gen y las variaciones en el número de copias de un gen (CNV). (Palmirotta et al., 2018)

La biopsia líquida permite detectar si existe enfermedad mínima residual tras la cirugía, puesto que, por ejemplo, en cáncer gástrico, pulmonar y colorrectal la cantidad de ctDNA es proporcional al *Residual Cancer Burden* (RCB). Los pacientes con niveles de ctDNA detectables tras la cirugía, tenían peores resultados de supervivencia. (Wan et al., 2017) De este trabajo se concluye que una concentración mayor de ctDNA se correlaciona con resultados clínicos y radiológicos más pobres. También permite estratificar a los pacientes ya que tener mayor cantidad de ctDNA post-operatorio implica una posible recaída. (Palmirotta et al., 2018) Además permite monitorizar al paciente durante la terapia, aunque entre distintos pacientes haya diferencias en los niveles de ctDNA, en un mismo individuo los niveles de ctDNA en cada tiempo se correlacionan bien con los cambios del tumor. Cabe destacar que la corta vida del ctDNA en los fluidos permite medir a tiempo real la carga del tumor en respuesta a la terapia. A su vez, permite la monitorización de la resistencia a un tratamiento, puesto que en el ctDNA se captura la heterogeneidad molecular asociada con la resistencia. Las múltiples alteraciones que dan lugar a la aparición de resistencias coexisten frecuentemente en diferentes lugares de una metástasis, pero pueden ser detectadas colectivamente en el ctDNA de una muestra única de plasma (Figura 2). (Corcoran y Chabner, 2019)

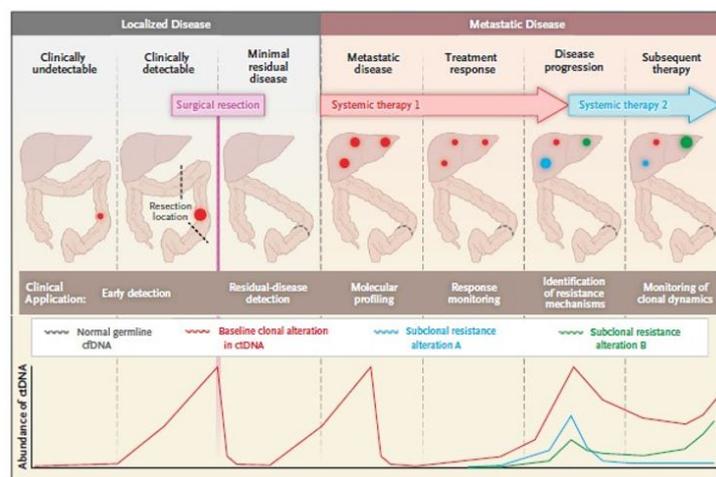


Figura 2: Esquema de la monitorización de un paciente mediante la medición del ctDNA (Corcoran and Chabner, 2019)

En definitiva, las ventajas que ofrece la biopsia líquida frente a la biopsia convencional son numerosas, probablemente la más importante de todas sea la fácil obtención, que cobra

mucha importancia en aquellos tumores en los cuales obtener una biopsia de tejido en ocasiones no es posible, cómo el de pulmón. Además la biopsia líquida representa toda la heterogeneidad del tumor, que sólo podría alcanzarse con la biopsia convencional si se tomaran muestras de tejido de diversas zonas del tumor. También permite, cómo ya se ha mencionado anteriormente, la monitorización del paciente mediante la toma de biopsias seriadas. Todo ello implica que el término biopsia líquida forme parte de la medicina personalizada. (Neumann et *al.*, 2018)

1.3. Importancia del análisis de copy number variation o CNV

En cáncer una de las alteraciones genéticas somáticas más comunes son los cambios en el número de copias (CNV, por sus siglas en inglés Copy Number Variation) de algunos genes. Una de las consecuencias de CNV es la amplificación de proto-oncogenes y la eliminación de los genes supresores de tumores (Beroukhim et *al.*, 2017). Pero no sólo tienen efecto sobre la dosis del gen si no que pueden tener otros efectos dependiendo de la región en la que se encuentre, cómo funciones cys-regulatorias. Hasta casi un 80 % de los cánceres alberga CNV (Kumaran et *al.* 2017).

Las CNV son duplicaciones y deleciones que pueden ser pequeñas (1-50 pb) o grandes (> 1Mpb) (Zhang et *al.*, 2016). No existen demasiados estudios de análisis de CNV en individuos sanos. Itsara A. et *al.*, (2009) publicaron los datos de un microarray de SNPs en 2500 individuos sanos, y demostraron que la mayoría albergaban CNV de al menos 100 pb. Sin embargo, hoy en día sabemos que a pesar de que hay alteraciones en el número de copias en el genoma de todos los individuos, algunas están directamente relacionadas con enfermedades como el cáncer o enfermedades neurodegenerativas. En las células cancerosas la variación en el número de copias puede llegar a afectar grandes regiones del genoma, y a menudo estas contienen genes involucrados en la proliferación celular (Sharma et *al.*, 2016). Por tanto, la formación y progresión del cáncer en ocasiones está relacionada con la variación en el número de copias.

En un estudio llevado a cabo por Li et *al.* (2016) consiguieron relacionar el número de copias de ciertos genes con cada uno de los estadios de la progresión del cáncer de mama. Por ejemplo que la alteración en el gen *ERBB2* ocurre en los estadios iniciales de este tipo de cáncer. Gao et *al.* (2017), realizaron un estudio en el que utilizaron el ctDNA obtenido de plasma de pacientes con cáncer gástrico, analizaron el número de copias de *ERBB2*, y los resultados obtenidos mostraron que *ERBB2* está sobreexpresado entre un 13 y un 22 % en cáncer gástrico. La amplificación de este gen promueve la tumorigenesis y está asociada con un peor pronóstico en muchos otros tumores. Por tanto es necesario conocer este dato para que aquellos que tengan amplificación de *ERBB2* se puedan beneficiar de la terapia dirigida con Trastuzumab, anticuerpo monoclonal anti-ERBB2.

Finalmente, en el estudio publicado por Smith et *al.* (2018) en el que analizaron 30 genes *drivers* de diferentes cánceres, se demostró que había 108 asociaciones significativas entre la variación en el número de copias de estos genes con el pronóstico de los pacientes, frente a las 23 asociaciones entre mutaciones y pronóstico. Así concluyeron que mediante la determinación del número de copias de un oncogen o de un gen supresor de tumores en un tumor primario, se podía estratificar mejor el riesgo del paciente, que con el análisis de mutaciones de un solo nucleótido. Así la detección de CNV en muestras de biopsia tisular o biopsia líquida puede ser de gran ayuda a la hora del diagnóstico, el pronóstico y en la elección del tratamiento (Liang et *al.*, 2015).

1.4. Técnicas de análisis

Tradicionalmente, tanto el análisis de las mutaciones como de la variación en el número de copias de un gen se ha detectado mediante PCR (Polimerase Chain Reaction) convencional. La concentración de ADN presente en muestras clínicas como sangre y tejidos tumorales es muy baja, aproximadamente unos 180 ng/ mL. Esta cantidad es indetectable por ninguna tecnología, por tanto es necesario llevar a cabo una amplificación mediante PCR. Posteriormente, los amplicones se detectan en un gel de agarosa, por tanto para el análisis de CNV, este método es cualitativo, detectando únicamente presencia o ausencia del gen respecto a un control endógeno. (Cao et al., 2017)

Tras la PCR convencional la técnica que comúnmente se utiliza para analizar variaciones genéticas es la PCR cuantitativa (qPCR), esta permite la visualización a tiempo real con el uso de fluoróforos, lo cual nos da información sobre la expresión de un gen, incluyendo duplicaciones y deleciones. En esta técnica se parte de ADN o de ARNm que se debe retrotranscribir a cDNA. A diferencia de la PCR convencional la qPCR detecta la amplificación al final de cada ciclo de amplificación ya que el termociclador de qPCR es capaz de detectar fluorescencia (Kuang et al., 2018). El análisis de CNV es difícil mediante PCR a tiempo real, ya que se cometen errores sistemáticos en la toma de medidas del Ct (Ciclo umbral), ciclo de PCR en el que aparece la intersección entre una curva de amplificación y una línea de umbral, este valor es inversamente proporcional a la cantidad de ADN diana (THERMO FISHER, 2019). También se cometen errores al normalizar las concentraciones de ADN usando muestras estándar, estos errores pueden llevar a resultados incorrectos.

1.4.1. Droplet digital PCR (ddPCR)

La ddPCR como la PCR convencional y la qPCR se basa en la reacción de amplificación catalizada por la Taq Polimerasa. Sin embargo, en la ddPCR los componentes de la reacción están encapsulados en miles de pequeñas gotas de aceite antes de que se dé la reacción de PCR. Además los datos se obtienen nada más finaliza la reacción con la ayuda de un lector, lo cual da una cuantificación absoluta sin la necesidad de utilizar curvas standard como en la qPCR, así obtenemos unos datos más precisos y reproducibles. Con la técnica ddPCR las concentraciones de ácidos nucleicos se obtienen contando el número de *droplets* positivas sin la necesidad de establecer Ct, así se obtienen resultados más precisos (Mazaika y Homsy, 2014). Esto fue comprobado en un artículo de Taylor. S et al. (2017) publicado en Nature Scientific Reports en el que se usaron muestras de ADN sintético para comparar las dos tecnologías. La tecnología ddPCR presenta numerosas ventajas frente a la PCR convencional y la qPCR (Tabla 2).

Tabla 2: Ventajas de la ddPCR frente a otras tecnologías

Ventajas ddPCR
No depende de patrones o referencias
Mayor precisión
Gran tolerancia a inhibidores
Capacidad de analizar mezclas complejas
Respuesta lineal al número de copias
Mayor sensibilidad
Técnica reproducible

La ddPCR es un método de cuantificación absoluta de ácidos nucleicos que se basa en la detección de señales fluorescentes y la enumeración de eventos binomiales (presencia de fluorescencia = 1, ausencia de fluorescencia = 0) (Quan et al. 2018). En la ddPCR primero se necesita una dilución del ADN a una cierta concentración. Comparado con las tecnologías de PCR convencionales la ddPCR es más sensible, precisa y con resultados reproducibles. Cabe destacar que los instrumentos de la ddPCR adoptan la distribución de Poisson, con ella puedes calcular el número de moléculas del ADN diana, de acuerdo con la ecuación: $\lambda = -\ln(1 - p)$, donde λ es la media de moléculas de ADN por reacción y la p es la fracción de reacciones positivas (Hndson et al., 2011). La concentración de ADN puede calcularse cómo: Concentración = λ/V , siendo V la media del volumen de una gota (Zhao et al., 2016)

El sistema de ddPCR de Bio-Rad's QX200 combina la tecnología de gotas de emulsiones de agua y aceite con microfluidos. El QX200 AutoDG™ es el generador de gotas y convierte las muestras en 20.000 gotas. La amplificación de la PCR se lleva a cabo en cada gotita en un termociclador (C1000 Touch). Después las gotitas se clasifican en único archivo con el lector QX200, el cual cuantifica la fluorescencia positiva. Por último, el software llamado QuantaSoft mide el número de *droplets* positivas y negativas para cada fluoróforo. (BIORAD 2019)

En esta técnica se utilizan sondas TaqMan, estas son secuencias específicas marcadas con un fluoróforo en el extremo 5' y con un *quencher* en el extremo 3'. Por tanto cuando la sonda está intacta, debido a la proximidad del *quencher* el fluoróforo no emite fluorescencia, pero en el paso de extensión de la ddPCR la enzima Taq polimerasa (con actividad exonucleasa 5' a 3'), deja al fluoróforo libre y entonces emite fluorescencia. Las ventajas de trabajar con este tipo de sondas son la alta especificidad, un alto ratio señal/ruido, y la capacidad de llevar a cabo reacciones multiplex.

Las aplicaciones más comunes de la *droplet* digital PCR en el ámbito biomédico son la detección de mutaciones raras y la cuantificación de ácidos nucleicos. Sin embargo, las aplicaciones de la ddPCR son numerosas, podemos detectar patógenos, variación en el número de copias (CNV), analizar la expresión de microRNA, cuantificar librerías para protocolos de secuenciación masiva (Next generation sequencing), análisis unicelulares y detección de anomalías cromosómicas (Cao et al., 2017).

2. Objetivos

El objetivo de este trabajo es la puesta a punto de la tecnología ddPCR para la determinación de CNV para distintos genes diana en muestras de diferentes tumores sólidos, tanto en muestras parafinadas como en muestras de biopsias líquidas. Así cómo, profundizar en el estudio de análisis y evaluación de los resultados.

3. Materiales y métodos

3.1. Materiales

3.1.1. Muestras

Las muestras utilizadas para poner a punto la técnica ddPCR provenían de diferentes tejidos; plasma, tumor parafinado y capa de linfocitos de la sangre. En concreto se

utilizaron, 27 muestras parafinadas, 23 muestras de capa de linfocitos de la sangre y 21 muestras de plasma. Las muestras procedían de pacientes con distintos tumores sólidos. Todas las muestras se anonimizan con un código interno del laboratorio. (Ver Anexos, Tabla 1)

Se testaron muestras de plasma y de capa de linfocitos de la sangre de pacientes con melanoma maligno (MM) del Hospital Gregorio Marañón (Madrid), gracias a una colaboración establecida entre ambos grupos en 2004. Cabe destacar que las muestras de tejido tumoral parafinado en este tipo de cáncer son difíciles de obtener puesto que es un tumor que se mide en mm y la mayoría de la muestra la requieren los patólogos para su análisis, por tanto no se disponía de este tipo de muestras. Las muestras de ADN de tejido parafinado procedían de pacientes diagnosticados con cáncer de colón, cáncer gástrico y cáncer de cabeza y cuello, del Hospital Clínico Universitario de Valencia. Las muestras de tumor se incluyen en parafina tras la cirugía y son almacenadas en el Departamento de Anatomía Patológica. También se utilizaron muestras de plasma de pacientes con cáncer de cabeza y cuello recogidas por el servicio de Biobanco del Instituto INCLIVA.

Para que un ensayo de ddPCR tenga validez, y para saber que la reacción ha transcurrido correctamente, se debe llevar a cabo con al menos un control negativo y uno positivo. Como control negativo de la validez del ensayo se utiliza agua estéril a la que añadimos todos los reactivos para llevar a cabo la reacción. Además, utilizaremos muestras de ADN, provenientes de diferentes materiales según el ensayo; tejido parafinado, ADN genómico y líneas celulares, con alteraciones genómicas conocidas como controles positivos y negativos de los resultados obtenidos. La selección de controles se explicará más en detalle en resultados.

3.1.2. Equipo para la ddPCR y Sondas CNV

Para llevar a cabo el experimento completo se necesitan 3 instrumentos, el QX200 AutoDG™ instrumento para la generación automática de las gotitas, el C1000 Touch™ *Thermal Cycler* con un módulo de reacción de 96 pocillos, para llevar a cabo la reacción de amplificación ddPCR y el QX200 *Droplet Reader*, para leer la fluorescencia de cada gotita. (Figura 3)



Figura 3: Equipo ddPCR. De izquierda a derecha; QX200 AutoDG™, termociclador C1000 Touch y lector QX200

Tanto el generador de *droplets* como el lector necesitan reactivos específicos: aceite, una botella de *Automated Droplet Generation Oil for Probes* de 140 mL, para generar las gotitas en emulsión y una botella de *ddPCR Droplet Reader Oil*, para la lectura de la fluorescencia. Además, son necesarios ciertos materiales consumibles como los DG32™

Cartridges, donde se produce la emulsión, puntas específicas para el QX200 AutoDG™, selladores de placas y las placas de 96 pocillos con paredes ultrafinas de polipropileno, este material permite la transferencia de calor, dónde se produce la reacción de amplificación (Figura 4).



Figura 4: Cartridges DG32™, reactivos ddPCR y aceites del generador de ‘droplets’ y del lector (BIORAD 2019)

Además, para llevar a cabo la reacción, se necesita ddPCR™ Supermix for Probes, al que se añaden sondas específicas que se unan a los genes dianas, así el lector puede determinar con posterioridad el número de copias de cada gen. Las sondas, diseñadas y validadas por Bio-Rad, que estaban disponibles en el laboratorio eran las mostradas en la Tabla 3.

Tabla 3: Sondas CNV disponibles en el laboratorio, fluoróforo, quencher, longitud, ensayos de referencia y enzima de restricción recomendada para cada una.

Nombre	Fluoróforo	Purificación	Quencher	Longitud (pb)	Ensayos de referencia	Enzima de restricción recomendada
CDKN2A	FAM	HPLC	Iowa Black	66	AP3B1, TERT, RPP30, EIF2C1	Hae III
PIK3CA	FAM	HPLC	Iowa Black	68	AP3B1, TERT, RPP30, EIF2C1	Hae III
EGFR	FAM	HPLC	Iowa Black	64	AP3B1, TERT, RPP30, EIF2C1	Hae III
AR	FAM	HPLC	Iowa Black	70	AP3B1, TERT, RPP30, EIF2C1	Hae III
ERBB2	FAM	HPLC	Iowa Black	66	AP3B1, TERT, RPP30, EIF2C1	Mse I
GSTT1	FAM	HPLC	Iowa Black	65	AP3B1, TERT, RPP30, EIF2C1	Hae III
TERT	HEX	HPLC	Iowa Black	70	AP3B1, TERT, RPP30, EIF2C1	Hae III
AGO1	HEX	HPLC	Iowa Black	70	AP3B1, TERT, RPP30, EIF2C1	Hae III

Las sondas específicas pueden ir marcadas con los fluoróforos FAM o HEX. Las marcadas con FAM son aquellas que hacen referencia a los genes diana que queremos estudiar. Aquellas sondas cuyo fluoróforo es HEX son sondas de genes que se conocen cómo controles endógenos, es necesario para los ensayos en los que lo que se desea es conocer el número de copias de un gen determinado. Estos genes cómo *TERT* o *AGO1* son genes

housekeeping, necesarios para mantener la funcionalidad básica de la célula y su expresión no varía en ningún caso. El ratio entre el analito y la concentración del gen de referencia indica el número de copias, de acuerdo con la fórmula $2 * (a/b)$, el 2 indica que siendo diploides tenemos 2 copias del gen de referencia, a es la concentración del analito y b la concentración del gen de referencia. De todas las sondas disponibles se utilizaron 5, la de *PIK3CA*, la de *ERBB2*, la de *GSTT1*, la de *TERT* y la de *AGO1*. *TERT* y *AGO1* se utilizan como controles endógenos.

El gen *GSTT1* codifica la proteína glutatión-S-transferasa theta 1, pertenece a la familia de las glutatión-S-transferasas, estas participan en la detoxificación de un gran número de especies reactivas de oxígeno (ROS) que se producen durante la síntesis de melanina y en procesos de estrés oxidativo. Se eligió estudiar *GSTT1* en primer lugar porque había muestras de pacientes con melanoma maligno que se habían testado previamente en el laboratorio mediante PCR convencional y qPCR, para tener una primera aproximación de los resultados de ddPCR (Ver anexos, Figura 1).

El gen *ERBB2* codifica un receptor tirosina quinasa de la familia de receptores de factores de crecimiento epidérmico. Este receptor no tiene dominio de unión a ligandos, para ello debe unirse a otro receptor y formar un heterodímero. La amplificación o la sobreexpresión de este gen se ha identificado en numerosos tipos de cáncer entre ellos el de mama (Li et al. 2016) y el gástrico (Liu et al., 2019). Se estudiaron las CNV de este gen en muestras de cáncer de colon y gástrico disponibles en el laboratorio, puesto que en la base de datos *Cancer Cell Line Encyclopedia* se ha comprobado que había variación en el número de copias de *ERBB2* en líneas celulares de estos tipos de cáncer (Ver anexos, Figura 2).

El gen *PIK3CA* codifica la subunidad catalítica de la fosfatidilinositol 3-quinasa, la cual utiliza ATP para fosforilar PtdIns, PtdIns4P y PtdIns(4,5)P2, la vía de señalización de PIK3 tiene un papel muy importante en diversas actividades de la célula como el crecimiento y la división celular. Se estudió este gen con muestras de cáncer de cabeza y cuello, puesto que la amplificación de este gen está relacionada con la metástasis ganglionar (Suda et al. 2012), y en muestras de cáncer gástrico. Además, en la *Cancer Cell Line Encyclopedia* se ha comprobado que había variación en el número de copias de *PIK3CA* en líneas celulares de estos tipos de cáncer (Ver anexos, Figura 3).

3.2. Métodos

3.2.1. Extracción de material genómico

Para cada tipo de muestras se utilizó un kit de extracción específico para el material de partida o un protocolo dedicado. Todos siguen el mismo principio, primero se realiza una lisis celular, luego se procede a la eliminación de contaminantes u otras sustancias presentes en el lisado, posteriormente se realiza el lavado del ADN y finalmente se procede a su elución.

3.2.1.1 Extracción de ADN a partir de tejido tumoral conservado en parafina

Para las muestras de tejido parafinado se siguió el protocolo del kit QIAamp DNA FFPE Tissue de la casa comercial QIAGEN, que utiliza membranas de sílice para la obtención del ADN. Partimos de 4 cortes de tejido parafinado de 20 μm de grosor. Los pasos a seguir fueron primero lavar con xileno y etanol cada muestra, para eliminar la parafina. Tras ello se incubó overnight a 56°C con proteinasa K y buffer ATL para que se diera

la lisis. Se calentaron las muestras a 90°C durante una hora, con este paso se revierte la reticulación de la formalina. A continuación los pasos siguientes permiten la unión del ADN a la membrana de la columna y la eliminación de otros componentes celulares o contaminantes, con este fin se añade buffer AL y etanol absoluto, se centrifuga y se transfiere el sobrenadante a un tubo nuevo con columna del kit, después se vuelve a centrifugar. Una vez el ADN se ha unido a la membrana se debe lavar para eliminar los contaminantes residuales, primero con buffer AW1 y luego con AW2. Cuando se seca la membrana se puede proceder al último paso, la elución del ADN con agua libre de DNAsas. El ADN se conservará a -20 °C hasta su utilización.

3.2.1.2 Extracción de ADN genómico a partir de muestras de sangre

Se utilizó el método salino para extraer el ADN de las muestras de sangre. Para ello, partimos de la capa de linfocitos de la sangre y un volumen aproximado de 1 mL. En un tubo se mezcla la capa de linfocitos de la sangre con STMT (sacarosa, Tris-HCl, MgCl₂, Tritón y agua destilada), una vez mezclado se centrifuga, se decanta el sobrenadante y se añade LB (Tris-HCl, EDTA, NaCl y agua destilada), se agita fuertemente para romper el pellet y se centrifuga igual que antes para quedarse sólo con el pellet. El siguiente paso consiste en la digestión proteica, para ello se añade LB, SDS y proteinasa K, se deja en agitación suave durante 24 horas. Al día siguiente se lleva a cabo la precipitación del ADN, para ello se añaden 400 µL de NaCl y se centrifuga, esta vez se recoge el sobrenadante y se añaden 2 volúmenes de etanol absoluto frío. Cuando se vortea aparece una especie de “medusa” que habrá que recuperar con ayuda de una pipeta, al final se lava con etanol al 70 % y cuando se seca, se eluye el ADN con agua y se deja en agitación fuerte overnight. El ADN se conservará a -20 °C hasta su utilización.

3.2.1.3 Extracción de ADN circulante a partir de muestras de plasma

La extracción de cfDNA de plasma la realizamos con dos kits: QIAamp Circulating Nucleic Acid Kit y DANAGENE Circulating DNA Minikit de dos casas comerciales distintas Qiagen y Danagen respectivamente.

El protocolo de Danagen consiste en centrifugar a máxima velocidad 1 ml de plasma durante 5 minutos para eliminar el posible ADN celular contaminante que provenga de células residuales. Tras recoger el sobrenadante se incuba a 55°C durante 15 minutos y se centrifuga otros 5 minutos. Se traspasa el sobrenadante a un nuevo tubo y se añaden tampón de Lisis PS y proteinasa K. Luego se deja incubar durante 15 minutos a 55°C para que la proteinasa K actúe. Se añade isopropanol, se mezcla bien y se pasa a una columna MicroSpin, se centrifuga, cuando se ha pasado toda la muestra por la columna se hace una última centrifugación. Se cambia la columna a otro tubo recolector y se añade tampón de desinhibición, tras lo que se centrifuga. Se elimina el líquido del tubo recolector y se añade a la columna tampón de lavado, después se centrifuga para eliminar el etanol residual. Por último se inserta la columna en un tubo de 1,5 ml y se eluye el ADN con agua. El ADN se conservará a -20 °C hasta su utilización.

En cuanto al protocolo de Quiagen primero se añaden proteinasa K y Buffer ACL, el cual contiene 1 µg de carrier de ARN a 1 ml de plasma. Es muy importante vortear bien en este paso para que la lisis se dé correctamente. Después el tubo se incuba a 60°C durante 30 minutos. Acto seguido se añade Buffer ACB al lisado, se vortea y se incuba 5 minutos en hielo. Se enciende el QuiAvac 24 Plus, una bomba de vacío, y se insertan en él las columnas. Se introduce la mezcla del lisado y del buffer ACB en la columna, y encendemos la bomba hasta que pase todo el líquido; cuando ya se ha pasado todo el

plasma, se procede al lavado, para ello se añade Buffer ACW1 a la columna y se enciende la bomba, cuando ha pasado todo el líquido se apaga, luego se añade Buffer ACW2 a la columna y se vuelve a encender la bomba hasta que pasa todo el líquido, luego se añade de etanol absoluto y se repite el proceso con la bomba. Por último se pasa la columna a un tubo nuevo y se eluye el ADN con agua. El ADN se conservará a -20°C hasta su utilización.

3.2.1.4 Extracción de ADN genómico a partir de líneas celulares

En este trabajo hemos utilizado líneas celulares como controles positivos en la determinación del número de copias. Las líneas celulares se encontraban en el congelador a -80°C , estas se descongelaron y se cultivaron con medio RPMI (Roswell Park Memorial Institute). Todos los días se cambió el medio. Además se realizaron 2 pases, puesto que necesitábamos un gran número de células para poder realizar los experimentos y para guardar, para ello se quita el medio y se lava con 2 mL PBS, luego se vacía el frasco. En el siguiente paso se añaden 2 mL de tripsina y se deja actuar 5 minutos en la estufa, para despegar las células, ya que son adherentes. Después se añaden 4 mL de medio RPMI rápidamente para inactivar la tripsina ya que este medio contiene suero fetal bovino y este inactiva la tripsina. Todo ello se centrifuga y se elimina el sobrenadante. Por último se resuspende el pellet en el que se encuentran las células con 2 mL de medio RPMI y se introduce en un frasco con 3 mL más de medio RPMI.

Para la extracción de ADN genómico de una línea celular no se utilizó un kit comercial, si no que se siguió el método del trizol, un protocolo de extracción basado en la utilización de fenol/cloroformo. Primero de todo se recogieron las células del medio de cultivo, lavando con PBS y despegándolas del frasco con la ayuda de un *cell scraper*. Se recuperaron con DMEM 10 % para centrifugar y así se obtuvieron los pellets compuestos por células, que fueron congelados a -80°C hasta el momento de la extracción del ADN. Para extraer el material genético se resuspendió el pellet en un buffer de digestión compuesto por Tris-HCl 1M pH 8, NaCl 4 M, SDS 10 %, EDTA 0,5 M, proteinasa K y agua miliQ y se incubó durante 18 horas a 50°C . Pasado ese tiempo se añadió 1 volumen de fenol:cloroformo:isoamílico (25:24:1) y se centrifugó, luego se recogió el sobrenadante y se añadió $\frac{1}{2}$ volumen de AcNH_4 7,5 M y 2 volúmenes de etanol absoluto frío. Tras centrifugar, se eliminó el sobrenadante y se lavó el ADN con etanol al 70 % frío. Por último se resuspendió el pellet con agua miliQ.

3.2.2. Cuantificación del ADN mediante Nanodrop y Qubit

El Nanodrop es un nano-espectrofotómetro de UV visible. Con este instrumento se cuantificó la concentración de ADN de doble cadena de cada muestra y de las digestiones de las mismas. Depositando 1 μL en el Nanodrop, el software calcula la concentración en $\text{ng}/\mu\text{L}$, y los ratios 260/280 y 260/230, los cuales indican la pureza del ADN. El ratio 260/280 debe ser aproximadamente de 1,8, mientras que el ratio 260/230 debe estar entre 2 y 2,2.

También se utilizó el fluorómetro Qubit, un método más sensible para la cuantificación del ADN. Para ello se preparan dos estándares con los que el propio instrumento realiza una curva patrón. Primero se prepara un mix con el Buffer y el Reagent del kit de Qubit, y luego se añade a cada tubo junto con el patrón en los estándares o las muestras de ADN. Una vez realizada la curva patrón se introdujeron las muestras en el Qubit para su medición.

3.2.3. Digestión del ADN genómico con enzimas de restricción

El ADN genómico necesita ser digerido por enzimas de restricción específicas para que el ensayo de determinación de número de copias mediante ddPCR tenga éxito y no se sature la reacción. Esto es necesario para separar las diferentes copias de un gen, asegurando un reparto adecuado de las copias en las diferentes gotitas, incluso cuándo las copias se encuentran en tándem. Así, se puede obtener un análisis óptimo del número de copias.

El ADN extraído a partir del tumor incluido en parafina se encuentra fragmentado y a priori no requiere digestión, aunque como explicaré más adelante hemos visto que puede ser digerido. En este trabajo se probaron diferentes cantidades de ADN total en la reacción para comprobar cuál daba mejores resultados. El ADN que proviene de los plasmas no se digiere porque el ADN circulante se encuentra altamente fragmentado con fragmentos de 50-1000 pb. El grado de fragmentación depende de muchos parámetros cómo el origen del ADN, el procedimiento de recolección de la capa de linfocitos de la sangre, el manejo etc. El ADN obtenido a partir del plasma tampoco se diluye, ya que la cantidad que obtenemos es mínima.

Se utilizó la enzima Hae III y la Mse I de la casa comercial TAKARA que corta en GG▼CC y CC▲GG y T▼TAA y AAT▲T respectivamente, puesto que eran las recomendadas por Bio-Rad, para las sondas testadas. Se requiere que el enzima no corte en el lugar que la sonda hibrida con el gen diana, por tanto nos aseguramos de ello buscando la secuencia de los exones del gen en Ensembl y haciendo un Blast2seq entre los exones del gen y la sonda (Figura 5). En este ejemplo, el alineamiento entre los exones del gen *GSTT1* y la sonda correspondiente se obtiene un E-value de 2e-44, el alineamiento es significativo.

Range 1: 37 to 123 [Graphics](#) ▼ Next Match ▲ Previous ↑

Score	Expect	Identities	Gaps	Strand
161 bits(87)	2e-44	87/87(100%)	0/87(0%)	Plus/Minus
Query 81	CTGTCCCAGCCCTGCCGCGCTGTTTACATCTTTGCCAAGAAGAACGACATTCCCTTCGAG	140		
Sbjct 123	CTGTCCCAGCCCTGCCGCGCTGTTTACATCTTTGCCAAGAAGAACGACATTCCCTTCGAG	64		
Query 141	CTGCGCATCGTGGATCTGATTAAAGGT	167		
Sbjct 63	CTGCGCATCGTGGATCTGATTAAAGGT	37		

Figura 5: Blast2Seq entre los exones de *GSTT1* y la sonda de CNV de *GSTT1* (NCBI, 2019)

Una vez conocida la región dónde se hibrida la sonda con el gen, se buscan dónde están los sitios de corte del enzima Hae III. Si no coincide ningún sitio de corte con la zona en la que se hibrida la sonda, esta enzima es adecuada para realizar la fragmentación del genoma para estudiar *GSTT1*. El mismo protocolo se siguió para la validación de la enzima Mse I y el gen *ERBB2*.

Para llevar a cabo la digestión con Hae III se añaden 5 µL de ADN (<1 µg), 2µL de Buffer compuesto por Tris-HCl, MgCl₂, NaCl y 2-mercaptoetanol, 1 µL de Hae III (10 U/µ) y 12 µL de agua. Por último se deja incubar durante 1 hora a 37°C.

Para llevar a cabo esta digestión con Mse I se añade ADN a menos de 1 µg, 2µL de Buffer compuesto por Tris-acetato, Mg-acetato, K-acetato, DTT y BSA, 2µL de BSA al 0,1 % y

1 μ L de Mse I (10 U/ μ) y agua hasta alcanzar los 20 μ L totales . Por último se deja incubar durante 1 hora a 60°C. Tras la digestión es necesario hacer una dilución 1:10, para reducir la concentración de sales, que podría inhibir la reacción de ddPCR.

3.2.4. Reacción ddPCR

La reacción de ddPCR se lleva a cabo en un volumen de 21 μ L. Los reactivos incluyen el Supermix de ddPCR para sondas (no dUTP) y las sondas para los diferentes genes. En una placa de 96 pocillos se introducen las muestras junto con el Supermix, la sonda para el gen diana, la sonda para el gen control y agua. Por tanto lo primero en realizarse es el Mix que contendrá el supermix, las sondas y el agua. Por muestra se necesitan 1 μ L tanto de la sonda para el gen diana, cómo de la sonda para el gen control endógeno, también 1 μ L de agua y 10 μ L de Supermix de ddPCR para sondas (no dUTP). El proceso de preparación de las placas se llevó a cabo en campana, previamente iluminada con luz UV y lavada con alcohol, para evitar las contaminaciones dada la alta sensibilidad de la técnica.

En cada pocillo añadimos 8 μ L de muestra y 13 μ L del Mix que se había preparado con anterioridad. Una vez sellada la placa con el sellador de placas a 180°C, se comprueba que no existen burbujas antes de introducir la placa en el robot AutoDG™ QX200, si las hubiera habría que centrifugar la placa, para intentar eliminarlas. Después, se prepara el robot introduciendo cajas de puntas, cartridges, dónde se producirá la emulsión, la placa sellada con las muestras y un bloque frio con la placa en la que se van a depositar las gotitas, tal y cómo se muestra en la figura (Figura 6).

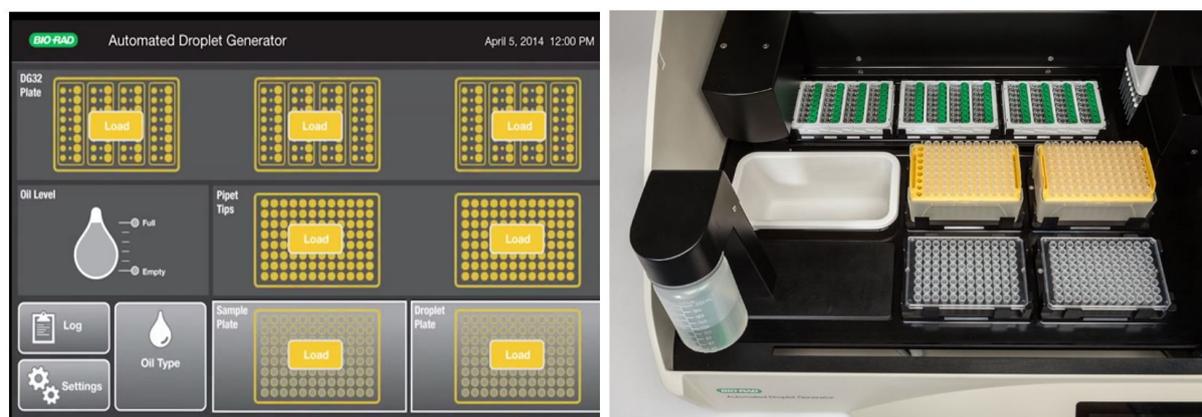


Figura 6: Montaje del AutoDG™ QX200 instrumento para la generación automática de gotitas, a la izquierda cómo se muestra en la pantalla del robot y a la derecha cómo se ve cuando está montado

El robot QX200 se encarga de generar las gotitas con el aceite. El ADN se distribuye al azar en las gotitas. Se producen mezclas uniformes tanto en tamaño como en volumen (1 nL), permitiendo la cuantificación específica y precisa. Cada gota contiene ADN y todos los reactivos necesarios para que se dé la reacción. Una vez obtenida la nueva placa, se vuelve a sellar y se introduce en el termociclador C1000 Touch dónde se lleva a cabo la amplificación. Se utilizará el programa de ddPCR que la compañía recomienda para CNV (Tabla 4). A lo largo de este trabajo se han optimizado el tiempo de annealing y el número de ciclos para comprobar cuál daba mejores resultados.

Tabla 4: Programa de PCR para CNV recomendado por Bio-Rad

step	°C	t	ramp	ciclos
1	95	10'	2°C/s	1
2	94	30"	2°C/s	40
3	60	1'	2°C/s	
4	98	10'	2°C/s	1
5	4	hold		1

Seguidamente tras acabar la reacción de amplificación se lleva a cabo la lectura de las gotitas. Se introduce la placa en el lector QX200, este analiza cada *droplet* individualmente usando un sistema basado en la detección de 2 fluoróforos (FAM y HEX). El inyector automático coge las gotitas de cada pocillo. Estas gotitas se espacian para la lectura de su fluorescencia en dos canales. Las positivas, aquellas que contienen al menos una copia del ADN diana, muestran un incremento en la fluorescencia comparado con las gotas negativas. El software llamado QuantaSoft mide el número de *droplets* positivas y negativas para cada fluoróforo en cada muestra. El software calcula la fracción de *droplets* positivas con un algoritmo de Poisson, así determina la concentración del ADN diana en unidades copias/ μ l de muestra.

Por último, se produce el análisis de los datos con el software, este proporciona el número de eventos totales por muestra (Figura 7), este es un dato importante a tener en cuenta, puesto que los expertos dictaminan que un número de eventos es aceptable a partir de 10.000, porque a partir de 10.000 aunque incrementemos el número de eventos vamos a disminuir muy poco el error.

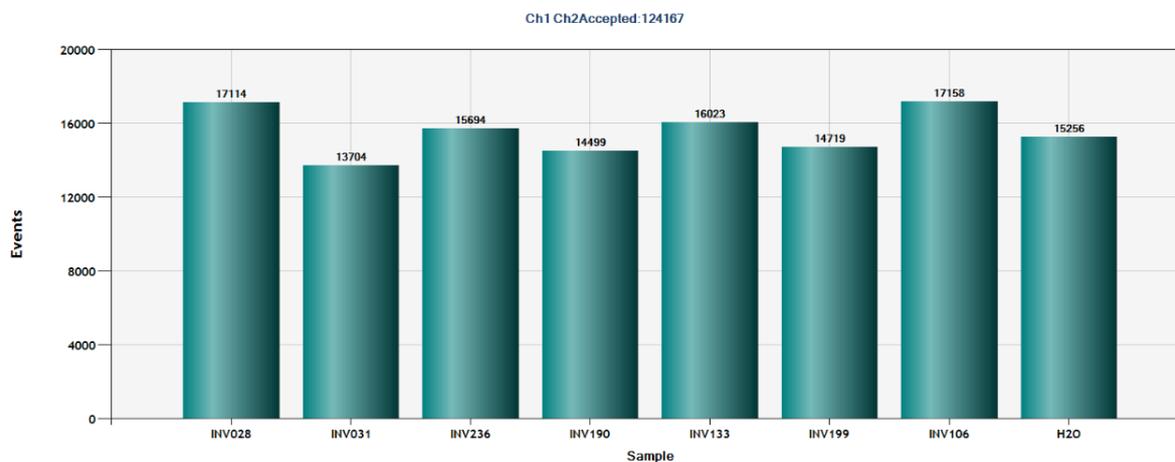


Figura 7: Gráfico de barras en el que se representa el número de eventos totales por muestra

El software permite ver también un gráfico 1-D, en el que se representa intensidad frente al número de gotitas, cada punto que se ve en el gráfico representa una gotita (Figura 8). Las que están por encima del “threshold” son las que se consideran positivas, se les asigna un valor de 1. Y las que están por debajo se puntúan como 0, se consideran negativas, porque no se ha detectado fluorescencia en el canal 1 (Ch1; fluoróforo FAM), ni el canal 2 (Ch2; fluoróforo HEX). El “threshold” lo delimita el usuario, por lo que puede estar sujeto a variaciones según el operario y la experiencia de quien realiza la ddPCR.

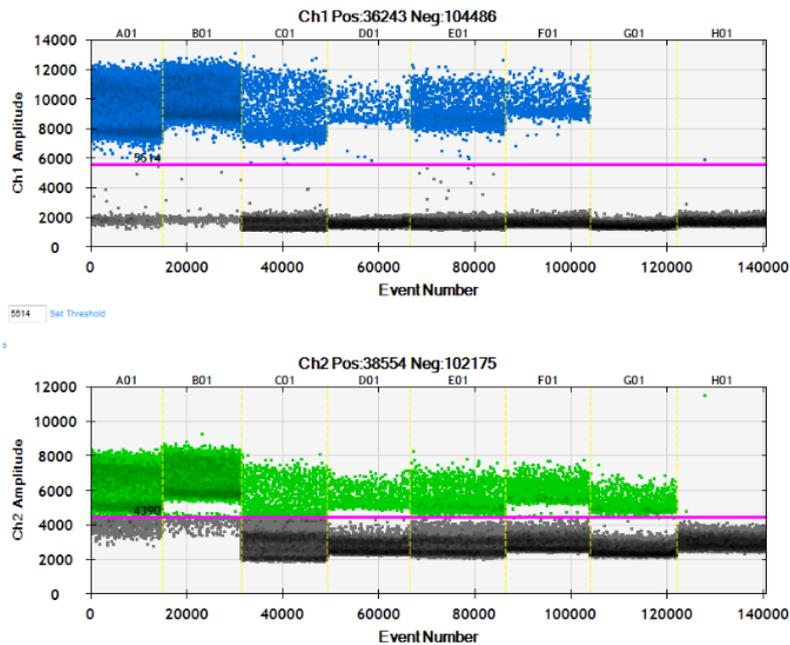


Figura 8: Gráfico 1D; en azul droplets positivas para el fluoróforo FAM, en verde droplets positivas para el fluoróforo HEX y en negro las droplets dobles negativas. (QUANTASOFT)

Además, en el programa se puede ver un gráfico 2D, en el cual el canal 1 de fluorescencia se representa frente al canal 2 de fluorescencia para cada gota (Figura 9). En azul se ve representado el fluorocromo FAM que va unido a la sonda para el gen diana, en verde se ve representado el fluorocromo HEX que va unido a la sonda para el gen control endógeno. En color naranja se representan los eventos positivos tanto para FAM como para HEX y en negro los eventos negativos para los dos fluorocromos.

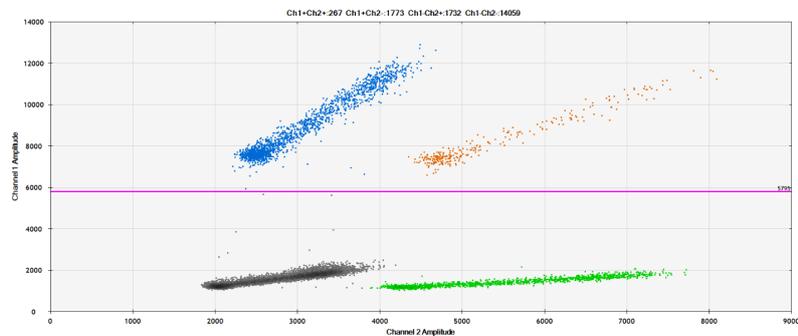


Figura 9: Gráfico 2D; en azul droplets positivas para el fluoróforo FAM, en verde droplets positivas para el fluoróforo HEX, en naranja las dobles positivas y en negro las droplets dobles negativas (QUANTASOFT)

Quantasoft crea un documento en formato csv del que se obtiene más información, cómo el número de eventos positivos, el número de eventos negativos para cada canal, la concentración (copias/ μ L) y el número de copias del gen entre otros. Además, el programa también proporciona un gráfico en el que detalla el número de copias de cada muestra y las barras de error de este resultado.

Los datos obtenidos fueron analizados mediante la información generada por el programa Quantasoft y para los análisis estadísticos se utilizó Statgraphics XVII.

4. Resultados y discusión

4.1. Digestión del ADN

Las muestras de ADN genómico procedente de muestras de la capa de linfocitos deben digerirse. Esta digestión permite una mayor exactitud al separar las copias de genes que están en tándem, reduce la viscosidad de la muestra y mejora la accesibilidad al ADN molde (BIORAD, 2019). En cuanto a las muestras de ADN procedentes de tejido tumoral parafinado, son muestras de ADN fragmentado, no es necesario en principio digerir. Sin embargo, se podría digerir para homogenizar los tamaños de los fragmentos de ADN. Las muestras de plasma, no hace falta digerirlas puesto que el ctDNA se encuentra fragmentado y en muy baja concentración en el plasma.

Así lo primero que se hizo fue testar ADN genómico digerido y sin digerir para el gen *GSTT1* y el gen endógeno *TERT* y de esta forma ver el efecto de la digestión en los resultados de ddPCR. En la reacción se utilizaron muestras con número de copias conocido, previamente testadas mediante qPCR (Ibarrola-Villava et al. 2012). En este primer ensayo las muestras analizadas fueron ADN genómico de los siguientes pacientes con melanoma: 04S1040, 04S1075 sin digerir, y las muestras 04S1040, 04S1046, 04S1075, 04S841, 04S659 digeridas. Se demuestra en el gráfico 1D que hubo saturación cuando las muestras estaban sin digerir (pocillos A1 y B1). Además, en este gráfico podemos observar que cada muestra se comporta de forma diferente, dado el distinto volumen de eventos positivos para FAM y HEX, y que no depende solamente de la concentración de ADN (Figura 10).

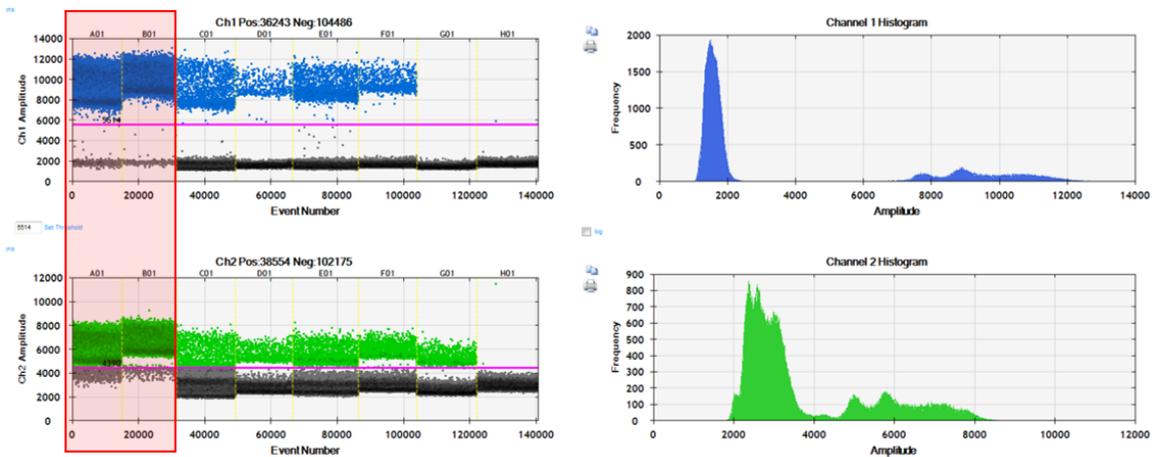


Figura 10: Gráfico 1D ensayo CNV de *GSTT1* siendo las muestras en la posición A1 y B1 ADN genómico sin digerir, y muestras digeridas en C1-G1.

Además, En los datos obtenidos del archivo .csv (Tabla 5) se observa que el número de dobles negativos es mayor de 10 % de las ‘droplets’ generadas en las muestras digeridas. Por lo tanto no es necesario disminuir la concentración de ADN. Sin embargo, en las muestras no digeridas el número de dobles negativos está por debajo del 10 %. En la muestra 04S1040 los dobles negativos representan el 0,97 % y en la 04S1075 representan un 0,14 %, con estas muestras no podríamos asegurar que verdaderamente el paciente tiene el número de copias que muestra el programa. Así, también se demuestra que cuando añadimos genómico sin digerir hay una saturación y que es mucho mejor utilizar el material genómico digerido.

Tabla 5: Archivo generado por Quantasoft, con datos del ensayo CNV de *GSTT1*

Sample	Target	Conc.	Positives	Negatives	Ch1+ Ch2+	Ch1+ Ch2-	Ch1- Ch2+	Ch1- Ch2-	Accepted Droplets	CNV
04S1040 madre	GSTT1	3390	14300	847	599	13701	700	147	15147	1.92
	TERT	3540	14401	746						
04S1075 madre	GSTT1	4600	16061	329	191	15870	305	24	16390	1.92
	TERT	5100	16175	215						
04S1040	GSTT1	143	2042	15789	1764	278	1226	14563	17831	2.76
	TERT	104	1504	16327						
04S1046	GSTT1	34.1	496	16879	473	23	864	16015	17375	1.11
	TERT	61.6	887	16488						
04S1075	GSTT1	158	2481	17277	2184	297	2149	15128	19758	2.03
	TERT	155	2446	17312						
04S841	GSTT1	58.9	862	16780	754	108	1695	15085	17642	0.93
	TERT	127	1803	15839						
04S659	GSTT1	0	0	18087	0	0	1336	16751	18087	0
	TERT	90.3	1336	16751						
H2O	GSTT1	0.06	1	18498	0	1	1	18497	18499	1
	TERT	0.13	2	18497						

Es muy importante llevar a cabo la digestión con la enzima adecuada que no corte en los puntos de unión de la secuencia diana. Si la digestión se lleva a cabo con una enzima no adecuada para el gen que se va a testar, esta cortará en los sitios de hibridación de la sonda con el gen y la reacción de ddPCR no se dará correctamente. Por lo tanto, no se producirá fluorescencia. Probamos a digerir el ADN extraído de la línea celular NCI-N87 con la enzima Hae III, la cual corta en la parte de la secuencia del gen *ERBB2* en la que la sonda hibrida con él. Los resultados fueron negativos para esta línea (Figura 11), puesto que a consecuencia de la no hibridación con la sonda, no había fluorescencia.

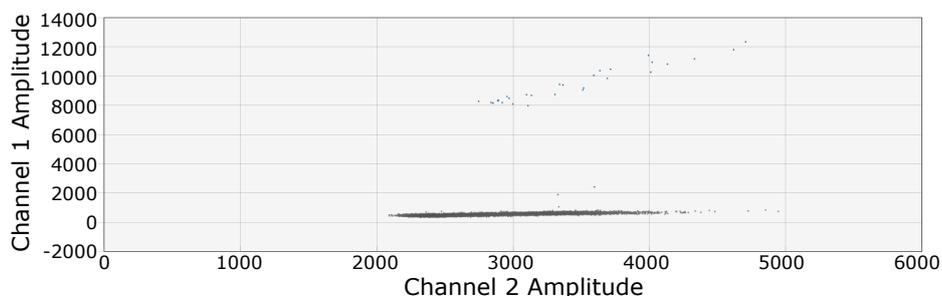


Figura 11: Gráfico 2D del ensayo CNV de *ERBB2*, resultado fallido para el ADN de la línea celular NCI-N87 por la utilización de una enzima de digestión incorrecta

Además, se probó a digerir el ADN procedente de parafinas. Aunque el material genético procedente de tejido parafinado ya está fragmentado, debido al proceso de conservación de la muestra, los fragmentos son de diversos tamaños. Por tanto, se digirió para homogenizar los tamaños de los fragmentos de ADN, y evitar el efecto lluvia o la dispersión de los eventos, en un intento de agrupar mejor los cústeres. Sin embargo, los resultados no fueron los esperados y no se obtuvieron grandes diferencias (Figura 12).

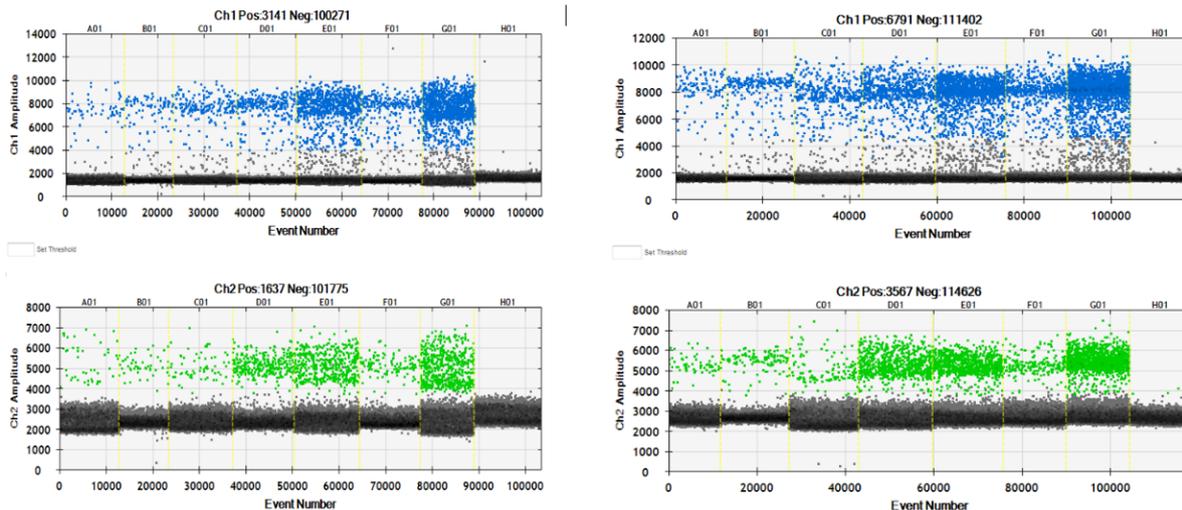


Figura 12: Gráficos 1D, a la izquierda ensayo de CNV de *PIK3CA* con ADN de tejido parafinado digerido, a la derecha ensayo de CNV de *PIK3CA* con las mismas muestras sin digerir

No existe una diferencia significativa entre las medias de positivos para FAM en las muestras de tejido tumoral parafinado digeridas y sin digerir. El promedio de positivos para FAM en muestras de tejido parafinado sin digerir es de 970,143 , mientras que el de tejido parafinado digerido es de 448,71. No obstante, la desviación estándar es mucho mayor en las no digeridas, 1032,63 frente a 475,27. La razón-F, que en este caso es igual a 1,47, es el cociente entre el estimado entre-grupos y el estimado dentro-de-grupos. Puesto que el valor-P de la razón-F es mayor o igual que 0,05, en concreto 0,25 no existe una diferencia estadísticamente significativa entre la media de Positivos FAM entre un nivel de muestra y otro, con un nivel del 5% de significación. Lo mismo sucede, para los positivos para HEX . El promedio de positivos para HEX en muestras de tejido parafinado sin digerir es de 509,14 , mientras que el de tejido parafinado digerido es de 233,86. No obstante, la desviación estándar es mucho mayor en las no digeridas, 521,64 frente a 225,84. El valor-P de la razón-F en este caso es 0,22. Por lo tanto, no existe una diferencia estadísticamente significativa entre la media de positivos HEX entre las muestras digeridas y las no digeridas, con un nivel del 5% de significación.

4.2. Determinación de la cantidad de ADN inicial

4.2.1. Cantidad adecuada de ADN inicial procedente de ADN genómico

Para conocer cuál es la cantidad adecuada de ADN genómico digerido para llevar a cabo la reacción de ddPCR, se realizó un gradiente de cantidades de ADN genómico digerido de la misma muestra (18S001). Este ensayo se llevó a cabo con la sonda del gen *GSTT1* y la del control endógeno *TERT*. En cada pocillo de la placa se introduce una cantidad de ADN total: 2,5, 5, 10, 20, 30, 40 y 50 ng de ADN genómico digerido y diluido 1:10. En el gráfico 1D generado por el programa Quantasoft (Figura 13), se comprueba el aumento de gotitas positivas conforme aumenta la cantidad de ADN.

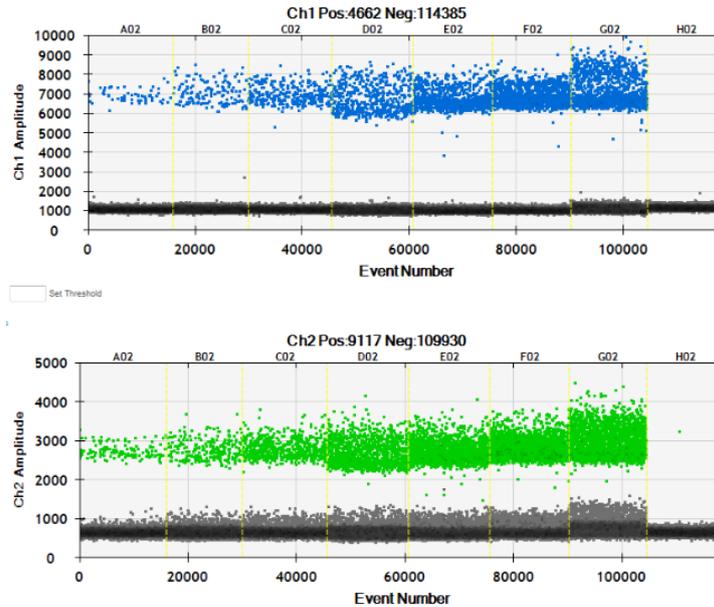


Figura 13: Gráfico 1D de ensayo de CNV de *GSTT1* en gradiente de cantidades de ADN genómico

En cualquier caso todas las cantidades de ADN genómico dieron como resultado el mismo número de copias. Cuando la concentración de ADN es baja, aumenta el tamaño del intervalo de confianza, y por tanto, las barras de error son más largas. Sin embargo, cuando la concentración es la adecuada las barras de error son de menor tamaño (Figura 14), como puede observarse a partir de 20 ng de ADN total en la reacción. .

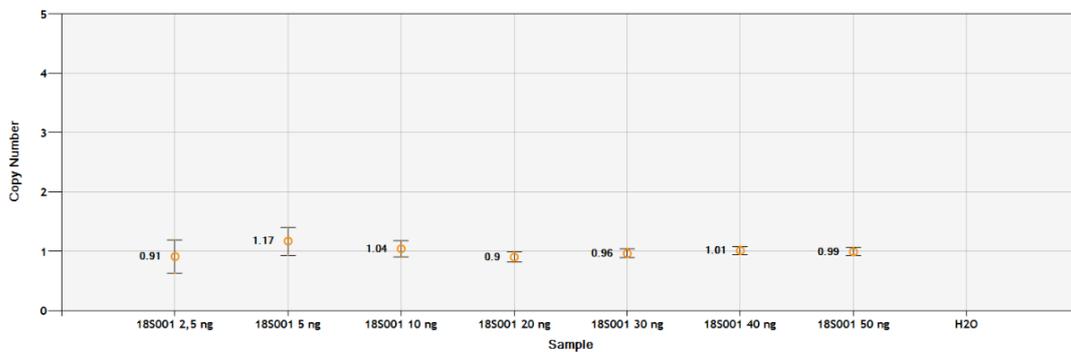


Figura 14: Gráfico de número de copias del ensayo para el gen *GSTT1* con el gradiente de concentración de ADN genómico total

Como siempre, se busca utilizar la menor cantidad de ADN posible y que de los mejores resultados, ya que las muestras son valiosas. En este caso las barras de error son aceptables cuando se utiliza una cantidad de ADN de 20 ng o más. Por tanto, consideramos que la cantidad de ADN genómico digerido óptima para llevar a cabo la reacción de ddPCR está entre 20 y 30 ng.

4.2.2. Cantidad adecuada de ADN inicial procedente de tejido tumoral parafinado

Con el fin de conocer la cantidad adecuada de ADN procedente de tejido tumoral parafinado, para llevar a cabo la reacción de ddPCR, se realizó un gradiente de cantidades de ADN de la misma muestra (INV199). Este ensayo se llevó a cabo con la

sonda del gen *ERBB2* y la del gen endógeno *TERT*. Se introdujo en cada pocillo una cantidad de ADN total de: 5, 10, 20, 40, 60, 80 y 100 ng. En el caso de ADN procedente de tejido tumoral parafinado el gradiente se hizo con mayor cantidad de ADN que el que se hizo con ADN genómico, puesto que este ADN está más fragmentado y es de menor calidad.

En el gráfico 1D generado, se observa con claridad el efecto lluvia, que se obtiene siempre que utilizamos tejido parafinado. Esto no se observa cuando se usaba ADN genómico o plasma (Figura 15). El efecto lluvia se explicará más adelante.

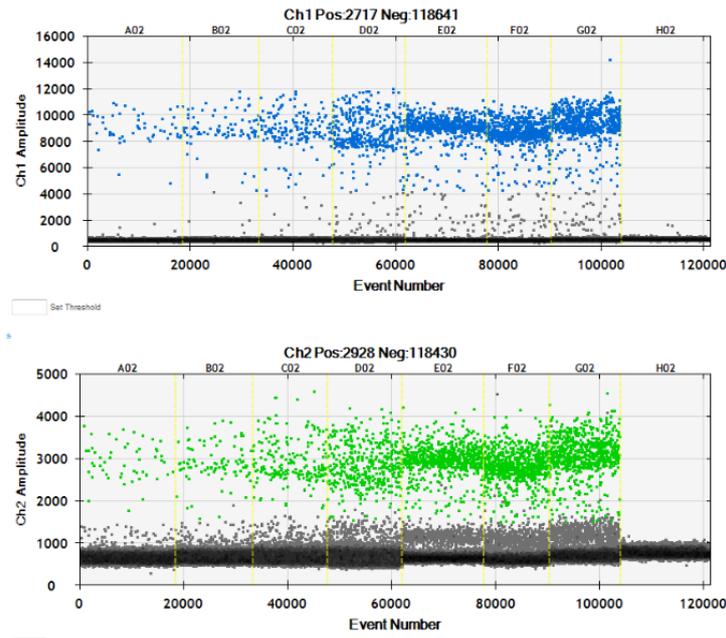


Figura 15: Gráfico 1D del ensayo de CNV para el gen *ERBB2* con el gradiente de concentración de ADN total procedente de tumores incluidos en parafina.

Al igual que en el caso anterior vemos que al aumentar la cantidad de ADN total en la reacción aumentan el número de eventos positivos. Al aumentar la cantidad las barras de error disminuyen, y cada vez el número de copias se aproxima con más exactitud a 2, el número de copias del gen *ERBB2* en esta muestra. Se afirma así que hay menos variación, en cuanto a número de copias entre las cantidades de ADN más altas (Figura 16).

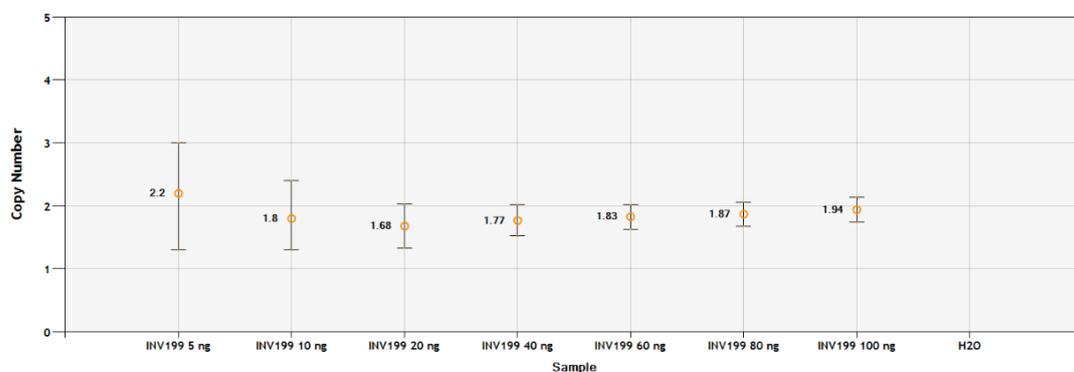


Figura 16: Gráfico de número de copias del ensayo de CNV para el gen *ERBB2* con el gradiente de concentración de ADN total procedente de tumores incluidos en parafina.

Pensamos que la cantidad de ADN procedente de tejido tumoral parafinado adecuada es de entre 80 y 100 ng, ya que el número de eventos positivos es aceptable, sobre 800 eventos y las barras de error son menores. Cabe decir, que la reacción no ha llegado a saturarse con la cantidad de 100 ng y por tanto podría añadirse más cantidad. Sin embargo, estas muestras proceden de biopsias convencionales, son muestras muy valiosas, y cómo siempre queremos utilizar la mínima cantidad de muestra que de los mejores resultados.

4.2.3. Cantidad adecuada de ADN inicial procedente de plasma

El ctDNA se encuentra a muy baja concentración en el plasma, por tanto para detectar alteraciones genéticas no se debe ni digerir ni diluir. En este trabajo se añadieron en todas las reacciones entre 160 y 240 ng de ADN procedente de plasma. No fue posible añadir más puesto que el máximo volumen de muestra que se puede añadir a la reacción es de 8 μ L. En el gráfico de número de copias vemos que las barras de error no son lo suficientemente cortas (Figura 17).

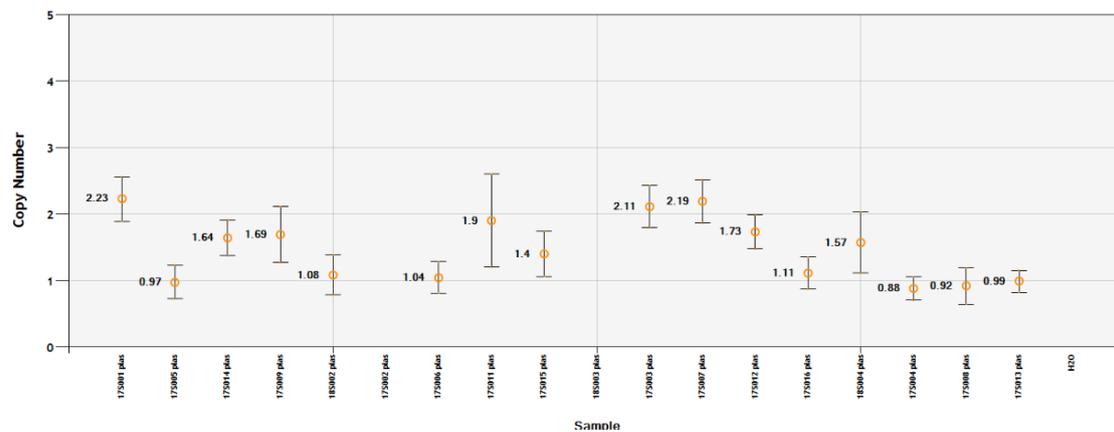


Figura 17: Gráfico de número de copias del gen *GSTT1* en muestras de cfDNA procedente de plasma

Por ello, sería necesario añadir la máxima cantidad de muestra posible en la reacción, puesto que el cfDNA en plasma está en muy baja concentración y es muy inestable.

4.3. Establecimiento del threshold

El hecho de separar bien aquellas gotitas positivas para FAM o HEX de aquellas negativas, implica al final una buena determinación del número de copias. El programa Quantasoft establece un threshold automático, pero consideramos que es mejor ajustar este umbral manualmente. El establecimiento manual del threshold representa un reto, puesto que existen gotitas que no son claramente positivas o negativas, si no que se encuentran a medio camino entre las dos, es lo que denominamos efecto lluvia (Figura 18). El software Quantasoft separa en dos canales aquellas gotitas con fluorescencia emitida por el fluoróforo FAM, de aquellas con fluorescencia emitida por el fluoróforo HEX. En este experimento se separan mejor las que tienen fluorescencia para FAM. La separación en el canal HEX es menos buena que en el canal FAM. Este fenómeno se observó en todos los ensayos de ddPCR realizados con el gen control endógeno *TERT*.

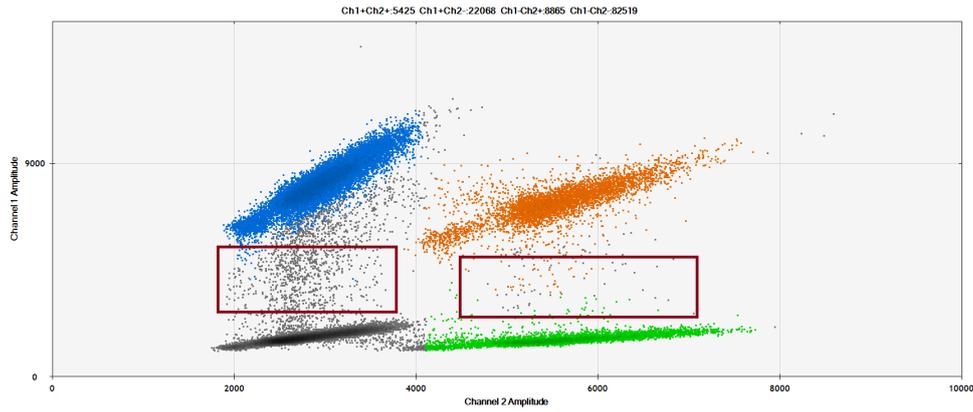


Figura 18: Gráfico 2D ejemplo de lluvia generada principalmente cuando el material genético utilizado proviene de tejido parafinado

El threshold lo establece el usuario que lleva a cabo la ddPCR por tanto este dependerá de la experiencia del operario. No obstante si se modifica el threshold (Figura 19) no afecta al número de copias en muestras de ADN genómico, siempre es el mismo (Tabla 6). Esto se debe a que el efecto lluvia es apenas detectable cuando el ADN es genómico, y la mayoría de gotitas están dentro de los clústers bien definidos.

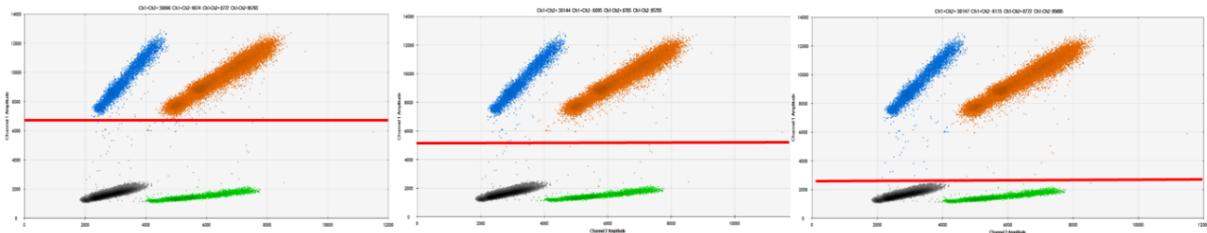


Figura 19: Tres threshold diferentes para el mismo ensayo con ADN genómico

Tabla 6: Número de copias del gen *GSTT1* por muestra de ADN genómico según los diferentes thresholds.

Muestras	Nº de copias threshold 1	Nº de copias threshold 2	Nº de copias threshold 3
04S1040 madre	1,97	1,96	1,98
04S1075 madre	1,91	1,90	1,91
04S1040 digerida	2,04	2,04	2,05
04S1046 digerida	1,10	1,10	1,10
04S1075 digerida	2,03	2,02	2,04
04S841 digerida	0,93	0,92	0,93
04S659 digerida	0	0	0

Si se modifica el threshold cuando se utilizan muestras de ADN procedente de tejido tumoral parafinado (Figura 20), el número de copias no varía, sí que varían los decimales en mayor medida en que lo hacían en el número de copias en ADN genómico (Tabla 7), debido a esa mayor dispersión en la nube de puntos.

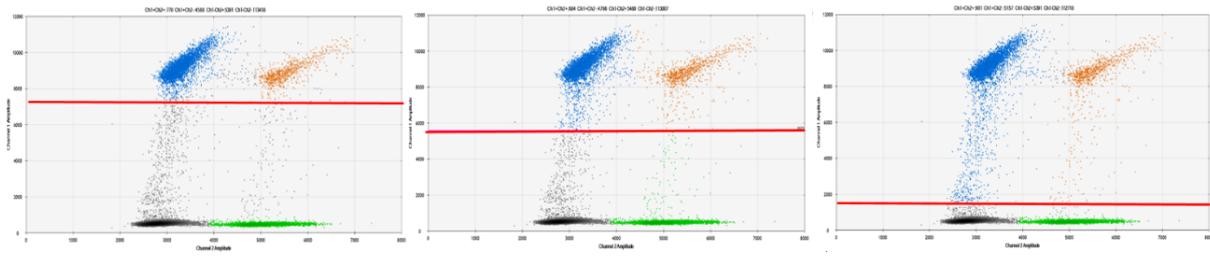


Figura 20: Tres threshold diferentes para el mismo ensayo con ADN procedente de parafinas

Tabla 7: Número de copias del gen *ERBB2* por muestra de ADN procedente de tejido parafinado con los diferentes thresholds

Muestras	Nº de copias threshold 1	Nº de copias threshold 2	Nº de copias threshold 3
INV028	2,05	2,19	2,41
INV031	0,85	0,89	0,99
INV236	2,18	2,22	2,34
INV190	1,93	1,96	2,13
INV133	1,64	1,71	1,82
INV199	1,31	1,36	1,48
INV106	1,10	1,10	1,30

Finalmente, destacar que se utilizaron las sondas de dos genes como controles endógenos *AGO1* y *TERT*. Así se podía comprobar si existía alguna diferencia en cuanto a los gráficos. En los gráficos 2D se pueden observar que cuando se utiliza la sonda para el gen *AGO1* se obtiene una mejor separación de los clusters, que utilizando la sonda para el gen *TERT* (Figura 21). Este fenómeno, se observa tanto en los ensayos para determinar el número de copias de *GSTT1*, como de *ERBB2* y *PIK3CA*.

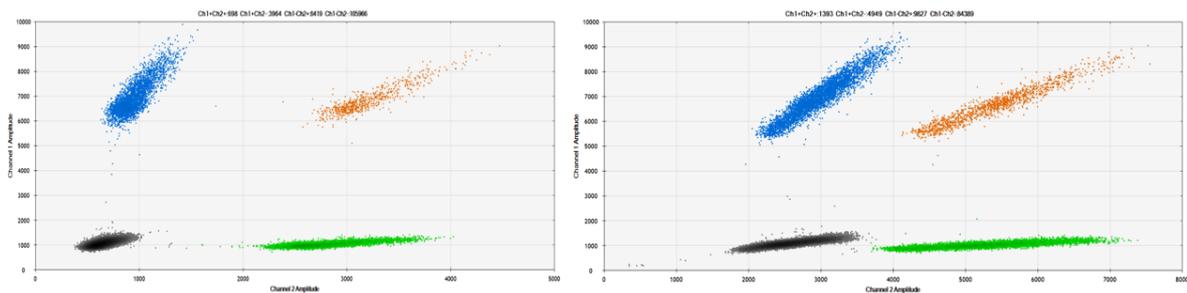


Figura 21: Izquierda; gráfico 2D utilizando la sonda para el gen *AGO1*, derecha; gráfico 2D utilizando la sonda para el gen *TERT*

Por lo tanto, se decidió intentar solucionar la mala separación obtenida cuando se utilizaba la sonda para el gen *TERT*, modificando el programa de ddPCR, como veremos a continuación.

4.4. Modificación del programa ddPCR

El próximo paso fue modificar el programa de la ddPCR, para ver si se conseguían separar mejor los clusters. Como cuando se utiliza el gen endógeno *TERT* se separan

peor, se utilizó este para llevar a cabo los siguientes ensayos. Así podría verse mejor el efecto que tenía el cambio en el programa sobre la separación de los clusters.

En este trabajo se modificaron tanto el número de ciclos de PCR, cómo el tiempo de annealing, con el fin de conseguir una mejor hibridación de las sondas con los respectivos genes. Se podría haber llevado a cabo un gradiente de temperaturas de annealing, pero los especialistas de la casa comercial BioRad no lo aconsejaron puesto que se trata de sondas ya probadas. Y por tanto, se entiende que deberían funcionar perfectamente a la temperatura que aconsejan. Sin embargo, puede cambiarse en experimentos futuros para una mayor seguridad.

4.4.1. Modificación del número de ciclos

Los pasos de desnaturalización, annealing y extensión se repiten durante la amplificación. En el programa de ddPCR para la determinación del número de copias de un gen recomendado por la casa comercial BioRad se programan 40 ciclos.

Se aumentó el número de ciclos a 50 para ver si así aumenta la separación entre los eventos positivos para *TERT* y los eventos negativos. Se lleva a cabo este experimento con muestras de ADN genómico digerido y con las muestras de cfDNA de plasma de pacientes con melanoma maligno. A pesar de haber cambiado el número de ciclos, cuando se comparan los gráficos 1D vemos que siguen sin separarse bien los eventos positivos para *TERT* de los eventos negativos para este. (Figura 22)

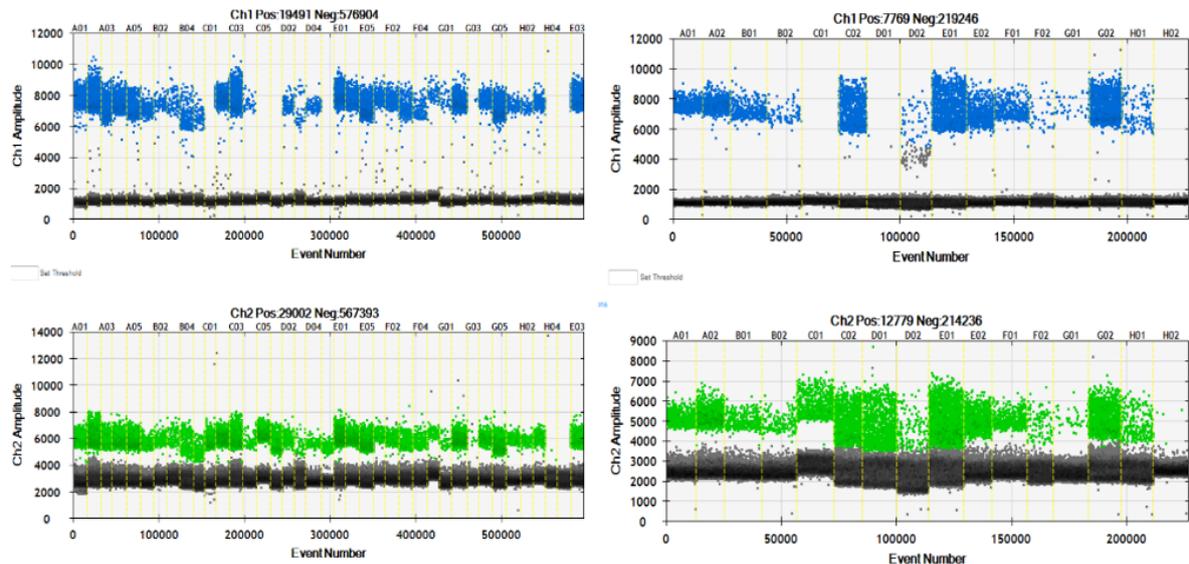


Figura 22: A la izquierda; gráfico 1D resultado de ddPCR realizada con 40 ciclos. A la derecha; gráfico 2D resultado de ddPCR realizada con 50 ciclos

4.4.2. Modificación del tiempo de annealing

También cambiamos el tiempo de annealing de un minuto a un minuto y medio, en este experimento se utilizaron muestras de tejido tumoral parafinado de pacientes con cáncer gástrico y colorrectal, para comprobar si disminuía el efecto lluvia y obteníamos de igual forma una mejor separación de los clusters.

No obstante, ni el efecto lluvia ni la separación de eventos positivos y negativos para *TERT* mejoraban con este cambio. En el gráfico 2D de cada experimento, se puede observar que en el que se modificó el programa también había menos cantidad de ADN (10 ng totales), por tanto se muestra que el efecto lluvia continua estando presente en la misma proporción a la cantidad de ADN (Figura 23). De esta forma, el usuario iba a tener el mismo problema a la hora de establecer el threshold tanto si el tiempo de annealing es de un minuto, cómo si es de un minuto y medio.

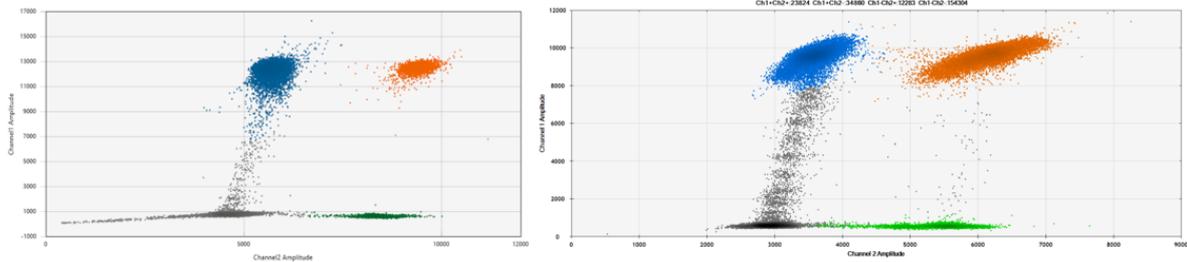


Figura 23: A la izquierda; gráfico 2D resultado de ddPCR realizada con un minuto y medio de tiempo de annealing. A la derecha; gráfico 2D resultado de ddPCR realizada con un minuto de tiempo de annealing.

Debido a los resultados anteriores se decidió aumentar el tiempo de annealing de un minuto y medio a dos minutos. Pero cómo se aprecia en el gráfico 2D comparado con el gráfico 2D del experimento en el que el tiempo de annealing era de 1 minuto, el efecto lluvia no mejora (Figura 24).

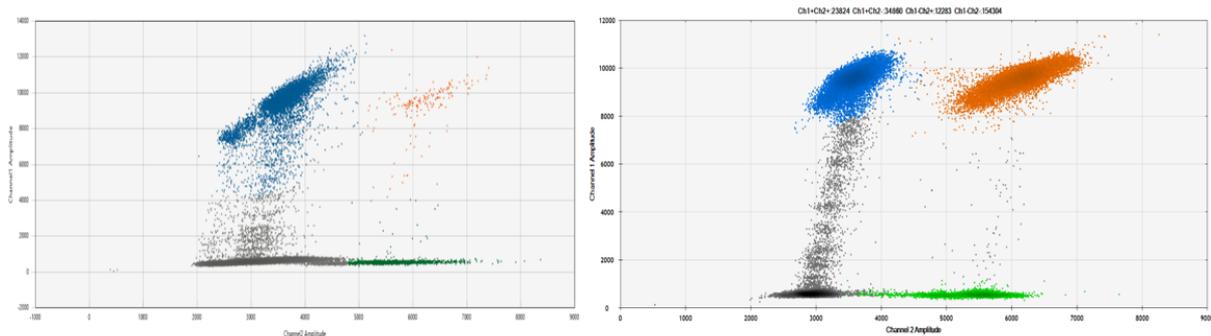


Figura 24: A la izquierda; gráfico 2D resultado de ddPCR realizada con dos minutos de tiempo de annealing. A la derecha; gráfico 2D resultado de ddPCR realizada con un minuto de tiempo de annealing.

4.5. Variación en el número de copias (CNV)

4.5.1. Evaluación del número de copias de los distintos genes de interés en pacientes con tumores sólidos

El programa Quantasoft genera un gráfico en el que representa el número de copias calculado para cada muestra, cada valor viene acompañado de una barra de error, cuanto menor es la barra de error más fiable es el resultado. Los resultados han sido consistentes cuando se han repetido muestras en diferentes ensayos y los resultados no cambian. Según el gen que se está testando, el paciente y el tipo de cáncer, el número de copias de un gen será uno u otro. A continuación, se explicarán los resultados obtenidos por gen.

En los estudios realizados para determinar el número de copias del gen *GSTT1* en pacientes con melanoma maligno, se observa que el número de copias de este gen varía entre 2, 1 y 0 (Figura 25). Lo normal es tener 2 copias del gen puesto que somos diploides. Algunos pacientes con este tipo de cáncer presentan deleciones de este gen. Esto puede explicarse, puesto que *GSTT1* codifica una glutatión-S-transferasa, las cuales participan en la detoxificación de un gran número de especies reactivas de oxígeno (ROS), que se producen por el estrés oxidativo causado por la radiación ultravioleta.

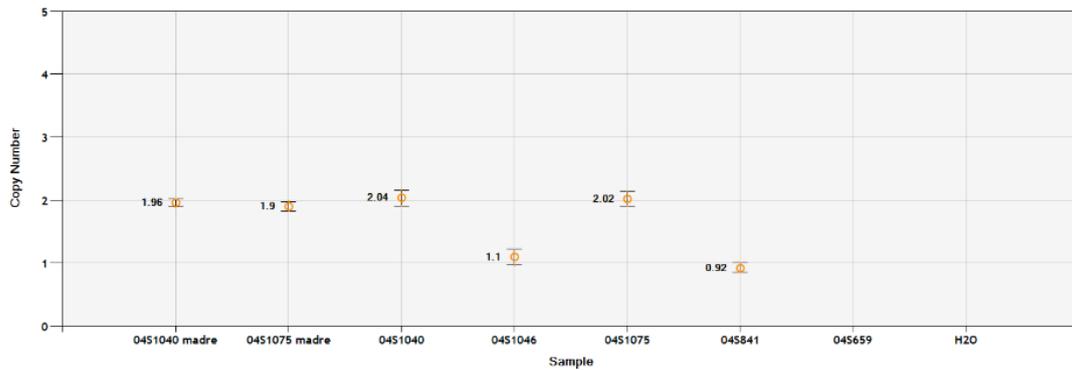


Figura 25: Número de copias del gen *GSTT1* en muestras de ADN genómico de pacientes con melanoma maligno

En todas las muestras el número de copias coincidía con el resultado obtenido en el estudio de Ibarrola-Villava et al (2012). A su vez las dos muestras que están tanto digeridas como sin digerir coinciden en el número de copias, por tanto se consiguió un 100 % de sensibilidad y especificidad.

En el gráfico 2D se puede observar que conforme el número de copias va disminuyendo el número de eventos para el *GSTT1* también lo hace. También disminuye el número de eventos dobles positivos para FAM/HEX, mientras que el número de eventos positivos para HEX (gen *TERT*) y el número de dobles negativos permanecen prácticamente constantes (Figura 26).

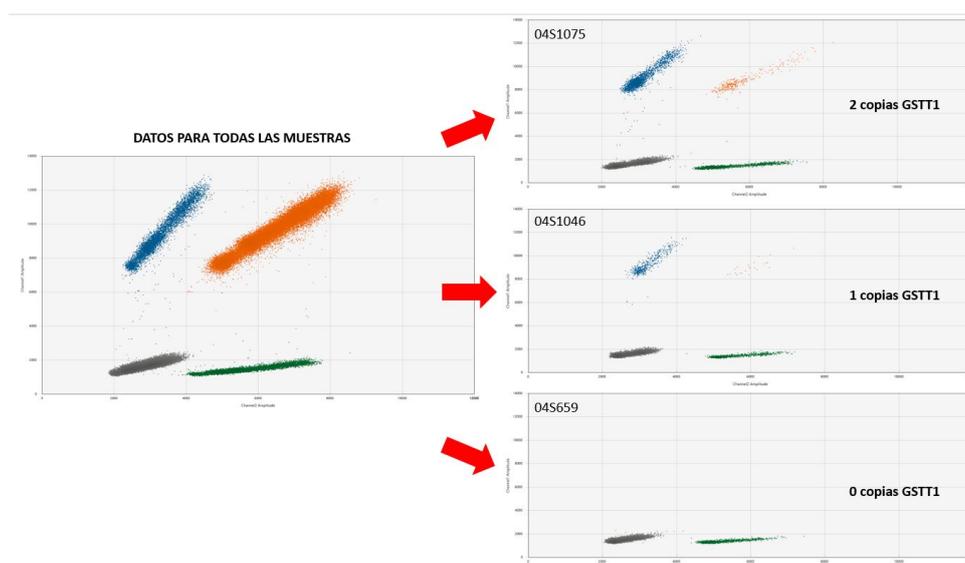


Figura 26: Gráficos 2D, para 0, 1 y 2 copias del gen *GSTT1*

El papel de los genes *GSTs* en melanoma no es claro, así el estudio de CNV en *GSTT1* o *GSTM1* no ha determinado un papel relevante de estos genes en el desarrollo del cáncer. Sin embargo, determinar el número de copias ha permitido establecer que cuando el número de copias es cero y además existen antecedentes fenotípicos específicos conocidos aumenta la susceptibilidad a desarrollar melanoma (Kanetsky et al 2001)

Por otra parte, en los estudios realizados para determinar el número de copias del gen *ERBB2* en pacientes con cáncer colorrectal, el número de copias va de 1 a 2 copias. Cabe destacar que en este estudio el número de pacientes es muy pequeño. En cáncer gástrico el número de copias de cada paciente es más diverso.(Figura 27)

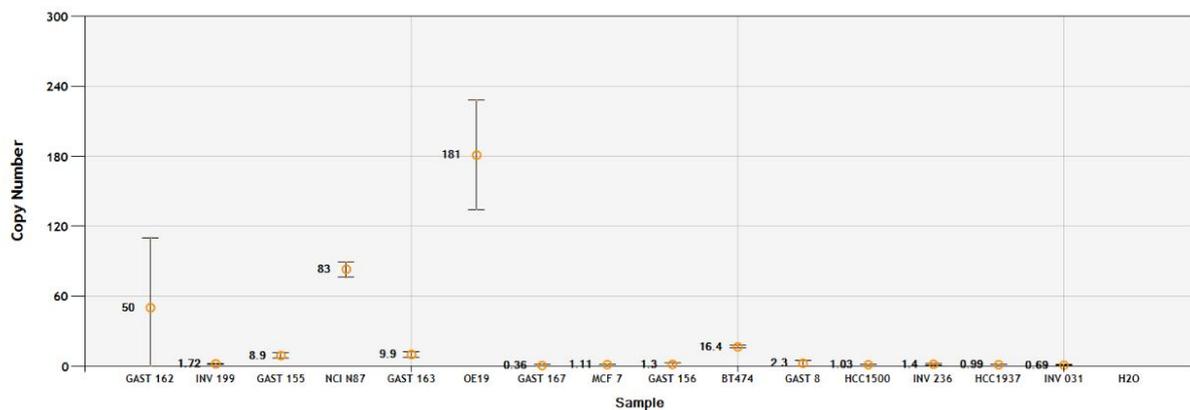


Figura 27: Número de copias del gen *ERBB2* en pacientes con cáncer colorrectal y en pacientes con cáncer gástrico.

Este gen tiene mucha importancia en algunos tipos de cáncer, sobretodo en cáncer de mama dónde aquellas pacientes con amplificación del gen *ERBB2* reciben un tratamiento basado en un anticuerpo humanizado (Trastuzumab) (Boekhout et al., 2011). Sin embargo, también se ha visto la importancia de la amplificación de este gen en cáncer colorrectal (Owen et al., 2018) y en cáncer gástrico, donde los pacientes en cuyo tumor *ERBB2* está amplificado pueden también beneficiarse de la terapia con Trastuzumab (Gao et al.,2017). Recientemente se ha destacado el estudio de CNV de *ERBB2* en ctDNA de plasma. El uso de técnicas no invasivas para guiar el tratamiento con Trastuzumab es un gran avance (Siravegna et al., 2019)

Finalmente, en cuanto al gen *PIK3CA*, hemos observado una variación en el número de copias de entre 2 y 12 en pacientes con tumores de cabeza y cuello (Figura 28). La amplificación de este gen se ha relacionado con un mal pronóstico, en este tipo tumoral, puesto que está relacionado con la metástasis ganglionar, de ahí la importancia de su evaluación como factor pronóstico. (Suda et al. 2012).

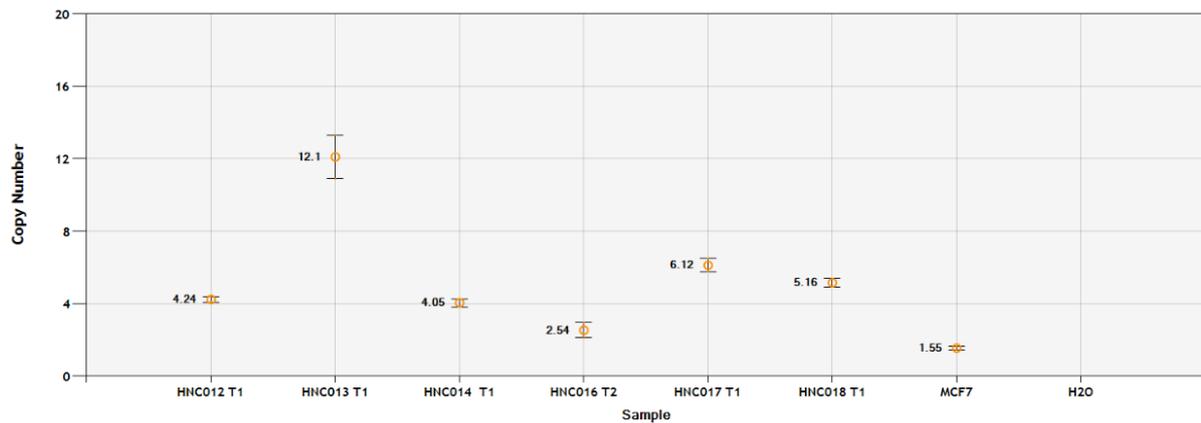


Figura 28: Número de copias del gen *PIK3CA* en pacientes con cáncer de cabeza y cuello

Asimismo, vemos una variación del número de copias de *PIK3CA* de entre 2 y 6 en pacientes con cáncer gástrico (Figura 29). Las muestras utilizadas habían sido analizadas anteriormente con inmunohistoquímica y qPCR.

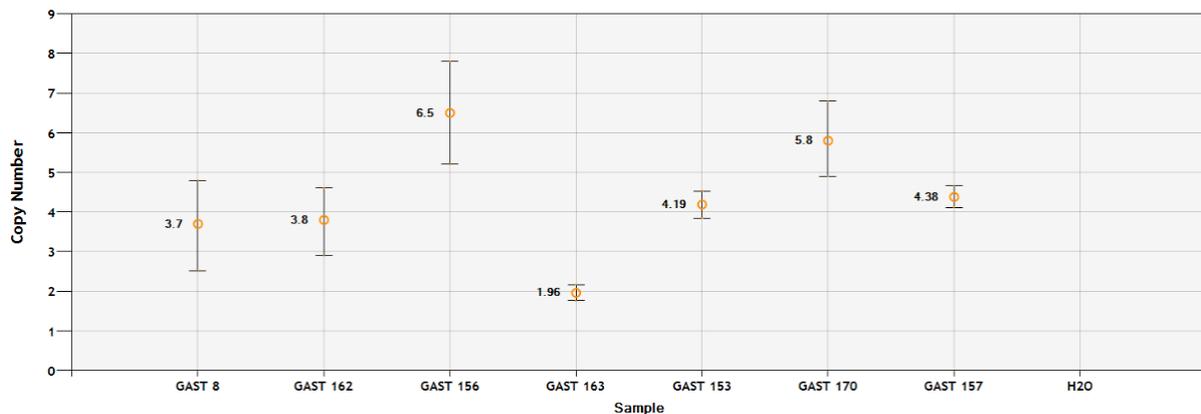


Figura 29: Número de copias del gen *PIK3CA* en pacientes con cáncer gástrico

Este gen tiene importancia clínica en este tipo de cáncer. Porque la amplificación, en este caso cuando hay más de 4 copias, del gen *PIK3CA* está significativamente asociada a un menor tiempo de supervivencia (Shi et al., 2012).

En todos los casos se requiere una evaluación patológica del porcentaje de celularidad tumoral ya que un porcentaje bajo de celularidad puede estar enmascarando los resultados.

4.5.2. Correlación entre CNV en muestras de plasma y CNV en muestras de ADN genómico

En un ensayo para determinar el número de copias de *GSTT1* se usaron muestras de ADN genómico digerido y sus correspondientes ctDNA extraídos a partir de plasma (Tabla 8).

Tabla 8: Número de copias del gen *GSTT1* por muestra en ADN genómico y ctDNA

Muestras	Nº de copias genómico	Nº de copias plasma
17S001	1,87	2,23
17S002	0	0
17S003	1,06	2,10
17S004	1,26	0,92
17S005	1,30	0,97
17S006	1,38	1,04
17S007	2,24	1,73
17S008	1,10	0,92
17S009	2,74	1,68
17S011	2,15	2,00
17S012	1,97	2,19
17S013	1,1	0,98
17S014	1,98	1,62
17S015	0,78	1,40
17S016	1,02	1,10
18S002	1,11	1,00
18S003	0	0
18S004	2,07	1,80

Existe una buena correlación entre el número de copias del gen *GSTT1* entre las muestras de ctDNA y las de ADN genómico, ya que existe una relación lineal y un buen coeficiente de correlación de Pearson (R). Como aparece en el gráfico (Figura 30) el coeficiente de determinación (R²) es igual a 0,6294, por tanto el coeficiente de Pearson (R) es igual a 0,793.

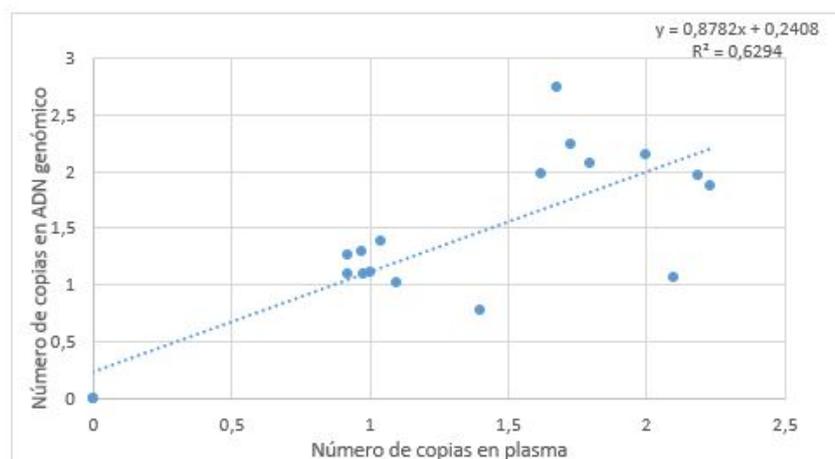


Figura 30: Correlación lineal entre el número de copias del gen *GSTT1* en muestras de ctDNA de plasma y en muestras de ADN genómico

Los valores absolutos del coeficiente de Pearson oscilan entre 0 y 1. Una correlación será perfecta positiva entre dos variables cuando una aumenta exactamente en la medida que aumenta la otra. En este caso, la correlación sería perfecta sí cuando el número de copias en la muestra de ctDNA de plasma aumenta en la misma medida que en las muestras de ADN genómico. Una vez se ha calculado el coeficiente de correlación, se debe determinar si este valor se debe a que las dos variables están relacionadas en realidad o sólo se debe

al azar. Se debe comprobar que el coeficiente de Pearson es significativo. Para ello se lleva a cabo un Test t-student. La t calculada es igual a 5,2, buscamos en la tabla para α 0,05 y 16 grados de libertad, en la tabla 1.74. Por tanto, se rechaza la Hipótesis nula con un riesgo (máximo) de equivocación de 0.05. Se concluye que el número de copias del gen *GSTT1* en ctDNA de plasma y en ADN genómico está relacionado significativamente.

Además, con ayuda del programa Statgraphics se realizó un gráfico de caja y bigotes, para poder comparar el número de eventos positivos para FAM y para HEX, siendo 1 las muestras de ADN genómico digerido y 0 las muestras de ctDNA proveniente de plasma (Figura 31).

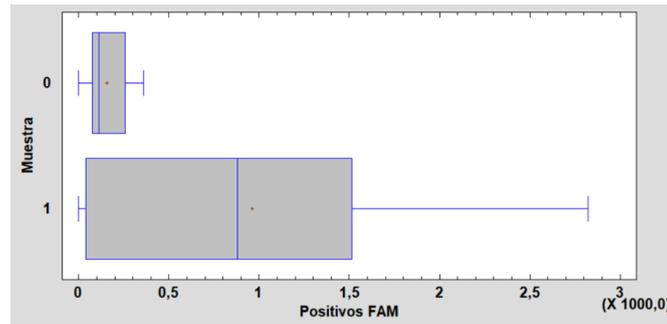


Figura 31: Gráfico de caja y bigotes en el que se representan los positivos para FAM tanto en muestras de ctDNA de plasma cómo en muestras de ADN genómico

La mediana de eventos positivos para FAM en ctDNA de plasma es de 113, y la media es 156,17. Mientras que en las muestras de ADN genómico digerido la mediana es 881 y la media 962,83. La prueba-F en la tabla ANOVA determinará si hay diferencias significativas entre las medias (ver Anexos, Tabla 2).

Puesto que el valor-P de la prueba-F es 0,0001, menor que 0,05, existe una diferencia estadísticamente significativa entre la media de eventos positivos para FAM entre los dos tipos de muestras, con un nivel del 5% de significación. Esto se debe a la baja cantidad de ctDNA que hay en el plasma.

Se comprobó también la diferencia que había entre los tipos de muestras en los eventos positivos para HEX (Figura 32). Se observa que las muestras de ctDNA de plasma hay menos positivos para HEX.

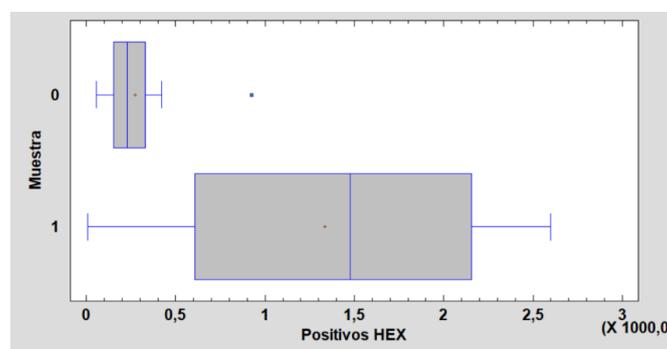


Figura 32: Gráfico de caja y bigotes en el que se representan los positivos para HEX tanto en muestras de ctDNA de plasma cómo en muestras de ADN genómico

La mediana de positivos para HEX en plasma es de 227, y la media es 270,72. Mientras que en las muestras de ADN genómico digerido la mediana es 1475,50 y la media es 1336,67. Se realizó la prueba-F en la tabla ANOVA (ver Anexos, Tabla 3).

Cómo el valor-P de la prueba-F es 0, menor que 0,05, existe una diferencia estadísticamente significativa entre la media de Positivos HEX entre un nivel de Muestra y otro, con un nivel del 5 % de significación.

4.5.3. Correlación entre CNV de muestras de plasma y CNV de muestras de ADN procedente de parafinas

Para determinar el número de copias del gen *PIK3CA* se realizó un ensayo con las muestras de pacientes con cáncer de cabeza y cuello; HNC-012, HNC-013 y HNC-014, tanto de tejido parafinado como de plasma (Tabla 9). También en este caso se comprobó si existía correlación entre los dos tipos de muestras (Figura 33).

Tabla 9: Número de copias del gen *GSTT1* por muestra tanto en ctDNA de plasma como en ADN procedente de tejido parafinado

Muestras	Nº de copias parafina	Nº de copias plasma
HNC-012	4,27	1,97
HNC-013	11,7	2,35
17S003	3,82	2,38

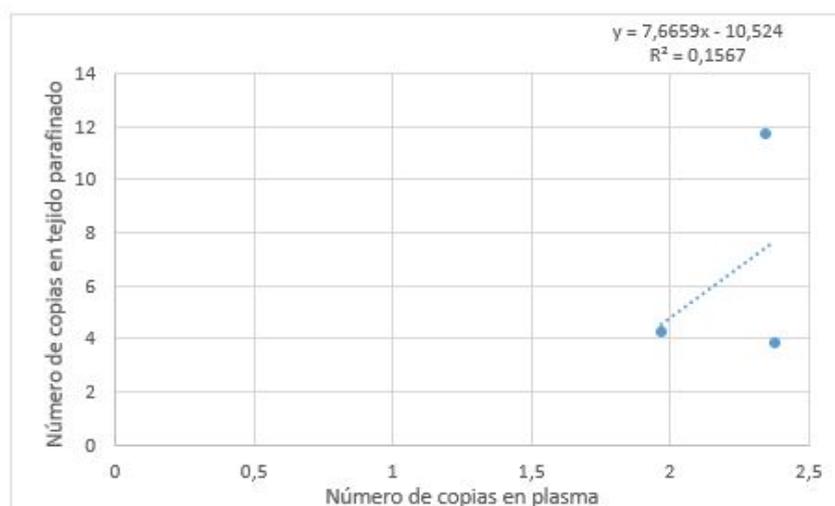


Figura 33: Número de copias del gen *PIK3CA* en muestras de ctDNA de plasma frente a número de copias de muestras de tejido parafinado

En este caso el coeficiente de determinación es 0,1567 y el coeficiente de correlación de Pearson es igual a 0,431. La t calculada es 0,431 buscamos en la tabla para α 0,05 y 1 grado de libertad ($t = 12,7$). Como la t calculada es menor está relación lineal se debe al azar y no existe una buena correlación entre las dos variables.

No hay una buena correlación, ya viendo los datos de la tabla se observa que todas las muestras difieren en el número de copias entre el tejido parafinado y el plasma. Sin embargo es arriesgado hacer conclusiones teniendo en cuenta que sólo se disponía de tres muestras, para hacer afirmaciones deberíamos ampliar el tamaño muestral.

5. Conclusiones

Las conclusiones a las que se han llegado una vez realizado este trabajo son:

- Una buena digestión del ADN genómico es fundamental para llevar a cabo la reacción de ddPCR.
- La cantidad adecuada de ADN por reacción cuando el ADN proviene de genómico digerido es de 20 ng totales.
- La cantidad adecuada de ADN por reacción cuando el ADN proviene de tejido tumoral parafinado es de entre 80 a 100 ng totales.
- De cfDNA procedente de plasma es necesario añadir la máxima cantidad posible.
- Aumentar el número de ciclos de ddPCR no hace que se separen mejor los eventos en el canal 2 en los ensayos que hemos testado.
- Aumentar el tiempo de annealing de 1 minuto a 2 minutos no hace que se separen mejor los clusters en los ensayos que hemos testado.
- Existe una correlación significativa entre el número de copias del gen *GSTT1* en ctDNA de plasma y ADN genómico en pacientes con melanoma maligno.
- No existe una correlación entre el número de copias del gen *PIK3CA* del ADN de tejido parafinado y del ctDNA en pacientes con cáncer de cabeza y cuello, posiblemente debido a la no presencia de ctDNA en el cfDNA.
- Se ha comprobado que cada muestra tiene un comportamiento diferente, esto también depende del porcentaje tumoral de la muestra.

6. Bibliografía

1. AMIN, M. (2017). AJCC cancer staging manual. New York: Springer
2. ARRUEBO, M., VILABOA, N., SÁEZ-GUTIERREZ, B., LAMBEA, J., TRES, A., VALLADARES, M. AND GONZÁLEZ-FERNÁNDEZ, Á. (2011). Assessment of the Evolution of Cancer Treatment Therapies. *Cancers*, 3(3), pp.3279-3330.
3. BEROUKHIM, R., ZHANG, X. AND MEYERSON, M. (2017). Copy number alterations unmasked as enhancer hijackers. *Nature Genetics*, 49(1), pp.5-6.
4. BOEKHOUT, A., BEIJNEN, J. AND SCHELLENS, J. (2011). Trastuzumab. *The Oncologist*, 16(6), pp.800-810.
5. BRAY, F., FERLAY, J., SOERJOMATARAM, I., SIEGEL, R., TORRE, L. AND JEMAL, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 68(6), pp.394-424.
6. CANCER RESEARCH UK. (2019). Types of cancer. [online] Available at: <https://www.cancerresearchuk.org/what-is-cancer/how-cancer-starts/types-of-cancer> [Accessed 30 Apr. 2019].
7. CANCER.ORG. (2019). Types of biopsies used to look for cancer. [online] Available at: <https://www.cancer.org/treatment/understanding-your-diagnosis/tests/testing-biopsy-and-cytology-specimens-for-cancer/biopsy-types.html> [Accessed 17 May 2019].
8. CAO, L., CUI, X., HU, J., LI, Z., RU, C., YANG, Q., LIN, M., HUILLI, Y., XU, F. (2017). Advances in digital polymerase chain reaction (dPCR) and its emerging biomedical applications. *Biosensors and Bioelectronics* 90, pp. 459-474.
9. CORCORAN, R. AND CHABNER, B. (2019). Application of Cell-Free DNA Analysis to Cancer Treatment. *Obstetrical & Gynecological Survey*, 74(4), pp.204-206.

10. FEITELSON, M., ARZUMANYAN, A., KULATHINAL, R., BLAIN, S., HOLCOMBE, R., MAHAJNA, J., MARINO, M., MARTINEZ-CHANTAR, M., NAWROTH, R., SANCHEZ-GARCIA, I., SHARMA, D., SAXENA, N., SINGH, N., VLACHOSTERGIOS, P., GUO, S., HONOKI, K., FUJII, H., GEORGAKILAS, A., BILSLAND, A., AMEDEI, A., NICCOLAI, E., AMIN, A., ASHRAF, S., BOOSANI, C., GUHA, G., CIRIOLO, M., AQUILANO, K., CHEN, S., MOHAMMED, S., AZMI, A., BHAKTA, D., HALICKA, D., KEITH, W. AND NOWSHEEN, S. (2015). Sustained proliferation in cancer: Mechanisms and novel therapeutic targets. *Seminars in Cancer Biology*, 35, pp.S25-S54.
11. GAO, J., WANG, H., ZANG, W., LI, B., RAO, G., LI, L., YU, Y., LI, Z., DONG, B., LU, Z., JIANG, Z. AND SHEN, L. (2017). Circulating tumor DNA functions as an alternative for tissue to overcome tumor heterogeneity in advanced gastric cancer. *Cancer Science*, 108(9), pp.1881-1887.
12. HANAHAAN, D. AND WEINBERG, R. (2000). The Hallmarks of Cancer. *Cell*, 100(1), pp.57-70.
13. HASSANPOUR, S. AND DEHGHANI, M. (2017). Review of cancer from perspective of molecular. *Journal of Cancer Research and Practice*, 4(4), pp.127-129.
14. ITSARA, A., COOPER, G., BAKER, C., GIRIRAJAN, S., LI, J., ABSHER, D., KRAUSS, R., MYERS, R., RIDKER, P., CHASMAN, D., MEFFORD, H., YING, P., NICKERSON, D. AND EICHLER, E. (2009). Population Analysis of Large Copy Number Variants and Hotspots of Human Genetic Disease. *The American Journal of Human Genetics*, 84(2), pp.148-161.
15. KANETSKY PA, HOLMES R, WALKER A, NAJARIAN D, SWOYER J, GUERRY D, HALPERN A, REBBECK TR (2001) Interaction of glutathione S-transferase M1 and T1 genotypes and malignant melanoma. *Cancer Epidemiol Biomarkers Prev* 10:509–513
16. KUANG, J., YAN, X., GENDERS, A., GRANATA, C. AND BISHOP, D. (2018). An overview of technical considerations when using quantitative real-time PCR analysis of gene expression in human exercise research. *PLOS ONE*, 13(5), p.e0196438.
17. KUMARAN, M., CASS, C., GRAHAM, K., MACKEY, J., HUBAUX, R., LAM, W., YASUI, Y. AND DAMARAJU, S. (2017). Germline copy number variations are associated with breast cancer risk and prognosis. *Scientific Reports*, 7(1)
18. LABCE.COM. (2019). The Advantages and Limitations of Liquid Biopsies - LabCE.com, Laboratory Continuing Education. [online] Available at: https://www.labce.com/spg1629022_the_advantages_and_limitations_of_liquid_biopsies.aspx [Accessed 7 May 2019].
19. LENZE, D., MÜLLER, H. AND HUMMEL, M. (2012). Considerations for the use of formalin-fixed and paraffin-embedded tissue specimens for clonality analysis. *Journal of Hematopathology*, 5(1-2), pp.27-34.
20. LI, X., LIU, C., HUANG, T. AND ZHONG, Y. (2016). The Occurrence of Genetic Alterations during the Progression of Breast Carcinoma. *BioMed Research International*, 2016, pp.1-5.
21. LIANG, L., FANG, J. AND XU, J. (2015). Gastric cancer and gene copy number variation: emerging cancer drivers for targeted therapy. *Oncogene*, 35(12), pp.1475-1482.
22. LIU, Y., YANG, M., JIANG, T., LAN, C., YUAN, H., LI, G., JIA, G., WANG, K. AND ZHAO, G. (2019). Quantitative Analysis of HER2 Amplification by Droplet Digital PCR in the Follow-Up of Gastric Cancer Patients Being Treated with Trastuzumab after Surgery. *Gastroenterology Research and Practice*, 2019, pp.1-8.
23. MAZAIKA, E. AND HOMS, J. (2014). Digital Droplet PCR: CNV Analysis and Other Applications. *Current Protocols in Human Genetics*, pp.7.24.1-7.24.13.
24. NATIONAL CANCER INSTITUTE. (2019). [online] Available at: <https://www.cancer.gov> [Accessed 1 May 2019].
25. NEUMANN, M., BENDER, S., KRAHN, T. AND SCHLANGE, T. (2018). ctDNA and CTCs in Liquid Biopsy – Current Status and Where We Need to Progress. *Computational and Structural Biotechnology Journal*, 16, pp.190-195.
26. PALMIROTTA, R., LOVERO, D., CAFFORIO, P., FELICI, C., MANNAVOLA, F., PELLÈ, E., QUARESMINI, D., TUCCI, M. AND SILVESTRIS, F. (2018). Liquid biopsy of cancer: a multimodal diagnostic tool in clinical oncology. *Therapeutic Advances in Medical Oncology*, 10, p.175883591879463.
27. QUAN, P., SAUZADE, M. AND BROUZES, E. (2018). dPCR: A Technology Review. *Sensors*, 18(4), p.1271
28. REYNOLDS, N. AND WAGSTAFF, A. (2004). Cetuximab. *Drugs*, 64(1), pp.109-118.
29. SHARMA, A., EILS, R. AND KÖNIG, R. (2016). Copy Number Alterations in Enzyme-Coding and Cancer-Causing Genes Reprogram Tumor Metabolism. *Cancer Research*, 76(14), pp.4058-4067.
30. SHI, J., YAO, D., LIU, W., WANG, N., LV, H., ZHANG, G., JI, M., XU, L., HE, N., SHI, B. AND HOU, P. (2012). Highly frequent PIK3CA amplification is associated with poor prognosis in gastric cancer. *BMC Cancer*, 12(1).

31. SIRAVEGNA, G., SARTORE-BIANCHI, A., NAGY, R., RAGHAV, K., ODEGAARD, J., LANMAN, R., TRUSOLINO, L., MARSONI, S., SIENA, S. AND BARDELLI, A. (2019). Plasma HER2 (ERBB2) Copy Number Predicts Response to HER2-targeted Therapy in Metastatic Colorectal Cancer. *Clinical Cancer Research*, 25(10), pp.3046-3053.
32. SMITH, J. AND SHELTZER, J. (2018). Systematic identification of mutations and copy number alterations associated with cancer patient prognosis. *eLife*, 7.
33. STEWART, C., KOTHARI, P., MOULIERE, F., MAIR, R., SOMNAY, S., BENAYED, R., ZEHIR, A., WEIGELT, B., DAWSON, S., ARCILA, M., BERGER, M. AND TSUI, D. (2018). The value of cell-free DNA for molecular pathology. *The Journal of Pathology*, 244(5), pp.616-627.
34. SUDA, T., HAMA, T., KONDO, S., YUZA, Y., YOSHIKAWA, M., URASHIMA, M., KATO, T. AND MORIYAMA, H. (2012). Copy Number Amplification of the PIK3CA Gene Is Associated with Poor Prognosis in Non-lymph node metastatic Head and Neck Squamous Cell Carcinoma. *BMC Cancer*, 12(1).
35. TAYLOR, S., LAPERRIERE, G. AND GERMAIN, H. (2017). Droplet Digital PCR versus qPCR for gene expression analysis with low abundant targets: from variable nonsense to publication quality data. *Scientific Reports*, 7(1).
36. WAN, J., MASSIE, C., GARCIA-CORBACHO, J., MOULIERE, F., BRENTON, J., CALDAS, C., PACEY, S., BAIRD, R. AND ROSENFELD, N. (2017). Liquid biopsies come of age: towards implementation of circulating tumour DNA. *Nature Reviews Cancer*, 17(4), pp.223-238.
37. WEINBERG, R. (2014). *The biology of cancer*. New York: Garland Science, Taylor & Francis Group.
38. ZHANG, N., WANG, M., ZHANG, P. AND HUANG, T. (2016). Classification of cancers based on copy number variation landscapes. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1860(11), pp.2750-2755.
39. ZHAO, Y., XIA, Q., YIN, Y. AND WANG, Z. (2016). Comparison of Droplet Digital PCR and Quantitative PCR Assays for Quantitative Detection of *Xanthomonas citri* Subsp. *citri*. *PLOS ONE*, 11(7), p.e0159004.

ANEXOS

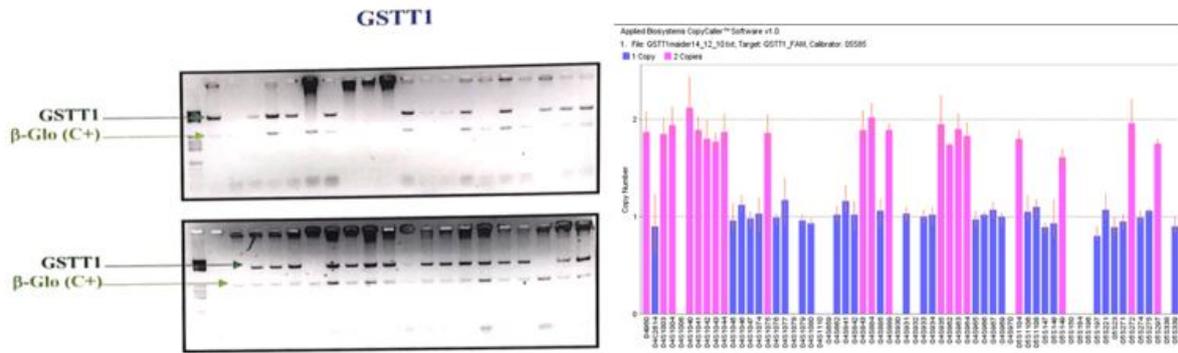


Figura 1. Resultados del análisis del número de copias de *GSTT1* por muestra en pacientes con melanoma mediante análisis por PCR convencional y qPCR junto con el programa de análisis Copy Caller (Ibarrola-Villava et al., 2012)

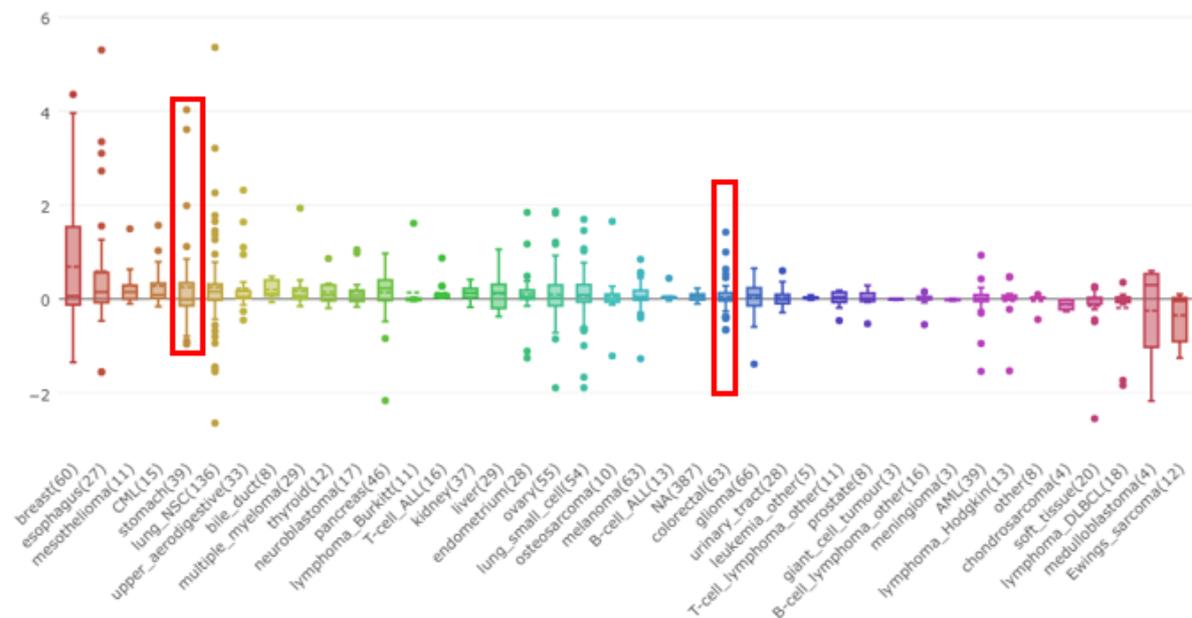


Figura 2. Gráfico de la variación en el número de copias de *ERBB2* en líneas celulares en diferentes tipos de cáncer. (THE BROAD INSTITUTE OF MIT & HARVARD, 2019)

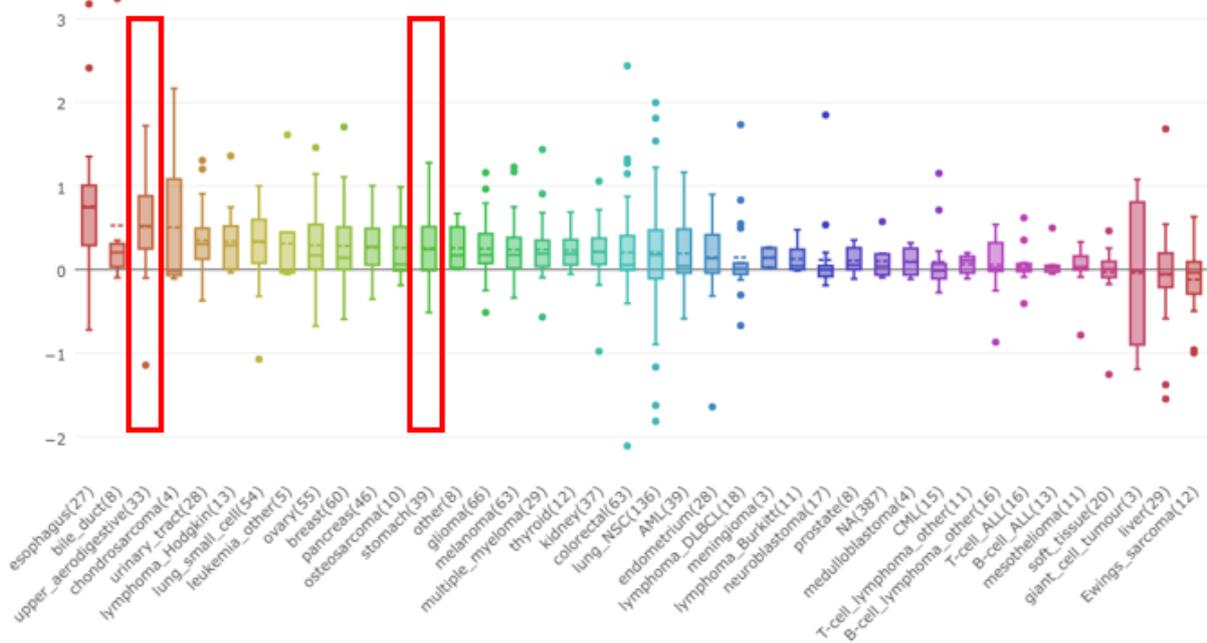


Figura 3. Gráfico de la variación en el número de copias de *PIK3CA* en líneas celulares en diferentes tipos de cáncer. (THE BROAD INSTITUTE OF MIT & HARVARD, 2019)

Tabla 1. Concentraciones, ratio 260/280, ratio 260/230, Tipo de cáncer y muestra de cada muestra

Muestras	Concentración ng/ μ L	260/ 280	260/ 230	Tipo de Cáncer	Tipo de Muestra	
04S1040	50.0	1.80	1.88	Melanoma Maligno	ADN genómico (capa de linfocitos)	
04S1075	50.0	1.70	1.97			
04S1040	28.5	1.84	0.43			
04S1046	18.3	1.69	0.40		ADN genómico digerido con la enzima Hae III (capa de linfocitos)	
04S1075	27.2	1.90	0.64			
04S841	30.9	1.74	0.28			
04S659	21.6	1.78	0.64			
17S001	18.2	1.58	0.21			
17S002	24.3	1.67	0.26			
17S003	37.2	1.91	0.41			
17S004	15.4	1.52	0.23			
17S005	38.0	1.88	0.40			
17S006	27.2	1.74	0.30			
17S007	35.2	1.96	0.52			
17S008	27.3	1.75	0.32			
17S009	18.5	1.63	0.27			
17S011	16.6	1.75	0.44			
17S012	23.4	1.61	0.20			
17S013	17.5	1.54	0.17			
17S014	22.1	1.71	0.26			
17S015	17.2	1.52	0.16			
17S016	26.5	1.73	0.29			
18S002	21.7	1.71	0.34			cfDNA de plasma
18S003	29.1	1.75	0.30			
18S004	28.6	1.62	0.19			
17S001	33.6	1.97	0.50			
17S002	24.5	1.81	0.60			
17S003	18.4	1.61	0.63			
17S004	14.3	1.67	0.58			
17S005	23.2	1.71	0.79			
17S006	17.0	1.67	0.86			
17S007	27.2	1.82	0.76			
17S008	21.0	1.93	0.80			
17S009	20.3	1.67	0.55			
17S011	32.7	1.70	0.70			
17S012	44.5	1.62	0.46			
17S013	17.6	1.55	0.80			
17S014	33.0	1.37	0.67			
17S015	13.2	1.79	0.74			
17S016	29.1	1.59	0.60			
18S002	38.8	1.59	0.53			
18S003	7.3	1.86	0.76			
18S004	21.5	1.98	0.48			

Muestras	Concentración ng/ μ L	260/ 280	260/ 230	Tipo de Cáncer	Tipo de Muestra
INV028	124.8	1.88	1.15	Cáncer colorrectal	ADN procedente de tejido tumoral parafinado
INV031	338.6	1.89	0.93		
INV236	235.2	2.10	1.98		
INV190	15.5	1.96	1.24		
INV133	211.9	1.94	0.38		
INV199	79.4	1.87	1.53		
INV106	8.9	1.05	0.12		
GAST 162	353.9	1.92	0.79	Cáncer gástrico	
GAST 155	132.2	1.82	0.53		
GAST 163	192.6	1.84	1.11		
GAST 167	34.9	1.70	0.33		
GAST 156	114.0	1.79	1.30		
GAST 8	213.6	1.96	0.97		
GAST 153	102.8	1.96	0.23		
GAST 170	187.9	1.94	0.44		ADN procedente de tejido normal parafinado
GAST 157	47.9	2.12	0.50		
NG155	25.4	1.80	0.36		
NG156	133.0	1.87	0.89		
NG162	52.2	1.83	1.10		
NG163	23.1	1.99	0.15	Cáncer de cabeza y cuello	ADN procedente de tejido tumoral parafinado
NG167	66.6	1.77	0.68		
HNC012-T1	44.6	1.87	1.83		
HCN013-T1	198.2	1.99	1.88		
HCN014-T1	81.8	1.92	1.82		
HCN016-T2	300.3	1.93	2.07		
HCN017-T1	50.1	1.91	1.88		
HCN018-T1	19.9	1.80	2.05		
HNC012-L1	39.8	2.16	0.50	cfDNA de plasma	
HNC013-L1	23.4	2.16	0.39		
HNC014-L1	19.7	1.90	0.45		

Tabla 2. ANOVA para Positivos FAM por Muestra en el ensayo de CNV para el gen *GSTT1*

Fuente	Suma de cuadrados	Gl	Cuadrado medio	Razón-F	Valor-P
Entre grupos	5,8564E6	1	5,8564E6	18,33	0,0001
Intra grupos	1,08656E7	34	319576	-	-
Total (Corr.)	1,6722E7	35	-	-	-

Tabla 3. ANOVA para Positivos HEX por Muestra en el ensayo de CNV para el gen *GSTT1*

Fuente	Suma de cuadrados	Gl	Cuadrado medio	Razón-F	Valor-P
Entre grupos	1,02261E7	1	1,02261E7	26,18	0,0000
Intra grupos	1,32831E7	34	390679	-	-
Total (Corr.)	2,35092E7	35	-	-	-