

LA ESTATMINA COMO MARCADOR DE ESTIMULACIÓN DE LA VÍA DE PI3K EN EL CÁNCER DE MAMA

FACULTAD DE CIENCIAS DE LA SALUD

**SECCIÓN DE MEDICINA
(DEPARTAMENTO DE CIENCIAS MÉDICAS BÁSICAS)**

TRABAJO DE FIN DE GRADO presentado por:

**Pedro Villavicencio Sánchez
Diego Francisco Rodríguez Fernández**

PROFESOR TUTOR:

Dr. Eduardo Salido Ruiz

ÍNDICE

1. Resumen.....	2
2. Abstract.....	2
3. Introducción.....	3
4. Hipótesis.....	8
5. Objetivos.....	8
6. Material y métodos.....	9
6.1. Estado mutacional de PI3K.....	10
6.1.1. Purificación de ADN.....	10
6.1.2. Cuantificación del ADN mediante nanoespectrofotometría.....	11
6.1.3. PCR de los exones 9 y 20.....	12
6.1.4. Pirosecuenciación.....	13
6.2. Expresión de estatmina.....	15
6.2.1. Bioquímica.....	15
6.2.1.1. Purificación de ARN.....	15
6.2.1.2. Retrotranscriptasa.....	17
6.2.1.3. Diseño de primers.....	19
6.2.1.4. Real Time PCR.....	22
6.2.2. Inmunohistoquímica.....	24
6.2.2.1. Preparación de las muestras.....	25
6.2.2.2. Medición.....	26
7. Resultados.....	27
7.1. Variables del test EndoPredict.....	27
7.2. Método DeltaDeltaCt.....	29
7.3. Expresión de estatmina en inmunohistoquímica.....	33
7.4. Expresión de estatmina y test Endopredict.....	37
8. Discusión.....	38
9. ¿Qué nos ha aportado este TFG?.....	40
10. Bibliografía.....	41

1. Resumen

La vía de PI3K/AKT (Fosfatidilinositol 3-quinasa) es una vía de señalización celular que regula múltiples procesos biológicos, incluyendo la proliferación celular, la diferenciación, la supervivencia y la migración celular. La sobreestimulación de esta vía se relaciona con una alta tasa de proliferación tumoral y empobrecimiento pronóstico en una gran variedad de tumores. Los pacientes que muestran sobreexpresión de esta vía celular pueden verse beneficiados mediante el uso de tratamientos dirigidos a frenar esta vía con fármacos específicos. Existen muchas causas por las que esta vía puede verse estimulada, siendo la más descrita en la clínica la mutación en el centro catalítico de la proteína PI3K, que condiciona una hiperestimulación de esta enzima. En la actualidad no existe ningún método de medir cuantitativamente la estimulación de esta vía intracelular, sin embargo en la literatura se ha descrito cierta asociación entre la estimulación de esta vía y la expresión de una proteína, la Estatmina. La Estatmina es una proteína reguladora del citoesqueleto cuya sobreexpresión puede asociarse en algunos tumores a un peor pronóstico. El objetivo de este proyecto es esclarecer la correlación entre la expresión de Estatmina y la activación de la vía de PI3K específicamente en el cáncer de mama, lo que podría ser de utilidad en el desarrollo de algún método de cuantificación de la activación de esta vía.

Palabras clave: estatmina, STMN1, PI3K, AKT, Endopredict, PTEN, cáncer de mama, luminal B.

2. Abstract

The PI3K/AKT (Phosphatidylinositol 3-kinase) pathway is a cellular signaling pathway that regulates multiple biological processes, including cell proliferation, differentiation, survival, and migration. Overstimulation of this pathway is associated with a high rate of tumor proliferation and poor prognosis in a variety of cancers. Patients showing overexpression of this cellular pathway can benefit from treatments aimed at inhibiting this pathway with specific drugs. There are many causes for the stimulation of this pathway, with the most commonly described in the clinic being the mutation in the catalytic center of the PI3K protein, which leads to hyperstimulation of this enzyme. Currently, there is no quantitative method to measure the stimulation of this intracellular pathway. However, the literature has described a certain association between pathway stimulation and the expression of a protein called Stathmin. Stathmin is a cytoskeleton regulatory protein whose

overexpression may be associated with poorer prognosis in some tumors. The objective of this project is to clarify the correlation between Stathmin expression and the activation of the PI3K pathway specifically in breast cancer, which could be useful in the development of a method to quantify pathway activation.

Keywords: stathmin, STMN1, PI3K, AKT, Endopredict, PTEN, breast cancer, luminal B.

3. Introducción

El cáncer de mama es el tumor maligno más incidente y prevalente en la mujer y en nuestro entorno lo padecerán 1 de cada 8 mujeres. En cuanto a la edad de aparición, se observan dos picos de incidencia: entre los 45 y 50 años, y entre 60 y 70 años. Los principales síntomas que refieren estas pacientes en un inicio son: masa palpable, dolor localizado, secreción sanguinolenta y otros síntomas inespecíficos. Cabe destacar entre los factores de riesgo de esta enfermedad, la dependencia hormonal (todo aquel proceso que suponga la exposición a estrógenos: nuliparidad, embarazo tardío, anticonceptivos orales...)

Es por ello que esta neoplasia tiene la peculiaridad de que en muchas ocasiones puede presentar receptores hormonales específicos, lo que habilita la vía de tratamiento hormonoterápico en muchos casos. En base a esto se suele clasificar a estos tumores en Luminal A (Receptores de estrógenos (RE) positivos, baja proliferación), Luminal B (RE positivos, alta proliferación), Her2 (Receptores HER2 positivos), Triple negativo (sin receptores hormonales específicos, peor pronóstico). En los de tipo Luminal A está indicado el tratamiento con hormonoterapia, mientras que los Triple negativo y HER2 amplificado requerirán quimioterapia. Sin embargo en los de tipo Luminal B el tratamiento variará en función de si presentan más características de diferenciación (candidato a hormonoterapia) o proliferación (además de hormonoterapia se pauta quimioterapia). Para medir estas características y establecer el grado de diferenciación/proliferación empleamos marcadores como el Ki67 que mide la proliferación celular. Si el Ki67 es menor de 20% (1) podemos asumir en la práctica clínica que dicho tumor muestra un fenotipo de diferenciación y por tanto será candidato a hormonoterapia sin quimioterapia. No obstante, en aquellos casos en los que el Ki67 es mayor del 20% no podemos afirmar con certeza cuál es la terapia más beneficiosa, por eso ante estos casos empleamos otras herramientas como el test de Endopredict. Sin embargo, en la práctica, ante la posibilidad de que se queden infratratadas pacientes por limitaciones de la determinación de Ki67, en el nuevo protocolo de actuación

en cáncer de mama del SCS, se ha decidido abrir la posibilidad de realizar el test Endopredict a casos con Ki67 por encima de 10%.

El test Endopredict (EP) es una prueba diseñada con el objeto de predecir el riesgo de recurrencia de enfermedad en aquellos cánceres de mama ER+ y HER2-. Esto se hace mediante el cálculo de un Score que se obtiene a partir de la cuantificación de la expresión de 11 genes (3 de control y 8 genes de interés). Según los resultados del EP, se puede clasificar a las pacientes en alto y bajo riesgo, lo cual determinará si se ven más beneficiadas de hormonoterapia y quimioterapia o solo de hormonoterapia respectivamente. Los 8 genes de interés que tiene el estudio se clasifican según su función biológica en 3 genes de proliferación (cuya expresión va ligada a mayor riesgo) y 5 genes de diferenciación (que cuanto más se expresan, mejor pronóstico) . La ventaja de este test, radica en que aporta valores cuantitativos que estiman la proliferación tumoral de cada caso, siendo capaz de predecir el porcentaje de recidiva tumoral en 10 años.

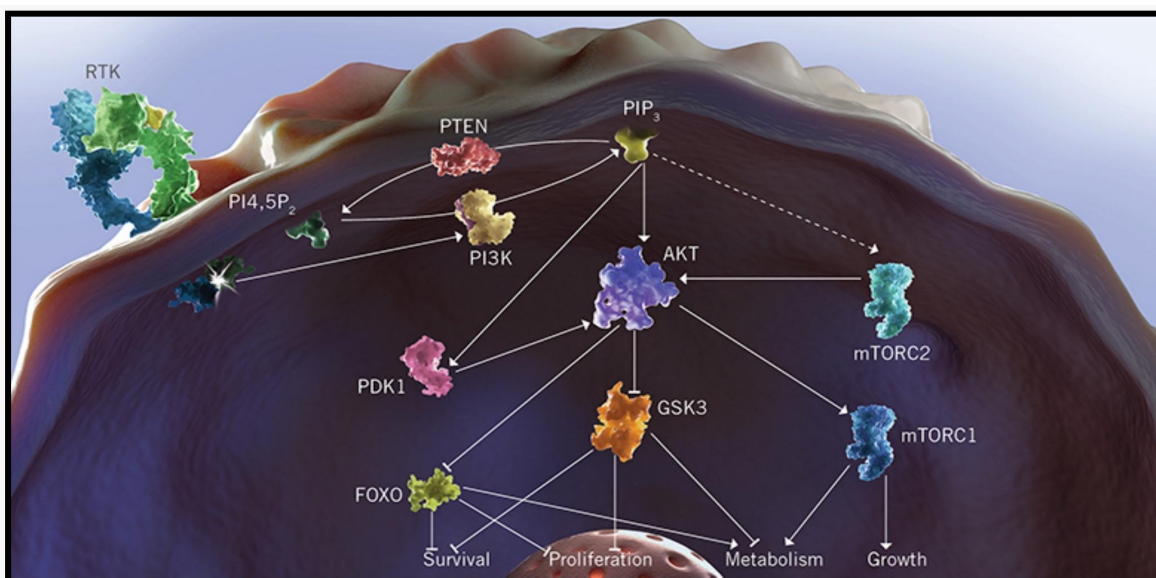
	Gene	Assigned biological processes
Proliferation associated Genes	UBE2C	Protein degradation, cell division
	BIRC5	Anti-apoptosis, cell division, cytokinesis, chromosome localization
	DHCR7	Cholesterol biosynthesis
Hormone Receptor associated Genes	STC2	Cell-to-cell communication
	AZGP1	Cell adhesion
	IL6ST	Various signal transduction pathways, cell proliferation, T cell proliferation
	RBBP8	DNA repair
	MGP	Transcriptional regulation

Genes de estudio en el test *Endopredict* (2).

En el proyecto de TFG “*Detección de mutaciones del gen PI3K en cáncer de mama mediante pirosecuenciación*” realizado por nuestros compañeros de laboratorio Ginés Koy Salido e Ismael Gómez Perez, se plantea la posibilidad de correlación entre las variables obtenidas del estudio del test EP y la mutación en el gen de PI3K, siendo ambas variables de peso en la influencia sobre el pronóstico del cáncer de mama, en el se encuentra una correlación relativa entre la estado mutacional de PI3K y algunas variables del test EP.

En este proyecto emplearemos el test EP como herramienta para medir el pronóstico molecular, lo que nos servirá para abordar el siguiente objetivo: búsqueda de un marcador de activación de la ruta PI3K/AKT que sirva como reportero de la misma y permita correlacionarla con las variables obtenidas del test EP como representación pronóstica de la enfermedad. En particular, hemos investigado la expresión del gen de la estatmina (a nivel de mRNA y de proteína) y su correlación con EP. Todas las muestras de nuestra base de datos son de cáncer de mama de tipo luminal B que han sido sometidas a la realización del test EP.

La vía de PI3K/AKT es una ruta de señalización intracelular importante que está involucrada en la regulación de varios procesos celulares, como la supervivencia celular, la proliferación, el crecimiento y la diferenciación. La vía de PI3K/AKT se activa cuando un receptor de membrana celular se une a un ligando específico, lo que desencadena la activación de una enzima llamada fosfatidilinositol-3-quinasa (PI3K). La PI3K convierte el fosfolípido fosfatidilinositol-4,5-bisfosfato (PIP₂) en fosfatidilinositol-3,4,5-trifosfato (PIP₃), que a su vez recluta la proteína quinasa AKT (también conocida como proteína cinasa B o PKB) a la membrana celular. Una vez que AKT se une a la membrana celular, se activa y comienza a fosforilar sus proteínas sustrato. Entre los objetivos de AKT se encuentran proteínas reguladoras clave de la apoptosis (muerte celular programada), como la proteína Bad, y proteínas reguladoras del ciclo celular, como la proteína p21. Además, AKT también puede activar la vía de mTOR, que está involucrada en la regulación del crecimiento y la síntesis de proteínas.



Esquema de la vía intracelular de PI3K/AKT (3)

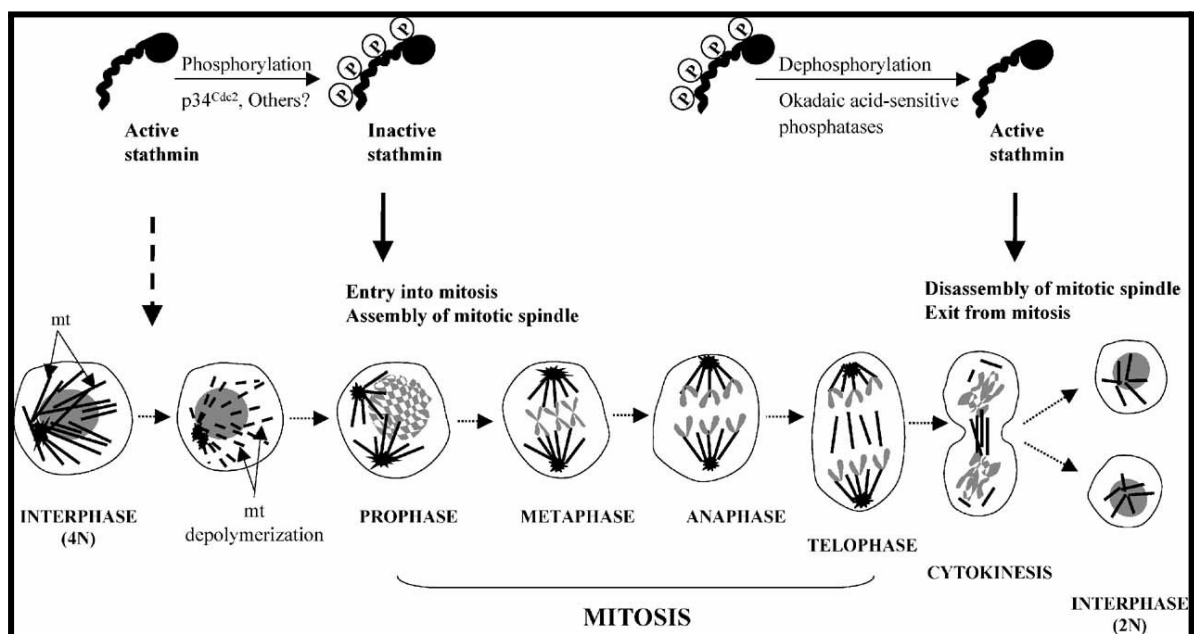
A modo de resumen, la estimulación de la vía de PI3K/AKT tiene como resultados finales la supervivencia, el crecimiento y la proliferación celular, procesos que empeoran el pronóstico de muchas neoplasias, y el cáncer de mama no es una excepción.

La vía de PI3K/AKT puede verse estimulada por numerosas causas: estimulación de receptores de membrana, factores de crecimiento (como el EGF), hormonas, estrés oxidativo o mutaciones en algunos de los componentes de la vía. De entre todas estas causas, la más directa resulta ser las mutaciones genéticas. Las mutaciones que más frecuentemente estimulan esta vía en cáncer de mama son las de la subunidad catalítica alfa del gen PI3K (PIK3CA). En particular, las mutaciones en dos sitios específicos en el exón 9 y el exón 20 de PIK3CA, conocidas como E545K y H1047R, respectivamente, son las mutaciones más comunes que se encuentran en varios tipos de cáncer (colorrectal, ovario, pulmón y fundamentalmente mama). Estas mutaciones activan la subunidad catalítica de PI3K aumentando el sustrato para los procesos posteriores y estimulando la vía, sin embargo existen otras mutaciones, como la de PTEN, cuya pérdida de función también resulta en activación de la vía. PTEN es una fosfatasa y actúa como regulador negativo de la vía de PI3K, catalizando la reacción inversa, por ello se le conoce como el freno de esta vía. Mientras que una mutación en PI3KCA habitualmente lo vuelve más activo (provoca una activación constitutiva), una mutación en el centro catalítico de PTEN lo incapacita para catalizar la reacción negativa de la vía, por lo que si la vía pierde su freno el resultado será una mayor estimulación.

La importancia clínica de la mutación en esta vía molecular reside en la presencia de tratamientos dirigidos para aquellos tumores que muestran indicios de mayor actividad, como las mencionadas mutaciones en PTEN o PI3KCA. Estos tratamientos consisten fundamentalmente en inhibidores de PI3K e inhibidores de mTOR. Los inhibidores de PI3K actúan bloqueando la actividad de la enzima PI3K y la señalización celular asociada. Ejemplos de inhibidores de PI3K incluyen alpelisib, idelalisib y copanlisib. Por otro lado, los inhibidores de mTOR, como everolimus y temsirolimus, también son utilizados para el tratamiento de tumores con mutaciones en PIK3CA, ya que mTOR es un componente clave de la vía de señalización celular regulada por PI3K. El uso de estas terapias dirigidas está aprobado en los pacientes que cumplan los criterios (mutación en PI3KCA/PTEN) y se suele asociar al tratamiento hormonal.

No obstante, la observación nos ha enseñado que no todas las pacientes con mutación en la vía de PI3K responden igual de bien al tratamiento dirigido, asociándose esta variabilidad de respuesta a la variabilidad en la estimulación de la propia vía PI3K (además de a otros factores vinculados al propio fármaco), pues dos pacientes con la misma mutación no tienen porqué tener la misma activación de esta vía, ya que hay otros factores además de la mutación que intervienen en el mismo fenómeno. Además, diferentes tumores presentan frecuencias alélicas distintas de las mutaciones en PI3K y no se conoce bien si esta variabilidad en la frecuencia alélica se correlaciona con variaciones en la activación de la ruta. Es ahí donde entra en juego la posibilidad de cuantificar la hiperestimulación de esta vía con otros marcadores situados “río-abajo” en la ruta y que puedan servir como reporteros de la misma, como por ejemplo la estatmina.

La estatmina, también conocida como Oncoproteína 18, es una proteína regulada por fosforilación codificada por el gen STMN1 y encargada de la regulación de la dinámica de los microtúbulos, orgánulos de vital importancia en el ciclo celular. Durante la interfase existen concentraciones bajas de tubulina libre por lo que la estatmina se encuentra desfosforilada y activada, de esta manera se produce el desensamblaje de los microtúbulos. Sin embargo, durante la mitosis la estatmina se fosforila para inactivarse, permitiendo que se ensamblen los microtúbulos y se forme el huso mitótico. Es decir, según las necesidades celulares va a regular la dinámica de los microtúbulos.



Funciones de la estatmina en el ciclo celular (4).

Dejando de lado las funciones intracelulares de la estatmina, su importancia radica en que está reconocida como una oncoproteína, pues su sobreexpresión anómala se asocia a mayor proliferación, crecimiento, diseminación tumoral y mayor resistencia al tratamiento. Esta asociación se cumple no solo para el cáncer de mama, sino para otros tumores tales como cáncer de ovario, cáncer colorrectal, cáncer de próstata y cáncer de pulmón (5-8), sin embargo la fuerza de asociación es variable según el tipo de tumor.

El rol de oncoproteína estatmina está fuertemente afianzado en la literatura, sin embargo no es tan sólida la asociación entre su expresión y la estimulación de la vía de PI3K. Nuestro artículo guía: *“Poor prognosis in carcinoma is associated with a gene expression signature of aberrant PTEN tumor suppressor pathway activity”* (9) encuentra una correlación negativa significativa entre la expresión de estatmina y expresión de PTEN (regulador negativo de la vía de PI3K), ambas medidas por inmunohistoquímica. Aquellos tumores que mostraban una menor expresión de PTEN (debido a una mutación en este gen) mostraban una mayor expresión de estatmina a la tinción inmunohistoquímica, sugiriendo la posibilidad de que la estatmina se sobreexpresara en aquellos pacientes con menor expresión de PTEN (y presumiblemente mayor estimulación de la vía de PI3K, pues como ya mencionamos, PTEN actúa como freno).

Es aquí donde surge nuestra decisión de ampliar este estudio a aquellos pacientes con mutación en la vía de PI3K para comprobar si se sigue cumpliendo esta asociación, pero centrándonos exclusivamente en el Cáncer de mama con el objetivo de eliminar la variabilidad entre diferentes tipos de tumores.

4. Hipótesis

La sobreexpresión de estatmina se correlaciona positivamente con la mutación en PI3K en el carcinoma de mama.

5. Objetivos

En el presente trabajo nos hemos planteado los siguientes objetivos:

- Estudiar la correlación de la expresión de estatmina con la mutación en PI3K en el carcinoma de mama tipo luminal B.
- Comparar otras variables intermedias derivadas del estudio de expresión génica Endopredict en busca de correlación.

- Iniciarnos en los fundamentos de un trabajo de investigación.
- Aprender a extraer información de diferentes portales online.
- Aprender a cuantificar la expresión de ARN mediante diferentes técnicas.
- Aprender a usar los diferentes métodos estadísticos para el estudio de correlación.

6. Material y métodos

Con el objetivo de estudiar la correlación entre la expresión de la estatmina y la activación de la vía de PI3K realizaremos estudios comparativos entre estas dos variables; expresión de estatmina y estado mutacional de PI3K en carcinoma de mama tipo luminal. Se sabe que las mutaciones principales de los dominios quinasa y helical de PI3K resultan en hiperestimulación de la vía PI3K-AKT. Por tanto, podemos usar la variable categórica PI3K (mutado / no mutado) para definir los grupos de carcinoma de mama en los que analizaremos la expresión de estatmina para investigar si dicha expresión de estatmina es un biomarcador de la activación de la ruta.

Por un lado nuestra variable independiente será el estado mutacional de PI3K, se trata de una variable cualitativa que tiene dos estados posibles: mutado y wild type (no mutado). Muchas de las muestras ya tenían la medición del estado mutacional de PI3K como resultado de la práctica clínica habitual, sin embargo para otras se realizó el estudio mediante amplificación de ADN y pirosecuenciación.

Por otro lado, nuestra variable dependiente será la expresión de estatmina. La cuantificación de la expresión de la estatmina se decidió medir por dos técnicas diferentes: la cuantificación de ARN mediante técnica de Retrotranscriptasa y qPCR y la tinción con inmunohistoquímica, para evaluar los dos niveles principales a los que se suele regular la expresión génica en células de mamífero: transcripcional y postraduccional. La medición de mRNA nos da una idea de la regulación transcripcional, mientras que la inmunohistoquímica nos informa de los niveles promedio (steady-state) de la proteína, lo que refleja no solo la regulación transcripcional sino también la posible postraduccional. La técnica de cuantificación de mRNA es más difícil de llevar a cabo y es muy dependiente de la experiencia del experimentador, pudiendo acumular mucha variabilidad y error dependiendo de la metodología. La técnica inmunohistoquímica, por otro lado resulta más sencilla de poner en práctica y acumula menos variabilidad, es directa (se utiliza un anticuerpo que se une a la proteína estatmina, siendo representativo de cuánta proteína se está expresando),

semicuantitativa (clasificamos en franjas de porcentajes según la cantidad de tejido que visualicemos teñido en el microscopio) y puede estar sujeta a interpretación (pues se trata de una medición visual estimada por el observador).

Nuestra base de datos consta de 100 pacientes aleatorizados que se subdividirán en mutados y wild type (no mutados). Se asumió una pérdida de casos por falta de datos o ausencia de muestras para realizar las posteriores técnicas de hasta 40 pacientes, según los requerimientos específicos de los procedimientos a realizar, esperando no calcular ningún resultado con un tamaño muestral menor a los 60 pacientes, número considerado ideal para obtener un buen balance entre fuerza estadística y puesta en marcha del trabajo considerando la variabilidad de las mediciones. En esta base de datos figura el estado del gen PI3K (mutado/wild type) de cada uno de los casos previamente estudiado en colaboración con nuestros compañeros de laboratorio Ginés Koy Salido e Ismael Gómez Pérez, además de otras variables de interés para estudios anexos como los valores relacionados con el cálculo del Test Endopredict (*EP clin, EP score...*).

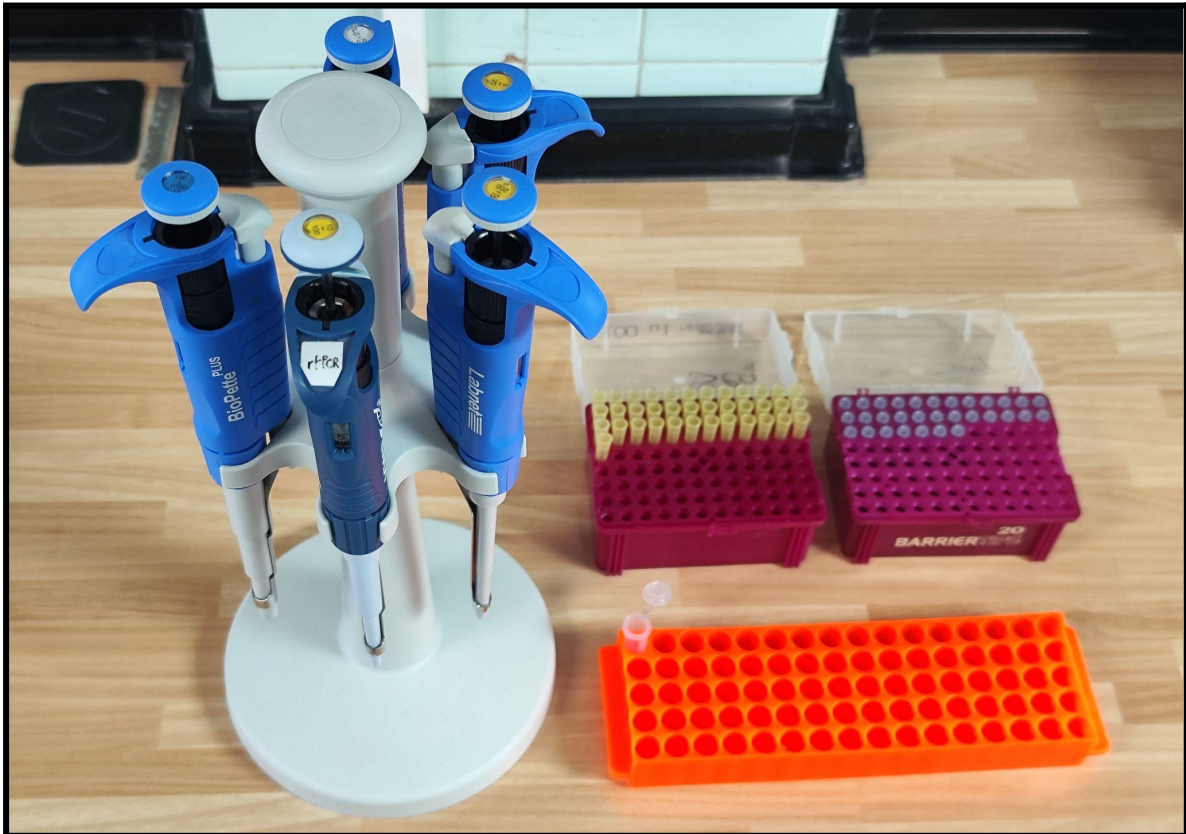
En un primer momento se realizó técnicas de amplificación de ADN y pirosecuenciación con el objetivo de conocer el estado mutacional de las muestras, aunque muchas de las muestras ya tenían esta variable estudiada como resultado de la práctica clínica habitual en estos tumores, repetimos el proceso en las muestras de última incorporación para aprender cómo se realiza el procedimiento. Posteriormente, para completar esta base de datos debíamos estudiar las variables relacionadas con la expresión de la estatmina: El valor de cuantificación derivado de la PCR realizada sobre ADNc obtenido de ARNm del gen STMN1 (gen codificador de la estatmina) y el valor de cuantificación del tejido teñido mediante técnica de inmunohistoquímica. Dividimos el procedimiento de medición de la variable “expresión de estatmina” en estas dos formas de medición; la bioquímica (mide ARNm) y la histológica (mide la proteína).

6.1. Estado mutacional de PI3K

6.1.1. Purificación de ADN

Para purificar ADN recurriremos al método de cromatografía por afinidad. La cromatografía por afinidad es una técnica de separación en la cual se utiliza una matriz de adsorbente que se une específicamente a una biomolécula o grupo de biomoléculas. En nuestro caso, esta matriz adsorbente se une a los grupos fosfatos del ADN permitiendo su

separación. Para realizar este método partimos del sobrenadante que resta tras el proceso de purificación de ARN (explicado más adelante). Las etapas del método son lisar (se omite esta fase, pues ya se realiza en la purificación del ARN), unión, lavar y eluir (11). En la fase de unión se une el componente de ADN a la matriz adsorbente en un tubo con una fase sólida de sílice, luego en la fase de lavado se añaden soluciones de NaCl y se centrifuga para eliminar todo el material que no sea ADN y para finalizar eluimos nuestro ADN final en un medio acuoso para su correcta conservación.



Equipo de pipeteo con diferentes niveles de precisión.

6.1.2. Cuantificación del ADN mediante nanoespectrofotometría

Una vez purificado el ADN, es necesario cuantificar su concentración final resultante. Para ello, emplearemos un nanoespectrofotómetro, un dispositivo que emplea la espectrofotometría para calcular la concentración de ADN en cada muestra. Éste dispositivo emite un haz de luz que atraviesa la muestra y alcanza un receptor. La muestra habrá absorbido parte de la luz en este proceso por lo que la cantidad de luz dispensada por el emisor y la cantidad que detecta el receptor mostrarán disparidad, en base a la cual la

máquina calculará la concentración de sustrato que hay en el medio (en nuestro caso, ADN disuelto en H₂O).

Para comenzar a cuantificar primero debemos calibrar la máquina asignándole un *blank* con la pipeta. Colocamos un microlitro del mismo medio donde se encuentra el sustrato que queremos cuantificar (en nuestro caso, agua) y le indicamos a la máquina que realice el blanco. La máquina habrá calculado cuánta luz absorbe el medio en el que se encuentra el ADN pudiendo tenerla en cuenta a la hora de medir cuánta luz absorbe el ADN. Posteriormente, limpiamos el punto sensor y depositamos un microlitro de muestra de nuestro ADN purificado en él. Luego, le indicamos a la máquina que comience la medición, acto seguido obtendremos el resultado. Procedemos a realizar lo mismo con todas las muestras, limpiando bien el medidor entre cada una de ellas y anotando la concentración en el lateral del tubo de cada paciente (esto servirá más adelante para estimar cuánto ADN hemos de pipetear en los pocillos para la PCR).



Equipo de nanoespectrofotometría nanoDrop 2000

6.1.3. PCR de los exones 9 y 20

Las mutaciones en el gen de PI3K tienen lugar en el centro catalítico y sus localizaciones principales se encuentran en dos exones, el exón 9 y el exón 20. La Reacción

en Cadena de la Polimerasa (PCR) obtendrá múltiples copias de los fragmentos de ADN correspondientes a los exones 9 y 20 para su posterior análisis molecular mediante pirosecuenciación. Para ello, nos valdremos de los primers (cebadores) específicos de los sitios de unión de cada uno de los exones, que flanquean la región diana (donde se producen las mutaciones) que se va a secuenciar mediante la unión a las cadenas complementarias del ADN. La amplificación se produce al calentar el ADN para separar las dobles cadenas, tras lo que se enfría la reacción para permitir que los cebadores se unan a las dos cadenas de ADN y después se elongan las secuencias a partir de los cebadores con la enzima ADN polimerasa. Este ciclo de calentamiento, enfriamiento y polimerización se desarrolla a través de varios ciclos, en cada uno de los cuales aumenta exponencialmente el número de copias del ADN diana.

Para realizar la PCR partimos del ADN purificado que será depositado en los correspondientes pocillos de la placa de PCR junto con su respectivo primer (primero lo realizamos para un exón y posteriormente para el otro) y la proporción adecuada de MásterMix. El MásterMix (mezcla madre) es la forma de llamar a la solución preparada con antelación que contiene todos los componentes necesarios restantes para la realización de la PCR (nucleótidos, ADN polimerasa, buffer, estabilizantes...). Como no es necesario cuantificar exactamente la cantidad de ADN que vamos a amplificar, sino que es para un estudio cualitativo del estudio mutacional, no será necesario normalizar la concentración de ADN que añadimos (siendo indiferente si, en algunas muestras hay más concentración y otras menos), a diferencia de lo que ocurrirá en el caso de la PCR cuantitativa (qPCR, explicada en el estudio de la expresión de estatmina).

6.1.4. Pirosecuenciación

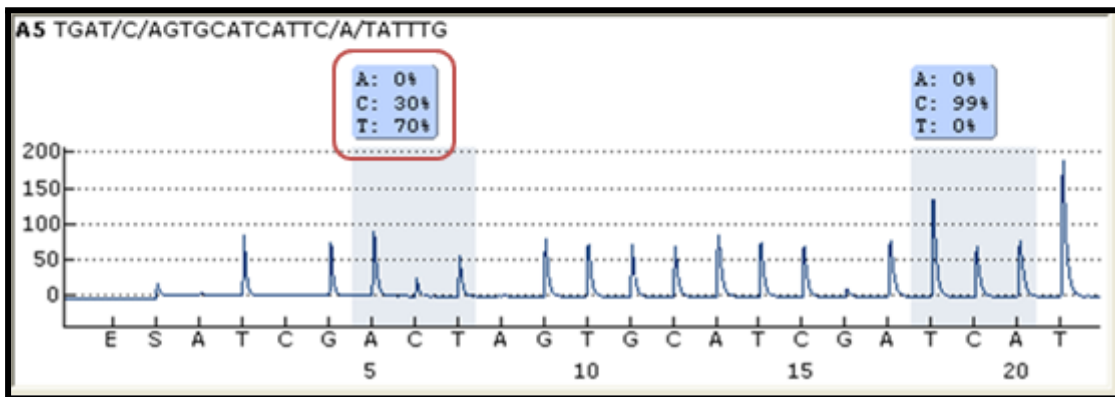
La pirosecuenciación utiliza la luminiscencia para detectar la incorporación de nucleótidos complementarios a los de la cadena molde de ADN (por ejemplo, productos de PCR desnaturalizados) a medida que se va sintetizando una nueva hebra. El objetivo final de este procedimiento es secuenciar fragmentos de ADN que contengan las regiones a analizar para luego cuantificar la frecuencia alélica de las diferentes mutaciones. En nuestro proyecto secuenciamos los productos de la PCR de las muestras de las 100 pacientes diagnosticadas de cáncer de mama. Dicho ADN se amplificó gracias al uso de cebadores específicamente diseñados para las regiones diana, es decir, aquellas donde se buscó cuantificar las mutaciones. Uno de estos cebadores viene con una molécula de biotina ya incorporada de

fábrica para la PCR, que servirá para poder capturar una sola hebra de ADN desnaturalizado y secuenciarla mediante esta técnica.

Para realizar la pirosecuenciación se prepara la máquina con los componentes necesarios para realizar la reacción de secuenciación (síntesis, extensión de bases), pero separando los nucleótidos en el cartucho de carga, de modo que la máquina sabe en todo momento qué nucleótido está añadiendo a la solución.

El fundamento por el cual se obtiene la secuenciación de las regiones diana de los pequeños fragmentos de ADN es gracias a una polimerasa que cataliza la incorporación de desoxinucleótidos trifosfato (dNTP) complementarios al ADN que se quiere analizar a partir de los cebadores que flanquean dichas regiones. Cuando esto ocurre, se libera una molécula de pirofosfato (PP) como producto de desecho. El PP actúa como sustrato por la ATP sulfurilasa, que cataliza la conversión de 5'-fosfosulfato a ATP. Este último es empleado por la luciferasa para reaccionar con la luciferina para generar, finalmente, luz. El pirosecuenciador, que funciona como una cámara oscura, detecta esta señal lumínica y la registra como picos de tamaño proporcional al número de nucleótidos añadidos a la nueva cadena creciente.

Como resultado obtendremos un “pirograma” cuantitativo que posteriormente tendremos que interpretar para discernir si existe una mutación o no. En el resultado se muestra una secuencia de nucleótidos que la máquina va añadiendo y una cantidad de luz captada por la máquina para cada uno. Si en el momento que se añade dicho nucleótido no se captó luz, significa que en la hebra de ADN en esa posición no existe el nucleótido complementario, de igual forma que si se capta luz al 100%, el nucleótido en esa posición de la hebra es con seguridad el nucleótido complementario al que liberó la máquina en ese momento. Por otra parte cuando se capta luz a un porcentaje inferior y se reparte esta luz entre dos nucleótidos, podemos asumir que en la solución existen hebras de ADN con que se unen a un nucleótido mientras que otras se unen a otro, por lo que existe material genético con una mutación en dicha posición. Es importante recordar que en la práctica la luz captada no siempre se corresponde al 100% y ello no implica que exista mutación, pues la solución puede estar contaminada y ser impura, por ello la cantidad de luz captada debe ser interpretada con razonamiento y la supervisión de personal experimentado para poder diferenciar las mutaciones con certeza.



Representación del resultado de la pirosecuenciación, donde se observa a la izquierda de la cadena una mutación, pues la captación de luz está dividida entre la dispensación del nucleótido C (30%) y T (70%).

6.2. Expresión de estatmina

6.2.1. Bioquímica

Este apartado consiste en la cuantificación de el ARN mensajero del gen STMN1 que codifica para la síntesis de estatmina, a mayor cantidad de ARN en tejido tumoral mayor expresión de estatmina, puesto que si encontramos niveles altos de ARNm significa que esta proteína se está sintetizando en mayor cantidad. Tomamos la muestra raspando el material tumoral con una aguja de un porta con un corte histológico del tumor, guiándonos por otro corte en el mismo plano teñido con Hematoxilina-Eosina para asegurarnos de raspar el mayor contenido del tumor, así como el menor contenido de tejido no tumoral posible. Posteriormente purificamos el ARN de esta muestra con el objetivo de eliminar el resto de material y restar únicamente ARN, que posteriormente será convertido a ADN usando la enzima Retrotranscriptasa (RT). Sobre este ADN final aplicamos técnicas de cuantificación como es la qPCR. Explicados en detalle cada uno de estos pasos:

6.2.1.1. Purificación de ARN

La purificación del ARN es un primer paso fundamental en muchos análisis (en nuestro caso tanto el test EP como la cuantificación de la expresión de estatmina) y plantea retos particulares debido a la fácil degradabilidad del ARN. Las RNasas suelen ser muy ricas en enlaces disulfuro, lo que las convierte en enzimas increíblemente estables. De esta manera, es fácil contaminar un laboratorio de secuenciación de ARN con RNasa, pero es notoriamente

difícil asegurar su completa erradicación. Esto se debe a la presencia de RNasa ambiental en fuentes microbianas en el aire, así como en la piel, el pelo o la saliva de los seres humanos. Por tanto, es inexorable la utilización de guantes durante el procedimiento.



Es necesario utilizar guantes a la hora de pipetear muestras de ARN para evitar la contaminación con RNasas.

En principio, escogemos la muestra macroscópica de tejido mamario (que contiene el tumor luminal B de cuyas células queremos extraer ARN y purificarlo) y su corte histológico correspondiente. Así, si observamos minuciosamente el corte histológico podemos distinguir una distribución heterogénea conformada por: zonas más basófilas (moradas), con células más agrupadas y más núcleos (de elección para el raspado que se realizará posteriormente); y zonas más eosinófilas (rosadas), con células más dispersas y menos núcleos.

Seguidamente, superponemos el corte histológico en la muestra macroscópica en aras de poder dilucidar qué zona es la idónea para raspar con la cuchilla (es decir, las zonas de carcinoma, basofílicas mencionadas anteriormente). Es beneficioso mirar previamente al

microscopio el corte histológico con el objetivo de evitar raspar áreas que contengan carcinoma *in situ*.

Más tarde, cogemos un tubo Eppendorf y lo introducimos en el ThermoMixer con el Buffer para disolver la parafina y el raspado. A partir de este momento, podemos estructurar la purificación del ARN en 5 pasos: lisis, unión, lavado, eluir y destrucción del ADN genómico restante. Cabe señalar que para el proceso de unión se requiere de microesferas magnéticas con afinidad por los ácidos nucleicos. Después de la fase de lavado, nos ayudaremos de un soporte magnético en el que sostendremos el tubo de microcentrifugación, para que las microesferas queden sujetas a su pared y poder, así, extraer el sobrenadante y los lavados sin dificultad respectivamente.

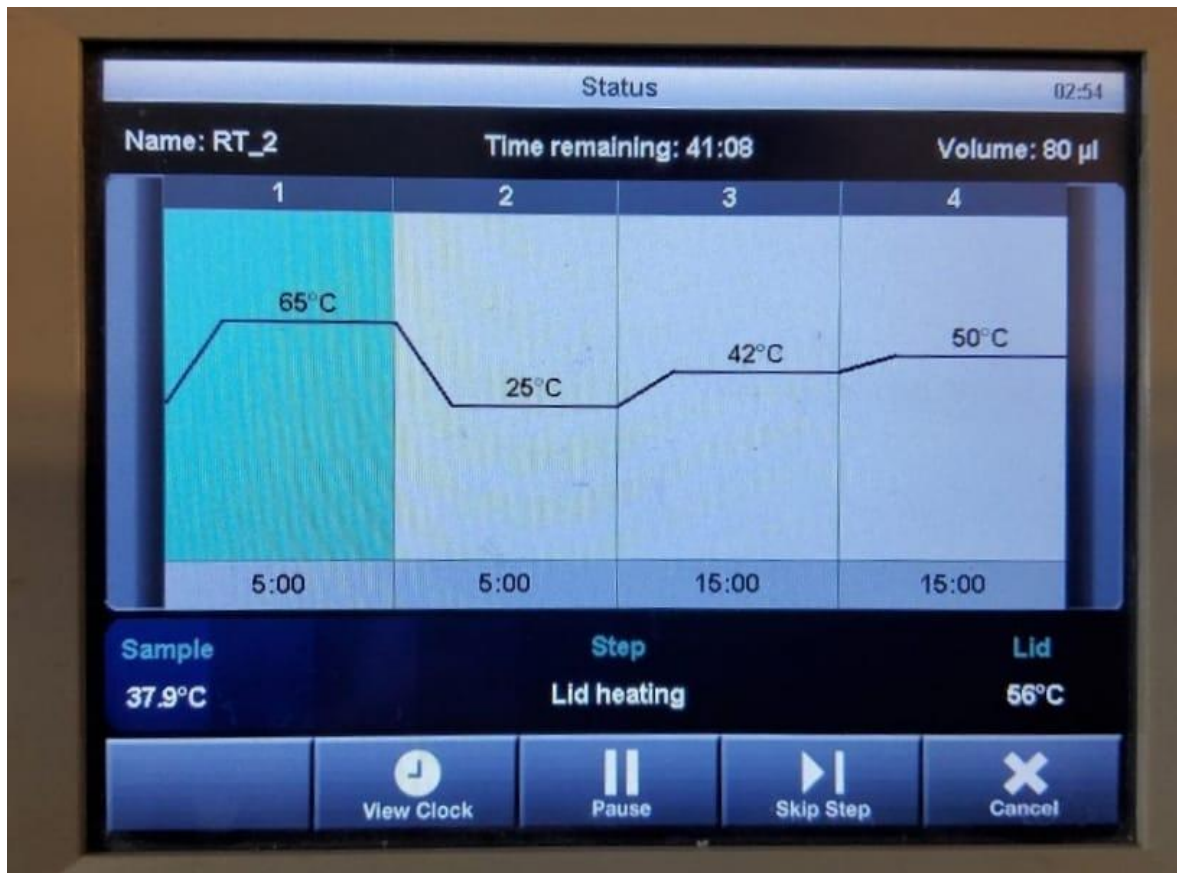
6.2.1.2. Retrotranscriptasa

Para pasar del ARN purificado al ADN complementario (ADNc) preparado para cuantificar debemos someterlo a una reacción mediada por la retrotranscriptasa. Para hacer esto primero se debe medir la concentración de ARN presente en cada una de nuestras muestras para de esta forma estandarizar y asegurar de esta forma, que en el ADN cíclico para la qPCR todas las muestras partan de la misma concentración de ARN total de partida. Esta vez, en lugar de usar un espectrofotómetro como el nanodrop, mediremos empleando un Fluorómetro Qubit 3.0, que destaca por tener mayor precisión que el anterior. La fluorometría consiste en añadir a una porción de nuestra muestra una molécula fluorescente de unión a ARN (199microl mmix + 1microl ARN) para posteriormente medir la cantidad de luz que esta emite, obteniendo de esta forma la concentración de forma más precisa. Tras calibrar el dispositivo se midieron las concentraciones de nuestras muestras y procedimos al siguiente paso.



Equipo de nanofluorimetría Qubit 3.0.

Preparamos en un tubo eppendorf los componentes necesarios para la reacción (llamándolo MasterMix) en los que se incluían nucleótidos, retrotranscriptasa y random primers (primers de 9 nucleótidos de longitud, de secuencia aleatoria, que se unen en numerosos puntos del ARN), y lo añadimos a nuestro ARN. En un primer momento cometimos el error de añadir el doble de la cantidad correspondiente de random primers a nuestro MasterMix, por lo que duplicamos del mismo modo las cantidades del resto de los componentes para que las concentraciones de todos los reactivos se igualaran en las proporciones adecuadas. Empleando la concentración previamente medida calculamos la cantidad de muestra que hay que añadir en cada pocillo para que todas tengan 50ng de ARN. Finalmente añadimos nuestra placa de muestras al termociclador para someterlos a los siguientes ciclos:



Diseño de ciclos correspondiente para la técnica retrotranscriptasa.

6.2.1.3. Diseño de primers

La Reacción en Cadena de la Polimerasa (PCR) permite obtener de un fragmento de ADN particular un gran número de copias, denominadas amplicones, para su posterior análisis molecular. Para llevar a cabo la misma, es necesario que fabriquemos los *primers* o cebadores (tanto el *forward* como el *reverse*). Se tratan de oligonucleótidos que flanquean los extremos del genoma que queremos amplificar y que sirven como punto de partida para la replicación. Sin lugar a dudas, los *primers* son un pilar fundamental en la determinación del éxito o el fracaso de la reacción de amplificación. Es, por tanto, que debemos considerar una serie de puntos en su diseño: seleccionar cebadores con una distribución de bases balanceada, desechar secuencias con estructuras secundarias significativas, y enfrentarlos el uno con el otro para comprobar la complementariedad y evitar el fenómeno *primer dimer*. Otro aspecto importante es la Temperatura de *melting* (T_m) que se calcula de la siguiente manera:

$$Tm = 2 \times (A + T) + 4 \times (G + C)$$



Temperatura a la cual la mitad de las moléculas del primer están apareadas con el ADN molde.

En nuestro caso, el diseño de primers surge de la necesidad de cuantificar cantidad de ARNm del gen *STMNI*, por lo que no necesitamos que los primers incluyan un exón concreto del gen, pudiendo elegir cualquier exón priorizando así otras características importantes de un buen primer como puede ser la especificidad de unión o la Temperatura de melting. Para el diseño del primer nos valemos de herramientas online de uso gratuito que nos ayudan a calcular la T_m así como la especificidad del primer, con el fin de evitar que estos *primers* se unan en cualquier otro punto no deseado del genoma. El resultado final consistió en una pareja de primers localizados entre los exones 3 y 4 con una T_m cercana a los 55°C , puesto que consideramos esa la temperatura adecuada para la solución y la termocicladora a emplear.

Primers STMN1:

Left Primer: AAGAAATTAGAAGCTGCAGA

Right Primer: TCGGTTCTCTTTATTAGCTT

Exón 3:

ATATCCAGGTGAAAGAACTGGAGAAGCGTGCCTCAGGCCAGGCTTTTGAG
CTGATTCTCAGCCCTCGGTCAAAGAATCTGTTCCAGAATCCCCCTTTCC
CCTCAAAGAAGAAGGATCTTTCCTGGAGGAAATTCAG **AAGAAATTAGAA**
GCTGCAGAAGAAAGACGCAAG

Exón 4:

TCCCATGAAGCTGAGGTCTTGAAGCAGCTGGCTGAGAAACGAGAGCACGA
GAAAGAAGTGCTTCAGAAGGCAATAGAAGAGAACAACA ACTTCAGTAAAA
TGGCAGAAGAGAACTGACCCACAAAATGG **AAGCTAATAAAGAGAACCGA**
GAGGCACAAATGGCTGCCAACTGGAACGTTTGCAGAGAGAAG

Longitud: 183 pb

1. Analysis Results

Sequence: 5' AAGAAATTAGAAGCTGCAGA 3'

Base Count	Molecular Weight	Extinction Coefficient	Oligo Type	µg/OD at 260 nm	Length (bp)	Tm (°C)	GC%	GC Clamp	Run Length (bp)	Secondary Structure	Primer Dimer	BLAST
A = 10, U = 0, G = 5, C = 2, T = 3, I = 0, Total = 20	6207.1	217.2	No Mod	28.6	20	56.0	35.0	2	3	None	No	Sequence

1. Analysis Results

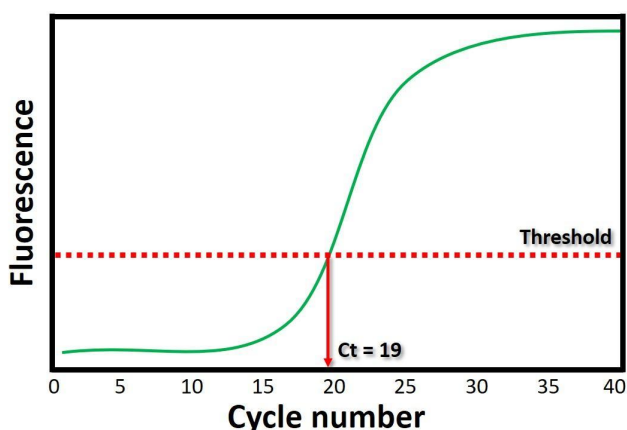
Sequence: 5' TCGGTTCTCTTTATTAGCTT 3'

Base Count	Molecular Weight	Extinction Coefficient	Oligo Type	µg/OD at 260 nm	Length (bp)	Tm (°C)	GC%	GC Clamp	Run Length (bp)	Secondary Structure	Primer Dimer	BLAST
A = 2, U = 0, G = 3, C = 4, T = 11, I = 0, Total = 20	6055.0	176.8	No Mod	34.2	20	54.9	35.0	2	3	None	No	Sequence

Documento explicativo del diseño de *primers* enviado al fabricante.

6.2.1.4. Real Time PCR

La Real Time PCR o Quantitative PCR (qPCR) es una técnica empleada para establecer cuánta cantidad de ADN hay en una muestra. Al igual que una PCR sencilla, se mezclan primers, nucleótidos, polimerasa y ADN en el medio adecuado y se somete a diversos ciclos con cambios de temperatura para que los primers específicos y la polimerasa amplifiquen una región exacta del ADN, pero con la peculiaridad de que con esta técnica podremos cuantificar el ADN amplificado. Para ello introducimos en las muestras un agente fluorescente, en nuestro caso se ha utilizado *SYBR Green*. Una molécula que tras su unión al ADN emite fluorescencia, de tal forma que el dispositivo será capaz de detectar cuánta cantidad de luz se está emitiendo, relacionándose ésta con la cantidad de ADN amplificado hasta el momento. De esta forma, se establece un valor de fluorescencia umbral a partir del cual el dispositivo registra el ciclo de la PCR en el que se alcanzó ese umbral, obteniendo así un valor de Umbral de Ciclo (Ct, en inglés Cycle Threshold), que es inversamente proporcional a la cantidad de ADN existente en la muestra (pues cuanto más bajo sea el Ct implica que antes se cruzó el umbral de luz por lo que mayor concentración de ADN existía en la muestra previa al comienzo de la qPCR). Hay que recalcar que el Ct es una unidad de medida de carácter exponencial, pues con cada ciclo de temperatura que se realiza en la termocicladora se duplica la concentración de ADN.



Una vez obtenemos el ADN complementario al ARN de nuestras muestras mediante la técnica de la retrotranscriptasa, se preparará un mastermix (MM) que contenga los nucleótidos, el SYBR Green (que emitirá fluorescencia una vez se sintetice el ADN), la polimerasa y los primers previamente diseñados y sintetizados. Posteriormente pipeteamos individualmente las proporciones adecuadas de dicho MM junto con el ADN de cada muestra en proporciones similares en la placa de PCR.

El valor de Ct de un gen de interés a estudio en una qPCR puede variar enormemente dependiendo de la termocicladora usada, la experiencia del experimentados, la técnica de

pipeteo y la propia variabilidad de los procedimientos anteriores para obtener el ADN. Por ello además de someter las muestras de ADN a la qPCR con los primers para el gen *STMNI* deberemos introducir en la misma carrera la misma cantidad de ADN por cada muestra pero con el primer de un gen de referencia. Los genes de referencia o gen *housekeeping* son genes que se expresan de manera constante y uniforme en diferentes tejidos y condiciones experimentales. Estos genes se utilizan comúnmente para normalizar los niveles de expresión génica de otros genes de interés en una muestra. Al normalizar la expresión de estos genes de interés con respecto a los niveles de expresión de un gen de referencia, se puede tener una idea más precisa de la cantidad relativa de ARN mensajero presente en una muestra y, por lo tanto, de la actividad de los genes de interés, de modo que el valor de Ct del gen de interés (*STMNI*) no sirve de nada si no se compara con la expresión del gen de referencia o *housekeeping*, que en nuestro caso se ha decidido usar el gen *RPL37A* puesto a que también es empleado como *housekeeping* en el endopredict y tiene una expresión muy estable.

Adicionalmente se preparan las muestras por duplicado en la misma tanda como medida adicional para combatir la variabilidad. El resultado final previo a ejecutar la qPCR debe ser de una placa que contenga por cada muestra dos pocillos de ADN con MásterMix y con primer del gen *STMNI* y otros dos pocillos de ADN con MásterMix y con primer del gen *RPL37A*.



Placa de qPCR cargada y lista para ser introducida en el termociclador



Dispositivo termociclador Stratagene MX3005P para qPCR con su sistema informático

6.2.2. Inmunohistoquímica

La inmunohistoquímica es la técnica de laboratorio que nos permite visualizar y detectar proteínas en una muestra de tejido mediante el uso de anticuerpos específicos. En este caso emplearemos la inmunohistoquímica como una técnica de medición (alternativa a la qPCR) de la cantidad de estatmina que expresan los tejidos tumorales.

Esta técnica de medición ha sido empleada previamente en numerosos estudios alrededor de la estatmina, una vez teñidas las muestras con el anticuerpo correspondiente, uno o varios observadores le asignan un valor en función al porcentaje de muestra que muestra positividad para la tinción. Esta técnica guarda similitud con otras formas de medición de inmunohistoquímica de mayor importancia en la clínica como es la medición de los receptores de estrógenos (ER) y de progesterona (PR) además del Ki67.

La principal ventaja de esta técnica es que es una medición directa sobre la expresión de la estatmina, pues los anticuerpos se unen específicamente a la proteína pudiendo

cuantificar su expresión. Además resulta una técnica más barata de realizar si se cuenta con el material necesario en el laboratorio, pues solamente requiere la tinción con el anticuerpo específico que se solicita al distribuidor del laboratorio. Por otra parte esta técnica tiene el inconveniente de que es observador dependiente, por lo que habrá que tener en cuenta un posible sesgo del observador.

6.2.2.1. Preparación de las muestras

Para comenzar se recopiló la información de etiquetado de todas las muestras de la base de datos para obtener los cortes histológicos en parafina almacenados en el departamento de Anatomía Patológica. En el archivo encontramos cortes histológicos conservados en parafina preparados para tinción de 31 muestras, faltando cortes histológicos para el resto de los casos, por lo que se acudió al archivo de cortes en bloque recuperándose 34 de los casos restantes. Dichos bloques fueron cortados y preparados para tinción con la ayuda del personal del laboratorio. Las muestras preparadas en portas fueron pasadas por un sistema automatizado de tinción cargado con tinción de anticuerpos Anti-Stathmin/Op18 Antibody, que previamente había sido solicitado al distribuidor con varios meses de antelación debida a la demora de entrega típica de este tipo de producto a laboratorios ubicados fuera de la península ibérica. Una vez obtenidas las muestras teñidas estas se dejaron secar durante uno o dos días tras lo cual estaban listas para ser cuantificadas al microscopio óptico.

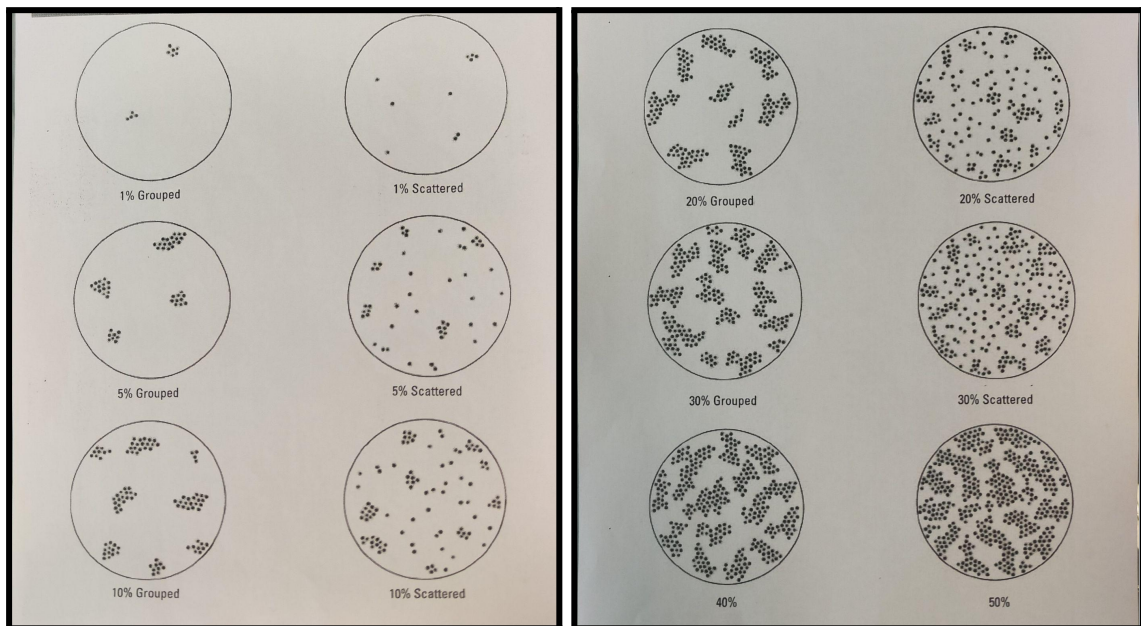


Dispositivo de tinción inmunohistoquímica automatizado *BenchMark Ultra*

6.2.2.2. Medición

Para la cuantificación de las muestras al microscopio es necesaria la utilización de un sistema de escala que combine el porcentaje de superficie de la muestra con tinción positiva con la intensidad de la tinción para obtener un valor de “Score”. Para ello hemos optado por emplear el sistema de Allred Score (10), que es un sistema semicuantitativo que toma en consideración el porcentaje de células positivas a la tinción en una puntuación del 0 al 5 (0, nada; 1, 0-1%; 2, 1-10%; 3, 10-33%; 4, 33-66%; and 5, 66-100%) y la intensidad de la tinción (0, nada; 1, débil; 2, intermedio; y 3, fuerte). Estos se suman para obtener la puntuación de allred (0-8). Además usamos imágenes de apoyo visual sobre las que se establecía una referencia a las diferentes proporciones para obtener una medición más precisa. Con ese mismo objeto también se optó por realizar la medición por parte de dos observadores diferentes que asignan cada uno un valor individual de porcentaje e intensidad

que posteriormente serían puestos en común hallando el valor medio, de esta forma se consigue combatir el sesgo del observador.



Referencias visuales para la cuantificación del porcentaje de células teñidas

La medición del porcentaje de tumor se debe realizar sobre el tejido tumoral por lo que en aquellas muestras más complejas se utilizó como referencia el corte histológico teñido con hematoxilina-eosina para ayudar a identificar el componente tumoral.

Para la obtención del Allred score la medición se ha de realizar sobre la zona tumoral con mayor tinción positiva para la estatmina (punto caliente), sin embargo también recogimos la proporción de tejido teñido en aquellas zonas tumorales con un nivel medio de tinción (punto intermedio) con el objeto de estudios adicionales.

7. Resultados

7.1. Variables del test EndoPredict

El test EP nos permite cuantificar el pronóstico de cada caso calculando la amplificación de genes de proliferación y de diferenciación, reflejando el pronóstico en forma de diferentes scores que emplearemos como variables para el posterior estudio de correlación. El score de proliferación (EP prolif) obtenido a partir de la expresión de los genes de proliferación (que se correlacionan a peor pronóstico), el score de diferenciación (EP dif)

calculado mediante la expresión de los genes de diferenciación (asociados a un mejor pronóstico), el score de risk biológico (EP bio) calculado a partir de los dos previos que refleja el pronóstico basado únicamente en datos biomoleculares, y por último el score clínico (EP clin) obtenido aplicando al EP bio modificaciones según parámetros clínicos (como la T y la N del tumor). Este último es el de mayor importancia en la práctica clínica, diferenciando si se trata de un tumor de alto o bajo riesgo en función de su valor.

Partimos de 100 muestras de pacientes diagnosticadas de cáncer de mama de tipo luminal B sometidas al test EndoPredict (EP). Dichas muestras ya habían sido sometidas al test endopredict, cuantificando la expresión de los genes necesarios para este estudio y aplicándose los datos en el algoritmo de la web de la empresa para obtenerse los valores de los scores. Sin embargo, con el objeto de entender el funcionamiento del test, decidimos recalcular el score de las muestras (según las fórmulas publicadas en un artículo en el que se basó el desarrollo comercial de EndoPredict) (12) empleando los valores de amplificación de los genes necesarios y los datos clínicos de cada paciente, tales como la T y la N del tumor.

Para ello utilizamos el programa Excel, en el que preparamos en varias hojas los algoritmos para obtener tanto el *score* EPbio como el *score* EPclin a partir de los datos introducidos. Dichas ecuaciones fueron obtenidas a partir del artículo, ya citado, que detalla el funcionamiento del test EP. Para obtener los *scores*, partimos de las concentraciones obtenidas en una qPCR de 8 genes de interés (GOI; *gene of interest*) relacionados con el cáncer de mama (*BIRC5*, *UBE2CA*, *DHCR7*, *RBBP8*, *IL6ST*, *AZGP1*, *MGP*, y *STC2*), y de otros 3 genes denominados “*housekeeping*” (*CALM2*, *OAZ1*, y *RPL37A*) que se emplearán como genes de referencia. La expresión relativa de cada GOI se denominó *Delta Cycle threshold* (DCt) y se obtienen empleando la siguiente fórmula que normaliza los valores en base a la expresión de los genes de referencia:

$$\Delta C_t(\text{GOI}) = 20 - C_t(\text{GOI}) + [C_t(\text{CALM2}) + C_t(\text{OAZ1}) + C_t(\text{RPL37A})]/3$$

Los DCt de cada uno de los genes se combinaron de la siguiente manera para obtener el *predictive unscaled risk score* (Su):

$$\begin{aligned}
s_u = & 0.41 \cdot \Delta C_t(BIRC5) - 0.35 \cdot \Delta C_t(RBBP8) \\
& + 0.39 \cdot \Delta C_t(UBE2C) - 0.31 \cdot \Delta C_t(IL6ST) \\
& - 0.26 \cdot \Delta C_t(AZGP1) + 0.39 \cdot \Delta C_t(DHCR7) \\
& - 0.18 \cdot \Delta C_t(MGP) - 0.15 \cdot \Delta C_t(STC2) - 2.63
\end{aligned}$$

El EP bio (S) final oscila entre 0 y 15 por lo que se emplearon los siguientes criterios para obtener el EP bio (S) a partir del S_u :

$$\begin{aligned}
s = 0, & \quad \text{if } 1.5 \cdot s_u + 18.95 < 0 \\
s = 15, & \quad \text{if } 1.5 \cdot s_u + 18.95 > 15 \\
s = 1.5 \cdot s_u + 18.95, & \quad \text{otherwise}
\end{aligned}$$

El EP clin (Sclin) por otra parte, es un score que combina el EP risk score y otros parámetros clínicos como el tamaño del tumor y la afectación ganglionar. Se obtiene a partir de esta fórmula:

$$s_{clin} = 0.35 \cdot t + 0.64 \cdot n + 0.28 \cdot s$$

En ella t hace referencia al tamaño del tumor (1: hasta 1 cm, 2: >1 a 2 cm, 3: >2 a 5 cm y 4: >5 cm) y n al número de ganglios afectados (1:negativo, 2: 1–3 ganglios positivos, 3: 4–10 ganglios positivos, y 4: >10 ganglios positivos).

Finalmente, comprobamos cómo nuestro valor de EP tenía un error medio de -0.62 con una desviación estándar de 0.084, lo que significa que en todas las muestras existe un error de cálculo negativo constante muy aproximado a la media. Lo mismo ocurrió con el EPclin, que presentaba un error medio de -0.19 con una desviación estándar de 0.071. La causa de este error debió ser o bien el empleo de una fórmula de cálculo desactualizada (dado que el único artículo donde aparece pública esta fórmula es del año 2011) o bien un error decimal debido a que sólo introducimos valores del orden de dos decimales en nuestra ecuación.

7.2. Método DeltaDeltaCt

Para recopilar toda la información se fueron recabando y almacenando los datos en una base en el programa Excel. Aquellas variables de las que ya disponíamos la información sobre el estado de mutación del PI3K fueron almacenadas directamente mientras que aquellas

que estaban pendientes de determinación se fueron almacenando según se obtuvo los resultados de la Pirosecuenciación. Por otro lado se fue recopilando los resultados obtenidos de las qPCR sobre las muestras con los primers para *STMN1* y para *RPL37A*. Se realizó un total de 10 carreras de qPCR con un resultado final de 82 casos cuantificados, con una media de entre 8 y 9 muestras por cada tanda de qPCR. Partiendo de las 100 muestras sobre las que se hizo estudio EP, durante la cuantificación hubo un total de 18 muestras que no se pudo cuantificar, siendo los principales motivos; la procedencia de los datos de las muestras de otros hospitales (imposibilitando la obtención de ADN) y la ausencia de material genético suficiente. Esta pérdida de muestras estaba dentro de nuestra estimación inicial.

El resultado final fue de una base de datos de 82 casos en los que figuraban las variables de estado mutacional (mutado/wild type), y de expresión de ARN del gen *STMN1* (junto con su gen de referencia) de cada una de ellas. En cada carrera se había amplificado cada muestra por duplicado, por lo que el primer paso a realizar fue obtener el valor medio de Ct de ambas réplicas de cada muestra.

Partiendo de esta base de datos se optó por realizar una técnica específica de comparación para la expresión de diferentes genes; el método DeltaDeltaCt ($\Delta\Delta Ct$). El método $\Delta\Delta Ct$ es una técnica ampliamente utilizada en biología molecular para analizar la expresión génica relativa de un gen objetivo en un grupo de testeo en comparación con un grupo de control. Este método se basa en la medición de la cantidad de ARN mensajero (ARNm) de interés (*STMN1*) y de un gen de referencia (*RPL37A*), y en la comparación de los niveles de expresión entre las muestras. Para ello se somete a los valores de Ct a una serie de fórmulas matemáticas, en las que se obtiene la diferencia entre el gen a estudio (GOI) y el gen de referencia (Ref) tanto en el grupo de mutados (A) como en el grupo de wild type (B), para posteriormente calcular la diferencia de las diferencias ($\Delta\Delta Ct$) y luego normalizar el resultado mediante la fórmula $2^{-(\Delta\Delta Ct)}$.

$$\begin{aligned} \Delta Ct A &= Ct_{GOI} A - Ct_{Ref} A \\ \Delta Ct B &= Ct_{GOI} B - Ct_{Ref} B \\ &\downarrow \\ \Delta \Delta Ct &= \Delta Ct A - \Delta Ct B \\ \text{normalized expression} &= \\ &2^{-(\Delta \Delta Ct)} \end{aligned}$$

Tras esto se debe aplicar cálculos de tendencia central sobre los valores de $2^{-(\Delta\Delta Ct)}$ de cada grupo, en nuestro caso se realizó la media aritmética, media geométrica y la mediana. Mientras que la media aritmética destaca al tener en cuenta el aporte de todas los casos a la tendencia central, la mediana y la media geométrica destacan a la hora de descartar aquellos casos que se alejan enormemente de la tendencia central (*outliers*), pudiendo no ser correctamente representativos de esta.

El resultado final es de un valor de media aritmética, geométrica y mediana de $2^{-(\Delta\Delta Ct)}$ para el grupo control (wild type) y para el grupo de estudio (mutados), que se interpreta mediante el cociente de ambos. Si el cociente de ambos valores de $2^{-(\Delta\Delta Ct)}$ es igual a 1 significa que en ambos grupos el gen de interés se expresa por igual, mientras que si es igual a 2 significa que el gen se expresa el doble en el grupo de estudio. Nuestros resultados fueron los siguientes.

$2^{-(\Delta\Delta Ct)}$	Wild type	Mutados
Media aritmética	1,786	1,644
Media geométrica	1	1,107
Mediana	0,823	1,039

Calculando el cociente en cada caso:

Media aritmética	1,644/1,786	0,92
Media geométrica	1,107/1	1,107
Mediana	1,039/0,823	1,262

La interpretación de estos resultados será la siguiente:

- Acorde al cálculo mediante **media aritmética** en los pacientes mutados en PI3K la expresión de estatmina (determinada por la cuantificación de ARN del gen STMN1) es inferior a la de los pacientes wild type para PI3K, siendo esta un **92%** (8% menor) en relación a los no mutados.
- Acorde al cálculo mediante la **media geométrica** los pacientes mutados en PI3K sobreexpresan la estatmina (determinada por la cuantificación de ARN del gen STMN1) en relación a los pacientes wild type, siendo esta un **110,7%** (10,7% mayor) en relación a los no mutados.
- Acorde al cálculo mediante la **mediana** los pacientes mutados en PI3K sobreexpresan la estatmina (determinada por la cuantificación de ARN del gen STMN1) en relación a los pacientes wild type, siendo esta un **126,2%** (26,2%) en relación a los no mutados.

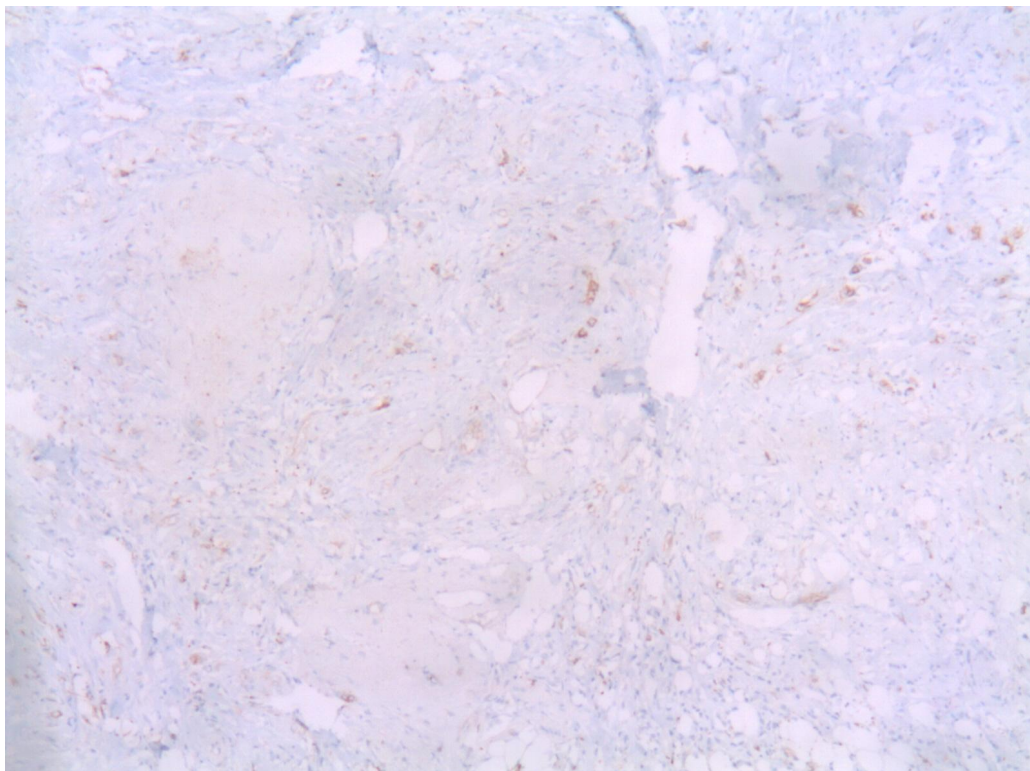
La gran variabilidad de estos resultados según el método de tendencia central que se utilice apunta a que con gran probabilidad no alcancen el nivel de significatividad requerido. No obstante, el hecho de que cuanto más despreicia los outliers la medida de tendencia central (mediana > media geométrica > media aritmética) más se sobreexpresa la estatmina, sugiere que este cálculo está sometido a una gran variabilidad en la medición de las variables y que con estudios posteriores de un mayor tamaño muestral, los resultados pueden llegar a alcanzar un nivel de significatividad superior.

Aunque nuestros resultados fueron mixtos, seguimos considerando que el método delta delta Ct es la mejor forma de relacionar estadísticamente las expresiones génicas de dos grupos diferentes. Sin embargo, para confirmar que no hay una relación significativa entre niveles bajos de Ct y la mutación en PI3K, realizamos una prueba U de Mann Whitney (asumiendo que la distribución de los valores de Ct no sigue la normalidad) del valor de ΔCt agrupando por el estado mutacional de PI3K. Los datos obtenidos demostraron que no hay

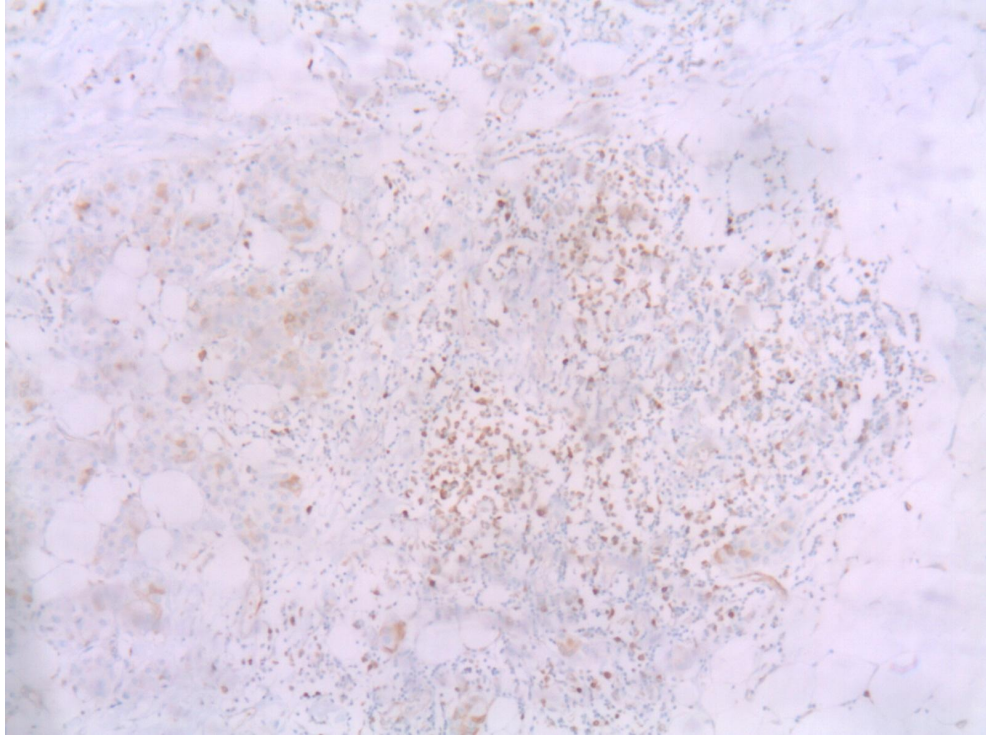
una correlación significativa entre estos dos parámetros, siendo la P de un valor de 0,74 (mayor a 0,05).

7.3. Expresión de estatmina en inmunohistoquímica

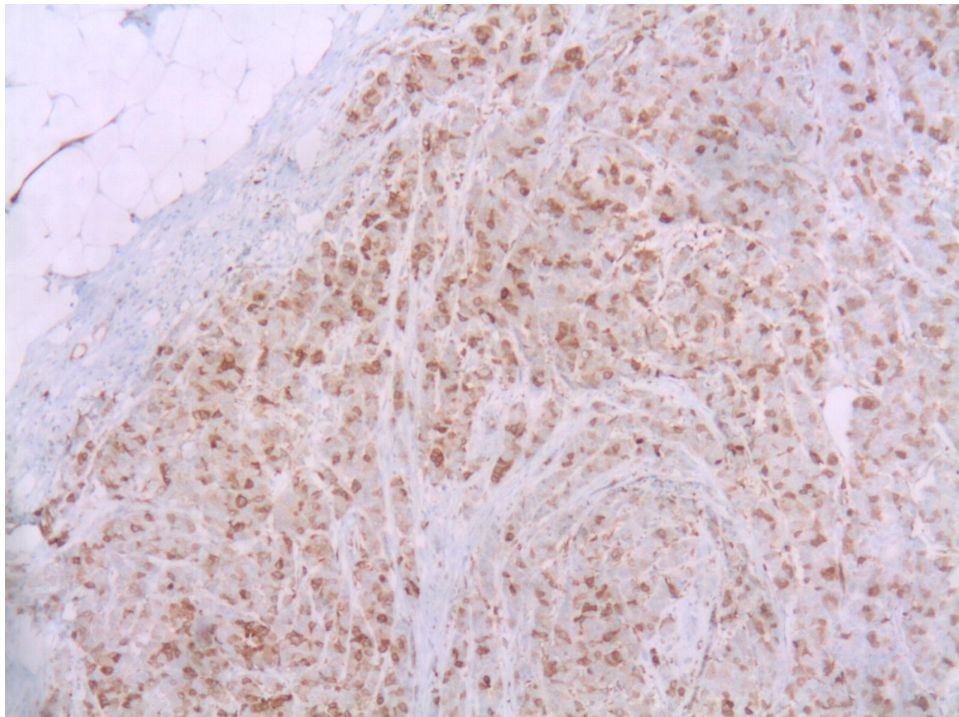
La medición de la tinción en inmunohistoquímica se realizó siguiendo el esquema de Allred, que tiene en cuenta el porcentaje de tinción en la zona de mayor tinción tumoral y la intensidad de dicha tinción. Este sistema establece un Score con una puntuación resultado de la suma de los parámetros anteriores, que va del 0 al 8. Además del porcentaje de tinción en zona de mayor tinción (que llamamos “tinción en punto caliente”) y la intensidad se midió como variable adicional para el estudio la tinción en la zona intermedia del tumor que no expresaba tanta estatmina (llamada “tinción en punto intermedio”). La tinción en punto caliente e intermedio se midieron con la unidad de porcentaje (%), representando el porcentaje de tinción observado a 10 aumentos en el microscopio óptico, mientras que la intensidad consistía en un valor representativo de la cantidad de tejido al que se adhiere el anticuerpo representado en una escala del 0 al 3.



Muestra que presenta un 5% de células teñidas, intensidad 1. Allred score 3.



Muestra que presenta un 15% de células teñidas, intensidad 2. Allred score 5.



Muestra que presenta >66% de células teñidas, intensidad 3. Allred score 8.

Durante la medición de esta variable se perdieron un total 17 casos en relación al total sobre los que se realizó la técnica de qPCR. Esta pérdida de seguimiento se debe mayoritariamente a la ausencia de los bloques necesarios para hacer los cortes histológicos, pues debido a la antigüedad de algunos de los casos, estos no se conservaron en el laboratorio. Nuestra n final para esta medición será de 65 casos, que se subdividen en 42 mutados y 23 wild type.

Una vez se asignaron los valores a las muestras, estas fueron agrupadas según la variable PI3K Status, que representa el estado mutacional de PI3K. Acto seguido se realizó una prueba T de Student para comparar las medias de ambos grupos asumiendo que ambos siguen una distribución normal. Estos fueron los resultados:

	PI3K_Status	N	Media	Desv. estándar	Media de error estándar
Allred_Score	Mutado	42	4,48	1,330	,205
	Wild Type	23	4,13	1,517	,316
Punto_Caliente	Mutado	42	19,45	16,516	2,549
	Wild Type	23	16,61	14,935	3,114
Punto_Medio	Mutado	42	8,67	8,336	1,286
	Wild Type	23	7,78	8,180	1,706
Intensidad	Mutado	42	1,83	,696	,107
	Wild Type	23	1,70	,822	,171

La principal variable representativa de la expresión de estatmina en la inmunohistoquímica es el Allred score, a pesar de esto, el estudio se extendió también a estas otras tres variables en busca de correlaciones adicionales. En esta tabla podemos observar como el grupo mutado representa un valor medio superior al grupo wild type para todas las variables que representan la expresión de estatmina, sin embargo completando el análisis de la prueba T de Student en la tabla inferior podemos observar que ninguna de estas diferencias resulta significativa debido a que la P en ningún momento es inferior a 0,05.

	n	Significación	Diferencia de medias	Diferencia de error estándar	95% de intervalo de confianza de la diferencia	
		Valor de P			Inferior	Superior
Allred_Score	63	,172	,346	,363	-,379	1,070
Punto_Caliente	63	,248	2,844	4,146	-5,441	11,128
Punto_Medio	63	,341	,884	2,148	-3,409	5,177
Intensidad	63	,239	,138	,193	-,247	,522

Con el objetivo de estudiar más correlaciones, convertimos la variable de Allred Score en una variable categórica denominada Allred Status. Para ello se acepta que aquellas tinciones con un valor de Allred Score superior a 3 son positivas, mientras que aquellas con valor igual o inferior a 3 serán negativas para la expresión de estatmina (un esquema similar al utilizado para semicuantificar la expresión de receptores de estrógenos). De este modo podemos realizar una tabla de contingencia entre las dos variables de PI3K Status y Allred Status para realizar una prueba Chi cuadrado.

		Allred_Status		Total
		Negativo	Positivo	
PI3K_Status	Mutado	11	31	42
	Wild Type	11	12	23
Total		22	43	65

	Valor	gl	Significación asintótica (bilateral)
Chi-cuadrado de Pearson	3,107	1	,078
Corrección de continuidad	2,216	1	,137
Razón de verosimilitud	3,057	1	,080
N de casos válidos	65		

Para este caso existe un valor de Chi-cuadrado de 3,107, lo que indica una discrepancia entre los datos observados y los datos esperados considerable. Sin embargo no podemos expresar significatividad pues el valor de P es de 0,078 que es mayor de 0,05. En

cualquier caso, un valor de p menor a 0,1 sugiere que hay una tendencia hacia la significación estadística, lo que incentiva a futuros investigadores a realizar nuevos estudios de correlación aumentando el tamaño muestral así como la precisión en la medición de las variables, quizá mediante análisis de imagen y cuantificación de la señal inmunohistoquímica.

Con el objetivo de evaluar la relación entre los datos obtenidos mediante la qPCR y la inmunohistoquímica, se llevaron a cabo varios análisis estadísticos no paramétricos. En primer lugar, se realizó una correlación de Spearman entre el Allred Score y el Ct de las mismas muestras. Se esperaba encontrar una alta correlación entre la expresión de estatmina, independientemente de la técnica empleada, dado que se trata de diferentes formas de cuantificación de una misma variable (la expresión de estatmina). Sin embargo, no encontramos una correlación significativa. Esta falta de correlación podría, en teoría, deberse a que el gen esté regulado a nivel transcripcional y postraduccional de una manera compleja. También es posible que la metodología de determinación de los niveles de ARNm y proteína no sean igualmente robustos o sensibles. Es decir que al menos uno de ellos, o bien no es lo suficientemente representativo del fenómeno natural a medir, o bien está sujeto a demasiada variabilidad siendo necesaria un tamaño muestral mayor para obtener representatividad.

En conclusión, nuestros resultados sugieren que la relación entre los datos obtenidos mediante la qPCR y la inmunohistoquímica no es tan clara como se podría esperar. La falta de una correlación significativa puede deberse a la complejidad del Ct como un dato no paramétrico. Es necesario continuar investigando en esta área para entender mejor la relación entre estas técnicas y su interpretación.

7.4. Expresión de estatmina y test Endopredict

Como objetivo adicional se planteó el estudio de correlación entre la expresión de estatmina y las variables derivadas del test EP, pues este se emplea como forma de cuantificar numéricamente el pronóstico en 10 años vista de las pacientes. Con esta correlación podremos esclarecer si la medición de la expresión de estatmina mediante inmunohistoquímica es lo suficientemente precisa como para servir como método adicional de cuantificación pronóstica.

Para ello se realizó pruebas de correlación entre las variables de expresión de stamina (Allred score y Allred status) y las variables derivadas del test EP (EP clin, EP bio, EP dif y

EP prolif). En primer lugar se realizó una prueba T de student para las variables del test EP agrupando en base a la variable categórica Allred status.

	Allred_Status	N	Media	Desv. estándar	Media de error estándar
EP_Clin	Negativo	22	3,306727	,5741231	,1224035
	Positivo	43	3,331744	,7083438	,1080214
EP_Bio	Negativo	22	5,681182	2,4936276	,5316432
	Positivo	43	5,977326	2,2242923	,3392014
EP_Dif	Negativo	22	22,711273	,8936082	,1905179
	Positivo	43	22,361488	1,0266806	,1565673
EP_Prol	Negativo	22	16,495227	1,1794219	,2514536
	Positivo	43	16,343000	1,0516433	,1603741

	n	Significación	Diferencia de medias	Diferencia de error estándar	95% de intervalo de confianza de la diferencia	
					Inferior	Superior
EP_Clin	63	,443	-,0250169	,1747370	-,3742010	,3241672
EP_Bio	63	,314	-,2961438	,6074918	-1,5101194	,9178318
EP_Dif	63	,090	,3497844	,2580175	-,1658225	,8653912
EP_Prol	63	,299	,1522273	,2872627	-,4218214	,7262759

En esta prueba observamos que la media es mayor en los positivos para el Allred Score de estatmina para el EP clin y en el EP bio, sin embargo se trata de una correlación sin significancia suficiente para rechazar la hipótesis nula. Por otra parte la media para el EP dif es menor en el grupo de los positivos en Allred status, siendo esta a favor de la hipótesis alternativa, pues el EP dif es un score inversamente proporcional al pronóstico. Esta última correlación muestra una cierta tendencia a la significancia, pues la p tiene un valor de 0,09, sin ser suficiente para rechazar la hipótesis nula.

Además se realizó análisis de correlación de Pearson entre el score de EP clin y el Allred score. Se obtuvo un valor de Pearson de 0,05, que al estar próximo al cero indica ausencia de correlación, además de una significancia muy alejada del 0,05.

8. Discusión

En esta investigación, hemos explorado la relación entre la expresión de estatmina y la activación de la vía de PI3K. Sin embargo, nuestros resultados no han sido concluyentes en

cuanto a una relación significativa entre estas dos variables. A pesar de que nuestros datos iniciales no muestran una correlación clara entre la expresión de estatmina y la activación de PI3K, estos resultados deben ser considerados con precaución debido al tamaño limitado de nuestra muestra y la complejidad de los datos. En primer lugar, hemos encontrado una ligera tendencia a una mayor expresión de estatmina en las muestras mutadas que en las wild type. Por otro lado, en la inmunohistoquímica también se ha observado esta misma tendencia. Sin embargo, no se halló relación entre los datos de ambas técnicas. Además, ninguna de las dos demostró una diferencia significativa. Quizás la estadística más interesante es la que relaciona el Allred Status y el PI3K Status, que parece indicar una tendencia más fuerte. Recordemos que la P a pesar de no ser inferior a 0,05, es inferior a 0,1 lo que puede indicar que si se amplía el tamaño muestral, esta podría alcanzar la significancia. Sin embargo se trata de una variable que es menos representativa que el Allred Score, pues esta última es continua mientras que la primera es nominal.

En cualquier caso, es necesario destacar que los estudios previos han sugerido una fuerte conexión entre la estatmina y la proliferación de ciertos tipos de cáncer, así como su asociación con la activación de la ruta PI3K/AKT. La estatmina se ha identificado como un importante regulador de la vía de PI3K en varios tipos de cáncer, incluyendo el cáncer de mama y el cáncer de pulmón. En consecuencia, la identificación de estatmina como un marcador potencial de la activación de la vía de PI3K ha generado un gran interés en la investigación oncológica.

Aunque nuestros datos no proporcionan una evidencia concluyente de esta relación, nuestros resultados proporcionan información valiosa para futuros estudios. En particular, sugieren que la expresión de estatmina no puede ser considerada como un marcador confiable para la activación de la vía de PI3K en nuestra población de estudio. Sin embargo, es importante destacar que este se ha centrado en una muestra pequeña y puede haber limitaciones en la generalización de nuestros resultados a otras poblaciones.

En conclusión, nuestra investigación sugiere que se necesitan estudios adicionales para explorar más a fondo la relación entre la expresión de estatmina y la activación de la vía de PI3K. Es importante destacar que la identificación de marcadores fiables para la activación de PI3K es de gran importancia clínica, ya que la activación anormal de PI3K es un evento común en el cáncer de mama (además de otros procesos oncológicos) y es una diana importante para la terapia antitumoral. Por lo tanto, se necesitan más estudios para identificar

marcadores precisos y confiables de la activación de PI3K, lo que puede tener implicaciones importantes en el tratamiento y pronóstico del cáncer.

9. ¿Qué nos ha aportado este TFG?

A lo largo de la realización de este proyecto se nos han presentado numerosos desafíos. Algunos de ellos los conseguimos superar con nuestro propio esfuerzo y razonamiento, mientras que para otros hemos requerido de la ayuda de nuestros compañeros del laboratorio y de nuestro tutor, aprendiendo de todos ellos nuevas habilidades y lecciones que demostrarán su utilidad en el futuro. Uno de los principales desafíos ha sido aprender a reconocer e identificar los errores críticos (critical point failures) donde cualquier técnica puede fracasar, para poder anticiparnos y actuar en ellos con especial cuidado. Otro principal desafío ha sido aprender a recopilar información de artículos y bases de datos científicas, tales como cBioPortal, UniProt, PubMed... El conocimiento de los métodos de trabajo del laboratorio demostrará su utilidad incluso si nuestro futuro se orienta a la práctica clínica, pues hemos aprendido que este es la base de casi todos los procedimientos que se realizan en el hospital.

Por otro lado, la virtud que más ha relucido en este TFG ha sido el trabajo en equipo y el compañerismo que ha destacado ante todo, enriqueciendo todos los demás aspectos de este proyecto. En definitiva, este proyecto nos ha entrenado y nos ha ayudado a crecer como personas, tanto en la dimensión profesional; aprendiendo los fundamentos y la importancia de los trabajos de investigación, como en la dimensión social conociendo a muchos compañeros que con el transcurso del proyecto se han convertido en amigos.

10. Bibliografía

1. Panal Cusati, M., Herrera de la Muela, M., Hardisson Hernaez, D., Choqueneira Dionisio, M., Román Guindo, A., & de Santiago Garcia, F. J. (2014). Correlación entre la expresión de Ki67 con factores clásicos pronósticos y predictivos en el cáncer de mama precoz. *Revista de senología y patología mamaria*, 27(4), 163–169. <https://doi.org/10.1016/j.senol.2014.07.005>
2. <https://www.dlongwood.com/productos/endopredict/>
3. Understanding the PI3K/AKT pathway. (2021). Recuperado 27 de octubre de 2022, de Genentech Oncology website: <https://www.genentechoncology.com/pathways/cancer-tumor-targets/pi3k.html>
4. Rubin CI, Atweh GF. The role of stathmin in the regulation of the cell cycle. *J Cell Biochem*. 2004 Oct 1;93(2):242-50. doi: 10.1002/jcb.20187. PMID: 15368352.
5. Lee JH, Kang MJ, Han HY, Lee MG. Expression of stathmin/op18 as a prognostic marker in colon cancer. *Int J Colorectal Dis*. 2010;25(11):1333-1338. doi:10.1007/s00384-010-0981-7
6. Lu Y, Lin YZ, LaPushin R, et al. Stathmin, a new microtubule-associated protein regulated by phosphorylation. *J Cell Biol*. 1992; 138(3): 657-666. doi: 10.1083/jcb.138.3.657
7. Sevinc AI, Seker M, Bilici A, et al. The relationship between stathmin 1 and survivin expressions and prognosis in non-small cell lung cancer. *J BUON*. 2013;18(2):444-450.
8. Taghavi N, Biramijamal F, Sotoudeh M, Khademi H, Malekzadeh R, Moaven O. Association of stathmin expression with esophageal squamous cell carcinoma and its prognosis. *J Gastroenterol Hepatol*. 2008;23(4):675-9. doi: 10.1111/j.1440-1746.2007.05244.x
9. Saal LH, Johansson P, Holm K, Gruvberger-Saal SK, She QB, Maurer M, Koujak S, Ferrando AA, Malmström P, Memeo L, Isola J, Bendahl PO, Rosen N, Hibshoosh H, Ringnér M, Borg A, Parsons R. Poor prognosis in carcinoma is associated with a gene expression signature of aberrant PTEN tumor suppressor pathway activity. *Proc Natl Acad Sci U S A*. 2007 May 1;104(18):7564-9. doi: 10.1073/pnas.0702507104. PMID: 17442748; PMCID: PMC1855396.
10. Obayashi S, Horiguchi J, Higuchi T, Katayama A, Handa T, Altan B, Bai T, Bao P, Bao H, Yokobori T, Nishiyama M, Oyama T, Kuwano H. Stathmin1 expression is

- associated with aggressive phenotypes and cancer stem cell marker expression in breast cancer patients. *Int J Oncol.* 2017 Sep;51(3):781-790. doi: 10.3892/ijo.2017.4085. Epub 2017 Jul 27. PMID: 28766688; PMCID: PMC5564402.
11. Purificación de ácidos nucleicos mediante columnas de centrifugado. En: Wikipedia, la enciclopedia libre [Internet]. 2022 Disponible en: https://es.wikipedia.org/w/index.php?title=Purificaci%C3%B3n_de_%C3%A1cidos_nucleicos_mediante_columnas_de_centrifugado&oldid=143540045
 12. Filipits M, Rudas M, Jakesz R, Dubsky P, Fitzal F, Singer CF, Dietze O, Greil R, Jelen A, Sevelda P, Freibauer C, Müller V, Jänicke F, Schmidt M, Kölbl H, Rody A, Kaufmann M, Schroth W, Brauch H, Schwab M, Fritz P, Weber KE, Feder IS, Hennig G, Kronenwett R, Gehrman M, Gnant M; EP Investigators. A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors. *Clin Cancer Res.* 2011 Sep 15;17(18):6012-20. doi: 10.1158/1078-0432.CCR-11-0926. Epub 2011 Aug 1. PMID: 21807638.