



UNIVERSIDAD REGIONAL AMAZÓNICA IKIAM

FACULTAD DE CIENCIAS DE LA VIDA

CARRERA DE INGENIERÍA EN BIOTECNOLOGÍA

**EVOLUCIÓN Y ANÁLISIS FUNCIONAL DE LA REGIÓN
N-TERMINAL DE LA ONCOPROTEÍNA GOLPH3 DEL APARATO
DE GOLGI EN MAMÍFEROS**

Proyecto de investigación previo a la obtención del Título de:

INGENIERA EN BIOTECNOLOGÍA

AUTORA: JESSICA ABIGAIL GALARZA CHICO

TUTORA: PhD. ERY ODETTE FUKUSHIMA

CO-TUTOR: Dr. GONZALO MARDONES

Napo-Ecuador

2022

DECLARACIÓN DE DERECHO DE AUTOR, AUTENTICIDAD Y RESPONSABILIDAD

Tena, 21 de octubre de 2022

Yo, Jessica Abigail Galarza Chico con documento de identidad N° 1750306340, declaro que los resultados obtenidos en la investigación que presento en este documento final, previo a la obtención del título son absolutamente inéditos, originales, auténticos y personales.

En virtud de lo cual, el contenido, criterios, opiniones, resultados, análisis, interpretaciones, conclusiones, recomendaciones y todos los demás aspectos vertidos en la presente investigación son de mi autoría y de mi absoluta responsabilidad.

Por la favorable atención a la presente, suscribo de usted,

Atentamente,

Firma:



Jessica Abigail Galarza Chico

CI:1750306340

AUTORIZACION DE PUBLICACION EN EL REPOSITORIO INSTITUCIONAL

Yo, Jessica Abigail Galarza Chico, con documento de identidad N° 1750306240 en calidad de autor/a y titular de los derechos morales y patrimoniales del trabajo de titulación:.. Evolución y análisis funcional de la región N-terminal de la oncoproteína GOLPH3 del aparato de Golgi en mamíferos, de conformidad con el Art. 114 del CÓDIGO ÒRGANICO DE LA ECONOMÍA SOCIAL DE LOS CONOCIMIENTOS, CREATIVIDAD E INNOVACIÒN, reconozco a favor de la Universidad Regional Amazónica Ikiam una licencia gratuita, intransferible y no exclusiva para el uso no comercial de la obra, con fines estrictamente académicos.

Así mismo autorizo a la Universidad Regional Amazónica Ikiam para que realice la publicación de este trabajo de titulación en el Repositorio Institucional de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación superior.

Tena, 21 de octubre del 2022

Firma:



Jessica Abigail Galarza Chico

CI:1750306340

CERTIFICADO DE DIRECCIÓN DE TRABAJO DE INTEGRACIÓN CURRICULAR

Certifico que el trabajo de integración curricular titulado: “*Evolución y análisis funcional de la región N-terminal de la oncoproteína GOLPH3 del aparato de Golgi en mamíferos*”, en la modalidad de: artículo original, fue realizado por: Jessica Abigail Galarza Chico, bajo mi dirección.

El mismo ha sido revisado en su totalidad y analizado por la herramienta de verificación de similitud de contenido; por lo tanto, cumple con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad Regional Amazónica Ikiam, para su entrega y defensa.

Tena, 21 de octubre de 2022

Firma:



Ery Odette Fukushima

C.I: 1759301714

Firma:

A handwritten signature in blue ink is shown. Below the signature is a horizontal dotted line.

Gonzalo Antonio Mardones Cofre

RUT: 12.128.650-5

DEDICATORIA

A mis padres Luis Galarza y Sandra Chico, a mis hermanas Melanie y Gabriela, a mi novio Nicolas, a mi gatita Noha, que siempre estará en mi corazón, a mis amigas Gaby Jiménez, Mafer Muñoz, Mariuxi Sánchez, a mi compañero de carrera Jefferson Pastuña, a mis amigos, que me acompañaron en Chile, Jan Castro, Sebastián Triviño, Rubén, Camila, Andrés, Charlotte, a mi amiga Tammy Pastor, que estuvo conmigo en esta aventura, a mis profesores, que me ayudaron a lo largo de mi carrera, Emmanuel Ambriz, Patricia Salerno, Gonzalo Mardones, Juan Opazo, Ery Fukushima, les dedico esta tesis, por darme su apoyo, a pesar de la distancia, sin ellos, este trabajo no habría tenido razón de ser. A la Amazonía Ecuatoriana, por su enseñarme, cuidado, protección, amor, sencillez y magia en cada rincón. A Valdivia, por ser, la ciudad más linda de Chile y la capital cervecera del país.

AGRADECIMIENTOS

A la ciencia, que nos muestra que, en el caos aparente de los fenómenos, la manifestación de sus leyes fundamentales. La biotecnología hace uso de estas leyes para generar herramientas que ayuden a la humanidad, sin olvidar el uso correcto de los recursos naturales, aquella ciencia que no se aísla del pueblo, sino que está dispuesta a servir al mismo. A Dios, por cuidar de mi cada día de mi vida. A mi padre, que desde pequeña me educó con un profundo respeto por la naturaleza, por ayudarme, apoyarme y nunca dejar de creer en mí. A mi madre, que me enseñó que todo esfuerzo es necesario para lograr cada objetivo, por protegerme, quererme y entenderme. A mis hermanas quienes han sido mi fuerza, a quienes espero haber sido un ejemplo y un apoyo en su vida. A mi tutora Ery Fukushima, por su apoyo incondicional y paciencia. A mis tutores, Gonzalo Mardones y Juan Opazo, quienes me brindaron la gran oportunidad de realizar esta investigación. Permitirme vivir mi primera experiencia fuera de mi país. Su apoyo y confianza en mi trabajo y su capacidad de guiarme cada día, no solo para este trabajo si no también en mi formación como estudiante. Les agradezco también el haberme facilitado los medios suficientes para el desarrollo de esta tesis. Destacando no solo lo material, si no el tiempo y la paciencia de enseñarme. A Charlotte, por ayudarme con el desarrollo y uso de las líneas celulares para este estudio. A todos mis amigos a quienes dedico esta tesis y a mis compañeros de carrera. Juntos hemos superado obstáculos y nos hemos emocionado por cada victoria individual.

Gracias a todos los que permitieron que este pequeño sueño sea posible.

INDICE GENERAL

CARATULA

DECLARACIÓN DE DERECHO DE AUTOR, AUTENTICIDAD Y RESPONSABILIDAD.....	i
AUTORIZACION DE PUBLICACION EN EL REPOSITORIO INSTITUCIONAL	ii
CERTIFICADO DE DIRECCIÓN DE TRABAJO DE INTEGRACIÓN CURRICULAR.....	iii
DEDICATORIA.....	iv
AGRADECIMIENTOS.....	v
INDICE GENERAL	vi
ÍNDICE DE TABLAS.....	viii
ÍNDICE DE FIGURAS.....	ix
ÍNDICE DE ANEXOS.....	x
Resumen.....	xi
Abstract.. ..	xii
1. Introducción.....	1
2. Materiales y métodos.....	3
2.2. Construcciones de ADNc recombinante.....	4
2.3. Expresión y purificación de la proteína GOLPH3 de las especies de mamíferos seleccionadas.....	4
2.4. Transfección y microscopía de inmunofluorescencia.....	6
2.5. Ensayo de proteólisis parcial.....	6
2.6. Ensayo de estabilidad térmica por fluorimetría de escaneo diferencial (DSF)	7
2.7. Ensayo de separación de fases líquido-líquido.....	7
2.8. Análisis estadísticos.....	8
3. Resultados.....	8
3.1. El segmento N-terminal de la oncoproteína GOLPH3 es más variable en comparación a su dominio C-terminal.	8
3.2. Las variantes de GOLPH3 recombinantes muestran leves diferencias en su estabilidad térmica	14

3.3. La estructura de las variantes del segmento N-terminal de GOLPH3 es similar al GOLPH3 del tipo salvaje.....	16
3.4. Las variaciones del segmento N-terminal de GOLPH3 favorecen su partición por separación de fases líquido-líquido.....	18
4. Discusión y conclusiones.....	21
Referencias.....	28

ÍNDICE DE TABLAS

Tabla 1:	Punto isoeléctrico de las variantes de GOLPH3.....	12
-----------------	--	----

ÍNDICE DE FIGURAS

Figura 1:	Árbol de máxima verosimilitud que muestra la filogenia de GOLPH3 de mamíferos	9
Figura 2:	Alineamiento de las secuencias de aminoácidos de GOLPH3 de Humano (<i>Homo sapiens</i>), Tití (<i>Callithrix jacchus</i>), Cabra (<i>Capra hircus</i>) y Dingo (<i>Canis lupus dingo</i>).	11
Figura 3:	Las variantes de GOLPH3 fusionadas a GFP se localizan en el aparato de Golgi.....	13
Figura 4:	Análisis de estabilidad térmica de las variantes del segmento N-terminal de GOLPH3 en mamíferos.....	15
Figura 5:	Análisis de proteólisis parcial de las variantes del segmento N-terminal de GOLPH3.....	17
Figura 6:	Ensayo de separación de fases líquido-líquido de las variantes de GOLPH3.....	20

ÍNDICE DE ANEXOS

Anexo A:	Metodología Complementaria
Imagen suplementaria 1:	Subclonamiento del ADNc de las variantes de GOLPH3 de Tití, Dingo y Cabra desde el vector pGFP a pGST-Parallel-1.
Imagen suplementaria 2:	Purificación de las variantes de GOLPH3 recombinantes mediante cromatografía en tándem.
Tabla S1:	Números de acceso de los genes de GOLPH3 utilizados en este estudio

Resumen

La fosfoproteína 3 de Golgi (GOLPH3) es considerada la primera oncoproteína del aparato de Golgi, pero con funciones propuestas que aún no se comprenden a nivel molecular. La estructura de GOLPH3 está compuesta por una región N-terminal considerada no estructurada de una longitud de alrededor de 50 aminoácidos, de función desconocida y un dominio C-terminal plegado con alrededor de 200 aminoácidos de longitud que contiene un sitio de interacción con el fosfatidilinositol-4-fosfato (PI4P). El presente estudio tiene como objetivo analizar la historia evolutiva de GOLPH3 en mamíferos e identificar variaciones que den aproximaciones sobre las posibles funciones del segmento N-terminal y su capacidad para promover la separación de fase líquido-líquido (LLPS, por sus siglas en inglés). Esto se realizó mediante un análisis *in silico* (análisis de secuencias de ADN y análisis filogenéticos) y uno *in vitro* [fluorimetría de barrido diferencial (DSF), proteólisis parcial y separación de fases líquido-líquido (LLPS)]. Como resultado se encontró que la región N-terminal presenta mayor divergencia que el dominio C-terminal en todas las especies de mamíferos analizadas. Además, se encontró que las variaciones en el segmento N-terminal de GOLPH3 de las especies de mamíferos seleccionadas no afectan su localización celular, por lo tanto, su interacción con PI4P. Se seleccionaron variantes de GOLPH3 de tres especies de mamíferos que junto con la variante de humano se produjeron de manera recombinante para su comparación *in vitro*. El comportamiento de las proteínas recombinantes en ensayos de DSF, proteólisis parcial y separación de fases líquido-líquido mostró que las proteínas mantienen una estructura y comportamiento similar a la variante de humano. Por lo tanto, GOLPH3 de las distintas especies de mamíferos analizadas tiene la capacidad de sufrir LLPS *in vitro*, pero con algunas diferencias que dependen del segmento N-terminal.

Palabras claves: Mamíferos, aparato de Golgi, GOLPH3, LLPS.

Abstract

Golgi phosphoprotein 3 (GOLPH3) is considered the first Golgi oncoprotein, with proposed functions that are not yet understood at the molecular level. The structure of GOLPH3 is composed of about 50 amino acids long N-terminal region considered unstructured, with an unknown function, and a folded C-terminal domain of about 200 amino acids long that contains an interaction site for PI4P. The aim of this study is to analyze the evolutionary history of GOLPH3 in mammals and identify variations that justify the possible functions of the N-terminal segment and its ability to promote liquid-liquid phase separation (LLPS). This was done using *in silico* (DNA sequence and phylogenetic analyses) and *in vitro* (differential scanning fluorimetry, limited proteolysis, and liquid-liquid phase separation) assays. As a result, it was found that the N-terminal region presents greater divergence than the C-terminal domain in all analyzed mammalian species. In addition, the variations in the N-terminal segment of GOLPH3 of the selected mammalian species do not affect its cellular localization, therefore, its interaction with PI4P. GOLPH3 variants from three mammalian species that were recombinantly produced along with the human variant were selected for *in vitro* comparison. The behavior of the recombinant proteins in DSF, partial proteolysis and liquid-liquid phase separation assays showed that the proteins maintain a structure and behavior similar to the human variant. Therefore, analyzed mammalian GOLPH3 variants show the capacity to undergo LLPS *in vitro* with some differences that depend on the N-terminal region.

Keywords: Mammals, golgi apparatus, GOLPH3, LLPS

1. Introducción

El aparato de Golgi cumple funciones relacionadas con la distribución de lípidos y proteínas en la ruta secretora en el proceso conocido como tráfico de membranas ((Glick & Nakano, 2009; Mellman & Warren, 2000). Una de estas funciones es la modificación covalente de lípidos y proteínas que sintetiza la célula en la ruta secretora, a través, por ejemplo, de glicosilaciones llevadas a cabo por enzimas denominadas glicosiltransferasas (Stanley, 2011; Xiang et al., 2013). En el proceso de tráfico de membranas participan una serie de proteínas citosólicas que son reclutadas a las membranas del aparato de Golgi para ejecutar diversas actividades, entre ellas su regulación (Glick & Nakano, 2009). Sin embargo, muchas de las proteínas involucradas en estos procesos no han sido identificadas y de algunas de las que se tiene información aún se desconoce el mecanismo específico de su participación (Wu et al., 2000; Yarwood et al., 2020). Conocer la identidad de las moléculas que realizan este trabajo puede ayudar a comprender las funciones del aparato de Golgi en la vía secretora (Taylor et al., 1997; Yarwood et al., 2020). Por ejemplo, un propuesto regulador del tráfico de membranas es la proteína fosfoproteína 3 de Golgi (GOLPH3/GPP34/GMx33 α /MIDAS) (Nakashima-Kamimura et al., 2005; Rahajeng et al., 2019).

GOLPH3 es una proteína que interactúa principalmente con el dominio citosólico de la red *trans*-Golgi (TGN), y su asociación con la membrana de Golgi depende de su unión a fosfatidilinositol-4-fosfato (PI4P) (Bishé et al., 2012; Rahajeng et al., 2019). En humanos, se ha descrito que varios tipos de tumores sólidos contienen de manera frecuente la amplificación de la región cromosomal 5p13, sitio donde se encuentra el gen que codifica para la proteína GOLPH3 (Scott et al., 2009). También se ha demostrado que la sola sobreexpresión de GOLPH3 en células no tumorales es suficiente para transformarlas a un fenotipo tumoral, y por ello que se considera a GOLPH3 como la primera oncoproteína de Golgi (Scott et al., 2009; Sechi et al., 2015). Por esta razón, se ha propuesto que podría utilizarse GOLPH3 como biomarcador de progresión tumoral (Feng et al., 2015; Scott et al., 2009; Sechi et al., 2015). Actualmente, no se conoce cuál es la función específica de GOLPH3 en células de mamíferos, pero se ha planteado que participa en varios procesos celulares que incluyen la síntesis de glicolípidos, la localización de glicosiltransferasas del aparato de Golgi, la modulación de la masa mitocondrial y la formación de vesículas del *trans*-Golgi,

entre otras (Ali et al., 2012; Isaji et al., 2014; Nakashima-Kamimura et al., 2005; Rahajeng et al., 2019; Rizzo et al., 2021; Sechi et al., s/f).

Una manera de avanzar en conocer las funciones en las cuales está involucrada GOLPH3 son los análisis evolutivos. En particular, las variaciones de los aminoácidos podrían entregar información clave en la relación estructura/función (Lichtarge & Wilkins, 2010). GOLPH3 tiene un parálogo, GOLPH3L, con ortólogos 1:1 en la mayoría de las especies de mamíferos (Wu et al., 2000). La estructura de GOLPH3 está compuesta por una región N-terminal considerada no estructurada, intrínsecamente desordenada, de una longitud de entre 40-60 aminoácidos, de función desconocida, y un dominio C-terminal plegado con alrededor de 200 aminoácidos de longitud que contiene el sitio de interacción con PI4P (Opazo et al., 2021; Wood et al., 2009). Los estudios de divergencia de aminoácidos de GOLPH3 demuestran que la región N-terminal posee una mayor variación que su dominio C-terminal (Opazo et al., 2021). Al respecto, estudios de divergencia evolutiva muestran que las regiones intrínsecamente desordenadas presentan un menor grado de conservación de secuencias y coevolución que los dominios conservados (Chemes et al., 2012; Szalkowski & Anisimova, 2011).

En condiciones fisiológicas una región intrínsecamente desordenada (IDR) no tiene una estructura secundaria ordenada, pero esta condición es imperativa para su función biológica (Marín et al., 2012). En particular, se ha propuesto que las IDR de algunas proteínas facilitan la formación de compartimentos subcelulares sin membrana, los cuales son esenciales para la fisiología celular en condiciones normales y patológicas (Alberti et al., 2019; Yoshizawa et al., 2020).

Las características líquidas de los compartimentos sin membrana indican que su formación es mediante separación de fases líquido-líquido (LLPS) (Alberti et al., 2019). LLPS es un proceso fisicoquímico en donde las soluciones de proteínas, ácidos nucleicos u otros solutos se condensan en una fase densa que coexiste con una fase diluida que depende de la concentración de las moléculas que se encuentran en la solución (Alberti et al., 2019; Banani et al., 2017). Se ha reportado que la IDR de algunas proteínas permite la formación de los denominados condensados biomoleculares dando lugar a que se separen en fases por sí solas *in vitro* en condiciones óptimas (Banani et al., 2017). Una de las características de las IDR es que por su baja complejidad de secuencia promueven la separación de fases en las proteínas (Han et al., 2012; Kato et al., 2012; Nott et al., 2015). Estas secuencias se caracterizan por tener secciones repetitivas de distintos aminoácidos como glicina, serina, glutamina, asparagina, glutamato o aspartato. Por lo tanto, la separación de fases dependerá de la composición

de aminoácidos y los patrones de la secuencia de la región intrínsecamente desordenada (Banani et al., 2017). Resultados no publicados de nuestro laboratorio muestran que la proteína GOLPH3 de humano sería capaz de experimentar separación de fase líquido-líquido *in cellulo* e *in vitro*. Así, el presente estudio tiene como objetivo analizar la historia evolutiva de GOLPH3 en mamíferos e identificar variaciones que den aproximaciones sobre las posibles funciones del segmento N-terminal. Mediante un análisis de la divergencia evolutiva de GOLPH3 en mamíferos se identificaron variantes de secuencias de aminoácidos del segmento N-terminal que se seleccionaron para estudiar algunas de sus propiedades *in vitro*. Utilizando ensayos de estabilidad térmica y proteólisis parcial, se identificó que las variaciones en la secuencia de aminoácidos del segmento N-terminal afectan la estabilidad estructural de GOLPH3, e influyen en la separación de fases líquido-líquido *in vitro*.

2. Materiales y métodos

2.1. Secuencias de ADN y análisis filogenéticos.

Se realizó una búsqueda de secuencias nucleotídicas de GOLPH3 en genomas disponibles de mamíferos en las bases de datos Ensembl v.102 (Yates et al., 2020) y en el Centro Nacional de Información Biotecnológica (NCBI) (Sharma et al., 2018). La secuencia humana (*Homo sapiens*) de GOLPH3 se usó como referencia para la búsqueda en otras especies de mamíferos. En los casos en que las secuencias fueron incompletas, se anotaron manualmente mediante la extracción de piezas genómicas incluyendo los genes flanqueantes 5' y 3' del gen de GOLPH3. Después de esta extracción, se realizó una curación de las secuencias anotadas mediante una comparación de las secuencias de exones conocidos con el fragmento genómico y se usó el programa blast2seq versión 2.2.24 con parámetros por defecto (Tatusova & Madden, 1999). El muestreo taxonómico incluyó especies representativas de todos los grandes grupos de mamíferos *i.e.*, monotremas, marsupiales y placentados. La identidad de las especies, así como los números de acceso de las distintas secuencias se encuentran en Anexos; *Tabla S1*. Las secuencias de aminoácidos de mamíferos encontradas se alinearon utilizando la estrategia L-INS-i definida por MAFFT v.7 (Katoh & Standley, 2013). Las relaciones filogenéticas fueron estimadas utilizando el método de máxima verosimilitud. Para seleccionar el modelo de sustitución de aminoácidos que mejor se ajusta (Blosum62+F+I+G4) se usó el programa IQ-Tree y se obtuvo el árbol de máxima verosimilitud (Trifinopoulos et al., 2016). Se evaluó el soporte para los nodos

con 1000 pseudoréplicas utilizando la rutina ultrarrápida (Hoang et al., 2018). Por último, se seleccionaron las especies que tuvieron una variación en sus aminoácidos del segmento N-terminal de GOLPH3 juzgada como interesante, que pueda influenciar en su estructura, estabilidad o movilidad.

2.2. Construcciones de ADNc recombinante.

Para la generación de las construcciones de ADNc recombinante se tomó la decisión de introducir sustituciones de cisteína por serina en los sitios C84, C108, C122 y C280. Un análisis empírico demostró que estas sustituciones reducen el número de cisteínas libres que podrían producir puentes de disulfuro artefactuales (Arolas et al., 2006). A las construcciones con estos cambios se les dio la nomenclatura C4S. Una vez seleccionadas las variantes de interés de GOLPH3, se procedió con la adquisición comercial del respectivo ADNc sintetizado, para luego ser clonado entre los sitios *EcoRI* y *XhoI* del plásmido pcDNA3.1(+)-N-eGFP (GenScript; Piscataway, NJ). Los plásmidos resultantes codifican cada variante de GOLPH3 fusionada a la proteína verde fluorescente (GFP), dejando a esta última en el amino terminal de la construcción. Estas construcciones fueron denominadas pGFP-GOLPH3-C4S-Xx, en el que Xx corresponde a las iniciales de la especie mamífera respectiva seleccionada, y permiten la expresión de las proteínas de fusión en células de mamíferos. Para la generación de una construcción que permita la expresión de las variantes de GOLPH3 en bacterias, se efectuó un subclonamiento desde cada vector pGFP-GOLPH3-C4S-Xx al vector pGST-Parallel-1 (Sheffield et al., 1999) mediante el uso de los sitios de restricción *EcoRI* y *XhoI*. Los plásmidos resultantes codifican cada variante de GOLPH3 fusionada a la proteína Glutación S-transferasa (GST), dejando a esta última en el amino terminal de la construcción. Estas construcciones fueron denominadas pGST-GOLPH3-C4S-Xx, en el que Xx corresponde a las iniciales de la especie mamífera respectiva seleccionada, y permiten la expresión de las proteínas de fusión en el citoplasma de las bacterias. La secuencia de nucleótidos de todas las construcciones recombinantes se confirmó mediante secuenciación en el core facility AUSTRAL-omics de la Universidad Austral de Chile (<https://australomics.cl/>; Ver metodología extendida).

2.3. Expresión y purificación de la proteína GOLPH3 de las especies de mamíferos seleccionadas

Las variantes del segmento N-terminal de GOLPH3 recombinante etiquetados en el N-terminal con GST, seguida de un sitio de escisión de la proteasa TEV, se expresaron y se purificaron según el protocolo de Ross (Ross et al., 2014a), con algunas modificaciones. Bacterias *E. coli*, cepa B834(DE3)pLysS (Novagen, Madison, WI), se transformaron con los respectivos plásmidos pGST-GOLPH3-C4S-Xx y la expresión de cada proteína recombinante se hizo en cultivos líquidos de 1.5 L de medio LB (Tryptona 10 g/L, extracto de levadura 5 g/L, NaCl 10 g/L), induciendo con IPTG 0.35 mM a 25 °C durante 16 horas a 200 rpm de agitación en una incubadora Innova 44 con órbita de 2.5 cm. Después de centrifugar los cultivos a 3000 x *g* por 40 min a 4 °C, los sedimentos de bacterias se resuspendieron en un tampón de homogeneización (Tris-HCl 50 mM, NaCl 0.5 M, EDTA 5 mM, β-mercaptoetanol 5 mM, azida de sodio 0.02% y fluoruro de fenilmetilsulfonilo [PMSF] 2 mM, pH 8.0) y se procedió a la lisis bacteriana mediante sonicación usando un sonicador digital por ultrasonido Misonix Sonicator S-400 durante tres periodos de sonicación (2 min, 2 min y 1 min) con dos tiempos de descanso entre periodos de 5 min para evitar el calentamiento de la muestra. Después de centrifugar el homogeneizado a 1833315 x *g* por 90 min a 4 °C, el sobrenadante se purificó en Glutación-Sefarosa 4B (GE Healthcare) equilibrada en tampón (Tris 50 mM, 150 mM NaCl, 0.5 M β-mercaptoetanol, EDTA 5 mM, azida de sodio 0.02%, pH 8.0). La proteína de fusión unida a Glutación-Sefarosa 4B fue eluída con tampón de elución (Tris 50 mM, NaCl 150 mM, EDTA 5 mM, L-glutación reducido 20 mM, azida de sodio 0.02%, pH 8.0) colectándose 5 fracciones de 2.3 mL. La proteína eluida se concentró hasta 4-5 mL y se sometió a cromatografía en una columna HiLoad 16/600 Superdex 200 pg (GE Healthcare) para la eliminación de impurezas de alto y bajo peso molecular. Las fracciones conteniendo GST-GOLPH3-Xx se juntaron (entre 9 y 12 mL) y la proteína GST de cada proteína de fusión fue escindida por incubación con 75 µL de la proteasa TEV (fusionada a poli-histidina), por 16 horas a 20 °C. La mezcla de digestión se concentró hasta 4-5 mL utilizando un filtro VivaSpin de 30 kDa y se sometió a cromatografía en la columna HiLoad 16/600 Superdex 200 pg. Las fracciones conteniendo GOLPH3-C4S-Xx se juntaron (entre 9 y 12 mL) para ser sometidas a cromatografía secuencial en Glutación-Sefarosa 4B y Ni-Sefarosa 6 FF (GE Healthcare) para remover de la muestra a la proteína GST escindida y la proteasa TEV, respectivamente. Las proteínas purificadas se mantuvieron en un tampón de almacenamiento (Tris HCl 50 mM, NaCl 150 mM, DTT 1 mM y EDTA 0.5 mM, pH 8.0) y se guardaron en alícuotas a -80 °C. (Para más detalle ver en metodología extendida).

2.4. Transfección y microscopía de inmunofluorescencia

Se sembraron 2.5×10^4 células MDA-MB-231 en cubreobjetos de vidrio redondos de 12 mm en una placa de 24 pocillos y se mantuvieron en medio de cultivo completo. Después de 24 h, las células se sometieron a transfecciones transitorias para expresar GFP-GOLPH3-C4S-Xx utilizando Lipofectamine 2000 (Life Technologies), de acuerdo con las instrucciones del fabricante. Después de 16 h, las células transfectadas se fijaron en paraformaldehído al 4 %, se permeabilizaron con 0.3% de Tritón X-100 por 5 min bloqueando en solución de gelatina al 0.2% por 15-30 min y se inmunomarcaron con anticuerpo monoclonal de ratón contra GM130 (clon 35/GM130 BD Biosciences). El anticuerpo secundario que se usó fue IgG anti-ratón de burro conjugado con Alexa-594 (Life Technologies) por 30 min a 37 °C. Los núcleos se tiñeron con 4',6-diamidino-2-fenilindol (DAPI) (Life Technologies). Las imágenes de microscopía de fluorescencia se adquirieron con un microscopio AxioObserver.D1 equipado con un objetivo de inmersión en aceite PlanApo 63x (NA 1.4) y una cámara digital AxioCam MRm (Carl Zeiss). Para preparar las figuras, las imágenes se procesaron con el software Image J (versión 1.44o; Wayne Rasband, NIH, <http://imagej.nih.gov>) y el software Adobe Photoshop CS3 (Adobe Systems, Mountain View, CA).

2.5. Ensayo de proteólisis parcial.

Este ensayo identifica sitios de mayor flexibilidad o despliegue local de una cadena polipeptídica (Fontana et al., 2004). Las variantes de N-terminal de GOLPH3 recombinantes se sometieron a proteólisis parcial. Se prepararon siete tubos de 1.5 mL con 7 µg de proteína recombinante, 1 µg de proteinasa K (0.001 M), el cual se llevó a un volumen final de 30 µL con tampón con glicerol al 10% (Tris-HCl 25 mM, NaCl 150 mM, glicerol 10%, β-mercaptoetanol 5 mM, pH 8.0). Cada mezcla de reacción se incubó a 25 °C por 7 tiempos diferentes: 0 (sin proteinasa K), 5, 15, 30, 60, 120 o 180 min. Se detuvo la reacción al someter la muestra a 95 °C después de mezclar con tampón de carga/muestra para electroforesis de proteínas (sulfato de dodecilo de litio (LDS), SERVA Blue G250 y rojo de fenol; Invitrogen 4X). Las muestras se analizaron mediante electroforesis en geles de poliacrilamida en condiciones denaturantes (SDS-PAGE), en geles en gradientes de 4 al 12% (NuPAGE Bis-Tris; Life Technologies). Para visualizar el resultado, los geles se tiñeron con solución que contiene azul brillante de Coomassie (SimplyBlue, ThermoFisher).

2.6. Ensayo de estabilidad térmica por fluorimetría de escaneo diferencial (DSF).

La estabilidad térmica de las variantes del segmento N-terminal de GOLPH3 versus la variante humana se evaluó mediante desnaturalización térmica en presencia de SYPRO Orange. Se usó 5 µg de cada proteína con 2 µL de una disolución 1:25 (v/v) de SYPRO Orange 20X (Sigma-Aldrich, St. Louis, MO) en un volumen final de 20 µL, y se realizaron tres repeticiones. Se determinó el cambio de fluorescencia en un rango de temperatura de 25-90 °C. El aumento de la temperatura fue de 1 °C por minuto. Las mediciones se realizaron en un termociclador de PCR cuantitativo en tiempo real modelo Rotor-Gene Q (QIAGEN). Para la lectura de fluorescencia se utilizó un filtro que permite excitación a 470 nm y emisión a 510 nm (Alexandrov et al., 2008). Los valores de T_m se determinaron mediante un cálculo de la primera derivada negativa de los datos sin procesar. Al inicio del ensayo las proteínas se encuentran en un estado conformacional en el que los aminoácidos hidrofóbicos se encuentran al interior de la molécula y los hidrofílicos al exterior. A medida que la temperatura aumenta, las proteínas adquieren un estado dinámico en el que los residuos hidrofóbicos se exponen, en este punto el SYPRO Orange se une a estas regiones emitiendo fluorescencia. El valor de T_m corresponde a la temperatura media entre el comienzo de la exposición de regiones hidrofóbicas de las proteínas, o comienzo de emisión de Sypro Orange, y el valor máximo de emisión de fluorescencia.

2.7. Ensayo de separación de fases líquido-líquido.

Se usó PEG DWBlock suite (NeXtal Biotechnologies; N° de catálogo 130904), un kit de 96 soluciones que contienen distintos tipos de polietilenglicol (PEG) como precipitante y una serie de otros componentes como tampones y sales que se utilizan normalmente para encontrar una o más condiciones que permiten la cristalización de una proteína. Se evaluó el comportamiento de GOLPH3-WT y GOLPH3-C4S de humano en las diferentes condiciones que ofrece el kit y se seleccionó una serie de condiciones que permitió una clara visualización de la separación de fases líquido-líquido (sulfato de magnesio 0.2 M, sulfato de magnesio 0.2 y fosfato de disodio 0.2 M; todas las condiciones con 20% de PEG 3500). Una vez seleccionadas las condiciones, se realizó una matriz adecuada para comparar las variantes de GOLPH3 recombinante con GOLPH3 humano de tipo salvaje y la variante C4S de humano. Se analizó también las diferentes sales seleccionadas en presencia de PEG 4000 este agente aglomerante fue

seleccionado debido a que presentaba gotas más grandes en comparación con PEG 3500. Las variantes de la proteína GOLPH3 fueron diluidas en un tampón (Tris-HCl 20 mM, NaCl 50 mM, DTT 1 mM, pH 8.0) para obtener las concentraciones de 1, 2, 4 y 8 μ M. Seguido, se prepararon soluciones que contenían agente de aglomeración, en este caso PEG 4000 al 20%, con distintas concentraciones de sulfato de magnesio (0.05, 0.1, 0.2 y 0.4 M). En una placa de 96 pocillos Intelli-Plate 96-2 Original, se dispensaron 80 μ L de cada solución con PEG 4000 al 20% conteniendo las distintas concentraciones de sulfato de magnesio. Utilizando el dispensador de nanovolúmenes para placas de 96 pocillos, modelo Mosquito LV (TTP-Labtech), se dispensaron 0.2 μ L de cada dilución de las variantes de la proteína GOLPH3 mezclándose luego con 0.2 μ L de las soluciones con PEG 4000 al 20% y sulfato de magnesio contenidas en cada respectivo pocillo reservorio. Las placas de 96 pocillos se sellaron con un film plástico transparente y se incubaron en incubadora marca Velp Scientifica, modelo FOC 225E, a 18 °C. La aparición de separación de fases líquido-líquido en las gotas fue inspeccionada visualmente utilizando una lupa (Nikon SMZ645). Para documentar el resultado, se tomaron imágenes de las gotas a las 2, 16 y 24 horas utilizando una cámara AxioCam acoplada a una lupa (ZEISS Axio Imager 2 Mat) utilizando el programa de adquisición de imágenes Zen core y Zen starter (Zeiss).

2.8. Análisis estadísticos.

Todos los experimentos y mediciones se realizaron al menos tres veces. El análisis de datos se realizó utilizando Microsoft Excel para MacBook Pro 2011 (Microsoft Corporation), los gráficos se confeccionaron usando R Studio. La significación estadística se determinó mediante la prueba *t-Student* de una cola.

3. Resultados

3.1. El segmento N-terminal de la oncoproteína GOLPH3 es más variable en comparación a su dominio C-terminal.

La recopilación de las secuencias y los análisis filogenéticos de GOLPH3 en mamíferos muestran que este gen es conservado. Los árboles filogenéticos son una estrategia para comprender las relaciones de homología, la historia duplicativa y los modos de evolución del gen de estudio. En este caso, el árbol de GOLPH3 en mamíferos no se desvió significativamente de las hipótesis filogenéticas propuestas por O'Leary *et al.*, 2013 (O'Leary *et al.*, 2013). Esto quiere decir que al ser un árbol de máxima verosimilitud realizado con un solo gen, presenta una filogenia con los tres grandes grupos de mamíferos, dando como indicio que GOLPH3 está presente en el ancestro común de mamíferos como copia única (**Figura 1**).

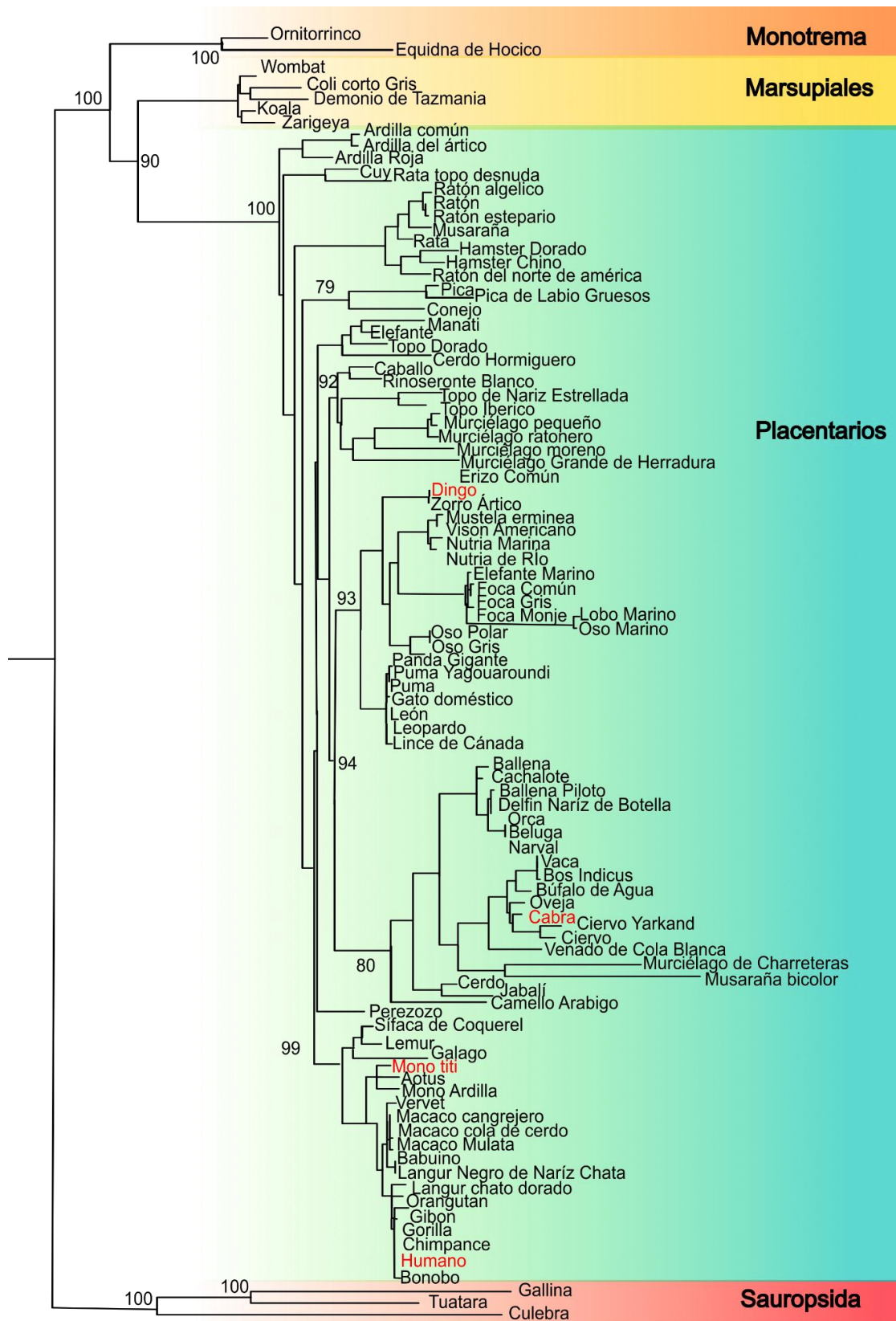


Figura 1. Árbol de máxima verosimilitud que muestra la filogenia de GOLPH3 de mamíferos.

Realizado por: Galarza, 2022

El alineamiento de las secuencias de mamíferos mostró que la región N-terminal presenta mayor divergencia que el dominio C-terminal en todas las especies de mamíferos analizadas. La estimación de la divergencia de la secuencia de aminoácidos para el segmento N-terminal y C-terminal por separado fue de 2.96% y de 2.28% respectivamente. Considerando que la región N-terminal posee de 40 a 60 aminoácidos, la divergencia es de alrededor del 56.48% de toda la proteína, mientras que el dominio C-terminal de 200 aminoácidos aproximadamente presenta una divergencia del 43.51%. El árbol de máxima verosimilitud separó a los tres grandes grupos de mamíferos (**Figura 1**), la divergencia para el segmento N-terminal fue de 2.48% para placentarios, 0.80% para marsupiales y 9.00% para monotremas. En cambio, para el segmento C-terminal fue de 2.23% para placentarios, 0.00% para monotremas y marsupiales. Estos resultados son consistentes con el hallazgo de Opazo, 2021 (Opazo et al., 2021) que menciona que las diferencias en la divergencia de la secuencia de aminoácidos se encuentran en su mayoría en el segmento N-terminal de la proteína y esto se conserva en todas las especies de mamíferos. Esto permite concluir que las variaciones de la región intrínsecamente desordenada de GOLPH3 surgieron por restricciones estructurales y funcionales durante la evolución de los mamíferos. Las variaciones en las secuencias pueden dar información sobre la estructura y función de las proteínas. Por ello, el alineamiento permitió identificar variaciones para los análisis subsecuentes, en donde se seleccionaron tres especies de mamíferos: el mono tití (*Callithrix jacchus*), Dingo (*Canis lupus dingo*) y Cabra (*Capra hircus*). El alineamiento de la secuencia de aminoácidos de GOLPH3 de estas especies con la secuencia de aminoácidos de GOLPH3 humano se muestra en la (**Figura 2**).

Figura 2. Alineamiento de las secuencias de aminoácidos de GOLPH3 de humano (*Homo sapiens*), tití (*Callithrix jacchus*), cabra (*Capra hircus*) y dingo (*Canis lupus dingo*).
Realizado por: Galarza, 2022

12

región N-terminal de las variantes de estas especies tienen variaciones que pueden influenciar en la estabilidad, función, localización o estructura de la proteína. Para estudiar estas variaciones se procedió a la generación de las proteínas recombinantes de GOLPH3, con el segmento N-terminal de cada especie (humano, tití, dingo y cabra) y el dominio C-terminal de GOLPH3-C4S-Humano.

Tabla 1. Punto isoeléctrico de las variantes de GOLPH3

Variante de GOLPH3	² pI (N-terminal) ³ GAMDPEF-1-50	pI (C-terminal) 51-298	pI GAMDPEF- GOLPH3
GOLPH3-WT-Humano	4.43	7.99	5.72
GOLPH3- ¹ C4S-Humano	4.43	7.99	5.72
GOLPH3-C4S-Tití	4.43	7.99	5.72
GOLPH3-C4S-Dingo	4.43	7.99	5.72
GOLPH3-C4S-Cabra	4.34	7.99	5.58

¹C4S. Corresponde a la denominación de las sustituciones de residuos de cisteínas por serinas que se diseñaron para la producción de proteínas recombinantes.

²pI. Punto isoeléctrico de la proteína.

³GAMDPEF-1-50. Corresponde a la secuencia GAMDPEF (resto de poli-histidina del sitio de corte con TEV) y los números mencionan los aminoácidos de la secuencia de GOLPH3.

Realizado por: Galarza, 2022

Una vez que se escogieron las secuencias de las especies a analizar, se procedió al estudio de las variantes de GOLPH3 comparándolas con la de humano. Primero, se analizó la distribución intracelular de las variantes por medio de microscopía de fluorescencia. Para esto, células HeLa (derivadas de carcinoma cervicouterino humano) fueron transfectadas con los respectivos plásmidos pGFP-GOLPH3. Después de 16 horas las células fueron procesadas para realizar una inmunofluorescencia indirecta. La distribución de cada variante de GOLPH3 fusionada a la proteína GFP se comparó con la distribución de la proteína de matriz de Golgi 130 (GM130). Se observó que todas las variantes de GOLPH3 colocalizan con GM130, indicando que todas se localizan en el aparato de Golgi, sugiriendo que las variaciones en el segmento N-terminal de GOLPH3 de las especies de mamífero seleccionadas no afectan su localización, por lo tanto, tampoco su interacción con PI4P (**Figura 3**).

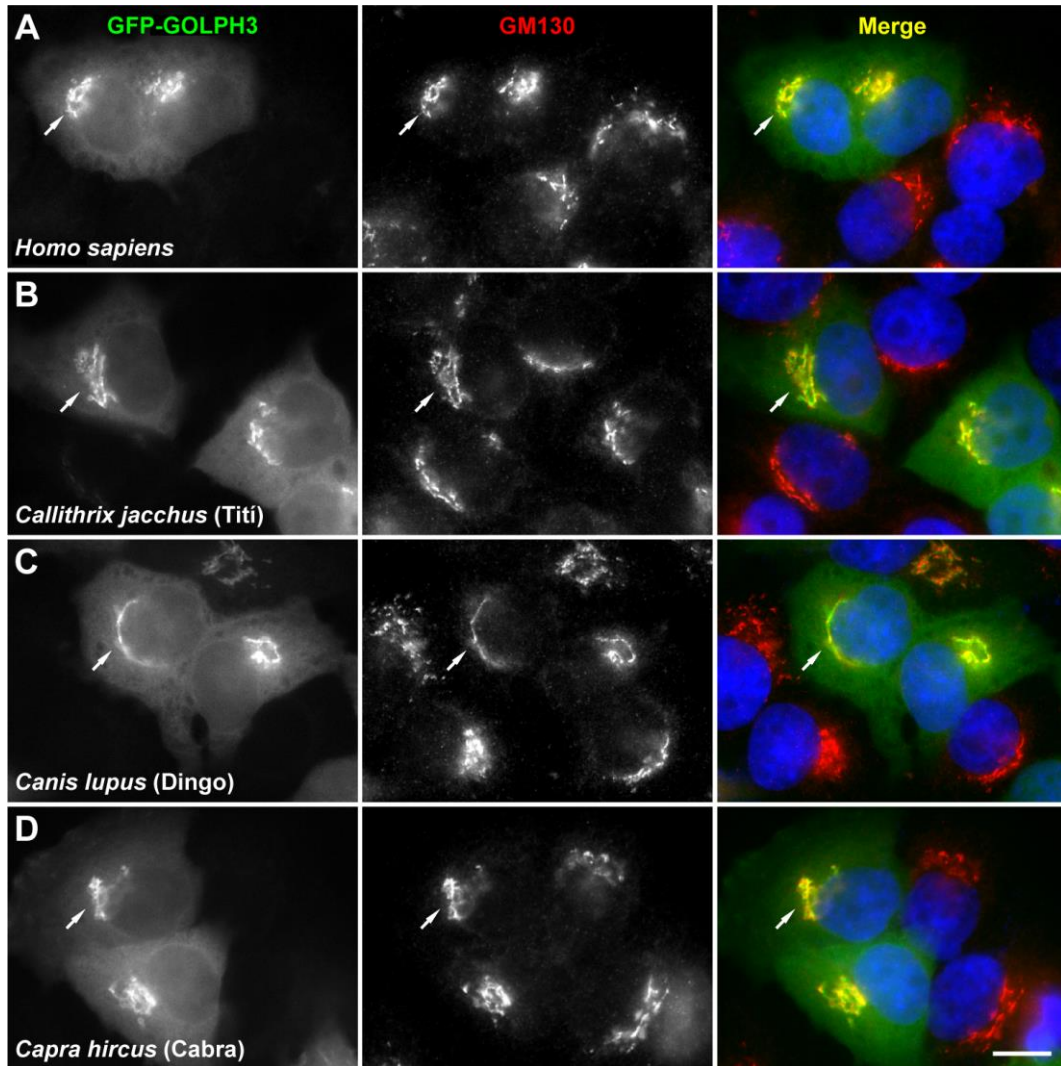


Figura 3. Las variantes de GOLPH3 fusionadas a GFP se localizan en el aparato de Golgi.

Realizado por: Galarza, Luchsinger, 2022

3.2. Las variantes de GOLPH3 recombinantes muestran leves diferencias en su estabilidad térmica.

Para determinar si existen diferencias entre las variantes de GOLPH3 en su capacidad de experimentar LLPS, éstas fueron producidas de forma recombinante en bacterias para su análisis *in vitro*. Para la producción y purificación de las variantes de GOLPH3 recombinantes, primero se realizó un subclonamiento del segmento de ADNc desde los respectivos plásmidos pGFP-GOLPH3 al vector pGST-Parallel-1 (Información suplementaria. S1). Este vector de expresión permite expresar en bacterias las proteínas de interés como proteínas de fusión con la proteína Glutación S-transferasa (GST). El subclonamiento se verificó mediante secuenciación del ADN en el core facility AUSTRAL-omics de la Universidad Austral de Chile. Una vez que se obtuvo estas construcciones se realizó la expresión y purificación de las proteínas recombinantes mediante el uso de varios pasos cromatográficos en tándem (Ver metodología extendida 2.3.1). Las condiciones cromatográficas fueron las mismas para todas las variantes de GOLPH3, sin embargo, las cantidades de proteína obtenidas para cada variante fueron diferentes. (Información suplementaria. S2). Como es el caso de la variante Tití de la que se obtuvo un menor rendimiento de la purificación, pero fue suficiente para realizar los siguientes análisis.

Antes de comparar la capacidad de experimentar LLPS de las variantes recombinantes, primero se analizó su integridad estructural. El primer análisis correspondió a la determinación de la estabilidad térmica mediante fluorimetría de barrido diferencial (DSF). Se compararon las variantes de GOLPH3-C4S de tití, dingo, cabra y humano entre ellas y con GOLPH3-WT-Humano, es decir, con la forma nativa de la proteína GOLPH3 (**Figura 4**). La medición de las T_m implica el seguimiento del desplegamiento de la proteína producido por la desnaturalización térmica. Se encontró que la T_m de GOLPH3-C4S-Cabra fue de 45.3 ± 0.3 °C y GOLPH3-C4S-Titi fue 45.5 ± 0.2 °C, valores que son significativamente menores que la T_m de GOLPH3-C4S-Humano, el cual fue 45.9 ± 0.2 °C ($p < 0.007$; $p < 0.045$; $n = 2$; **Figura 4**). Esto sugiere que las variaciones en el N-terminal de GOLPH3-C4S-Titi (S36G) y GOLPH3-C4S-Cabra (A33G, S35N, S36D) hacen un poco menos estable a estas proteínas, en comparación con la de humano, siendo la de Tití un poco más estable que la de cabra. El valor de T_m de GOLPH3-C4S-Dingo fue de 45.0 ± 0.4 °C ($p < 0.018$; $n = 2$; **Figura 4**), significativamente más bajo que todas las otras variantes, indicando que las variaciones en el N-terminal de GOLPH3-C4S-Dingo (N21S, A28V, A33G, S35/36G) hacen que la proteína presente

una mayor inestabilidad térmica en comparación a las otras variantes analizadas. Las variantes analizadas poseen el cambio C4S en su estructura debido a que estas sustituciones reducen el número de cisteínas libres que podrían producir puentes de disulfuro artefactuales que darían lugar a una proteína mal plegada no funcional en los estudios tanto *in vitro* e *in cellulo* (Arolas et al., 2006). Para verificar la influencia de estas sustituciones, se realizó el ensayo de DSF con la proteína GOLPH3-WT-Humano y se encontró que ésta posee un T_m de 49.9 ± 0.6 °C (**Figura 4**), un valor alto en comparación a GOLPH3-C4S-Humano. Sin embargo, el comportamiento de la proteína GOLPH3-WT-Humano a lo largo del ensayo demuestra que las proteínas recombinantes que se sintetizaron mantienen una estructura estable similar a la nativa y por lo tanto, no deberían haber perdido sus propiedades funcionales.

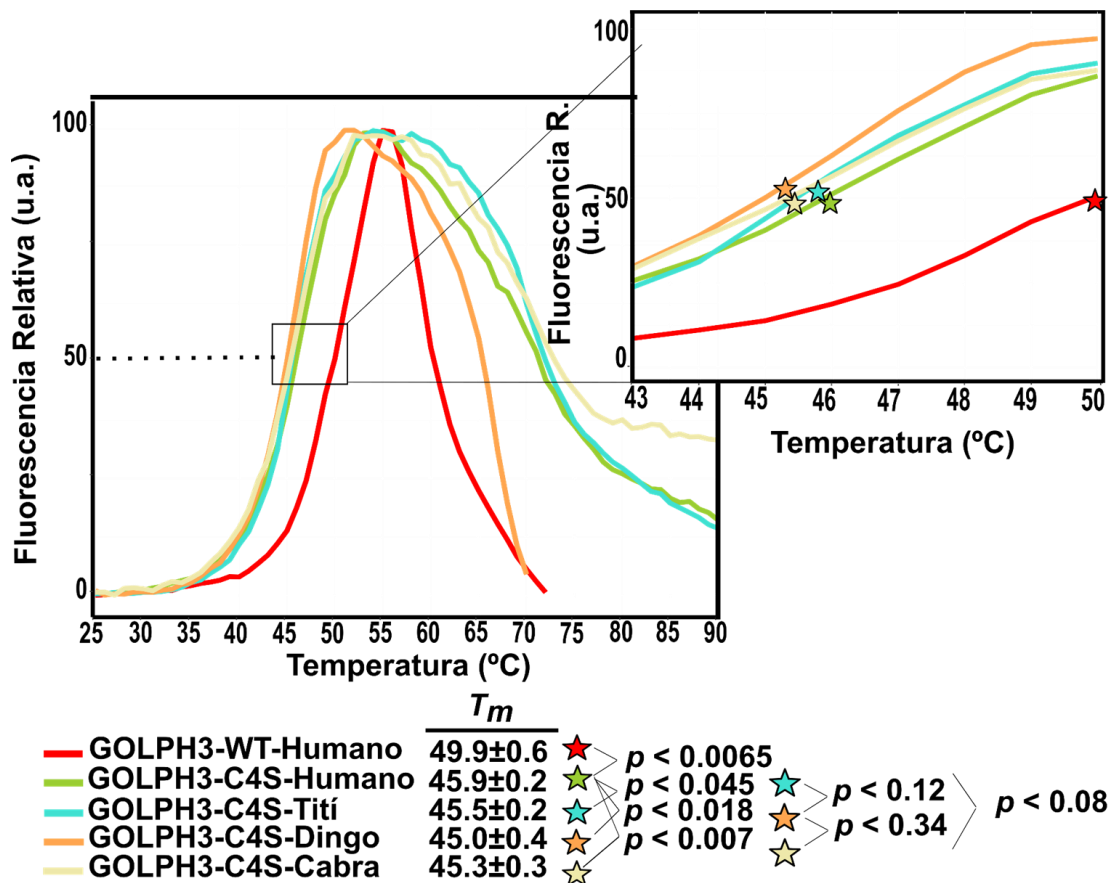


Figura 4. Análisis de estabilidad térmica de las variantes del segmento N-terminal de GOLPH3 en mamíferos.

Realizado por: Galarza, 2022

3.3. La estructura de las variantes del segmento N-terminal de GOLPH3 es similar al GOLPH3 del tipo salvaje

Para continuar evaluando la integridad estructural de las variantes de GOLPH3 recombinantes, se realizó un ensayo de proteólisis parcial. La proteólisis en presencia de proteinasa K genera intermediarios estables que se originan por digestión de sitios accesibles y ausencia de digestión en regiones compactas en donde la proteasa no tiene acceso. Se realizó una cinética de exposición a la proteinasa K por hasta 3 horas a 25 °C, la que dio como resultado 3 fragmentos principales y 2 fragmentos difusos para GOLPH3-WT-Humano. Las masas moleculares de los fragmentos principales fueron aproximadamente de 7.2, 10.6 y 16.4 kDa, donde probablemente corresponden a los fragmentos que se generan por digestión de las zonas expuestas, y la de los fragmentos difusos de 32.5 kDa y 14.1 kDa que pueden corresponder a intermediarios de los fragmentos que se tiñen de manera más intensa (**Figura 5A**). Con las masas moleculares aproximadas de los fragmentos principales, se realizó un análisis para determinar las posibles zonas de digestión de la proteinasa K. En el caso de GOLPH3-WT-Humano se hizo una estimación del sitio donde pudo haber sido el corte de la proteinasa K. Esto se realizó mediante un análisis de los pesos moleculares de los fragmentos observados en el último carril del gel, utilizando el software ImageJ (**Figura 5A; Carril 8**). La elección de los posibles fragmentos se realizó mediante una aproximación empírica en base a la estructura cristalina de GOLPH3 que permite visualizar posibles regiones expuestas donde pudo haber digestión por la proteinasa K, y el cálculo del peso molecular de cada posible fragmento obtenido, utilizando el servidor ProtParam (<https://web.expasy.org/protparam/>; **Figura 5F, 5G**). Los posibles sitios de corte permiten predecir tres posibles fragmentos. El primer fragmento de 7.1 kDa podría tener 60 aminoácidos. Este segmento se podría generar por el primer posible sitio de corte por la proteinasa K (R59 o L60). Como el segmento N-terminal de GOLPH3 es desordenado no ha sido posible obtener su estructura cristalina. Sin embargo, al ser una sección muy expuesta se asume que la proteinasa K podría digerir primero este segmento (**Figura 5F, 5H**). El segundo segmento de 16.4 kDa, posiblemente con 144 aminoácidos, y el tercer fragmento de 10.9 kDa, posiblemente con 94 aminoácidos, podrían haber sido generados por corte de la proteinasa K en el segundo posible sitio de corte (P203 o L204) (**Figura 5G, 5H**). La digestión parcial de GOLPH3-C4S-Humano mostró un patrón muy similar al de GOLPH3-WT-Humano (**Figura 5B**), sugiriendo que las sustituciones C4S no afectan de manera apreciable la estructura de la proteína y sería apropiado utilizarla para estudios de separación de fase líquido-líquido *in vitro*. La digestión parcial de las variantes del segmento N-terminal de GOLPH3 muestra leves

diferencias respecto a las variantes humanas. GOLPH3-C4S-Titi y GOLPH3-C4S-Dingo presentan una mayor resistencia a la digestión en los primeros 5 min (**Figura 5C y 5D**). En el caso de GOLPH3-C4S-Cabra se observa que a los primeros 5 min de digestión con proteinasa K hay una fragmentación similar a GOLPH3-C4S-Humano (**Figura 5E**). Sin embargo, a las 3 horas de digestión, GOLPH3-C4S-Cabra presentó una degradación completa del fragmento difuso de 32.5 kDa, a diferencia de las otras variantes (**Figura 5E**). Todas las variantes de GOLPH3 analizadas presentan patrones de digestión muy similares a las 3 horas de incubación con la proteinasa K, indicando que las variaciones del segmento N-terminal no causan cambios conformacionales que produzcan cambios en la sensibilidad a la proteinasa K.

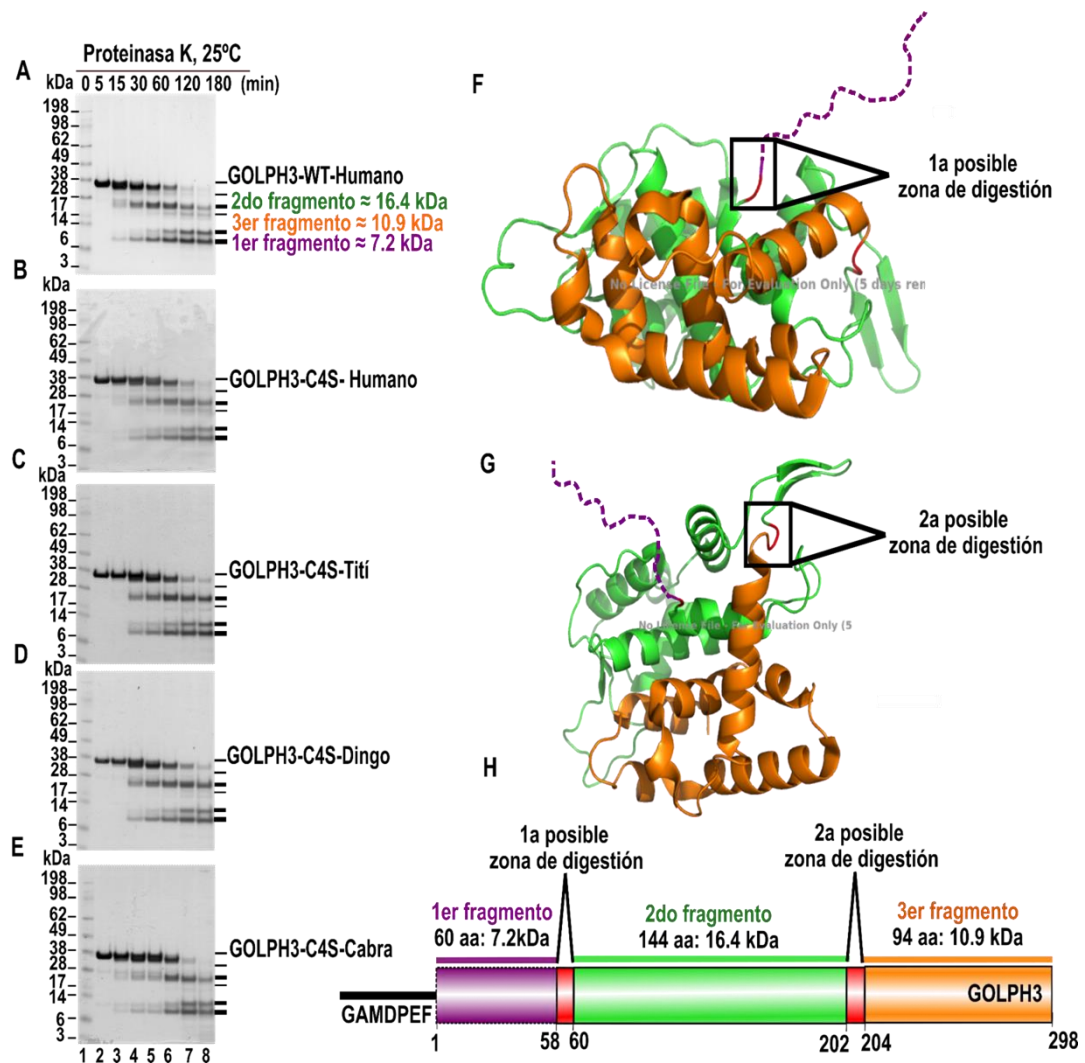


Figura 5. Análisis de proteólisis parcial de las variantes del segmento N-terminal de GOLPH3.

Realizado por: Galarza, 2022

3.4. Las variaciones del segmento N-terminal de GOLPH3 favorecen su partición por separación de fases líquido-líquido.

GOLPH3 tiene un dominio C-terminal conservado y con una estructura cristalina definida (Wood et al., 2009). Además, su segmento N-terminal posee una región intrínsecamente desordenada (IDR), que por lo tanto carecería de estructura tridimensional definida en condiciones fisiológicas (Opazo et al., 2021). Se ha descrito que las IDR impulsan las transiciones de fase (Han et al., 2012; Kato et al., 2012; Nott et al., 2015). El resultado combinado de los ensayos de fluorimetría de barrido diferencial (DSF) y proteólisis parcial permitió concluir que las proteínas recombinantes purificadas presentan estabilidad estructural muy similar y por lo tanto pueden ser usadas para ser comparadas en su capacidad de experimentar LLPS. Para poner a prueba que GOLPH3 posee IDR y dominios de baja complejidad se utilizó la herramienta de investigación de arquitectura modular simple (SMART) (Mészáros et al., 2018) y IUPred2 (Letunic & Bork, 2018), respectivamente. La herramienta SMART predijo dos dominios de baja complejidad (**residuos 28-34 y 37-56; Figura 6A**), mientras que IUPred2 reveló que todo el segmento N-terminal es una IDR (**residuos 1-51; Figura 6B**).

Para evaluar la capacidad de las variantes de GOLPH3 de experimentar LLPS *in vitro*, se usaron concentraciones crecientes de las variantes recombinantes de GOLPH3 que fueron mezcladas con una solución que contenía polietilenglicol 4000 (PEG 4000) al 20% y variación de sulfato de magnesio (una sal de alta fuerza iónica) a concentraciones de 0.4, 0.2 y 0.1 mM (imágenes con 0.1 mM no mostradas), que permite la partición de GOLPH3-WT-Humano por separación de fases líquido-líquido *in vitro* (resultados no publicados; **Figura 6C**). De manera interesante, se observó también la partición por separación de fases líquido-líquido de todas las otras variantes de GOLPH3, aunque con diferencias notables (**Figura 6, D-G**). Se realizaron diagramas de fase empíricos que proporcionaron información sobre la dinámica de separación de fases de las proteínas en las condiciones antes mencionadas (**Figura 6 H-L**). Para GOLPH3-WT-Humano la LLPS se observó a una concentración de proteína desde 1 μ M en presencia de concentraciones mayores a 0.2 M de MgSO₄ (**Figura 6 C-H**), a diferencia de GOLPH3-C4S-Humano, que no presentó LLPS en estas condiciones (**Figura 6 D-I**). Esto indica que las sustituciones de cisteína por serina influyen en la capacidad de experimentar LLPS *in vitro*. Por otro lado, observamos que las gotas generadas por LLPS de GOLPH3-C4S-Titi presentaron un tamaño mayor a una concentración de 0.2 M de MgSO₄, y el tamaño disminuyó a medida que se elevó la concentración de la sal. Para la LLPS de GOLPH3-C4S-Titi se requirió concentraciones de proteína menores a

2 μM para la generación de gotas grandes (**Figura 6E-J**). El único cambio en GOLPH3-C4S-Titi es la serina 36 por glicina que debería mejorar las interacciones hidrofóbicas, a pesar de tener el cambio de C4S en su estructura, debiendo ser capaz de experimentar LLPS más fácilmente que GOLPH3-C4S-Humano. Sucedió lo contrario con las variaciones en el N-terminal de GOLPH3-C4S-Dingo cuyos cambios de N21S, A28V, A33G, S35/36G produjeron gotas grandes a mayor concentración de sal y menor concentración de proteína (**Figura 6F-K**). La variación de asparagina por una serina permite que esta región cambie de un estado hidrofílico a uno menos hidrofílico, generando una mayor estabilidad en las gotas a pesar de la baja concentración de proteína (0.5 μM). Además, el cambio de alanina por valina produce aún más hidrofobicidad a la proteína. Al poseer una inserción de siete glicinas seguidas entre los aminoácidos 30-36, se espera que la proteína recombinante de Dingo presente un N-terminal con mayor movilidad. Finalmente, en el caso de GOLPH3-C4S-Cabra las variaciones de esta proteína (A33G, S35N, S36D) permitieron la formación de gotas en todas las concentraciones de proteína (0.5 μM hasta 4 μM), y en presencia de las dos concentraciones de sulfato de magnesio (0.2 y 0.4 M), pero el tamaño de las gotas fue en su mayoría similar (**Figura 6G-L**). Las variaciones de serina por ácido aspártico hizo que cuatro aminoácidos con carga negativa se encuentren de manera flanqueante (D36-E37-D38-D39), lo que indica que los aminoácidos que tienen regiones cargadas confirieron estabilidad a las gotas. Estos resultados muestran que el segmento N-terminal promueve la separación de fases líquido-líquido *in vitro* en presencia de agentes de aglomeración (PEG 4000) y una sal de alta fuerza iónica (MgSO_4).

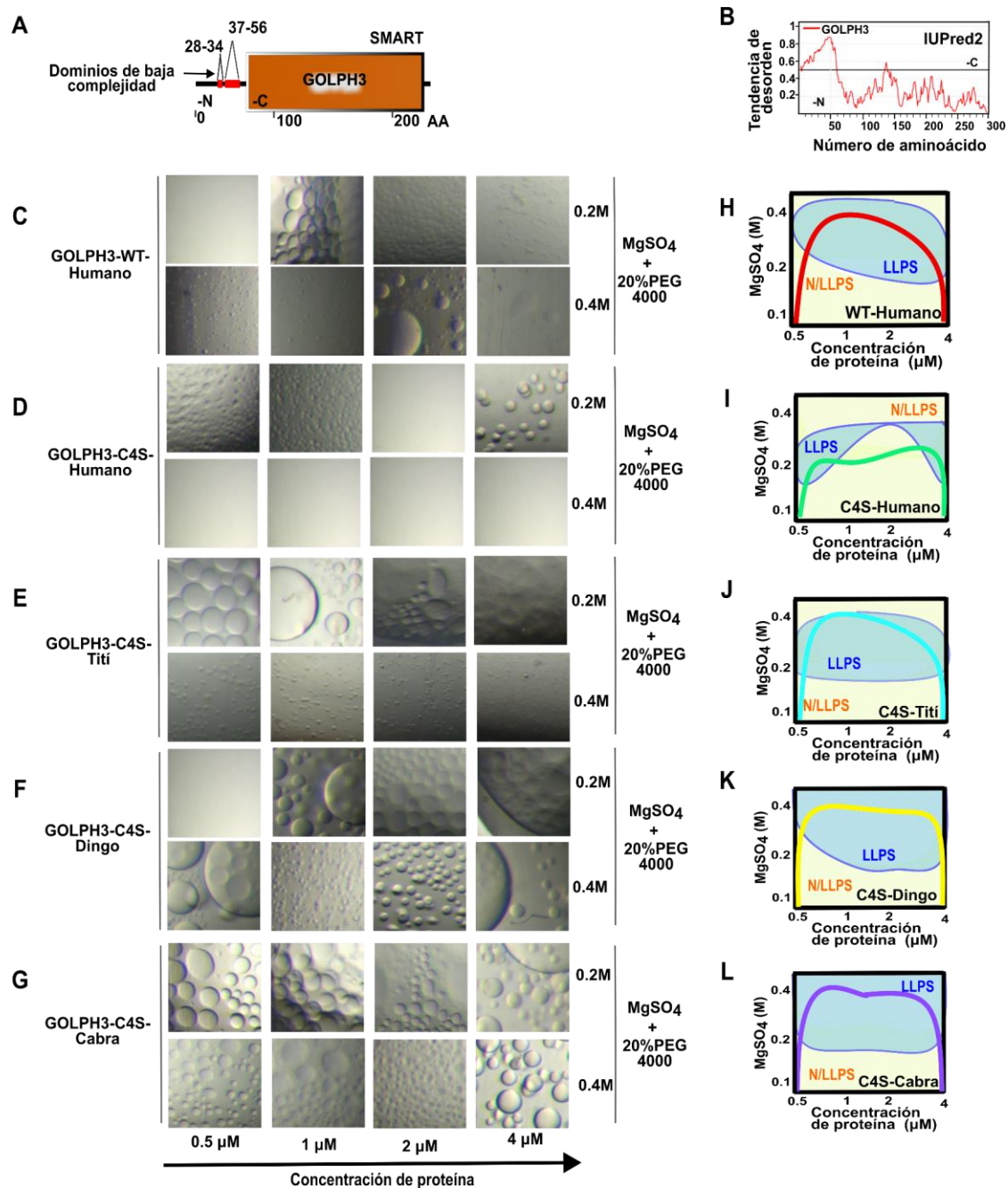


Figura 6. Ensayo de separación de fases líquido-líquido de las variantes de GOLPH3.
Realizado por: Galarza, 2022

4. Discusión y conclusiones

La secuencia de GOLPH3 se encuentra conservada en la mayoría de los mamíferos, sin embargo, su función molecular y estructural aún no se ha caracterizado en su totalidad. La estructura cristalina de GOLPH3 pudo ser resuelta gracias a la eliminación del N-terminal ya que este segmento carece de estructura secundaria definida (Wood et al., 2009). A través del estudio evolutivo de GOLPH3 hemos encontrado variaciones aminoácidas que permiten avanzar en nuestro entendimiento de la función del segmento N-terminal. En la actualidad, las bases de datos de información genómica completa permiten obtener una visión de la variación de las proteínas en distintos grupos taxonómicos (Tørresen et al., 2019). En el caso de la oncoproteína GOLPH3, al poseer ortólogos 1:1 en la mayoría de los mamíferos, estudios de su evolución dan aproximaciones de cómo GOLPH3 ha variado a lo largo de la historia evolutiva (Opazo et al., 2021; Wu et al., 2000). En esta tesis, las secuencias que fueron curadas manualmente permitieron determinar que GOLPH3 se encuentra conservada en todos los mamíferos analizados, pero con ligeras variaciones. El alineamiento de las especies de mamíferos y los análisis de divergencia mostraron que el segmento N-terminal de GOLPH3 presenta una mayor variabilidad en comparación al dominio C-terminal. Los análisis que se realizaron en esta investigación complementan la información proporcionada por Opazo, 2021 que menciona que en aves existe una clara divergencia en el segmento N-terminal de GOLPH3. No obstante, se deberían realizar más análisis evolutivos para poner a prueba si toda la familia de GOLPH3 (GOLPH3 y GOLPH3L) en mamíferos presenta este mismo patrón. Estas diferencias en la región N-terminal de GOLPH3 podrían haber surgido como consecuencia de restricciones estructurales y funcionales que permitieron la adaptación de la proteína a las diferentes condiciones que se exponen las especies de mamíferos tales como pH, temperatura o altura.

En el presente trabajo, se seleccionaron tres especies de mamíferos (Tití, Dingo y Cabra; **Figura 2**) que presentan variaciones de interés en la región intrínsecamente desordenada (IDR) de GOLPH3. GOLPH3 es una proteína que interactúa principalmente con el dominio TGN del aparato de Golgi mediante la unión a PI4P (Bishé et al., 2012; Rahajeng et al., 2019). Mediante la inmunofluorescencia indirecta (**Figura 3**) se comprobó que las variaciones en el segmento N-terminal de las proteínas recombinantes no afectan su distribución celular. Por lo tanto, este ensayo muestra que la especificidad de GOLPH3 respecto a su unión a PI4P se conserva tanto en levaduras, moscas y ahora otros mamíferos (Dippold et al., 2009; Rahajeng et al., 2019). Las proteínas recombinantes GOLPH3-C4S-Titi, GOLPH3-C4S-Dingo, GOLPH3-C4S-

Humano y GOLPH3-WT-Humano no presentaron diferencias en su punto isoeléctrico a pesar de las variaciones presentes en sus regiones N-terminal (Tabla 1). La excepción fue GOLPH3-C4S-Cabra que tiene una diferencia menor de su punto isoeléctrico. Consistente con lo anterior, la solubilidad de las variantes de GOLPH3 no se vio afectada para la producción y purificación de las proteínas recombinantes en los análisis *in vitro* (Información suplementaria S2).

Una manera de evaluar si las proteínas recombinantes purificadas mantienen una estructura similar a la nativa es por medio de ensayos de escaneo diferencial (DFS) y proteólisis parcial. DFS es una herramienta de alta sensibilidad que requiere muy poca concentración de proteína y permite una estimación de la estabilidad estructural mediante función de la sensibilidad térmica de una proteína (Pantoliano et al., 2001). Este ensayo puede ser realizado con un termociclador con capacidad de cuantificar fluorescencia para un PCR cuantitativo en tiempo real, como el Rotor-Gene Q (QIAGEN). Esta técnica es de aplicación fácil y rápida, permitiendo la reproducibilidad a muy bajo costo. Sin embargo, al comparar esta técnica con la calorimetría de escaneo diferencial (DSC), DSF posee desventajas respecto a la precisión de obtener datos cuantitativos de parámetros termodinámicos más específicos. DSF usa un fluoróforo para la reacción, el cual añade un parámetro extra para las ecuaciones clásicas de Gibbs–Helmholtz o Boltzmann, complicando el cálculo para ciertos parámetros termodinámicos de interés (cálculo de la entalpía, entropía o energía libre de Gibbs). A pesar de ello, se pueden obtener estos valores teniendo en cuenta los datos de T_m , ΔC_{pu} , ΔH_u y la intensidad de fluorescencia a una determinada temperatura (Senisterra et al., 2008). En esta investigación el parámetro de interés correspondió a la T_m y DSF sobresale en accesibilidad y rendimiento en comparación con otros métodos como el dicroísmo circular (CD) y DSC (Bruylants et al., 2005; Greenfield, 2007). La T_m corresponde al valor negativo de la primera derivada de los datos crudos de la intensidad de fluorescencia durante la desnaturalización térmica de la proteína. Este parámetro permite comparar proteínas con mutaciones debido a mayor T_m significa que una proteína presenta una mayor estabilidad estructural. Al comprar por DSF a GOLPH3-WT-Humano y las variantes GOLPH3-C4S-Humano, GOLPH3-C4S-Titi, GOLPH3-C4S-Dingo y GOLPH3-C4S-Cabra se obtuvo valores de T_m claramente menores en las variantes generadas, lo que indica que el cambio de cisteínas por serinas (C4S) afectan la estabilidad estructural de GOLPH3 (**Figura 4**). Al momento de comparar entre las variantes de GOLPH3-C4S para determinar si las variaciones del segmento N-terminal afectan la estabilidad estructural de la proteína, no se obtuvieron diferencias significativas, aun cuando se observa una tendencia de GOLPH3-C4S-Dingo hacia una

T_m menor. GOLPH3-C4S-Dingo presenta una inserción de glicina entre los aminoácidos 10-16 lo que promueve que esta proteína tenga una IDR con mayor movilidad puesto que glicina posee mayor grado de libertad en sus ángulos de torsión phi y psi (Lopez & Mohiuddin, 2020; Marsh, 1958). Por otro lado, una interpretación alternativa es que las sustituciones en el N-terminal de las variantes de GOLPH3 podrían favorecer la unión de SYPRO Orange ya que pueden tener zonas de mayor hidrofobicidad generando sitios con mayor afinidad por el fluoróforo.

Otra manera de evaluar la estabilidad estructural de una proteína es determinar su sensibilidad a la degradación por proteasas como la proteinasa K. Los resultados de los ensayos de proteólisis parcial mostraron que las variantes de GOLPH3 no tienen diferencias notables en la sensibilidad a la proteinasa K, indicando que tampoco deberían existir diferencias estructurales importantes. GOLPH3-WT-Humano y las variantes de GOLPH3-C4S presentaron los mismos fragmentos resistentes a la degradación (**Figura 5A-E**). La excepción fue GOLPH3-C4S-Cabra que presentó una degradación completa del fragmento difuso de 32.5 kDa (**Figura 5E**). Esto pudo haber sido favorecido por las sustituciones A33G, S35N y S36D (**Figura 2**) y un punto isoeléctrico menor (**Tabla 1**), lo que indica que estas variaciones con asparagina y ácido aspártico podrían hacer que GOLPH3 sea menos estable. El carboxilato aniónico (RCOO^-) del ácido aspártico y la carboxamida (RNH_3^+) de la asparagina pueden formar puentes salinos los cuales tienen influencia en la estabilidad de la conformación plegada, y éste es entrópicamente desfavorable para las proteínas (Biedermann & Schneider, 2016). Cabe recalcar que el análisis de los posibles sitios de digestión para explicar los fragmentos estables se realizó mediante un análisis de los pesos moleculares observados en los geles de poliacrilamida y la información del peso molecular derivado de la secuencia de los posibles fragmentos. Se concluyó que los posibles sitios de digestión con la proteinasa K serían en R56:L57, a razón de que esta región es la zona final del N-terminal que presenta un sitio expuesto, y P200:L201, puesto que, en este lugar se encuentra una horquilla beta expuesta de GOLPH3 (**Figura 5F-H**). Sin embargo, para tener una idea específica de los sitios de corte se deberían transferir los fragmentos a una membrana de fluoruro de polivinilideno (PVDF) y analizarlas mediante secuenciación de aminoácidos N-terminal aplicando degradación de Edman automatizada mediante un secuenciador de proteínas o alternativamente, mediante espectrometría de masas (Ross et al., 2014).

El objetivo de realizar los ensayos de DSF y proteólisis parcial fue definir si la estabilidad estructural de las proteínas recombinantes purificadas son similares a GOLPH3-WT-Humano, ya que en nuestro laboratorio esta proteína se pudo cristalizar y resolver su estructura, por lo tanto presenta una estabilidad y funcionalidad correcta. Sin embargo, también se pudieron haber realizado análisis a nivel funcional de las proteínas variantes. Un ejemplo es el ensayo de unión a fosfoinosítidos, en específico a PI4P, que implica evaluar si las variantes de GOLPH3-WT y GOLPH3-C4S recombinantes son capaces de unirse a PI4P inmovilizado en una membrana (Dippold et al., 2009; Dowler et al., 2000; Rahajeng et al., 2019). Sin embargo, se eligió evaluar primero con DFS y proteólisis parcial ya que brindan información estructural, y microscopía de fluorescencia para la información funcional. Una vez que se comprobó que las proteínas recombinantes purificadas mantienen propiedades similares a la proteína nativa se estudió su capacidad de experimentar LLPS.

Los condensados biomoleculares sin membrana, como el nucleolo, los cuerpos de Cajal o los cuerpos P, son generalmente esféricos e internamente dinámicos, compuestos por un conjunto de proteínas y otras moléculas (Banani et al., 2017; Brown et al., 2011). Se ha descrito que la formación de estos condensados biomoleculares puede ser favorecida por la presencia de proteínas con regiones intrínsecamente desordenadas (IDR) mediante separación de fases (Banani et al., 2017). Los segmentos IDR de las proteínas experimentan mayor tasa de divergencia de la secuencia a lo largo de la evolución (Brown et al., 2011). Las IDR poseen regiones enriquecidas de residuos que promueven el desorden, como alanina, arginina, glicina, glutamina, serina, ácido glutámico y lisina (Uversky et al., 2015). Por lo tanto, la composición de aminoácidos, la hidrofobicidad o la carga de las IDR juegan un papel importante para la separación de fases líquido-líquido (LLPS) de una proteína. Observamos que el segmento N-terminal intrínsecamente desordenado de GOLPH3 posee cambios evolutivos (adición o pérdida de residuos cargados, hidrofóbicos o aromáticos) que modulan la LLPS *in vitro*. Las variantes escogidas para este estudio poseen variaciones que intervienen en la dinámica de la LLPS. GOLPH3-WT-Humano posee cuatro cisteínas en su estructura. La eliminación de las cisteínas en las variantes C4S restringe la formación de puentes disulfuro artefactuales que podrían cambiar la solubilidad con el medio. El caso de GOLPH3-C4S-Titi resulta interesante ya que sólo posee una variación en un sitio potencialmente fosforilable de la proteína, que es serina 36 por glicina. El cambio por una glicina resultó en la presencia de pequeñas gotas a una concentración relativamente alta de sulfato de magnesio lo que no se vio en GOLPH3-C4S-Humano (**Figura 6DE;IJ**). La capacidad de glicina de tener mayor grado de libertad en sus ángulos de torsión phi

y psi proporciona al segmento N-terminal mayor movilidad, por lo tanto, es posible que facilite la separación de fases. Para GOLPH3-C4S-Dingo se ve favorecida la LLPS aún más que GOLPH3-C4S-Titi (**Figura 6EF;JK**). La presencia de gotas relativamente grandes en la mayoría de las condiciones indica que las variaciones de Dingo (N21S, A28V, A33G, S35/36G) hacen que posea propiedades más hidrofóbicas. Los dominios hidrofóbicos son ricos en valina, glicina, serina, leucina y prolina, por lo que la proteína disminuye su capacidad de formar enlaces de hidrógeno con las moléculas de agua del medio circundante y esto impulsa la LLPS (Miao et al., 2003; Pak et al., 2016). Al tener GOLPH3-C4S-Dingo glicinas seguidas, es posible que el aumento de la fuerza iónica o cambios ambientales desestabilicen las gotas haciendo que estas desaparezcan con el tiempo y no lleguen a un estado de maduración (Miao et al., 2003; Pak et al., 2016; Voorhees, 1992). Sin embargo, la LLPS no sólo depende de la acumulación de residuos hidrofóbicos en los IDR sino también de residuos cargados como es el caso de GOLPH3-C4S-Cabra. De hecho, la presencia de gotas relativamente grandes en presencia de 0.2 y 0.4 M de MgSO₄, y en todas las concentraciones de proteína, indica que las variaciones del N-terminal de GOLPH3-C4S-Cabra permiten una rápida LLPS *in vitro* (**Figura 6G;H**).

La variación S36G permite la acumulación de residuos cargados en el segmento N-terminal (D36-E37-D38-D39) que tienen un efecto hasta en el punto isoeléctrico de la proteína (Tabla 1). Esta variación puede formar interacciones electrostáticas en presencia de sal de alta fuerza iónica que permiten la estabilidad de las gotas (Nott et al., 2015). Por otra parte, la variación S35N puede contribuir a impulsar la separación de fases y eso podría ocurrir a través de interacciones dipolares de su cadena lateral (Banani et al., 2017; Crick et al., 2013). Por ende, las variaciones del segmento N-terminal de GOLPH3-C4S-Cabra tiene grupos de residuos cargados que promueven de gran manera la LLPS. Un aspecto interesante es que todas las variantes estudiadas generaron separación de fases *in vitro*, aunque el tamaño o cantidad de gotas varió, sugiriendo que otras regiones de la proteína podrían participar en la LLPS. Al realizar un análisis fuera del segmento N-terminal, es posible encontrar un motivo descrito en otros estudios previos sobre LLPS. El dominio C-terminal de GOLPH3 posee el motivo GYTSFW (76-81) expuesto hacia el solvente que podría ser de importancia. Se sabe que los residuos de triptófano, tirosina y fenilalanina (residuos aromáticos) flanqueados por serina o glicina (residuos alifáticos) son importantes para la separación de fases de algunas proteínas. Los residuos aromáticos se caracterizan por tener un enlace pi y los residuos alifáticos por tener enlaces peptídicos expuestos, por ende, se sugiere que las interacciones catión-pi que podrían establecerse entre estos aminoácidos son fuerzas

impulsoras para la LLPS (Burke et al., 2015; Kato et al., 2012; Li et al., 2018; Lin et al., 2015). Para comprobar que este motivo de GOLPH3 es de importancia para la LLPS se debería generar otra proteína recombinante que tenga mutaciones de estos aminoácidos y someterla a las mismas condiciones evaluadas en este estudio.

En conclusión, este estudio sugiere que la región intrínsecamente desordenada de GOLPH3 facilita la LLPS *in vitro* mediante fuerzas impulsoras como interacciones electrostáticas entre grupos de aminoácidos con carga, residuos hidrofóbicos e interacciones catión- π . Para completar el estudio se deberían evaluar la LLPS *in cellulo* y determinar si tiene características líquidas en un medio más complejo, y esto se podría realizar mediante ensayos de recuperación de fluorescencia después del fotoblanqueo (FRAP) después de transfectar los plásmidos que permiten la expresión en células de mamífero de las variantes de GOLPH3 fusionadas a la proteína fluorescente verde.

Referencias

- Alberti, S., Gladfelter, A., & Mittag, T. (2019). Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. En *Cell* (Vol. 176, Issue 3, pp. 419–434). Cell Press. <https://doi.org/10.1016/j.cell.2018.12.035>
- Alexandrov, A. I., Mileni, M., Chien, E. Y. T., Hanson, M. A., & Stevens, R. C. (2008). Microscale Fluorescent Thermal Stability Assay for Membrane Proteins. *Structure*, 16(3), 351–359. <https://doi.org/10.1016/j.str.2008.02.004>
- Ali, M. F., Chachadi, V. B., Petrosyan, A., & Chengs, P. W. (2012). Golgi phosphoprotein 3 determines cell binding properties under dynamic flow by controlling golgi localization of core 2 N-acetylglucosaminyltransferase 1. *Journal of Biological Chemistry*, 287(47), 39564–39577. <https://doi.org/10.1074/jbc.M112.346528>
- Arolas, J. L., Aviles, F. X., Chang, J. Y., & Ventura, S. (2006). Folding of small disulfide-rich proteins: clarifying the puzzle. En *Trends in Biochemical Sciences* (Vol. 31, Issue 5). <https://doi.org/10.1016/j.tibs.2006.03.005>
- Banani, S. F., Lee, H. O., Hyman, A. A., & Rosen, M. K. (2017). Biomolecular condensates: Organizers of cellular biochemistry. En *Nature Reviews Molecular Cell Biology* (Vol. 18, Issue 5, pp. 285–298). Nature Publishing Group. <https://doi.org/10.1038/nrm.2017.7>
- Biedermann, F., & Schneider, H. J. (2016). Experimental Binding Energies in Supramolecular Complexes. En *Chemical Reviews* (Vol. 116, Issue 9). <https://doi.org/10.1021/acs.chemrev.5b00583>
- Bishé, B., Syed, G. H., Field, S. J., & Siddiqui, A. (2012). Role of phosphatidylinositol 4-phosphate (PI4P) and its binding protein GOLPH3 in hepatitis C virus secretion. *Journal of Biological Chemistry*, 287(33). <https://doi.org/10.1074/jbc.M112.346569>
- Brown, C. J., Johnson, A. K., Dunker, A. K., & Daughdrill, G. W. (2011). Evolution and disorder. En *Current Opinion in Structural Biology* (Vol. 21, Issue 3). <https://doi.org/10.1016/j.sbi.2011.02.005>
- Bruylants, G., Wouters, J., & Michaux, C. (2005). Differential Scanning Calorimetry in Life Science: Thermodynamics, Stability, Molecular Recognition and Application in Drug Design. *Current Medicinal Chemistry*, 12(17). <https://doi.org/10.2174/0929867054546564>
- Burke, K. A., Janke, A. M., Rhine, C. L., & Fawzi, N. L. (2015). Residue-by-Residue View of In Vitro FUS Granules that Bind the C-Terminal Domain of RNA Polymerase II. *Molecular Cell*, 60(2). <https://doi.org/10.1016/j.molcel.2015.09.006>
- Chemes, L. B., Glavina, J., Alonso, L. G., Marino-Buslje, C., de Prat-Gay, G., & Sánchez, I. E. (2012). Sequence Evolution of the Intrinsically Disordered and Globular Domains of a Model Viral Oncoprotein. *PLoS ONE*, 7(10). <https://doi.org/10.1371/journal.pone.0047661>

- Crick, S. L., Ruff, K. M., Garai, K., Frieden, C., & Pappu, R. v. (2013). Unmasking the roles of N- and C-terminal flanking sequences from exon 1 of huntingtin as modulators of polyglutamine aggregation. *Proceedings of the National Academy of Sciences of the United States of America*, 110(50). <https://doi.org/10.1073/pnas.1320626110>
- Dippold, H. C., Ng, M. M., Farber-Katz, S. E., Lee, S. K., Kerr, M. L., Peterman, M. C., Sim, R., Wiharto, P. A., Galbraith, K. A., Madhavarapu, S., Fuchs, G. J., Meerloo, T., Farquhar, M. G., Zhou, H., & Field, S. J. (2009). GOLPH3 Bridges Phosphatidylinositol-4- Phosphate and Actomyosin to Stretch and Shape the Golgi to Promote Budding. *Cell*, 139(2). <https://doi.org/10.1016/j.cell.2009.07.052>
- Dowler, S., Currie, R. A., Campbell, D. G., Deak, M., Kular, G., Downes, C. P., & Alessi, D. R. (2000). Identification of pleckstrin-homology-domain-containing proteins with novel phosphoinositide-binding specificities. *Biochemical Journal*, 351(1). <https://doi.org/10.1042/0264-6021:3510019>
- Feng, Y., He, F., Wu, H., Huang, H., Zhang, L., Han, X., & Liu, J. (2015). GOLPH3L is a novel prognostic biomarker for epithelial ovarian cancer. *Journal of Cancer*, 6(9), 893–900. <https://doi.org/10.7150/jca.11865>
- Fontana, A., de Laureto, P. P., Spolaore, B., Frare, E., Picotti, P., & Zamboni, M. (2004). Probing protein structure by limited proteolysis. *Acta Biochimica Polonica*, 51(2). https://doi.org/10.18388/abp.2004_3573
- Glick, B. S., & Nakano, A. (2009). Membrane traffic within the Golgi apparatus. En *Annual Review of Cell and Developmental Biology* (Vol. 25). <https://doi.org/10.1146/annurev.cellbio.24.110707.175421>
- Greenfield, N. J. (2007). Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nature Protocols*, 1(6). <https://doi.org/10.1038/nprot.2006.204>
- Han, T. W., Kato, M., Xie, S., Wu, L. C., Mirzaei, H., Pei, J., Chen, M., Xie, Y., Allen, J., Xiao, G., & McKnight, S. L. (2012). Cell-free formation of RNA granules: Bound RNAs identify features and components of cellular assemblies. *Cell*, 149(4). <https://doi.org/10.1016/j.cell.2012.04.016>
- Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, 35(2). <https://doi.org/10.1093/molbev/msx281>
- Isaji, T., Im, S., Gu, W., Wang, Y., Hang, Q., Lu, J., Fukuda, T., Hashii, N., Takakura, D., Kawasaki, N., Miyoshi, H., & Gu, J. (2014). An oncogenic protein Golgi phosphoprotein 3 up-regulates cell migration via sialylation. *Journal of Biological Chemistry*, 289(30). <https://doi.org/10.1074/jbc.M113.542688>
- Kato, M., Han, T. W., Xie, S., Shi, K., Du, X., Wu, L. C., Mirzaei, H., Goldsmith, E. J., Longgood, J., Pei, J., Grishin, N. v., Frantz, D. E., Schneider, J. W., Chen, S., Li, L., Sawaya, M. R., Eisenberg, D., Tycko, R., & McKnight, S. L. (2012). Cell-free formation of RNA granules: Low complexity sequence domains form dynamic fibers within hydrogels. *Cell*, 149(4). <https://doi.org/10.1016/j.cell.2012.04.017>

- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30(4). <https://doi.org/10.1093/molbev/mst010>
- Letunic, I., & Bork, P. (2018). 20 years of the SMART protein domain annotation resource. *Nucleic Acids Research*, 46(D1). <https://doi.org/10.1093/nar/gkx922>
- Li, H. R., Chiang, W. C., Chou, P. C., Wang, W. J., & Huang, J. rong. (2018). TAR DNA-binding protein 43 (TDP-43) liquid-liquid phase separation is mediated by just a few aromatic residues. *Journal of Biological Chemistry*, 293(16). <https://doi.org/10.1074/jbc.AC117.001037>
- Lichtarge, O., & Wilkins, A. (2010). Evolution: a guide to perturb protein function and networks. *Current Opinion in Structural Biology*, 20(3), 351–359. <https://doi.org/10.1016/J.SBI.2010.04.002>
- Lin, Y., Protter, D. S. W., Rosen, M. K., & Parker, R. (2015). Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Molecular Cell*, 60(2). <https://doi.org/10.1016/j.molcel.2015.08.018>
- Lopez, M. J., & Mohiuddin, S. S. (2020). Biochemistry, Essential Amino Acids. En *StatPearls*.
- Marín, M., Thallmair, V., & Ott, T. (2012). The intrinsically disordered N-terminal region of AtREM1.3 remorin protein mediates protein-protein interactions. *Journal of Biological Chemistry*, 287(47), 39982–39991. <https://doi.org/10.1074/jbc.M112.414292>
- Marsh, R. E. (1958). A refinement of the crystal structure of glycine. *Acta Crystallographica*, 11(9). <https://doi.org/10.1107/s0365110x58001717>
- Mellman, I., & Warren, G. (2000). The road taken: Past and future foundations of membrane traffic. En *Cell* (Vol. 100, Issue 1). [https://doi.org/10.1016/S0092-8674\(00\)81687-6](https://doi.org/10.1016/S0092-8674(00)81687-6)
- Mészáros, B., Erdős, G., & Dosztányi, Z. (2018). IUPred2A: Context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Research*, 46(W1). <https://doi.org/10.1093/nar/gky384>
- Miao, M., Bellingham, C. M., Stahl, R. J., Sitarz, E. E., Lane, C. J., & Keeley, F. W. (2003). Sequence and structure determinants for the self-aggregation of recombinant polypeptides modeled after human elastin. *Journal of Biological Chemistry*, 278(49). <https://doi.org/10.1074/jbc.M308465200>
- Nakashima-Kamimura, N., Asoh, S., Ishibashi, Y., Mukai, Y., Shidara, Y., Oda, H., Munakata, K., Goto, Y. I., & Ohta, S. (2005). MIDAS/GPP34, a nuclear gene product, regulates total mitochondrial mass in response to mitochondrial dysfunction. *Journal of Cell Science*, 118(22). <https://doi.org/10.1242/jcs.02645>
- Nott, T. J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowietz, A., Craggs, T. D., Bazett-Jones, D. P., Pawson, T., Forman-Kay, J. D., & Baldwin, A. J. (2015). Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles. *Molecular Cell*, 57(5). <https://doi.org/10.1016/j.molcel.2015.01.013>
- O’Leary, M. A., Bloch, J. I., Flynn, J. J., Gaudin, T. J., Giallombardo, A., Giannini, N. P., Goldberg, S. L., Kraatz, B. P., Luo, Z. X., Meng, J., Ni, X., Novacek, M. J., Perini, F. A., Randall, Z. S., Rougier, G. W., Sargis, E. J., Silcox, M. T., Simmons, N. B., Spaulding, M., ... Cirranello, A. L.

- (2013). The placental mammal ancestor and the Post-K-Pg radiation of placentals. *Science*, 339(6120). <https://doi.org/10.1126/science.1229237>
- Opazo, J. C., Vandeweghe, M. W., Gutierrez, J., Zavala, K., Vargas-Chacoff, L., Morera, F. J., & Mardones, G. A. (s/f). Evolution of GOLPH3 genes Independent origins of the Golgi phosphoprotein 3 oncoprotein gene 3 4. *Original Research Article*, 1(2). <https://doi.org/10.1101/2021.01.31.429031>
- Pak, C. W., Kosno, M., Holehouse, A. S., Padrick, S. B., Mittal, A., Ali, R., Yunus, A. A., Liu, D. R., Pappu, R. v., & Rosen, M. K. (2016). Sequence Determinants of Intracellular Phase Separation by Complex Coacervation of a Disordered Protein. *Molecular Cell*, 63(1). <https://doi.org/10.1016/j.molcel.2016.05.042>
- Pantoliano, M. W., Petrella, E. C., Kwasnoski, J. D., Lobanov, V. S., Myslik, J., Graf, E., Carver, T., Asel, E., Springer, B. A., Lane, P., & Salemme, F. R. (2001). High-density miniaturized thermal shift assays as a general strategy for drug discovery. *Journal of Biomolecular Screening*, 6(6). <https://doi.org/10.1089/108705701753364922>
- Rahajeng, J., Kuna, R. S., Makowski, S. L., Tran, T. T. T., Buschman, M. D., Li, S., Cheng, N., Ng, M. M., & Field, S. J. (2019). Efficient Golgi Forward Trafficking Requires GOLPH3-Driven, PI4P-Dependent Membrane Curvature. *Developmental Cell*, 50(5). <https://doi.org/10.1016/j.devcel.2019.05.038>
- Rizzo, R., Russo, D., Kurokawa, K., Sahu, P., Lombardi, B., Supino, D., Zhukovsky, M. A., Vocat, A., Pothukuchi, P., Kunnathully, V., Capolupo, L., Boncompain, G., Vitagliano, C., Zito Marino, F., Aquino, G., Montariello, D., Henklein, P., Mandrich, L., Botti, G., ... Luini, A. (2021). Golgi maturation-dependent glycoenzyme recycling controls glycosphingolipid biosynthesis and cell growth via GOLPH3. *The EMBO Journal*, 40(8). <https://doi.org/10.15252/emboj.2020107238>
- Ross, B. H., Lin, Y., Corales, E. A., Burgos, P. v., & Mardones, G. A. (2014a). Structural and functional characterization of cargo-binding sites on the μ 4-subunit of adaptor protein complex 4. *PLoS ONE*, 9(2). <https://doi.org/10.1371/journal.pone.0088147>
- Ross, B. H., Lin, Y., Corales, E. A., Burgos, P. v., & Mardones, G. A. (2014b). Structural and functional characterization of cargo-binding sites on the μ 4-subunit of adaptor protein complex 4. *PLoS ONE*, 9(2). <https://doi.org/10.1371/journal.pone.0088147>
- Scott, K. L., Kabbarah, O., Liang, M. C., Ivanova, E., Anagnostou, V., Wu, J., Dhakal, S., Wu, M., Chen, S., Feinberg, T., Huang, J., Saci, A., Widlund, H. R., Fisher, D. E., Xiao, Y., Rimm, D. L., Protopopov, A., Wong, K. K., & Chin, L. (2009). GOLPH3 modulates mTOR signalling and rapamycin sensitivity in cancer. *Nature*, 459(7250), 1085–1090. <https://doi.org/10.1038/nature08109>
- Sechi, S., Frappaolo, A., Belloni, G., Colotti, G., & Giansanti, M. G. (s/f). Oncotarget 3493 www.impactjournals.com/oncotarget The multiple cellular functions of the oncoprotein Golgi phosphoprotein 3. In *Oncotarget* (Vol. 6, Issue 6). www.impactjournals.com/oncotarget/
- Senisterra, G. A., Hong, B. S., Park, H. W., & Vedadi, M. (2008). Application of high-throughput isothermal denaturation to assess protein stability and screen for ligands. *Journal of Biomolecular Screening*, 13(5). <https://doi.org/10.1177/1087057108317825>

- Sharma, S., Ciufu, S., Starchenko, E., Darji, D., Chlumsky, L., Karsch-Mizrachi, I., & Schoch, C. L. (2018). The NCBI BioCollections Database. *Database*, 2018(2018). <https://doi.org/10.1093/database/bay006>
- Sheffield, P., Garrard, S., & Derewenda, Z. (1999). Overcoming Expression and Purification Problems of RhoGDI Using a Family of “Parallel” Expression Vectors. *Protein Expression and Purification*, 15(1), 34–39. <https://doi.org/https://doi.org/10.1006/prep.1998.1003>
- Stanley, P. (2011). Golgi glycosylation. *Cold Spring Harbor Perspectives in Biology*, 3(4). <https://doi.org/10.1101/cshperspect.a005199>
- Szalkowski, A. M., & Anisimova, M. (2011). Markov models of amino acid substitution to study proteins with intrinsically disordered regions. *PLoS ONE*, 6(5). <https://doi.org/10.1371/journal.pone.0020488>
- Tatusova, T. A., & Madden, T. L. (1999). BLAST 2 SEQUENCES, a new tool for comparing protein and nucleotide sequences. *FEMS Microbiology Letters*, 174(2). [https://doi.org/10.1016/S0378-1097\(99\)00149-4](https://doi.org/10.1016/S0378-1097(99)00149-4)
- Taylor, R. S., Fialka, I., Jones, S. M., Huber, L. A., & Howell, K. E. (1997). Two-dimensional mapping of the endogenous proteins of the rat hepatocyte Golgi complex cleared of proteins in transit. *Electrophoresis*, 18(14). <https://doi.org/10.1002/elps.1150181416>
- Tørresen, O. K., Star, B., Mier, P., Andrade-Navarro, M. A., Bateman, A., Jarnot, P., Gruca, A., Grynberg, M., Kajava, A. v., Promponas, V. J., Anisimova, M., Jakobsen, K. S., & Linke, D. (2019). Tandem repeats lead to sequence assembly errors and impose multi-level challenges for genome and protein databases. En *Nucleic acids research* (Vol. 47, Issue 21). <https://doi.org/10.1093/nar/gkz841>
- Trifinopoulos, J., Nguyen, L. T., von Haeseler, A., & Minh, B. Q. (2016). W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Research*, 44(W1). <https://doi.org/10.1093/NAR/GKW256>
- Uversky, V. N., Kuznetsova, I. M., Turoverov, K. K., & Zaslavsky, B. (2015). Intrinsically disordered proteins as crucial constituents of cellular aqueous two phase systems and coacervates. *FEBS Letters*, 589(1). <https://doi.org/10.1016/j.febslet.2014.11.028>
- Voorhees, P. W. (1992). Ostwald ripening of two-phase mixtures. *Annual Review of Materials Science*, 22(1). <https://doi.org/10.1146/annurev.ms.22.080192.001213>
- Wood, C. S., Schmitz, K. R., Bessman, N. J., Setty, T. G., Ferguson, K. M., & Burd, C. G. (2009). PtdIns4P recognition by Vps74/GOLPH3 links PtdIns 4-kinase signaling to retrograde Golgi trafficking. *Journal of Cell Biology*, 187(7). <https://doi.org/10.1083/jcb.200909063>
- Wu, C. C., Taylor, R. S., Lane, D. R., Ladinsky, M. S., Weisz, J. A., & Howell, K. E. (2000). GMx33: A Novel Family of trans-Golgi Proteins Identified by Proteomics. En *Traffic* (Vol. 1). Munksgaard International Publishers.
- Xiang, Y., Zhang, X., Nix, D. B., Katoh, T., Aoki, K., Tiemeyer, M., & Wang, Y. (2013). Regulation of protein glycosylation and sorting by the Golgi matrix proteins GRASP55/65. *Nature Communications*, 4. <https://doi.org/10.1038/ncomms2669>

- Yarwood, R., Hellicar, J., Woodman, P. G., & Lowe, M. (2020). Membrane trafficking in health and disease. *DMM Disease Models and Mechanisms*, 13(4).
<https://doi.org/10.1242/dmm.043448>
- Yates, A. D., Achuthan, P., Akanni, W., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Azov, A. G., Bennett, R., Bhai, J., Billis, K., Boddu, S., Marugán, J. C., Cummins, C., Davidson, C., Dodiya, K., Fatima, R., Gall, A., ... Flicek, P. (2020). Ensembl 2020. *Nucleic Acids Research*, 48(D1), D682–D688. <https://doi.org/10.1093/nar/gkz966>
- Yoshizawa, T., Nozawa, R. S., Jia, T. Z., Saio, T., & Mori, E. (2020). Biological phase separation: cell biology meets biophysics. En *Biophysical Reviews* (Vol. 12, Issue 2, pp. 519–539). Springer. <https://doi.org/10.1007/s12551-020-00680-x>

ANEXOS

Anexo A. Metodología extendida.

Transformación de bacterias competentes

Las bacterias competentes fueron descongeladas a 4 °C. Se colocó 1 µL de ADN plasmidial con 20 µL de bacterias y la mezcla se mantuvo en hielo por 30 min. Después de este tiempo, se realizó una incubación durante 40 s a 42 °C. Luego se agregó 200 µL de medio SOC (Tryptona 2% w/v, extracto de levadura 0.5% w/v, NaCl 8.6 mM, KCl 2.5 mM, MgSO₄ 20 mM, Glucosa 20 mM) sobre la mezcla de transformación. Seguido, la mezcla se plaqueó en agar-LB (Agar 1.5 % w/v, Tryptona 10 g/L, extracto de levadura 5 g/L, NaCl 5 g/L) con ampicilina 100 µg/ml. El plaqueo varió dependiendo del tipo de procedimiento previo; si la transformación correspondió a un producto de ligación se plaqueó el total de la mezcla de transformación, y si la transformación fue con ADN plasmidial sobre enrollado se plaqueó 5 µL de la mezcla de transformación.

Purificación de DNA plasmidial mediante minipreparaciones

Las bacterias portadoras del vector pGFP-GOLPH3-C4S-Xx se cultivaron en 2 mL de medio LB durante 16 h a 37 °C. Seguido se tomó 1.4 mL del cultivo y se sedimentaron a 16000 x g durante 10 min. El sobrenadante se descartó y el sedimento se resuspendió con 250 µL de solución P1 (Tris 50 mM, EDTA 10 mM, RNasa 10 µg/mL, pH 8.0). Luego se procedió a añadir 250 µL de solución P2 (NaOH 200 mM, SDS 1%). En seguida se agregó 350 µL de solución P3 (hidrocloruro de guanidina 4.09 M y ácido acético 2.12 M). Se centrifugó a 16000 x g por 10 min, y el sobrenadante se cargó en una columna con hidroxiapatita y se centrifugó a 16000 x g por 1 min. La columna se lavó con 700 µL de tampón PE (Tris 10 mM, etanol 80%, pH 7.5). Finalmente el DNA se eluyó con tampón TE (Tris 10 mM, EDTA 1 mM, pH 7.5).

Separación de fragmentos de DNA en geles de agarosa.

Los geles de agarosa fueron preparados al 1% p/v. La agarosa se disolvió en tampón TAE (Tris base 0.03 M, ácido acético glacial 0.1% v/v, EDTA 0.01 M). Los geles se sometieron a 100 V por 40 min. Para visualizar la migración electroforética de las

muestras de ADN, se adicionó a las muestras 1/3 de su volumen de solución de carga (glicerol 25%, SDS 0.5%, azul de bromofenol 0.025%, xilen cianol FF 0.05%, EDTA 12 mM). Para tener una referencia del tamaño molecular de las bandas, también se cargó en el gel 5 µL de estándar (Mass Ruler Mix, Thermo Scientific). Una vez finalizada la electroforesis, el resultado se visualizó en un transiluminador con luz UV.

Subclonamiento de las variantes de GOLPH3 en pGST-Parallel-1.

Las tres variantes de GOLPH3 de mamíferos y GOLPH3 humano (WT) fueron subclonados en el vector pGST-Parallel-1 para expresión en bacterias. Para comenzar con el subclonamiento, al vector dador del inserto y el vector aceptor se sometieron a digestión con las enzimas de restricción *EcoRI* (1 µL) y *XhoI* (1 µL) durante 16 h a 37 °C. La reacción se realizó en tampón NEB-3 (Tris-HCl 50 mM, NaCl 100 mM, MgCl₂, 10 mM, DTT 1 mM, pH 7.9), y 10 µg de albúmina de suero bovino. El volumen final de reacción fue 50 µL. La mezcla de digestión se sometió a una electroforesis preparativa para separar los fragmentos de interés. Para extracción del inserto y el vector del gel de agarosa se usó el kit *QIAquick Gel Extraction*. Seguido, la ligación entre el inserto y el vector se realizó en una proporción de 4:1, respectivamente. Esta mezcla se procedió a transformar en bacterias ultra competentes XL10 Gold y B834(DH3) pLysS (Metodología 2.2.1)

Purificación de proteína GOLPH3 recombinante

La purificación de cada proteína recombinante comenzó con un preinóculo de 20 mL de medio LB (Tryptona 10 g/L, extracto de levadura 5 g/L, NaCl 10 g/L) con Ampicilina (100 µg/mL) y 20 µL del respectivo stock de la cepa B834(DE3)pLysS de *E. coli*, previamente transformada con el respectivo vector pGST-Parallel-1, mantenido en congelador a -80 °C con glicerol al 10%. . Estas construcciones consistieron en Glutación-S-transferasa (GST) fusionado al extremo N-terminal de las variantes de GOLPH3, y separados por un sitio de digestión con la proteasa del virus del grabado del tabaco (TEV). El preinóculo se incubó a 37 °C durante 16 h con agitación constante (220 rpm). Después, los 20 mL se utilizaron para inocular 1.5 L de medio LB con Ampicilina 100 µg/mL. Este inóculo se agitó a 37 °C por 6 h a 200 rpm. Seguido, se disminuyó la temperatura a 25 °C durante 30 min para proceder a la adición de IPTG 0.35 mM. El inóculo se incubó con IPTG por 16 h bajo agitación constante a 200 rpm. Las bacterias se sedimentaron a 3500 x g por 40 min a 4 °C en una centrifuga refrigerada. Después, se procedió a resuspender el sedimento con un tampón de homogeneización (Tris 50 mM, 500 mM NaCl, 0.5 M β-

mercaptoetanol, EDTA 5 mM, azida de sodio 0.02% y 2 mM PMSF, pH 8.0). La resuspensión se sonicó por 5 min en hielo con ciclos de 2 seg de sonicación y 2 seg de pausa. Seguido, se centrifugó a $1833315 \times g$ por 90 min a 4 °C. El sobrenadante se pasó por filtros de 0.45 μ m y se cargó en una columna de 1.5 cm de diámetro con 1 mL de resina Glutación-Sefarosa. Esta columna se equilibró previamente con 2.5 mL de agua destilada y 5 mL de tampón de lavado (Tris 50 mM, NaCl 0.5 M β -mercaptoetanol 5 mM, EDTA 5 mM, azida de sodio 0.02%, pH 8.0). Una vez que el filtrado pasó por la columna durante 3 h, se tomó una alícuota de 7,8 μ L y se procedió a lavar con 250 mL de tampón de lavado. Todas las variantes de GOLPH3 recombinante se eluyeron en fracciones de 2.33 mL con un tampón de elución (Tris 50 mM, NaCl 150 mM, EDTA 5 mM, L-glutación reducido 20 mM, azida de sodio 0.02%, pH 8.0). Seguidamente, se unieron las fracciones de proteína en un pool y se concentró por ultrafiltración con filtros VivaSpin de 30 kDa a $3000 \times g$ a 4 °C, el tiempo suficiente hasta alcanzar un volumen de 5 mL. El concentrado de la proteína recombinante se cargó en una columna HiLoad 16/600 Superdex 200 pg (GE Healthcare) mediante AKTA-Prime para cromatografía de exclusión molecular. El flujo fue de 1 mL/min en un tampón de cromatografía (Tris 25 mM, NaCl 150 mM, β -mercaptoetanol 5 mM, azida de sodio 0.02%, pH 8.0). Se colectaron fracciones de 2,5 mL y se analizaron alícuotas de 13 μ L que se mezclaron con 5 μ L de tampón de carga para electroforesis de proteínas mediante geles Nu-PAGE en gradiente de 4-12% (ThermoFisher). Inmediatamente, se juntó en un pool las fracciones que contenían la proteína recombinante y se agregó 75 μ L de proteasa TEV (50 μ g) para clivar la GST de las variantes de GOLPH3. Se incubó este pool durante toda la noche a 4 °C y al día siguiente se concentró hasta 5 mL con un filtro VivaSpin de 10 kDa y se sometió nuevamente a cromatografía de exclusión molecular en una columna HiLoad 16/600 Superdex 200 pg (GE Healthcare), con las mismas condiciones indicadas anteriormente. Después, se procedió a remover el fragmento de GST mediante cromatografía de afinidad. El pool se pasó dos veces a través de una columna de 1 mL de Glutation-Sefarosa, previamente equilibrada con un tampón de equilibrio (Tris-HCL 50 mM, NaCl 150 mM, pH 8.0). Finalmente, para remover la proteasa TEV, se añadió al pool 7 mM $MgCl_2$ y se centrifugó durante 10 min a $3000 \times g$ a 4 °C. Se preparó una columna con 500 μ L de resina de Níquel-Sefarosa pre-equilibrada con tampón de equilibrio (Tris 50 mM, 150 mM NaCl, 0.5 M β -mercaptoetanol, EDTA 5 mM, azida de sodio 0.02%, pH 8.0) y se pasó el pool de la muestra una sola vez. Enseguida, se añadió DTT y EDTA para alcanzar una concentración de 1 mM y 5 mM, respectivamente. Las proteínas purificadas se almacenaron a -80 °C.

Separación de proteínas en geles de poliacrilamida

Para el análisis de la purificación de las variantes de GOLPH3 recombinantes se usaron geles en gradiente de 4-12% en condiciones denaturantes, sometidos a 110 mA durante 45 min en una cámara para electroforesis, usando un tampón de electroforesis MES (MES 50 mM, Tris base 50 mM, SDS 0.1%, EDTA 1 mM, pH 7.3) y tampón de carga Nu-PAGE LDS (Invitrogen). Se usó un estándar de tamaño molecular SeeBlue® (Invitrogen). Una vez finalizada la electroforesis los geles se tiñeron con solución SimplyBlue (Invitrogen) durante 2 días, luego se recuperó el colorante y se lavaron los geles con agua por un día. Los geles se deshidrataron con etanol 20% y finalmente se secaron entre dos hojas de papel celofán durante 1 día.

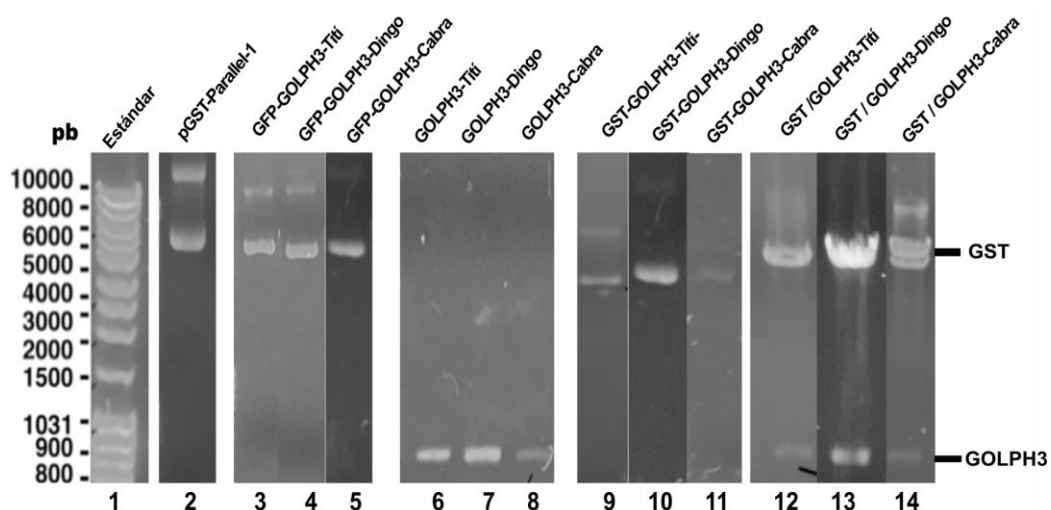


Imagen suplementaria 1. Subclonamiento del ADNc de las variantes de GOLPH3 de Tití, Dingo y Cabra desde el vector pGFP a pGST-Parallel-1.

Realizado por: Galarza, 2022

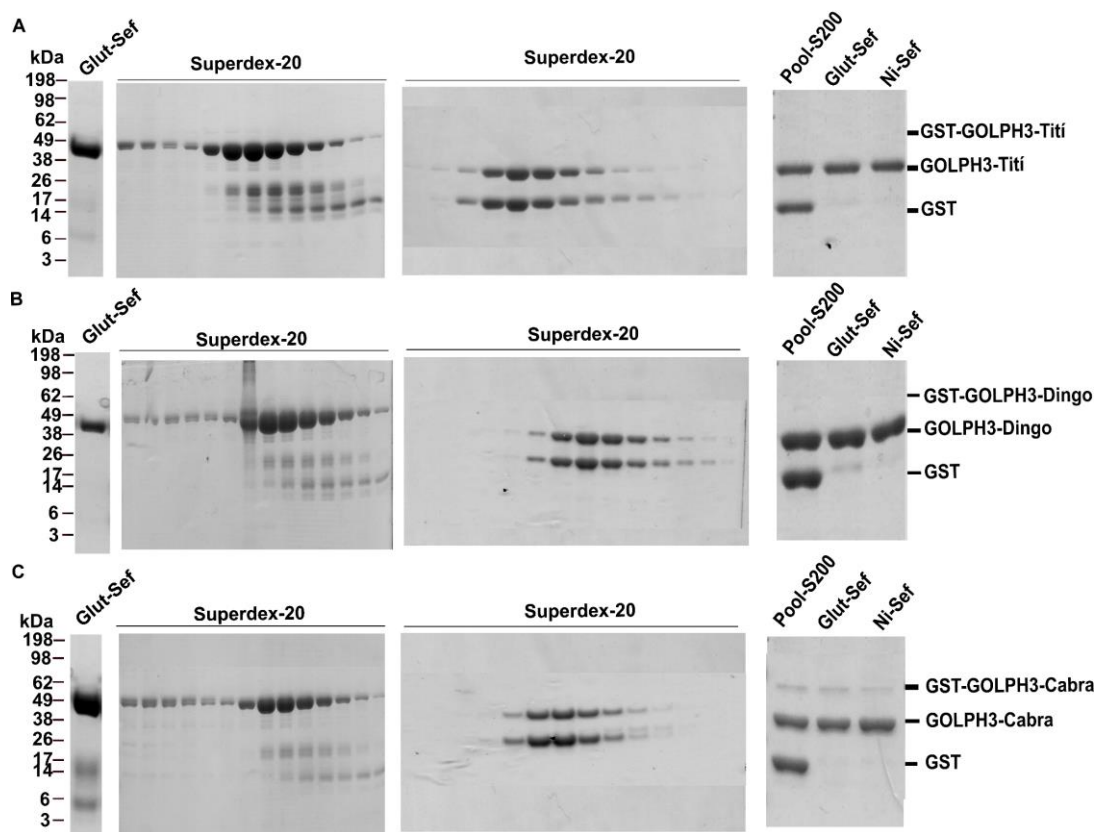


Imagen suplementaria 2. Purificación de las variantes de GOLPH3 recombinantes mediante cromatografía en tándem.

Realizado por: Galarza, 2022

Tabla S1. Números de acceso de los genes de GOLPH3 utilizados en este estudio

Gen	Nombre Científico	Nombre Común	GenBank/Ensembl ene/NCBI	Orden
GOLPH3	<i>Rhinopithecus bieti</i>	Langur negro de nariz chata	XP_017744944.1	Primates
GOLPH3	<i>Saimiri boliviensis</i>	Mono ardilla	ENSSBOG00000027748	Primates
GOLPH3	<i>Pan paniscus</i>	Bonobo	XP_024784496	Primates
GOLPH3	<i>Otolemur garnettii</i>	Galago	ENSOGAG00000014257	Primates
GOLPH3	<i>Pan troglodytes</i>	Chimpancé	ENSPTRG00000016768	Primates
GOLPH3	<i>Propithecus coquereli</i>	Sifaca de Coquerel	XP_012495980	Primates
GOLPH3	<i>Macaca fascicularis</i>	Macaco cangrejero	ENSMFAG00000045311	Primates

GOLPH3	<i>Hylobatidae</i>	Gibón	ENSNLEG00000028554	Primates
GOLPH3	<i>Rhinopithecus roxellana</i>	Mono dorado de nariz chata	ENSRROG00000036602	Primates
GOLPH3	<i>Gorilla</i>	Gorila	XP_018868895.2	Primates
GOLPH3	<i>Prolemur simus</i>	El lémur grande de bambú	ENSPSMG00000008704	Primates
GOLPH3	<i>Aotus nancymae</i>	Martejas	ENSANAG00000004528	Primates
GOLPH3	<i>Macaca mulatta</i>	Macaco	ENSMMUG00000013957	Primates
GOLPH3	<i>Callithrix jacchus</i>	Mono tití	ENSCJAG00000043003	Primates
GOLPH3	<i>Microcebus murinus</i>	Lemur ratón	XP_012593897	Primates
GOLPH3	<i>Papio anubis</i>	Babuino verde oliva	ENSPANG00000014158	Primates
GOLPH3	<i>Pongo abelii</i>	Orangután	ENSPPYG00000015373	Primates
GOLPH3	<i>Macaca nemestrina</i>	Macaco de cola de cerdo	ENSMNEG00000040145	Primates
GOLPH3	<i>Cercocebus atys</i>	Mangabey	ENSCATG00000040208	Primates
GOLPH3	<i>Chlorocebus sabaeus</i>	Vervet	XP_007959483	Primates
GOLPH3	<i>Mus Spretus</i>	Ratón del mediterraneo	MGP_SPRETEiJ_G0020577	Roedores
GOLPH3	<i>Mus Musculus</i>	Ratón	ENSMUST00000059680	Roedores
GOLPH3	<i>Peromyscus maniculatus bairdii</i>	Ratón de américa del norte	ENSPEMG00000028341	Roedores
GOLPH3	<i>Mus Pahari</i>	Musaraña	MGP_PahariEiJ_G0016060	Roedores
GOLPH3	<i>Muspicilegus</i>	Ratón estepario	ENSMSIG00000014371	Roedores
GOLPH3	<i>Heterocephalus glaber</i>	Hembra de rata topo desnuda	ENSHGLG00000003685	Roedores

GOLPH3	<i>Rattus Norvegicus</i>	Rata	ENSRNOG00000012186	Roedores
GOLPH3	<i>Urocyon parryi</i>	Ardilla	ENSUPAG00010002022	Roedores
GOLPH3	<i>Sciurus vulgaris</i>	Ardilla Roja	ENSSVLG00005009950	Roedores
GOLPH3	<i>Cavia Porcellus</i>	Cuy	ENSCPOG00000036449	Roedores
GOLPH3	<i>Octodon Degus</i>	Degu	ENSODEG00000001746	Roedores
GOLPH3	<i>Cricetulus Griseus</i>	Hamster Chino	ENSCGRG00001006205	Roedores
GOLPH3	<i>Mesocricetus Auratus</i>	Hamster dorado	ENSMAUG00000013910	Roedores
GOLPH3	<i>Ochotona princeps</i>	Pika	XP_004583690.1	Lagomorfos
GOLPH3	<i>Oryctolagus cuniculus</i>	Conejo	ENSOCUG00000011535	Lagomorfos
GOLPH3	<i>Ochotona curzon iae</i>	Pica de labios negros	XP_040833365	Lagomorfos
GOLPH3	<i>Sorex araneus</i>	Musaraña bicolor	XP_004620366	Soricomorfo
GOLPH3	<i>Condylura cristata</i>	Topo de nariz estrellada	XP_004678129.1	Soricomorfo
GOLPH3	<i>Talpa occidentalis</i>	Topo Ibérico	XP_037385417	Soricomorfo
GOLPH3	<i>Erinaceus europaeus</i>	Erizo común	ENSEEUG00000002510	Erinaceomorpha
GOLPH3	<i>Balaenoptera musculus</i>	Ballena azul	XP_036702093	Cetáceos
GOLPH3	<i>Delphinapterus leucas</i>	Beluga	ENSDLEG00000014012	Cetáceos
GOLPH3	<i>Monodon monoceros</i>	Narval	ENSMMNG00015000226	Cetáceos
GOLPH3	<i>Physeter catodon</i>	Cachalote	ENSPCTG00005010056	Cetáceos
GOLPH3	<i>Tursiops truncatus</i>	Delfín nariz de botella	XP_019785749	Cetáceos
GOLPH3	<i>Orcinus orca</i>	Orca	XP_004266079	Cetáceos
GOLPH3	<i>Globicephala melas</i>	Ballena piloto	XP_030699547	Cetáceos

GOLPH3	<i>Lipotes vexillifer</i>	Delfín Chino	XP_007453423	Cetáceos
GOLPH3	<i>Phocoena sinus</i>	Vaquita marina	ENSPSNG00000007340	Cetáceos
GOLPH3	<i>Canis lupus dingo</i>	Dingo	ENSCAFG00020017280	Carnívoros
GOLPH3	<i>Vulpes lagopus</i>	Zorro artico	XP_041604662	Carnívoros
GOLPH3	<i>Vulpes vulpes)</i>	Zorro rojo	XP_025839882	Carnívoros
GOLPH3	<i>Puma yagouaroundi</i>	Puma yagouaroundi	XP_040350715	Carnívoros
GOLPH3	<i>Panthera leo)</i>	Leones	ENSPLOG00000015195	Carnívoros
GOLPH3	<i>Panthera pardus</i>	Leopardo	ENSPPRG00000004427	Carnívoros
GOLPH3	<i>Lynx canadensis</i>	Lince de Canadá	XP_030187786	Carnívoros
GOLPH3	<i>Puma concolor</i>	Puma	XP_025780857	Carnívoros
GOLPH3	Gato	Felis Catus	XP_023095005	Carnívoros
GOLPH3	<i>Panthera tigris altaica</i>	Tigre	XP_015399880	Carnívoros
GOLPH3	<i>Mustela erminea</i>	Mustela erminea	XP_032193573	Carnívoros
GOLPH3	<i>Lontra canadensis</i>	Nutria de río	XP_032705751	Carnívoros
GOLPH3	<i>Enhydra lutris kenyon</i>	Nutria marina	XP_022348882	Carnívoros
GOLPH3	<i>Neovison Vison</i>	Visón Americano	ENSNVIG00000004457	Carnívoros
GOLPH3	<i>Mirounga leonina</i>	Elefante marino	XP_034854801	Carnívoros
GOLPH3	<i>Phoca vitulina</i>	Foca común	XP_032271129	Carnívoros
GOLPH3	<i>Leptonychotes weddellii</i>	Foca de Weddell	XP_006737357	Carnívoros
GOLPH3	<i>Halichoerus grypus</i>	Foca gris	XP_035961828	Carnívoros
GOLPH3	<i>Neomonachus schauinslandi</i>	Foca Monje	XP_021551627	Carnívoros
GOLPH3	<i>Zalophus californianu</i>	Lobo marino	XP_027461643	Carnívoros
GOLPH3	<i>Callorhinus ursinus</i>	Oso marino ártico	XP_025723221	Carnívoros

GOLPH3	<i>Ursus maritimus</i>	Oso polar	XP_040486452	Carnívoros
GOLPH3	<i>Ursus arctos horribilis</i>	Osos gris	XP_026362541	Carnívoros
GOLPH3	<i>Ailuropoda melanoleuca</i>	Panda gigante	ENSAMEG00000026105	Carnívoros
GOLPH3	<i>Pteropus vampyrus</i>	Zorro volador	ENSP VAG00000004773	Quirópteros
GOLPH3	<i>Myotis lucifugus</i>	Murciélago pequeño	ENSM LUG00000007410	Quirópteros
GOLPH3	<i>Myotis myotis</i>	Murciélago ratonero	XP_036166234	Quirópteros
GOLPH3	<i>Eptesicus fuscus</i>	Murciélago moreno	XP_008160017	Quirópteros
GOLPH3	<i>Sturnira hondurensis</i>	Murciélago de charreteras	XP_036919022	Quirópteros
GOLPH3	<i>Rhinolophus ferrumequinum</i>	Murciélago grande de herradura	XP_032966622	Quirópteros
GOLPH3	<i>Camelus dromedarius</i>	Camello árabe	XP_031299510	Artiodáctilo
GOLPH3	<i>Catagonus wagneri</i>	Jabalí	ENSCWAG0000000016 6	Artiodáctilo
GOLPH3	<i>Bos taurus</i>	Vaca	ENSBTAG000000032848	Artiodáctilo
GOLPH3	<i>Capra hircus</i>	Cabra	ENSCHIG000000003544	Artiodáctilo
GOLPH3	<i>Sus scrofa</i>	Cerdo	ENSSSCG000000031705	Artiodáctilo
GOLPH3	<i>Moschus moschiferus</i>	Ciervo	ENSM MSG00000014988	Artiodáctilo
GOLPH3	<i>Cervus hanglu yarkandensis</i>	Ciervo yarkand	ENSCHYG000000026387	Artiodáctilo
GOLPH3	<i>Odocoileus virginianus texanus</i>	Venado de cola blanca	XP_020764333	Artiodáctilo
GOLPH3	<i>Ovis aries</i>	Oveja	XP_004017118	Artiodáctilo
GOLPH3	<i>Bubalus bubalis</i>	Búfalo de agua	XP_006055870	Artiodáctilo
GOLPH3	<i>Bos indicus x Bos taurus</i>	Cruce	XP_027375798	Artiodáctilo

GOLPH3	<i>Bos indicus</i>	Cebú	XP_019838332	Artiodáctilo
GOLPH3	<i>Vicugna paco</i>	Alpaca	XP_006199703	Artiodáctilo
GOLPH3	<i>Camelus ferus</i>	Camello salvaje	XP_032326511	Artiodáctilo
GOLPH3	<i>Dasypus novemcinctus</i>	Armadillo	XP_023442149	Xenarthra
GOLPH3	<i>Choloepus didactylus</i>	Perezoso	XP_037655138	Xenarthra
GOLPH3	<i>Equus caballus</i>	Caballo	XP_001500663	Perisodáctilo
GOLPH3	<i>Ceratotherium simum simum</i>	Rinoceronte blanco	XP_004422758	Perisodáctilo
GOLPH3	<i>Equus asinus</i>	Burro	XP_014721805	Perisodáctilo
GOLPH3	<i>Loxodonta africana</i>	Elefante	XP_003407956	Afrotheria
GOLPH3	<i>Chrysochloros asiática</i>	Cape golden mile	XP_006873563	Afrotheria
GOLPH3	<i>Orycteropus afer afer</i>	Cerdo hormiguero	XP_007949797	Afrotheria
GOLPH3	<i>Vombatus ursinus</i>	Wónbat	ENSVURG00010004728	Marsupiales
GOLPH3	<i>Monodelphis domestica</i>	Colicorto gris	XP_007487565	Marsupiales
GOLPH3	<i>Sarcophilus harrisii</i>	Demonio de Tazmania	XP_031801744	Marsupiales
GOLPH3	<i>Phascolarctos cinereus</i>	Koala	XP_020850833	Marsupiales
GOLPH3	<i>Trichosurus vulpecula</i>	Zarigüeya	XP_036625171	Marsupiales
GOLPH3	<i>Ornithorhynchus anatinus</i>	Ornitorrinco	ENSOANT00000075599. 1	Monotrema
GOLPH3	<i>Tachyglossus aculeatus</i>	Equidna de hocico	XP_038599429	Monotrema
GOLPH3	<i>Trichechus manatus latirostris</i>	Manatíes	XP_004385280	Sirenios

Realizado por: Galarza, 2022