UNIVERSITAT DE BARCELONA

Identification of pTINCR: a new ubiquitin-like protein that promotes epithelial differentiation and suppresses tumor growth

Olga Boix Sánchez

UNIVERSITAT DE BARCELONA

VHIO VALL D'HEBRON Instituto de Oncología

Doctoral thesis

Identification of pTINCR: a new ubiquitin-like protein that promotes epithelial differentiation and suppresses tumor growth

Author:

Olga Boix Sánchez

Thesis directors:

María Abad, PhD and Enriqueta Felip, PhD

Programa de Doctorado en Biomedicina

Facultad de Medicina, Universidad de Barcelona

Julio 2022

UNIVERSITAT DE BARCELONA

VHIO VALL D'HEBRON Instituto de Oncología

Identification of pTINCR: a new ubiquitin-like protein that promotes epithelial differentiation and suppresses tumor growth

Thesis presented by

Olga Boix Sánchez

to qualify for the degree of Doctor in Biomedicine by the University of Barcelona

The work presented in this thesis has been performed in the Cellular Plasticity & Cancer Group at the Vall d'Hebrón Institute of Oncology (VHIO)

Dra. María Abad Méndez	Dra. Enriqueta Felip Font	Dr. Roger Gomis Cabré
Director	Director	Tutor
	DocuSigned by: Enriqueta Felip 55EA5D8D229E462...	DocuSigned by: R. Gomis 89DA32636C4C473...

Olga Boix Sánchez

Doctoranda

“Sería un error pensar que se necesita ser un héroe endurecido para lograrlo. No es así. Se necesita un corazón que esté dispuesto a morir y nacer, y morir y nacer una y otra vez.”

Mujeres que corren con los lobos

ACKNOWLEDGMENTS

Cuando llevas tanto tiempo visualizando este día, sintiéndolo tan cerca pero tan lejos, y de repente llega parece que no te lo crees. ¿De verdad estoy cerrando esta etapa de mi vida? La mezcla de sentimientos es tal que resulta muy difícil plasmarlo en papel. Tras 6 años se hace complicado intentar visualizarte haciendo otra cosa y en otro lugar. Y aunque realmente no sé bien que vendrá después ni hacia donde girará mi vida, si algo he aprendido en esta época es a relativizar las cosas y a vivir lo que tengo presente. Y si me centro en el presente, tengo muy claro que este momento no habría llegado de no ser por las personas que me han acompañado durante el viaje. Viaje que, aunque lleno de anécdotas, alegrías y MOMENTAZOS, he de decir que no ha sido para nada fácil. Son esas personas que han estado ahí y han tirado de mi cuando yo sola no podía, a las que les debo este día, este trabajo y esta etapa.

A María, porque este trabajo no existiría si no fuera por ti. Gracias por confiar en mi y en mis capacidades para iniciar esta aventura desde 0, creando algo único. Gracias por transmitirme esa pasión por la ciencia, la ilusión por cada resultado positivo y el aprendizaje en cada resultado negativo. El camino ha sido turbulento, con mejores y peores momentos. Muchas anécdotas, risas, fiestas y abrazos tremendamente sinceros; así como tardes eternas de correcciones, discrepancias y “”””discusiones””””” en tu despacho. Me llevo mucho de ti. Muchas gracias de parte de un pollo que se lanza a volar fuera del nido.

A mis compañeros de laboratorio, porque sin duda he tenido la suerte más grande del mundo trabajando con ellos, y es que las penas con alegría son siempre menos penas. De primera generación, gracias a Elena, por cada conversación científica y no científica, cada sabio consejo y cada vivencia. Porque desde el primer día robando bolis del armario de administración y teniendo un ataque de risa que supe lo que se venía contigo. Gracias por preocuparte tanto y por preguntar siempre, por ser como eres. Te echo de menos cada día y espero que me guardes un capítulo en tu futuro libro “la crisis del hombre moderno”. A Iñaki, por cada charla, confesión y discusión sobre la vida, la física y los gatos. Por estar ahí siempre, a pesar de todo, compartiendo alegrías y penas. Por toda la ayuda dentro y fuera del lab. No sabes lo mucho que ha supuesto para mi este último año saber que si lo necesitaba podía contar contigo. Desde el Vietnam más profundo y turbio, ¡gracias! A Emanuela, la tercera PhD F0 del MAL. Gracias por cada risa, anécdota y confidencia vivida. Por estar siempre dispuesta a ayudar y por aportar a este grupo esa esencia especial, con tu forma tan única de ver la vida, la ciencia y la cocina. ¡Forza e coraggio que la siguiente eres tú! Chicos, vuestras casas en Cambridge con habitación de invitados por favor.

A la segunda generación de PhDs, pero tan hijas del MAL como las que más. A Marion, la chica SUMO, por toda tu ayuda y apoyo desde tu llegada, currando y luchando por este proyecto. Desde el minuto uno que tuve muy claro la increíble mente que me había “tocado” de estudiante y, a mí que me cuesta tanto, me ha sido muy fácil confiar y delegar contigo. Espero haber sido una mentora a la altura y disculpa por alguna que otra mala cara que te hayas podido llevar indirectamente por culpa de un paper peleón. A Alba, por tu comprensión, preocupación y tu saber y querer escuchar. Porque pocas personas tienen realmente una empatía tan sincera como la tuya. Gracias por transmitirme eso y permitirme abrirme, con lo difícil que es para mí, en más de una

ocasión. Ha significado mucho. Gracias a Sandra, por traernos aire fresco, buen rollo y salseo al lab, por hacer tan buen match. També donar les gràcies a Lluís, per tota l'ajuda i calma que aportes i que aquest laboratori tant necessita. Y a Marta, por saber llevarnos con toda la paciencia del mundo, por tu sinceridad y por toda tu ayuda con este proyecto.

Sin dudarlo puedo decir que lo mejor de estos años en VHIO han sido las personas increíbles que he ido conociendo y que han sido una parte importantísima en esta etapa, a ellas también les debo un enorme gracias. Empezando por nuestros vecinos los ITAG: Noe, Imma, Jara y Albert, por salir al rescate en cada momento que lo hemos necesitado. Gracias a Alena, por las conversaciones científica tan inspiradoras, por toda tu preocupación y por cada intercambio de fotos y babas con los peques. Especialmente, a Judit, por todos los abrazos de ánimo y la calma que me transmites, por la compañía en muchísimos fines de semana. A Mery, ¡empezamos a la vez y terminamos a la vez amiga! Muchas gracias por cada momento vivido desde los inicios, cada risa y cada confidencia con unas birras (bueno, yo Coca-Cola) ¡Qué vivan esas *calçotadas*! Gracias por ser como eres, tan sincera, tan directa y tan única. A Andrea, el sol del 4th Floor. Ya te lo dije una vez, que eres un ser de luz maravilloso. Al menos a mí, verte por el pasillo, que me sonrías y me abraces me ha iluminado muchísimos días en los que ya entraba a primera hora con un nubarrón en la cabeza. Gracias compi de senescencia por cada conversación íntima, consejo y por tu forma de ver la vida tan bonita. A Anna, es increíble lo mucho que me has llegado en este último año. Gracias por soportar más de una llorera en las escaleras de emergencia, por tranquilizarme y ahuyentar ataques de ansiedad inminentes. Por hacerme ver que existe gente dispuesta a soltar las pipetas a la primera de cambio y pararse a escuchar. Muchísimo ánimo en esta última etapa que te queda, ¡qué sí se llega *xiki*! Y ambas sabemos lo que es realmente importante en esta vida. Puedes con todo y con todos.

Y aunque realmente me faltarían páginas si me pongo a agradecer todo lo que quisiera, no puedo dejar de mencionar a muchas otras personas de VHIO que me han marcado y que se han ganado un huequito en mi corazón. Gracias enormemente a Ale, por ser mi confidente más sincero de dramas entre cafés. Nos han faltado muchos, pero estoy segurísima de que tendremos todo el tiempo del mundo para hacerlos. A Faiz, por esa gran acogida, por todos los abrazos y los "Olguitas tú vales mucho" de borrachera. A Josep María, por preocuparte tanto por mi paper, mi salud mental y el tamaño de mi bíceps. ¡Hoy si me preguntas por fin podré darte buenas noticias! A Queralt, por cada conversación por los pasillos y tu felicidad contagiosa, por tu insistencia intentando convencerme de "una cerveza más y lo que surja". Y mil gracias a todas aquellas personas que me han demostrado de una u otra forma que el mundo está lleno de gente buena, dispuesta a echar un mano con la ciencia y con la vida, y que siempre han tenido un momento para un ¿qué tal? por los pasillos. Gracias Gemma, Alfonso, Tian, Sara, Enrique, Alex, Nico, Irene, Toni, Sandra, Laia, Andrea, Mireia, Beto y a tantísima gente que seguro me dejo.

Agradecer también a toda la gente que me ha ayudado científicamente durante estos años. A Carmen, por acogerme con tanto cariño en su laboratorio como a una más, enseñándome mano a mano el increíble mundo de SUMO. Gracias por cada reunión

científica y por todo lo que has aportado a este proyecto. A Santi, por todas las horas que has echado para ayudarme con este trabajo, tanto cuando estuve allí como en la distancia. Muchas gracias a ti y a Sandra, por la amistad tan bonita creada en tan poco tiempo. A Manuel Serrano y su laboratorio, por ayudarnos siempre que lo hemos necesitado. Gracias Manolo por toda tu aportación a este proyecto, por cada consejo, idea e inspiración. Gracias también a toda la gente que ha participado, colaborado y ayudado de alguna manera u otra en esta tesis.

Saliendo de Barcelona, tengo que, por supuestísimo, dar las gracias a Pablo, Miriam, Sario, Jose, Laura R, Silvia, David, Carol y Laura G por los que han sido los mejores años de mi vida. Gracias por cada risa incontrolable día a día, año tras año, bien fuera en clase, en una sala de estudio, en un examen, de fiesta o de resaca. Gracias por formar esta tan bonita familia Biotec. Os echo muchísimo de menos.

No me olvido "dels amics del cole, els de tota la vida", esas personas que, aunque no veas en mucho tiempo sabes que no van a cambiar y cuando os volvéis a ver no parece que haya pasado ni un segundo. La *terreta* no sería mi casa sin vosotros y porque cada vez que nos juntamos es sanador. Estoy segura de que llegaremos a los 60 celebrando cenas de *cabasset* los días de Roà y rezando por "los Fallen". Sois muchos como para nombraros a todos, así que un gracias general y *¡AU!* Pero en especial a Vicky, la que sin duda es mi pilar más grande e inamovible. Gracias por estar siempre ahí a pesar de la distancia, te quiero. ¡Os quiero!

Gracias a María, Aitor y Ander: psicóloga, coach y nutricionista. Por darme las herramientas que me ayudan a navegar cuando el mar está revuelto.

Un gracias muy especial a mi bebé O'Malley. Porque, aunque sé que cada día estoy más cerca de convertirme en la loca de los gatos, su cariño y compañía han conseguido mantenerme cuerda este último año.

Y ahora el gracias más grande y necesario, el que debo dar a toda mi familia por haber estado, y porque estará, siempre ahí. Gracias a mi abuela, a mi yaya y mi yayo, por tanto cariño. A todos mis tíos, tías y primos, ¡qué no son pocos! A mi cuñado, por entrar en nuestras vidas con ese buen humor y positividad, por hacer tan feliz a mi hermana. A mi sobri, el hombrecito de la casa, por toda la alegría que estás trayendo a nuestras vidas desde que llegaste. Te prometo que pronto estaré más cerca y más presente, viéndote crecer, aquí tienes tía para rato. A mi hermana, por ser el mejor ejemplo a seguir que podía haber tenido. Por tantas risas (y alguna que otra discusión) compartiendo 10m^2 de habitación, ¡te quiero *sis!* **Gracias a mis padres**, por todo el apoyo y la confianza que me han demostrado durante toda mi vida. Por la educación y los valores que me han transmitido y que me han convertido en la persona que soy. Gracias por enseñarme lo que es el sacrificio y el amor. Cada cosa que hago es para devolveros al menos un poquito de todo lo que me habéis dado.

A Barcelona. Me has costado mucho, y aunque ahora crea que no, sé que en un futuro también te daré las gracias.

RESUMEN

En los últimos años se ha demostrado que muchas regiones hasta ahora consideradas no codificantes en realidad codifican pequeñas proteínas bioactivas menores de 100 aminoácidos, denominadas microproteínas. Hasta la fecha, solo un subconjunto de éstas se ha caracterizado funcionalmente, demostrándose que desempeñan funciones esenciales en la regulación de diversos procesos biológicos, tales como la reparación del ADN, el "splicing" del ARNm o el metabolismo celular. Además, la creciente evidencia apunta hacía un papel crucial de las microproteínas también en la respuesta a estrés, como son la lesión tisular o la iniciación de tumores.

La adquisición de la identidad celular a través del proceso de diferenciación es esencial para el correcto funcionamiento de los tejidos. En los epitelios, el establecimiento de la polaridad celular juega un papel fundamental durante la morfogénesis del tejido y en el mantenimiento funcional y diferenciado de las células epiteliales adultas. Por otro lado, la pérdida de esta identidad epitelial ha demostrado ser un paso inicial crucial de la tumorigénesis, considerándose dicha polaridad celular un mecanismo clave para la supresión de tumores.

En esta tesis doctoral nuestro objetivo fue encontrar y caracterizar nuevas microproteínas con un papel importante en la identidad celular y el cáncer. Hemos identificado pTINCR, una microproteína de 87 aminoácidos codificada por el gen anotado erróneamente como no codificante *TINCR*. pTINCR es una microproteína conservada evolutivamente que se expresa en piel y en otros tejidos epiteliales. Mediante estudios de ganancia y pérdida de función, hemos demostrado que pTINCR es un regulador clave de la diferenciación de células epiteliales *in vitro* y promueve la diferenciación epidérmica *in vivo*. Además, el supresor de tumores p53 aumenta la regulación de pTINCR tras el daño celular, efecto que resulta necesario para la respuesta de diferenciación inducida por daño al ADN. En consonancia con su función supresora de tumores, la expresión de pTINCR se pierde en los carcinomas de células escamosas cutáneas (cSCC) humanos y su sobreexpresión reduce el crecimiento de xenoinjertos derivados de pacientes de cSCC. Además, la expresión de *TINCR* se correlaciona con un mejor pronóstico en varios cánceres epiteliales.

A nivel molecular, pTINCR es una nueva "Ubiquitin-like protein" (UBL) capaz de unir SUMO a través de su motivo de interacción a SUMO (SIM), y regular así la SUMOilación celular. Adicionalmente, hemos identificado dos interactores de pTINCR, NONO y CDC42, y demostramos que pTINCR modula su SUMOilación. Por un lado, probamos que pTINCR desencadena un patrón de "splicing" alternativo relacionado con las células epiteliales, probablemente a través de su interacción con NONO, proteína nuclear implicada en la regulación transcripcional y el "splicing" del ARNm. Por otro lado, nuestros resultados demuestran que pTINCR promueve la activación de CDC42, una Rho-GTPasa crítica para la remodelación del citoesqueleto de actina y el establecimiento de la polaridad celular, desencadenando una cascada de eventos que dan lugar la diferenciación epitelial.

Nuestros resultados amplían aún más el conocimiento sobre el campo aún inexplorado de las microproteínas y sugieren que el microproteoma oculto en los ARN no codificantes puede ser una fuente de nuevos reguladores de la identidad celular relevantes para el cáncer

SUMMARY

Recent advances in computational analyses, peptidomics and ribosome profiling have revealed that many regions previously annotated as non-coding are in fact translated into a myriad of bioactive proteins, largely overlooked until now. These small proteins, shorter than 100 amino acids, are called microproteins, micropeptides or SEPs (from sORF-encoded peptides). To date, only a subset of them have been functionally characterized, and they have been shown to play essential functions regulating a plethora of fundamental processes such as DNA repair, RNA splicing, or cell metabolism. Importantly, mounting evidence suggest that microproteins also have a role in response to stress, such as oxidative stress, viral and bacterial infection, tissue injury and even tumor initiation.

The acquisition of cell identity by cell differentiation is essential for proper tissue function. In epithelial tissues, cell polarity plays a pivotal role during tissue morphogenesis and in the maintenance of the specialized functions and differentiated status of epithelial cells. On the other hand, loss of epithelial cell identity is nowadays considered an essential initial step in tumorigenesis, and epithelial cell polarity has been positioned as a key tumor suppressive mechanism in epithelial cancers.

In this doctoral thesis, we aimed to find and characterize novel microproteins with an important role in cell identity and cancer. We have identified pTINCR, an 87-amino acid microprotein encoded by *TINCR*, a gene misannotated as a long non-coding RNA. pTINCR is an evolutionary conserved microprotein expressed in skin and several epithelial tissues. By gain- and loss-of-function studies, we have demonstrated that pTINCR is a key regulator of epithelial cell differentiation *in vitro* and promotes epidermal differentiation *in vivo.* Additionally, pTINCR is upregulated upon cellular damage by the tumor suppressor p53 and it is required for the DNA damage-induced differentiation response. Consistent with its tumor suppressive role, pTINCR expression is lost in human cutaneous squamous cell carcinomas (cSCCs), and its overexpression reduces malignancy in cSCC patient-derived xenografts (PDXs). Moreover, the expression of *TINCR* correlates with better prognosis in several epithelial cancers.

At the molecular level, pTINCR is a novel ubiquitin-like protein (UBL) which can bind to SUMO through its SUMO interacting motif (SIM) and regulate SUMO conjugation. Importantly, we have identified two pTINCR binding partners, NONO and CDC42, and showed that pTINCR enhances their SUMOylation. NONO is a nuclear protein involved in transcriptional regulation, pre-mRNA splicing and nuclear retention of defective RNAs. We have shown that pTINCR triggers a specific splicing program related to epithelial cells, probably through its interaction with NONO. On the other hand, CDC42 is a Rho-GTPase critical for actin cytoskeleton remodeling and the establishment of cell polarity. We demonstrate that pTINCR binds to CDC42 and promotes its SUMOylation and activation, triggering an epithelial pro-differentiation cascade.

In summary, we have identified a novel UBL, pTINCR, that regulates epithelial differentiation and has tumor suppressor activity. Our results further expand our knowledge on the yet unexplored field of microproteins and suggest that the microproteome hidden in previously assumed non-coding RNAs can indeed be a source of new regulators of cell identity relevant for cancer.

ABBREVIATIONS

3D	Three Dimensional
ActD	Actinomycin D
alt-ORF	Alternative Open Reading Frame
AS	Alternative Splicing
bp	Base Pair
BrdU	Bromodeoxyuridine
CEEA	Comité Ético de Experimentación Animal
CHX	Cyclohexamide
circRNA	Circular RNA
cSCC	Cutaneous Squamous Cell Carcinoma
DDR	DNA-Damage Response
DNA	Deoxyribonucleic acid
dORF	Downstream Open Reading Frame
DOX	Doxycycline
DOXO	Doxorubicin
DSC	Desmocollin
DSG	Desmoglein
DSP	Desmoplakin
DTT	Dithiothreitol
ECM	Extracellular Matrix
EGF	Epidermal Growth Factor
EMT	Epithelial-to-Mesenchymal Transition
epSC	Epithelial Stem Cell
ER	Endoplasmic Reticulum
ERK	Extracellular signal-regulated kinase
ESC	Embryonic Stem Cell
FBS	Fetal bovine serum
FC	Fold Change
FPKM	Fragments Per Kilobase
GAP	GTPase-activating proteins
GDI	Guanine nucleotide dissociation inhibitors
GDP	Guanosine-5'-diphosphate
GEF	Guanine nucleotide exchange Factor
GO	Gene Ontology
gRNA	guide RNA
GSEA	Gene Set Enrichment Analysis
GST	Glutathione S-transferase
GTP	Guanosine-5'-triphosphate
HR	Hazard Ratio
iPSC	Induced Pluripotent Stem Cell
kb	Kilobase
kDa	Kilo-Dalton
KO	Knock-Out
KRT	Keratin

LC-MS	Liquid Chromatography-Mass Spectrometry
LIF	Leukemia Inhibitory Factor
LLPS	Liquid-liquid Phase Separation
lncRNA	Long non-coding RNA
LPS	Lipoploysaccharide
MAPK	Mitogen-activated protein kinase
MEF	Mouse Embryonic Fibroblast
MFE	Minimum Free Energy
mRNA	Messenger RNA
MS	Mass Spectrometry
NES	Normalized enrichment score
ORF	Open Reading Frame
P/S	Penicillin / Streptomycin
PBS	Phosphate-buffered saline
PDX	Patient-Derived Xenograft
PEP	Posterior Error Probability
PG	Plakoglobin
PI	Propidium iodide
PKP	Plakophilin
PSI	Percent Spliced Index
RE	Recycling Endosomes
RLU	Relative Light Units
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
RT-qPCR	Real time quantitative polymerase chain reaction
SCC	Squamous Cell Carcinoma
SD	Standard Deviation
SDS	Sodium Dodecyl Sulfate
SEM	Standard Error of the Mean
SENP	SUMO-specific proteases
SEP	sORF-Encoded Protein
SIM	SUMO Interactive Motif
smORF	Small/short Open Reading Frame
SNP	Single Nucleotide Polimorfisim
TCGA	The cancer genome atlas
TGN	Trans-Golgi Network
UBL	Ubiquitin-Like
uORF	Upstream ORF
UTR	Untranslated region
UV	UltraViolet
WB	Western blot
WT	Wild Type

INDEX

A new breakthrough for an old dogma

In September 1957, in a lecture at the Society for Experimental Biology in London, it was stated for the first time the central dogma of molecular biology: *"Once information has got into a protein it can't get out again"* [1]. This sentence settled by Francis Crick determined that genetic information passes in a one and only direction and it cannot be transferred back. At that time, the oversimplified thought that "DNA makes RNA, and RNA makes protein" conceived only three types of "genetic flows", starring DNA to DNA (replication), DNA to RNA (transcription) and RNA to protein (translation). This idea narrowed how scientists understood life for more than a decade. It was not until 1970 when this dogma was broken due to the discovery in retroviruses by David Baltimore of the reverse transcriptase enzyme [2]. This finding revolutionized the molecular biology field, providing experimental evidence of genetic information flowing also in the direction RNA to DNA. Skepticism aside, the process of reverse transcription was finally accepted to be an essential biological process, not only in retroviruses but also in eukaryotic cells. However, the idea of genetic information also "flowing" in any other different direction remained heretical until 1982, when Stanley Prusiner coined the "prion" concept [3]. The existence of a protein-based but nucleic acid-independent infectious agent with the ability to self-replicate, store and transmit (conformational) genomic information challenged once again an already outdated dogma.

The most underrated molecule of all has been perhaps the RNA, overshadowed since the 1960s by the also widely held dogma of "the junk DNA". The fact that more than 95% of our genome did not code for proteins kept RNA behind the scenes for many decades, simply considered a bridge between DNA and proteins. It was not until 2012, after the Encode research project announced that about 80% of the human genome do transcribe functional non-protein coding RNAs [4], when researchers started to understand the complete biological significance of the RNA molecule. Luckily, today we know that RNA actually does much more than just deliver genetic information: it regulates gene expression (riboregulators), acts as enzymes (ribozymes), behaves as structural molecules (ribonucleoproteins) and even actively transfers genetic information (RNA-containing extracellular vesicles).

As knowledge evolve unstoppably, more established "dogmas" are becoming updated. In 1996, a study in soybean identified two small peptides encoded by an evolutionary conserved long non-coding RNA, emerging the novel concept of "micropeptide" [5]. Decades later, the development of high-throughput technologies has confirmed a surprising truth: many of those assumed non-coding RNAs can in fact code for small functional proteins. This largely overlooked biological underworld of small proteins has revolutionized once again the field of molecular biology, demonstrating that the genome complexity is far greater than we previously thought.

You may agree, then, that there are very few (if none) unalterable principles in science and that the questioning of old established dogmas is what has (and still will) led to incredible new breakthroughs.

INTRODUCTION

1. Small ORF-encoded microproteins

1.1. The hidden microproteome

The laborious task of genome annotation of different model organisms began back in the 90's. Classically performed by bioinformatic algorithms, novel putative protein-coding genes started to be predicted, identified and annotated. To ensure the identification of meaningful open reading frames (ORFs) within the genome, the annotation process relied on the features of already known genomic sequences and, given the magnitude of the task, certain filters were assumed when annotating coding sequences. As an instance, scientists at that time realized that setting up a 300-bp (100-amino acids) cut-off drastically reduced the noise of false positives predictions [6, 7]. However, this *ad hoc* cut-off of 100-amino acids had the considerable disadvantage of being deliberately underestimating small ORFs that do code for functional proteins. Moreover, ORF annotation was strained by the assumptions that eukaryotic genes were monocistronic and that translation should start with an ATG codon. During the past two decades, and thanks to the development of advanced computational tools and OMICS technologies, we have discovered that there are thousands of non-canonical ORFs located in previously assumed non-coding regions that indeed code for small active proteins [8-11]. This surprising discovery has opened a new level of biological complexity.

A frequent challenge in any emerging field is to stablish a consensus nomenclature. In the case of these non-canonical ORFs, a broadly accepted classification is done according to their genomic location (**Fig. 1**). Generally, the ORFs located in previously assumed non-coding transcripts and smaller than 100 amino acids are categorized as short or small open reading frames (sORFs/smORFs). Additionally, non-canonical ORFs found in mRNAs are named alternative ORFs (alt-ORFs) when overlapping the canonical ORF but presenting different amino acid sequence; upstream ORFs (uORF) when located in 5'UTR sequences; and downstream ORF (dORFs), when located in 3'UTRs [12-14]. The protein products of these non-canonical ORFs are named micropeptides, microproteins or sORFs-encoded peptides (SEPs).

Importantly, the origin of micropeptides is different from other small bioactive peptides such as natural peptides, neuropeptides, peptide hormones or antimicrobial peptides, which are usually encoded as larger precursor proteins and cleaved posttranslationally.

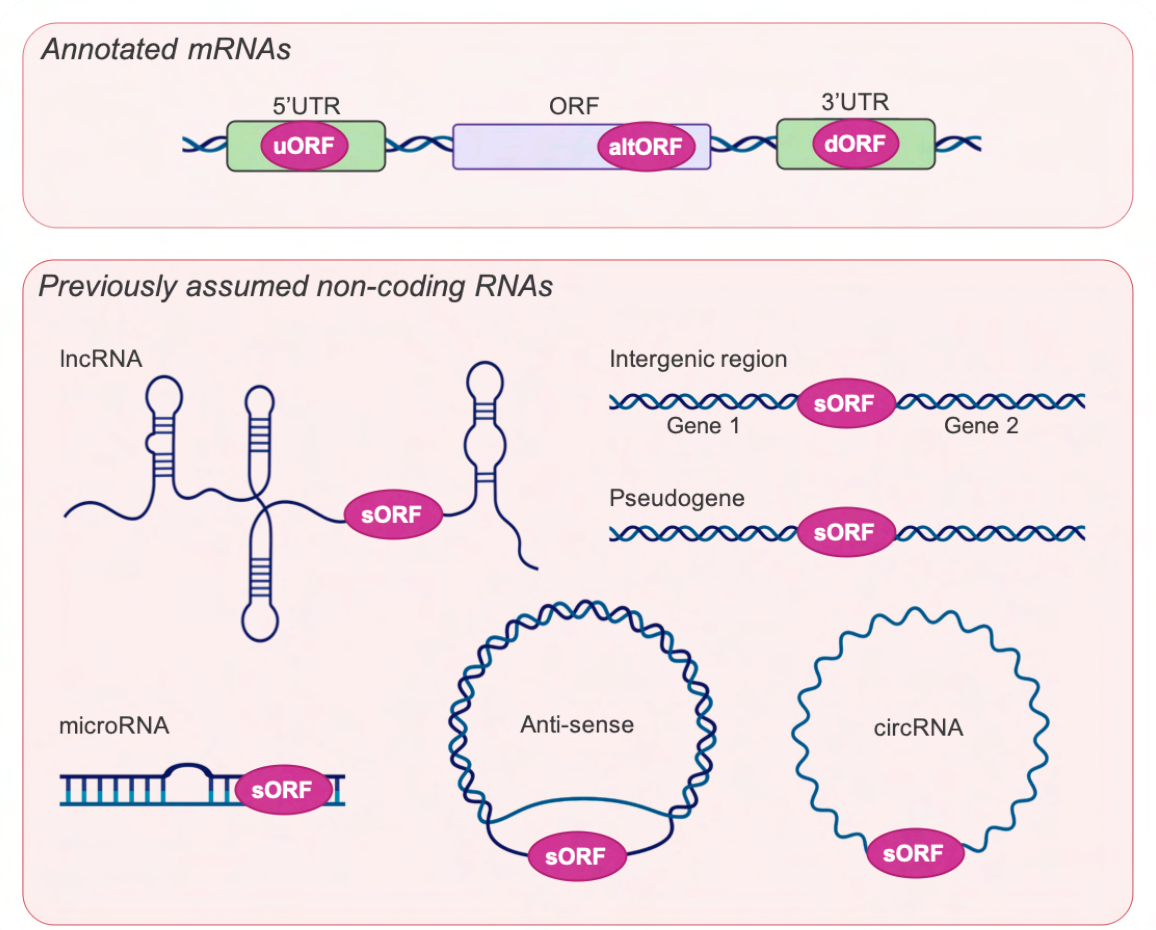


Figure 1. Classification of non-canonical ORFs.
sORFs are smaller than 100 codons and located in: 1) previously annotated mRNAs (upper panel) overlapping the main-ORF (alt-ORFs) or in UTRs, named upstream ORFs (uORFs) when located in the 5′ UTR or downstream ORFs (dORFs) when located in the 3′ UTR; 2) previously assumed non-coding RNAs (lower panel) such as lncRNAs, microRNAs, intergenic regions, pseudogenes, anti-sense RNAs or circRNAs

1.2. Methodology to identify non-canonical ORFs and microproteins

The discovery of non-canonical ORFs has generated the necessity of re-evaluating the coding potential of all known genomes. At present, there are several methods for the identification of non-canonical ORFs and microproteins, which can be separated into computational or experimental approaches (**Fig. 2**).

1.2.1. Computational approaches

Some good indicators of functionality within a ORF are 1) the conservation at the amino acid level across evolution, and 2) the presence of motifs or domains described in previously identified functional proteins. Based on these criteria, bioinformatic algorithms have been developed for the identification of microprotein-coding sORFs, which usually present low level of nucleotide sequence conservation. An example of an algorithm that operates using phylogenetic conservation and codon substitution frequencies is PhyloCSF, a powerful tool that can be integrated into the UCSC Genome Browser for

free and easy access [15]. Other additional features can be added to evaluate computationally potential sORFs, such as the presence of a Kozak sequences (motifs that signal for protein translation initiation in most eukaryotic mRNAs) or consensus N-terminal signal peptides of protein localization. This approach, however, overlooks "young" sORFs coming from *de novo* coding genes, given their lack of evolutionary conservation.

1.2.2. Experimental approaches (OMICS)

1.2.2.1. Translatomics

Translatomics refer to the study of all open reading frames being actively translated in a cell or organism, and the collection of those ORFs is known as the translatome. In the context of the "microtranslatome", ribosome profiling (or Ribo-seq) has become a common tool to search for novel encoding sORFs. Ribo-seq technique is based on the measurement of ribosome-protected RNA fragments, which are isolated, sequenced and afterwards analyzed using computational algorithms such as RibORF, that scores translatable ORFs based on the frequency of pausing and the uniformness of the ribosomes across codons [16]. This methodology has been used to expand and explore the emerging field of microproteins, but also has provided evidence for the use of non-AUG start codons and alternative stop codons [17, 18].

Evidently, ribosome occupancy alone does not necessarily confirm the coding potential of the identified region since translation requires the binding to RNA of many other factors besides ribosomes. Therefore, the ability of Ribo-seq to identify microproteins is limited and usually requires complementary methods for further confirmation.

1.2.2.2. Proteomics

Proteomics refers to the large-scale study of the proteome, being mass spectrometry (MS) the gold-standard technique for the detection and quantification of the proteins present in a biological sample. However, MS-based strategies have required a prior optimization and modification in order to be implemented in the field of microproteins. Technically, MS analyzes the specific amino acidic spectra collected after enzymatic digestion of proteins by aligning those spectra to databases that contain information of already known proteins. However, most microproteins remain nonannotated and, hence, they cannot be identified following common past protocols. To overcome this, customized databases including microproteins have been designed. Based in RNA-Seq data, scientists have generated *in silico* the potential "microproteome" from specific annotated transcriptomes and it has been used as a custom database to successfully identify microproteins by proteomics [19, 20]. However, microproteins have challenged the MS methodology in many other ways. For instance, digested microproteins are rapidly degraded due to the lack of protective secondary structures, being most of them lost during the experimental process. Moreover, due to their short length their abundance is very low compared to bigger proteins and the number of unique tryptic peptides useful to distinguish them from other proteins is reduced, strongly limiting their identification. Nevertheless, given that MS provides a direct evidence of protein existence and additional information about posttranslation modifications, it has emerged as a very

powerful and appealing method for the discovery of microproteins and massive advances are being made in order to adapt it to this new era [21].

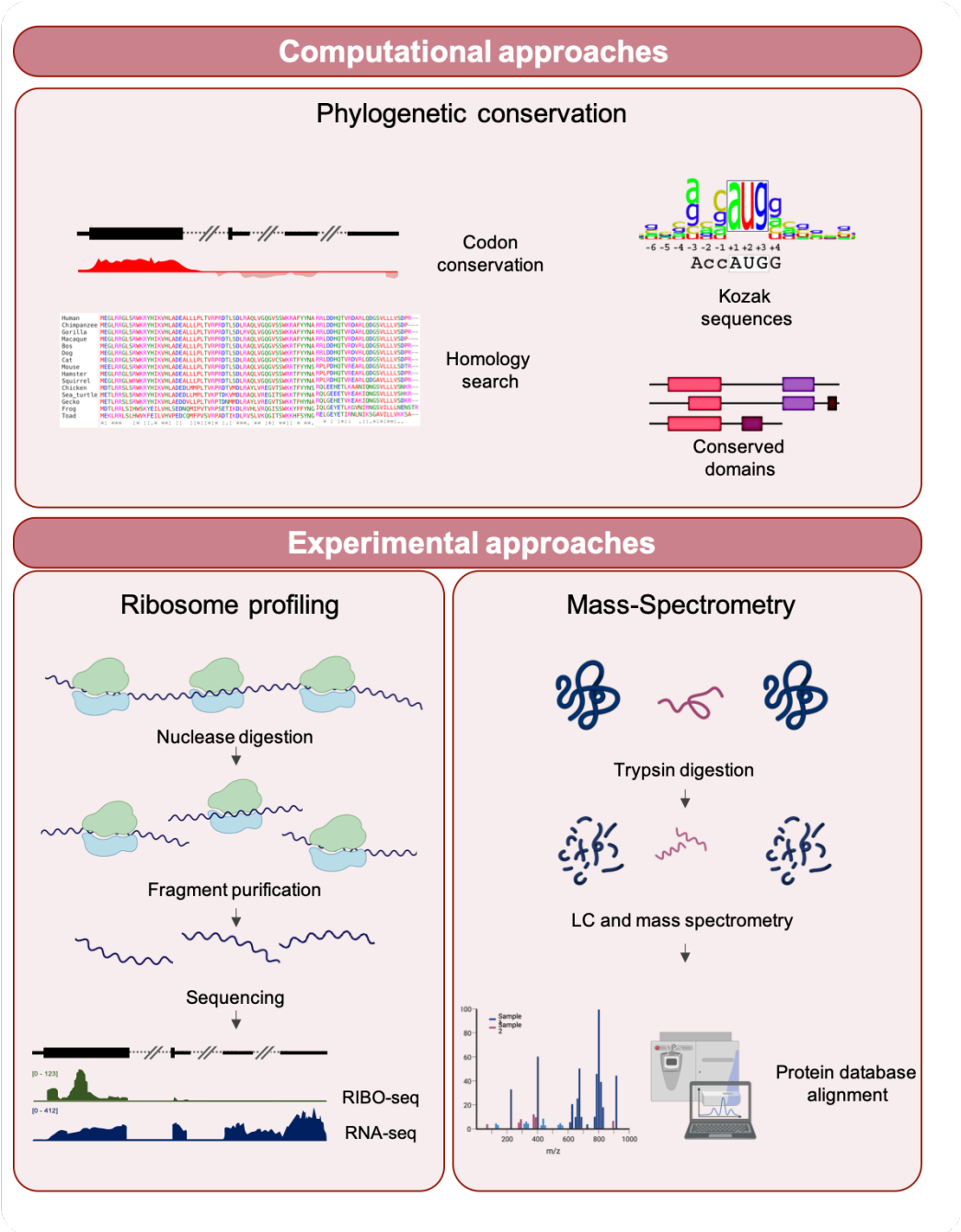


Figure 2. Methods for the identification of microproteins.
Methods for microprotein discovery can be divided into computational or experimental approaches. Computational approaches evaluate phylogenetic codon conservation and analyze some characteristics of protein-coding genes, such as the presence of Kozak sequences or conserved domains. Experimental approaches render direct evidence for protein expression, either by the evaluation of active translation of transcripts using Ribo-Seq or by direct detection and identification of the microproteins by mass spectrometry analysis.

Undeniably, microproteins represent a huge new field that requires equally huge efforts to be explored and understood. Each technique described above presents its own advantages and disadvantages. While codon conservation strongly highlights biological relevance in a cost-effective way, it does not provide sufficient evidence about real translation. On the other hand, experimental approaches do provide direct evidence of ORF translation, but this is not always an indicative of functionality. Therefore, likely the best strategy is the combination of various approaches by integrating bioinformatics prediction, translatomics support and proteomics confirmation. Besides, none of these methods evaluates the biological function of microproteins, and further functional and molecular characterization is always needed case by case.

Despite that to date only a small fraction of microproteins have been identified and functionally characterized (from Arabidopsis Thaliana or Drosophila to humans), the scientific community is already aware that these small sized bioactive molecules bear unique features that make them critical in several cellular processes, both in physiological and pathological scenarios.

1.3. Microproteins' Functions

1.3.1. Microproteins in physiological processes

Because of their reduced size, microproteins present certain functional limitations, including those that would involve enzymatic activities. However, this feature makes them ideal for other purposes, such as fine-tuning fundamental processes by binding and regulating bigger protein complexes.

For example, microproteins have been related to DNA and RNA regulation. The MRI-2 microprotein stimulates non-homologous end joining (NHEJ) by interacting with Ku heterodimeric proteins during DNA-repair [22], while CYREN microprotein exhibits the opposite function [23]. NoBody microprotein participates in mRNA decapping by directly regulating decapping proteins localized into P-bodies, and HOXB-AS3 micropeptide modifies RNA splicing by antagonizing the hnRNP A1-mediated regulation of pyruvate kinase M (PKM) [24, 25].

The regulin microproteins family exemplifies the importance of microproteins also in cellular metabolism and muscle performance. Myoregulin, Phospholamban, Sarcolipin, Endoregulin, Another-Regulin and DWORF interact with the Calcium-ATPase SERCA and regulate intracellular calcium dynamics in a tissue specific manner [26-28]. Similarly, pTUNAR microprotein guides neural differentiation and neurite formation through the regulation of intracellular calcium, also possibly by interacting with SERCA2 [29]. Other examples are Mitoregulin and MOXI, both regulating mitochondrial metabolism and energy homeostasis by stabilizing mitochondrial respiratory supercomplexes and enhancing fatty acid β-oxidation [30-32].

Some other curious examples could be the uORF within the 5'UTR of ASNSD1 mRNA named ASDURF, which is assembled into the chaperone complex PAQosome and participates in the biogenesis of several human protein complexes [33]; Kastor and Polluks polypeptides, which cooperatively regulate mitochondrial sheath formation and

spermatogenesis [34]; or Cdk2ap1$^{\Delta N}$, an N-terminally truncated protein variant driven by an alternative promoter controlled by a retrotransposon that mediates preimplantation embryo development in mouse [35].

1.3.2. Microproteins in cellular stress and disease

Mounting evidence highlights the importance of microproteins during stress response [36-38]. Some particular examples can be SPAR, which regulates the skeletal muscle regeneration following injury [39]; PIGBOS, a microprotein located in the endoplasmic reticulum (ER) that acts on the unfolded protein response [40]; or Humanin, a microprotein widely expressed in several human tissues with several properties such as neuroprotector, antiaging, antiapoptotic, anti-inflammatory and antifibrogenic [41-45].

A rapid response and adaptation to unfavorable scenarios is critical for the restoration of cellular homeostasis upon stress, and translational regulation is described as a more effective mechanism for this purpose than transcriptional regulation. Importantly, microproteins may be important regulators of the stress response. Diverse stimuli (including hypoxia, nutrient starvation or oxidative stress) lead to a translational reprogramming that reduces the levels of overall protein synthesis [46] while allowing the translation of a subset of stress-response proteins that promote cell survival. As an example, uORFs have been seen to become translated upon certain stimuli to subsequently regulate the translation of downstream cistrons mRNAs to further cope with the stress [47-49]. Similarly, microproteins have been implicated in infection and innate immunity responses, with some studies reporting the upregulation of sORFs upon influenza virus-infection [50] or in lipopolysaccharide (LPS)-treated mouse macrophages [51], enhancing host-protection. This mechanism of translational regulation has also been postulated as crucial during tumor initiation, a process in which an unexpected shift to unconventional sORF translation has been reported [52]. This finding and others undoubtedly point to a key participation of microproteins in tumorigenesis.

1.3.3. Microproteins in cancer

Compelling evidence demonstrate that microproteins play important roles in the development of different human tumors. Indeed, microproteins have been associated to most of hallmarks of cancer [12]. Their specific molecular functions on proliferation, metabolism, cell motility, apoptosis or angiogenesis (among others), postulate them as potential novel “oncogenes” or tumor suppressors. A summary of micropeptides with their described functions in cancer is presented in **Figure 3**.

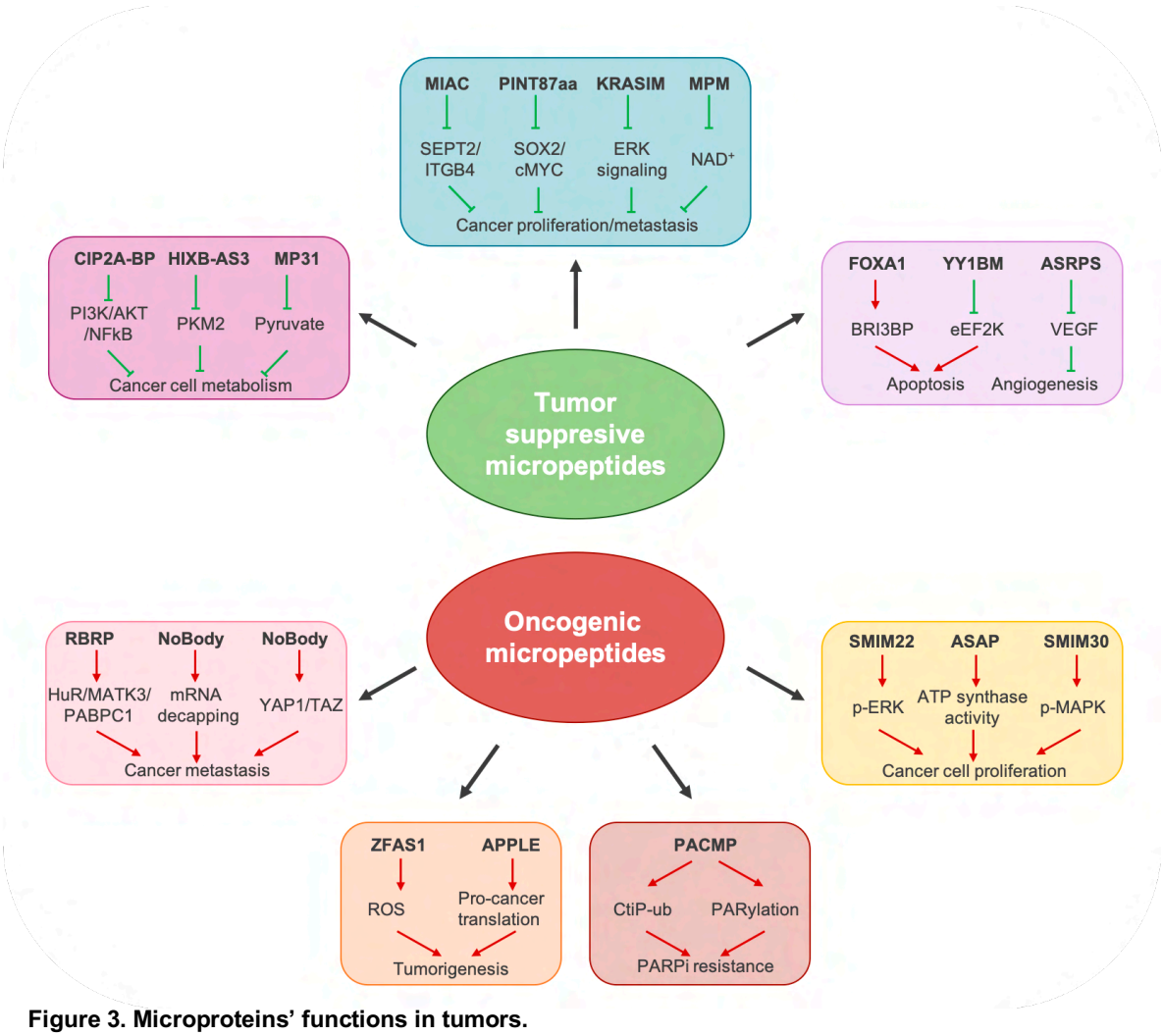


Figure 3. Microproteins' functions in tumors.
Selected examples of microproteins and their molecular mechanism in different cancer phenotypes, including proliferation, apoptosis, angiogenesis, metastasis and therapy resistance. *Top*: functions and mechanisms of the tumor suppressive microproteins. *Bottom*: functions and mechanisms of oncogenic microproteins.

Despite the limited current understanding of the role of microproteins in cancer, they represent an exciting goldmine for understanding and combating this complex disease. Characterized microproteins have shown to be fine-tuning regulators of specific pathways in cancer cells (**Fig. 3**). Moreover, given their small size is easy to visualize scenarios in which microproteins may act masking regulatory sites, inducing conformational changes on other proteins or even mimicking certain binding domains. These features make them particularly suitable for the future development of targeted-based therapies. Moreover, sORFs are usually dysregulated at the transcript level in cancer tissues, having their encoded micropeptides promising clinical diagnostic and prognostic value [53]. In fact, although little is known about secreted micropeptides and their levels in body fluids, some studies are highlighting their possible used as specific and cost-effective biomarkers [54-56].

Beyond their functions and expression, genetic variations in sORF regions and their association with disease have also awaken great interest among researchers. Some

studies have revealed that mutations occurring within non-coding regions can act as driver mutations in cancer [57, 58], evidence that could be extrapolated to their possible encoded-microproteins. In the same line, microproteins have been demonstrated to be a great new source of tumor-specific antigens (neoantigens) [59-61], increasing the number of possible targetable epitopes for cancer vaccination and immunotherapy.

In conclusion, microproteins have been revealed as a new class of bioactive molecules that regulate a plethora of cellular processes. The discovery of the microproteome has expanded our vision on the coding capacity of the genome, adding a new biological layer of physio-pathological complexity and having a huge impact on basic, translational, and clinical research.

2. Epithelial tissues

Epithelia are one of the four types of existing animal tissues, along with the connective, muscle and nervous tissues. Cells within epitheliums are tightly packed and linked, forming one or more continuous layers that line the internal and external body surfaces (surface epithelium) and secreting organs (glandular epithelium). Epithelial tissues have a variety of specialized functions, mainly related with protection, secretion, absorption, transportation, and sensory reception. They can be classified according to the shape of the cells forming the tissue or by the number of layers present in the tissue (**Fig. 4**). Regarding epithelial cell shape, there are three types: squamous (flat, thin and sheet-like appearance), cuboidal (boxy-like appearance) or columnar (rectangular appearance). On the other hand, there are two principal types of tissue arrangements: simple epithelium (a single layer of cells) or stratified epithelium (a multilayer of cells). Additionally, in some tissues a single layer of irregularly shaped columnar cells can be found giving the appearance of a multi-cellular layer, this type of tissue is called pseudostratified. Altogether, the combination of the different shapes and arrangements gives rise to all types of epithelial tissues in our body.

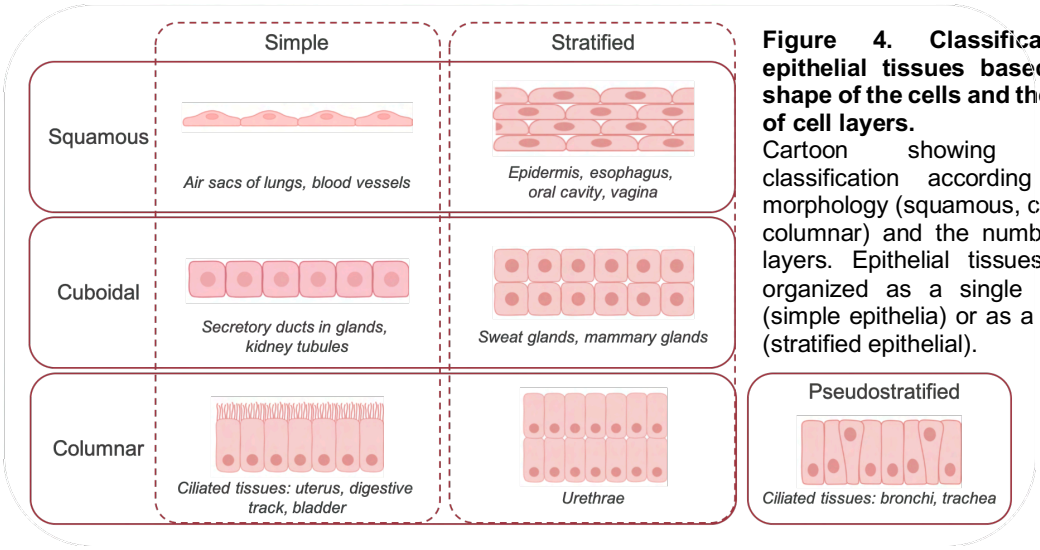


Figure 4. Classification of epithelial tissues based on the shape of the cells and the number of cell layers. Cartoon showing epithelia classification according to cell morphology (squamous, cuboidal or columnar) and the number of cell layers. Epithelial tissues can be organized as a single cell layer (simple epithelia) or as a multilayer (stratified epithelial).

2.1. Epithelial cell identity

Epithelial cells are the building blocks of epithelial tissues. They are typically characterized by having an asymmetric distribution of membrane-bound proteins, cytoskeletal components and organelles. This intrinsic feature, called cell polarity, defines three regions in their surface structure with very different functions: the apical surface, the basal surface and the lateral surface.

The apical region is located towards the lumen of the organ which epithelium covers or towards the external environment. This surface often contains specialized structures that alter its shape, determining the function of the tissue. For example, microvilli present in intestinal epithelial cells increase the surface area of absorption and fluid transport; or cilia present in epithelial cells of the respiratory tract, which are motile structures that facilitate the transport of substances. The basal region is the bottom edge of the cell in close contact with the underlying basement membrane or the extracellular matrix (ECM), separating epithelial cells from the surrounding connective tissue. The lateral region is located on the sides of the cell, where connections with neighboring cells are stablished. Given that both lateral and basal regions act as surfaces for interaction with surrounding structures, they are usually described together as the basolateral region.

The establishment of this polarity is an essential part of the differentiation program that determines cell behavior and tissue specialization [62]. This highly complex process involves a variety of molecules and molecular mechanism dynamically interplaying with each other.

2.1.1. Intercellular junctions

Cell-cell and cell-ECM contacts are crucial for the maintenance of a differentiated epithelium. In epithelial cells, cell-to-cell adhesion marks and initiates the establishment of the lateral surface, which leads to the subsequent specification of the basal and apical surfaces by the formation of the whole repertoire of tidily compact intercellular connections that maintains the tissue integrity. These intercellular junctions allow the formation of both, extracellular connections between adjacent cells and intracellular connections with the cytoskeleton. In differentiated epithelial tissues, these intercellular connections comprise tight junctions, adherens junctions, desmosomes and gap junctions [63, 64] (**Fig. 5**).

Tight junctions are the most apical intercellular junctions in mammalian epithelial cells. They are formed by molecules that engage and seal the epithelium, impeding the transmembrane diffusion of substances and fluids and creating a selective barrier. These molecules include transmembrane proteins (such as occludins, claudins and junction adhesion molecules), which bind to the actin cytoskeleton through peripheral membrane proteins (such as ZO-1). Tight junctions are critical to preserve cell polarity since they prevent the lateral diffusion of membrane proteins between the apical and basolateral surfaces [65, 66]. Basally to tight junctions are adherens junctions, which ensure the folding and shape of epithelial tissues. They are formed by cadherin transmembrane proteins (such as E-cadherin) that associate with cytoplasmic catenins (such as β-catenin) and join the actomyosin cytoskeleton. In addition to be a physical linkage

between cells, adherens junctions also participate in the transduction of cellular signals and regulation of gene transcription [67, 68]. Finally, some tissues can also present desmosomes, which contain cadherin-like proteins (such as Desmogleins) and Desmocollins) bounded to linker proteins (such as Desmoplakin), which associate intracellularly to keratin intermediate filaments. They provide strong cell-to-cell adhesion, ensure tissues mechanical strength and mediate cell signaling. This type of junctions is particularly abundant in epidermis [69].

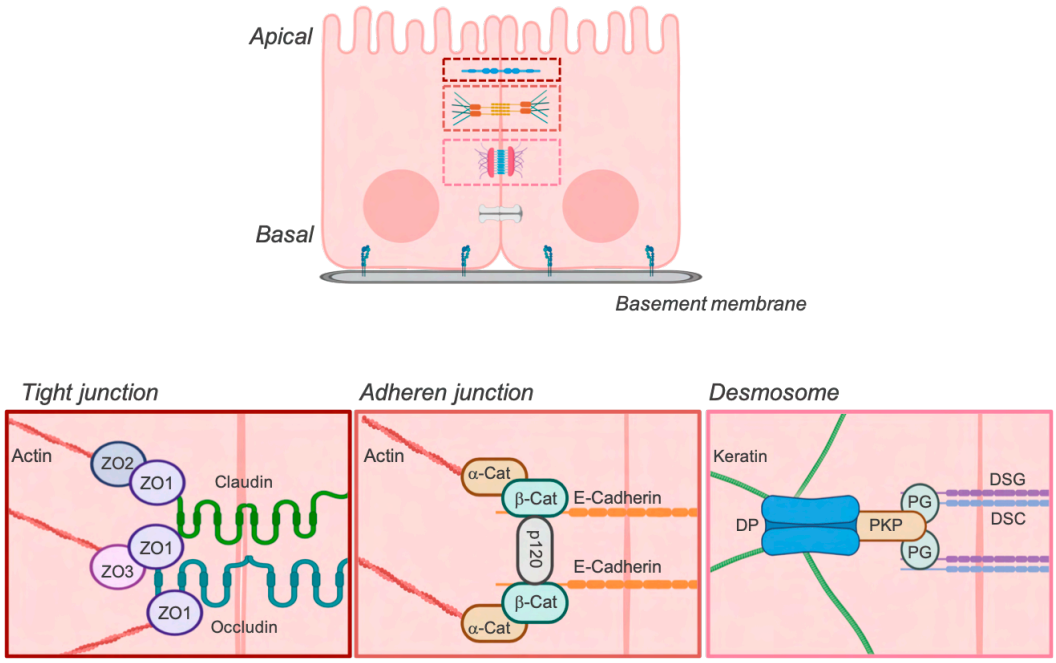


Figure 5. Intercellular junctions of epithelial cells.
Tight junctions (dark red*)* form a belt at the apical pole of the cell, sealing cells together and creating the epithelial barrier. They consist of transmembrane proteins (claudins and occludins) scaffolded by cytoplasmic occludens proteins (ZO1, ZO2 and ZO3), which are linked to actin cytoskeleton. *Adherens junctions* (pink*)* support cell-cell adhesion and regulate signal transduction. Classic cadherins (such as E-cadherin) are recruited to the intercellular junction through their binding to p120 catenin and β-catenin (β-Cat), which associate with the actin cytoskeleton through α-catenin (α-Cat). *Desmosomes* (light pink*)* form a very strong connection between neighboring cells. They are composed by members of the cadherin family (Desmogleins (DSGs) and Desmocollins (DSCs)) linked to Plakoglobin (PG) and Plakophilins (PKPs), which directly bind keratin intermediate filaments through Desmoplakin (DP).

2.1.2. Protein sorting and trafficking

Following intercellular junction formation, differentiated epithelial cells also develop a selective packaging, sorting, and transport machinery to preserve the polarized distribution of membrane-related proteins [70, 71]. This spatial distribution of proteins (and phospholipids) is regulated by two principal sorting machineries, the trans-Golgi Network (TGN) and the recycling endosomes (RE). Newly synthesized membrane proteins are transported from the endoplasmic reticulum (ER) to the Golgi apparatus, and directly delivered to the proper cell-surface domain through the TGN. On the other hand, already-polarized cargo can be internalized by endocytosis when needed,

undergoing an indirect but rapid recycling pathway that will sort them back to their target surface. This asymmetric vesicle trafficking and protein delivery is directed by specific sorting motifs contained within the proteins themselves, which direct their destination. Importantly, coordinated tubular transport across the cytosol to the specific areas of the cell surface is also pivotal for this process, and is controlled by the actin and microtubule cytoskeletons.

2.1.3. Actin cytoskeleton

Remarkably important for epithelial differentiation is the role of the actin cytoskeleton, a dynamic network composed by actin filaments (F-actin) and associated actin binding proteins. We have already mentioned its involvement in the formation of intercellular junctions (binding protein components of tight and adheren junctions) or in the asymmetric transport of membrane-proteins (guiding loaded vesicles to the target localization), but actin cytoskeleton has a crucial function *per se* safeguarding epithelial identity [72]. Indeed, actin cytoskeleton undergoes dramatic rearrangements during epithelial morphogenesis, whose outcomes define differential apical, lateral o basal actin organization. Actin filaments can be visualized as stress fibers, bundles of actin filaments that connect two adhesion points; or as a cortical roughly meshwork of filaments that underlies the plasma membrane of the entire cell. As newly cell-to-cell junctions are formed across the lateral cell membrane, actin cytoskeleton starts to be remodeled into a circumferential actin ring that coincides and connects the junctions [73]. In fully polarized epithelial cells, the stress-fiber actin network is substituted by this cortical ring of actin cables (**Fig. 6**). The most important function of this actin belt-like organization is to generate the mechanical force (cortical tension) determinant for epithelial cell shape, which in the end is what sustains the entire architecture of epithelial tissues.

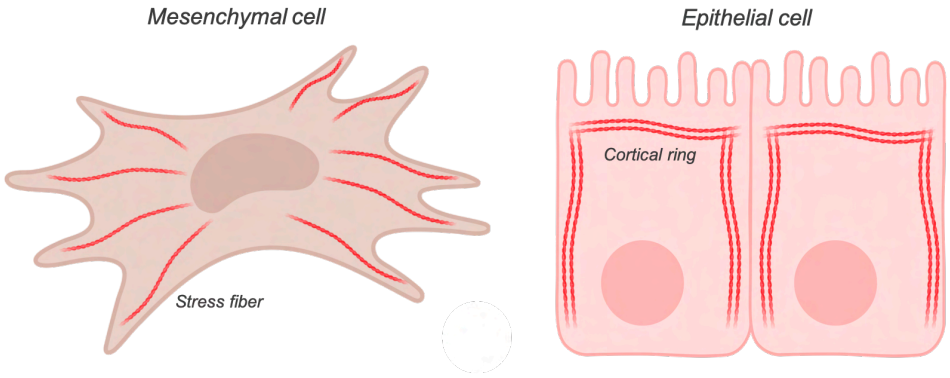


Figure 6. Structural organization of actin cytoskeleton in mesenchymal and epithelial cells.
Left, actin stress fibers in a mesenchymal cell, which shows actin cytoskeletal protrusions at the leading edges. *Right*, cortical actin assembly in epithelial cell beneath the apical and lateral membrane.

2.2. Molecular players underlying epithelial differentiation

2.2.1. Calcium signaling

Calcium signaling involves the use of calcium ions (Ca^{2+}) as second messengers in a signal transduction process that integrate external signals and trigger intracellular responses. During epithelial morphogenesis, and also later on in the adult life, a vast range of environmental local signals must be handled and coordinated in order to regulate different tissue processes, such as differentiation, cell division, self-renewal, regeneration and death. Ca^{2+} signaling contributes to the regulation of all these collective behaviors.

The "Ca^{2+} switch" model was postulated as an essential early step in the biogenesis of epithelial polarity, junction formation and epithelial differentiation. This model states that upon Ca^{2+} exposure, nonpolarized cells exhibiting little cell-to-cell contacts rapidly switch to a polarized state with well-formed desmosomes, tight and adherens junctions [74]. For instance, upon Ca^{2+} exposure it has been shown that aPKC protein locates at the cell-cell borders and triggers the phosphorylation of vinculin, a crucial event during the assembly of the epithelial junctional complex [75]. Also, deregulated intracellular Ca^{2+} levels are also known to alter ZO-1/actin binding, as well as subcellular localization of occludins, therefore disrupting the integrity of tight junction [76]. A key determinant of this Ca^{2+} signaling mechanism is the spatial-specific release of intracellular Ca^{2+} ions, which is achieved by polarized localization of Ca^{2+} pumps, channels, and receptors. Indeed, polarized epithelial cells display Ca^{2+}-integrating elements at the apical or the lateral plasma membrane, where they regulate intercellular junctions' formation [77].

2.2.2. Spindle orientation

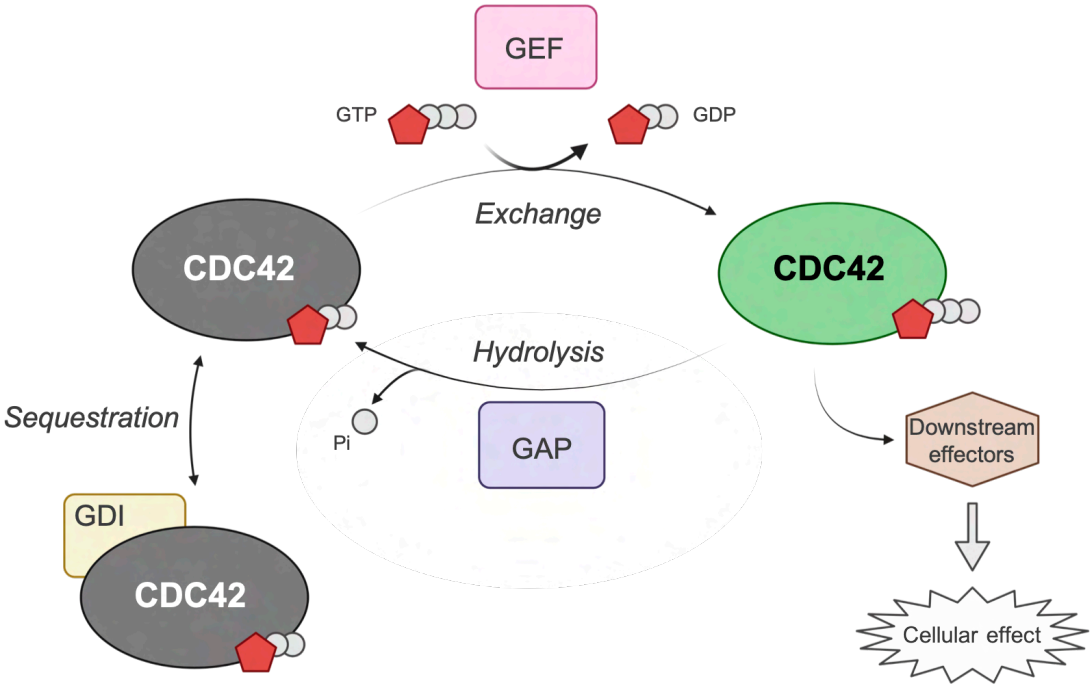
Mitotic spindle orientation is a highly conserved, dynamic, and complex process that plays a crucial role in tissue morphogenesis. This process is particularly important in epithelial organs, in which mitotic spindle orientation determines tissue architecture: while in simple epithelia spindle normally orients parallel to the basal lamina and compels cells to divide in the plane of the epithelial sheet, stratified epithelia require perpendicular divisions for their layering [78]. Moreover, mitotic spindle positioning is also determinant for epithelial cell fate decisions [79]. Epithelial stem cells can divide symmetrically (spindle parallel to the basal lamina) or asymmetrically (spindle perpendicular to the basal lamina) in order to self-renew or to differentiate, respectively. This process of fate decision is key for the maintenance and regeneration of most adult epithelial tissues [80, 81].

2.2.3. CDC42-PAR6-aPKC axis

Cell polarity protein complexes are pivotal in recognizing the initial polarization cues. Therefore, its function is determinant for the identity of epithelial cells. There are three major polarity complexes: the PAR complex, the "Crumbs" complex and the Scribble complex. Among these, the PAR complex is indispensable during epithelial polarity establishment, being required when the other complexes are not. Many of the functions

related with epithelial differentiation lead by the PAR complex are performed in close association with the cell division control protein 42-homolog (CDC42).

CDC42 is an evolutionarily conserved GTPase of the Rho family. Rho GTPases reversibly transit between an active (GTP-bound) and an inactive (GDP-bound) state, acting as molecular switches whose on/off state can be controlled spatially and temporally. Spatial activation of Rho GTPases can be regulated either positively by guanine nucleotide exchange factors (GEFs) or negatively by GTPase-activating proteins (GAPs). Additionally, guanine nucleotide dissociation inhibitors (GDIs) can sequestrate Rho GTPases to keep them in their inactive GDP bound state [82] (**Fig. 7**).



CDC42 activity by three classes of proteins: GEFs, GAPs, and GDIs.
CDC42 acts as a molecular switch that transit between an inactive GDP-bound state and an active GTPbound state. GEFs promote GDP/GTP exchange while GAPs catalyze the hydrolysis of GTP. The family of GDIs sequesters inactive CDC42 to avoid their targeting to the plasma membrane and activation. Once bound to GTP, CDC42 activates numerous downstream effectors to trigger different cellular responses.

CDC42 regulates multiple cell functions by controlling actomyosin cytoskeleton dynamics, including cell shape, proliferation, invasion, migration, and differentiation [83]. It plays a pivotal role in inducing cortical cytoskeletal rearrangement through the activation of a large number of effector proteins [84]. For example, it is known to control the activation of members of the Wiskott–Aldrich syndrome protein (WASp) family, which mediate cortical actin polymerization by directly binding and subsequently activating the actin nucleation activity of the actin related protein (Arp) 2/3 complex [85-87]. Given the CDC42-dependent regulation of actin cortex assembly during early steps of epithelium

morphogenesis [73], it is considered an essential orchestrator of epithelium formation (**Fig. 8**).

At initial steps of epithelial cell polarity stablishment, CDC42 is recruited on forming junctions at the apical area through its binding to adhesion proteins (such as E-cadherin or β-catenin) [88]. There it activates and acts in collaboration with the PAR complex, constituted by PAR3 and PAR6 (two scaffolding proteins) and the atypical serine–threonine kinases PKC (aPKC) [89-91]. Once assembled, this complex promotes the maturation of forming junctional structures, including adherens junctions and tight junctions, and ensures their integrity once established [92-95]. In lung epithelial cells, CDC42 was proven essential for the maturation of primordial junctions through the activation of PAK4 and aPKC [96]. In agreement, inhibition of CDC42 in epithelial Caco-2 and MDCK cells has been reported to cause defects on adherens junctions' assembly [97, 98]. Similarly, it has been demonstrated that CDC42 regulates adheren junctions' stability by controlling the binding of α-catenin with β-catenin and the consequent interaction with the actin cytoskeleton in endothelial cells [99]. Studies on embryoid bodies demonstrated that CDC42 is required for epithelial polarity formation during mammalian development. Depletion of CDC42 at early stages of embryo differentiation caused decreased phosphorylation and activation of aPKC, finally resulting in defective cell-to-cell junctions' formation [100]. However, studies using dominant-negative and constitutively active mutants of CDC42 have revealed that both, inhibition and overstimulation of this protein lead to abnormal junctions' formation [96, 101-103], suggesting that a tight regulation of CDC42 activity is required for the establishment of epithelial polarity.

As a trafficking regulator, CDC42 has been shown to guide polarized trafficking in epithelial cells by regulating actin dynamics [104-106]. Inactivation of CDC42 has been linked to defects on vesicle formation [107] and mis-sorting of basolateral membrane proteins to the apical membrane [108, 109], highlighting its importance in regulating polarized biosynthesis, endocytosis and recycling. At the cell cortex, CDC42 regulates exocytosis [110], while when located at the Golgi apparatus it promotes vesicle formation and targeted transport [111-113].

During cell division, orientation of the mitotic spindle is regulated by the interaction of astral microtubules with the cortical actin. In this context, CDC42 regulation of actin cytoskeleton plays an essential role by properly positioning the mitotic spindle. For example, CDC42 regulate the alignment of chromosomes during prometaphase and metaphase [114] and knockdown of CDC42 disrupts the cortical actin structures formed during metaphase, inducing general spindle misorientation and defective chromosome segregation [115]. Interestingly, CDC42 is also required for proper spindle positioning in polarized cells undergoing asymmetric cell division [116-118]. During asymmetric stem cell division, CDC42 associated with the PAR complex exhibits a polarized positioning and segregate differentially into the two daughter cells, determining which one of the two daughters will commit to differentiate. Depletion or inhibition of CDC42, PAR3 or aPKC, disrupts spindle orientation during cell division and leads to mispositioning of the nascent apical surface [119, 120].

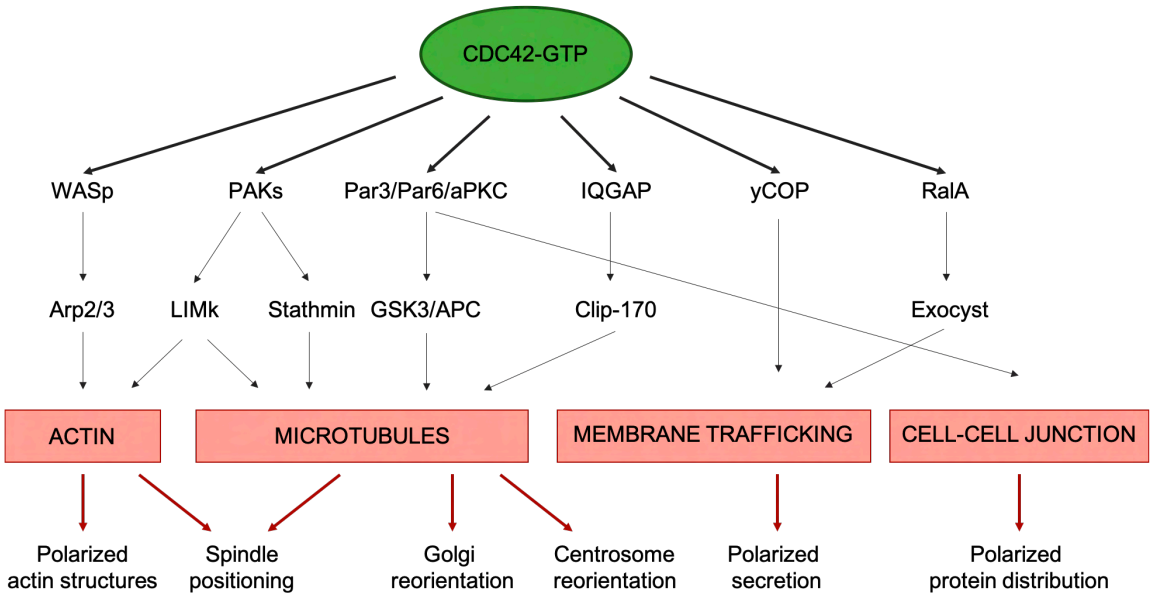


Figure 8. Signaling pathways controlled by CDC42.
Epithelial cell polarization requires spatial and temporal regulation of several cellular processes, such as actin and microtubule cytoskeletons organization, membrane traffic regulation or cell-to-cell contacts formation. Activation of CDC42 orchestrates these cellular events through its several effectors.

2.3. The skin

The epidermis is a highly specialized stratified epithelium with multiple essential protective functions: water loss prevention, toxins exclusion, mechanical stress resistance and immune responses [121]. This tissue is primarily composed by keratinocytes, which stratify into several layers with structural and functional differences in a process known as keratinocyte terminal differentiation (**Fig. 9**). First, the "stratum basale" (basal layer) contains proliferative keratinocytes (skin stem cells or basal cells) sitting and attached to the basal lamina. Suprabasally the "stratum spinosum" (spinous or squamous layer), assembled by early differentiated keratinocytes highly connected through desmosomes. Following, "the Stratum Granulosum" and the "Stratum Lucidum" (granular and clear layers), two thin layers composed of flat keratinocytes actively secreting lamellar bodies (containing lipids, keratins and cornified envelope proteins such as Involucrin) into the extracellular space. And finally, the "stratum corneum" (cornified layer), the outermost layer of the epidermis in which dead keratinocytes crosslink with secreted lipids and proteins to become part of the skin barrier [121]. Polarity in this tissue is, thus, presented at both cellular level (the basal layer is inherently polarized, with a lower surface anchored to the basement membrane and an opposite apical surface) and at tissue level (with an asymmetric distribution of cells with different signaling activity, protein expression and cytoarchitectural organization across layers).

The maintenance of this well-organized architecture requires a fine-tuned balance between self-renewal and differentiation. The process of keratinocyte terminal differentiation is a tightly regulated program that involves multiple events occurring

simultaneously at organellar and molecular levels (**Fig. 9**). Briefly, the appearance of hyaline-filled granules [122], nuclear condensation and extrusion [123, 124] and mitochondria alterations [125, 126] are typical organelle-related processes triggered during keratinocyte terminal differentiation. Molecularly, this process starts with the asymmetric cell division of basal cells, in which one of the cell daughters enter the spinous layer and commit to differentiate [121]. Early in differentiation, keratinocytes remodel its intracellular proteins and intercellular junctions, acquiring new and specialized cell-to-cell junctions (such as desmosomes, adherens and tight junctions) and expanding a cortical ring of actin cytoskeleton coupled to these intercellular junctions [73, 127, 128]. As keratinocytes stratify, they progressively cease proliferation and switch the transcriptional expression of basal markers (KRT5 and KRT14) to specific differentiation markers (KRT1, KRT10, Involucrin, Filaggrin) [129-131].

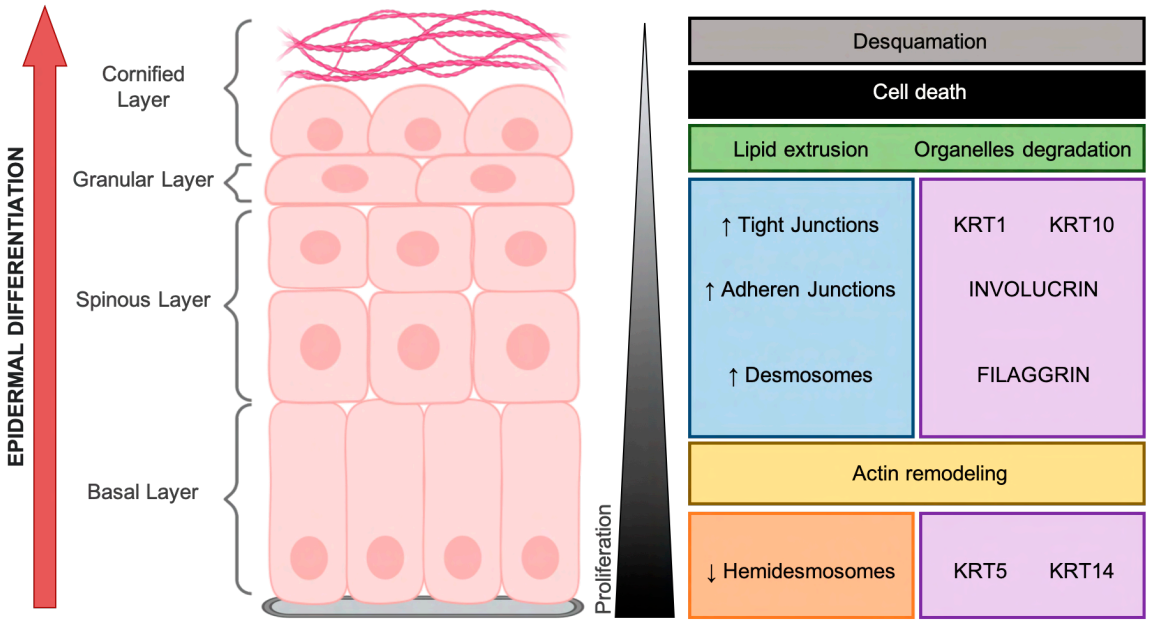


Figure 9. Epidermal architecture.
The epidermis is composed of stratified cell layers, which undergo programmed epidermal differentiation to maintain skin homeostasis. Four main layers are illustrated, the basal layer, the spinous layer, the granular layer and de cornified layer. During keratinocyte differentiation, a specific transcriptional program and cytoarchitecture are triggered in each of the four layers. These differentiation-dependent changes in the composition and organization of keratinocytes are needed to drive tissue morphogenesis and ensure the specific functions of each layer.

Members of the Rho family of GTPases are critical players in epidermal homeostasis and differentiation. As an instance, conditional depletion of Rac1 in mouse skin leads to follicular stem cell exhaustion and hyperproliferation of the epidermal stem cell compartment [132]. On the other hand, differentiation defects have been reported in mice presenting deficiencies in RhoA and their effectors, which are essential regulators of cytoskeleton polarization during the early steps of keratinocyte stratification [73, 133, 134]. Finally, CDC42 has also been proven indispensable for proper epidermal development and barrier formation. Conditional depletion of *CDC42* in KRT5^{+} mouse skin

cells demonstrated its requirement for hair follicle stem cells differentiation and cell-to-cell junctions' formation through the regulation of β-catenin [135]. An additional study using the same CDC42 KO mice also reported an hyperproliferative phenotype and defective deposition of the basement membrane components in epidermal keratinocytes [136]. When CDC42 deficiency is restricted to the epidermal compartment by conditional targeting $KRT14^{+}$ cells, a severe dysfunction of the epidermal barrier has been described, which finally resulted in neonatal lethality. Epidermis of CDC42-KO mice presented high proliferation, increased thickening, reduced apoptosis and a generalized impairment of tight junctions, adherens junctions and desmosomes formation [137].

2.4. Epithelial cells and stress

Epithelial tissues constitute the first line barrier to many internal and external stressors, making them highly susceptible to mutational injury. Given the continuous genetic impact suffered, epithelial cells require accurate and robust protective mechanisms to sustain the continuously challenged homeostasis. One well-known guardian of epithelium integrity is the tumor suppressor p53. After damage, p53 becomes stabilized and activated, triggering multiple cellular processes that prevent the division of DNA-damaged cells, including cell cycle arrest, cellular senescence, apoptosis or differentiation [138].

2.4.1. p53-dependent induced apoptosis

One common outcome of transient p53 activation after DNA damage or oncogenic signals is the induction of cell cycle arrest, which leads to DNA repair. However, when damage is sustained or irreversible, induced p53 activation predominantly leads to apoptosis as a way to completely prevent damaged cells from growing and progressing to cancer. In continually renewing tissues, such as lung epithelial cells or cells at the base of the crypts in intestinal epithelium, cell replacement is usually preferred as an alternative to cell cycle arrest in response to cigarette smoke exposure [139, 140] or ionizing radiation [141, 142], respectively. A paradigmatic example of constantly damaged epithelium is the epidermis, which integrity is continuously threatened by ultraviolet radiation. In this tissue, p53-induced apoptosis exerts a crucial protecting role by reducing the presence of DNA-damaged cells and the risk of subsequent malignant transformation [143]. Indeed, loss of this protective mechanism by p53 mutations is a common early event during the formation of skin-related malignancies, such as basal and squamous cell carcinoma [144].

2.4.1. p53-dependent induced senescence

Cellular senescence is a permanent state of cell cycle arrest usually triggered upon cellular stress, and therefore considered an important tumor suppressor mechanism [145]. It can be induced by several triggers, including DNA damage, DNA replication stress, oncogene activation or oxidative stress, among others [146-150]. Like apoptosis, senescence-associated growth arrest is regulated mainly by p53. However, and unlike apoptotic cells, senescent cells do not die, they remain metabolically active and produce a prominent secretome (SASP, senescence-associated secretory phenotype) that

modulates the microenvironment for as long as they persist in the tissue, having both beneficial and detrimental effects [145, 151]. In the epidermis, lack of senescent cells impairs wound healing and contribute to tumorigenesis [152, 153]. However, when chronically present, senescent cells also impair tissue homeostasis, regeneration and contribute to skin aging phenotypes [154, 155].

2.4.2. p53-dependent damage induced differentiation

Although induction of apoptosis or senescence in potential neoplastic cells are considered the main mechanisms by which p53 prevents tumorigenesis, damage-induced differentiation has been postulated as an alternative response to mutational injury [156-160]. In the last years, this process has started to be explored in skin epidermis and other epithelia. A study performed in mouse epidermis has shown that in response to potential cancer-causing mutations, skin stem cells preferentially opt to differentiate instead of self-renew [161]. Given that differentiated skin cells decrease their proliferation and eventually undergo cell death to become part of the epidermal barrier, this protective strategy arises as an alternative mechanism that impedes dangerous mutation from spreading without disrupting tissue function. Several studies have supported this evidence [162-164], highlighting this damage- or oncogene-induced differentiation response as an emerging protective mechanism against oncogenic alterations in epithelial tissues. Although the implications of p53 in this novel mechanism remains controversial [165], some studies do directly link this damage-induced differentiation response to p53 [166, 167] or with pathways in which p53 is involved, such as the DNA-damage response (DDR) pathway [163].

3. Epithelial tumors

Maintenance of normal tissue homeostasis requires a dynamic equilibrium between cell proliferation, cell death, tissue morphogenesis and differentiation. Sustained alterations of this equilibrium originate a permissive environment for the progression of diseases, such as cancer. In epithelial tissues, this homeostasis is continuously challenged by different internal and external insults, which is why epithelial tumors (carcinomas) comprise approximately 90% of all human cancers [168]. Carcinomas are divided into four major subtypes: adenocarcinoma (arose from glandular cells), transitional cell carcinoma (which originates in stretchy cells in the urinary tract epithelium), basal cell carcinoma (developed in cells of the basal layer of the epithelium) and squamous cell carcinoma (SCC). SCCs arise from stratified epithelia including lungs, esophagus, cervix, or oral cavities and, more rarely pancreas, thyroid, bladder, or prostate. However, the most common type refers to skin. Cutaneous SCCs (cSCC) result from the uncontrolled proliferation of keratinocytes, which develope actinic keratoses lesions that gradually progress into invasive carcinomas [169]. The mutational burden of cSCC is extremely high, being mutations in the *TP53* gene the most common and a prevalent risk factor of this disease [169, 170]. The incidence of SCCs is sharply increasing mainly due to the actual elevated exposure to carcinogens, such as ultraviolet radiation related to

sun exposure, smoking, alcohol consumption or human papilloma virus (HPV), therefore SCCs represent today a major cause of death worldwide [171].

3.1. Loss of epithelial cell identity: a hallmark of epithelial tumors

Tumorigenesis is understood as a process of loss of cell differentiation [172-176], in which somatic cells acquire malignant properties such as uncontrolled proliferation, evasion of apoptosis, immunosurveillance, dysregulated metabolism and epigenetics and invasive potential, among others. These events have been generalized as the hallmarks of cancer [177]. In carcinomas, loss of epithelial identity by dedifferentiation is considered a crucial initial step [178, 179] associated with the dysregulation of common epithelial differentiation-controlling pathways. The loss of p53 expression is well-known to promote cell proliferation and apoptosis resistance, leading to tumor formation. However, it has been also highlighted its impact on enhancing symmetric cell divisions [180]. Restoration of p53 expression is sufficient to reduce the number of symmetric cell divisions and promote cell differentiation, a mechanism that may involve the regulation of the Notch signaling pathway [181]. Notch signaling pathway, which promotes keratinocyte differentiation [182, 183], is commonly mutated in around half of human cSCCs [184-186]. Indeed, deletion of *Notch1* gene in mouse epidermis accelerates and increases skin tumorigenesis [187]. The *TP63* gene is expressed in the basal compartment of the mature epidermis, where it is required for the maintenance of the proliferative potential of basal keratinocytes by negatively regulating Notch activity [188-190]. This gene, downregulated during terminal differentiation, is frequently amplified in cSCC together with SOX2, a transcription factor that controls pluripotency of stem cells and inhibit terminal epidermal differentiation [191].

The loss of the differentiation status is closely related with epithelial cell polarity integrity, which is considered today a non-canonical tumor suppressor given its participation in the establishment and maintenance of the epithelial tissue organization as a whole. Indeed, the disruption of epithelial cell polarity is recognized as a hallmark of epithelial carcinomas [179, 192-194], implying the alteration of its main determinants: cell-to-cell adhesion, 3D intracellular organization and asymmetric cell division (**Fig. 10**).

Disruption of cell-to-cell adhesion is principally given by the disassembly of the junctional complex, including adherens and tight junctions. A classic event on carcinoma progression that often occurs in later stages of tumorigenesis is the loss of E-cadherin, which allows transformed cells to detach from the tissue and acquire invasive features (epithelial-to-mesenchymal transition, EMT) [178]. Additional to their structural function, intercellular junctions are key centers for mechanotransductional signaling and transcriptional programs regulation. Indeed, they commonly regulate many pathways related to epithelial proliferation. For example, disruption of adherens junctions results in an increase of the cytosolic and nuclear β-catenin and upregulation of the Wnt signaling pathway, which drives epithelial dedifferentiation, cell proliferation and EMT, [195-197]. Moreover, defects on the integrity of junctional complexes are directly connected with alterations in the cortical actin ring that assists epithelial tissue architecture. Indeed, transformed cells are characterized for having a reduced membrane stiffness and softer shape, which allow their extrusion from the tissue during the initial steps of metastasis

[198, 199]. Cytoskeleton alterations generally affect key features of cell polarity, such as vectorial trafficking and mitotic control [194, 200]. Indeed, proper spindle orientation requires the interaction of the mitotic spindle with cortical actin sites and is indispensable to maintain a proper balance of asymmetric and symmetric cell division [201]. In the context of the cancer stem cell theory, a shift from asymmetric towards symmetric divisions due to alterations on the mechanism that determines spindle orientation is thought to lead to an excessive proliferation of the stem compartment, contributing to the appearance of cancer-initiating cells and the formation of tumors.

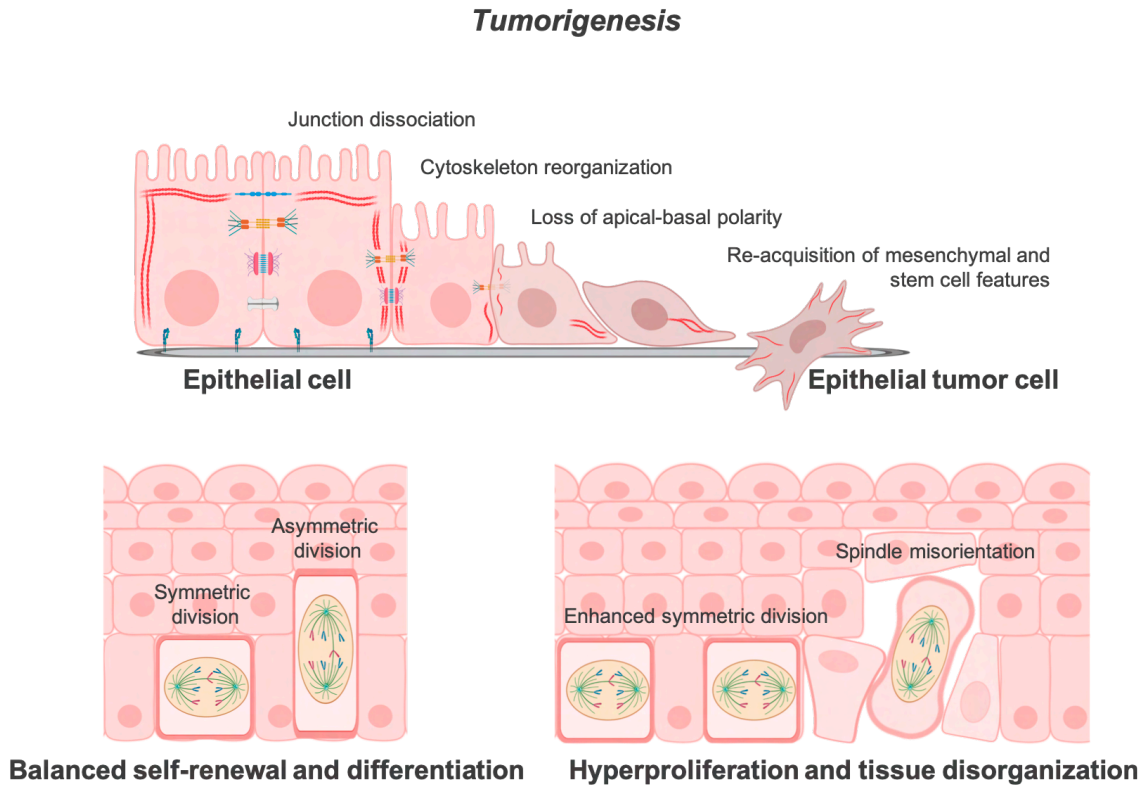


Figure 10. Process of epithelial tumorigenesis driven by alterations on epithelial cell polarity. *Top*: the process of tumorigenesis by dedifferentiation involves cell-to-cell contact dissociation, cytoskeleton reorganization, loss of apical-basal polarity and reacquisition of mesenchymal and stem-cell features. Transformed epithelial cells undergo extensive rearrangements of actin cytoskeleton, shifting from cortical actin towards stress fibers, which mediate shape changes, front-back polarity, invasive behavior, and migration. *Bottom*: Misorientation of the mitotic spindle organization during tumorigenesis leads to alterations in the ratio of asymmetric/symmetric cell divisions, which unbalance the processes of differentiation and self-renewal. These alterations trigger the hyperproliferative and disorganized phenotype typical of epithelial tumors.

Interestingly, the proteins that precisely control epithelial cell polarity crosstalk with signaling pathways regulating cell growth, frequently by means of independent downstream effectors. Because of this, they are considered both tumor suppressors and oncoproteins [202-207]. A good example of this dualism is CDC42, which participation in the maintenance of the apical-basolateral polarity has been extensively discussed in this introduction. However, given the plethora of effectors acting downstream CDC42 activation, its biological function spreads to a variety of different cellular processes

including proliferation, migration, adhesion or transcription regulation [208]. Active CDC42 increases the activity of PI3K, which triggers cortical actin reorganization during epithelial differentiation [209]. However, it also enhances symmetric cell division of hematopoietic stem cells, blocking leukemia cell differentiation and favoring its progression [210]. In the intestinal epithelium, CDC42 is known to constrain hyperproliferation of intestinal stem cells and subsequent tissue hyperplasia [211]. On the other hand, the CDC42 effector kinase PAK is known to phosphorylate both Raf and MEK, enhancing ERK signaling and promoting proliferation [212]. Intriguingly, although the association of CDC42 with the polarity protein complex PAR3/PAR6/aPKC is an essential initial step for cell polarization during epithelial morphogenesis, it can also cooperate with the transforming growth factor β (TGFβ) receptor to promote TGFβ-induced EMT [213]. Indeed, elevated expression of CDC42 is commonly observed in invasive cells [214-217], where it promotes aberrant actin remodeling, disruption of cell-cell cohesion and metastatic behaviors [207, 218-221]. Finally, despite its function as an epithelial identity safeguard, studies using dominant-negative CDC42 mutants have reported its ability to enhance anchorage-independent growth and EGFR- and Ras-mediated cell transformation [222-224].

4. Ubiquitin and ubiquitin-like proteins

Maintenance of cell homeostasis requires a tight control of protein activity, function, stability, and localization. To achieve this, a widely evolutionary conserved strategy implies the posttranslational modifications of proteins, which can involve chemical conjugation of small elements (such as phosphorylation, acetylation, or methylation) as well as covalent binding of small proteins (such as ubiquitination). The 76-amino acid protein Ubiquitin is considered a molecular tag that mainly target proteins to the proteasome for degradation, although it has been also described to regulate other processes including endocytosis, membrane protein trafficking, cell signaling, and DNA repair [225]. The process of ubiquitylation involves the covalent binding of Ubiquitin via its C-terminal diglycine motif (GG-motif) to a lysine residue in the target protein through a cascade of reactions carried out by different enzymes, referred to as E1 activating enzymes, E2 conjugating enzymes, and E3 ligases. Additionally, deubiquitylating enzymes (DUBs) can revert this process by hydrolyzing the bond between Ubiquitin and its substrates.

Ubiquitin-like proteins are a family of proteins that share evolutionary features with Ubiquitin, not in sequence homology but in common three-dimensional structure [226]. Every member of this family is characterized by presenting a conserved tertiary structure, generally referred to as the β-grasp fold, consisting of five-stranded β-sheet partially wrapped around a central α-helix (**Fig. 11)**. They can be divided into two subclasses based on their conjugating capacity [225]. Type I UBLs can either exist freely or covalently attached to other proteins. Like Ubiquitin, this subclass of UBLs present a diglycine motif through which they accomplished the conjugation process via an enzymatic cascade similar to that of ubiquitylation. Type II UBLs do present the ubiquitin-

like domain but lack the GG-motif, therefore they cannot ligate other proteins. Instead, they have been described as adaptor proteins (binding noncovalently to ubiquitin or other UBLs via an interaction motif or domain), escorting proteins (guiding polyubiquitinated proteins to the proteasome) or as protein-protein interaction scaffolds [225, 227, 228]. At present, several UBLs have been characterized and demonstrated to modify other proteins, showing diverse roles in several cellular processes [226].

4.1. SUMO proteins

The small-ubiquitin-related modifier (SUMO) is a family of proteins belonging to the type I subclass of UBLs. Of the 5 SUMO paralogs identified in the human genome, all known SUMO modifications are performed by SUMO-1, SUMO-2, or SUMO-3, which are ubiquitously expressed in all tissues. Expression of SUMO-4 and SUMO-5 seems to be restricted to specific tissues and their processing and conjugating capacities remained unclear [229, 230]. SUMO-2 and SUMO-3 share nearly identical sequence (97% of similarity) and cannot be immunologically distinguished (thus referred to as SUMO-2/3), but they significantly differ from SUMO-1 presenting only 45% of sequence identity. Another discriminating feature between paralogs is their ability to form mono- or poly-SUMOylated chains on target proteins, being SUMO-2/3 capable of stablishing polymeric chains while SUMO-1 exclusively forms monomeric chains or acts as a chain terminator on SUMO-2/3 polymers. In either case, the highly dynamic SUMOylation process can be rapidly reverted by specific proteases named SENPs (sentrin-specific proteases). Regarding their substrates, SUMO-1 and SUMO-2/3 present a considerable overlap (including proteins such as p53, B23 or PML proteins) [231, 232], however paralog-specificity for many other SUMOylated proteins has also been reported. This aspect of preferential SUMO paralog modification is still poorly understood but some determinants have been discovered. One of them is the presence of highly conserved hydrophobic stretches on SUMO substrates, called SUMO-interacting motifs (SIMs) [226]. SIMs consist of a hydrophobic core flanked by N- or C-terminal acidic or phosphorylatable serine residues which mediate noncovalent interactions between SUMO and SIM-bearing proteins. To date, hundreds of SUMO targets are known to contain noncovalent SUMO binding sites, separately or in tandem, to facilitate and regulate the binding of proteins to a particular SUMO paralog [233, 234].

The highly dynamic nature of SUMOylation and its ability to selectively target and rapidly mobilize a wide range of proteins and enzymes turns it into one of the most effective stress sensors in the cell, being involved in many biological processes, such transcriptional regulation, protein stability and transport, apoptosis, cell growth, and even cellular identity.

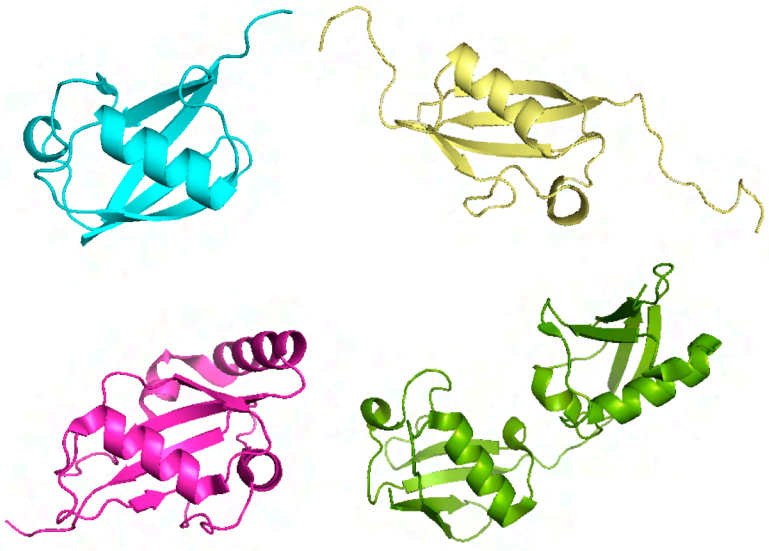


Figure 11. Tertiary structure of Ubiquitin and selected UBLs.
Cartoon depicting structural modeling of the conserved β-grasp fold of Ubiquitin (blue, PDB: 1UBQ), SUMO (yellow, SUMO1 PDB: 1A5R), ATG8 (pink, PDB: 1EO6) and ISG15 (green, PDB: 1Z2M). ISG15 has a tandem ubiquitin folding. Files were downloaded from PDB and diagrams were made using PyMol.

4.2. SUMOylation and differentiation

Accurate coordination of the molecular events triggered during cell fate specification within stem cells is pivotal for ensuring correct cell and tissue differentiation. In this context, SUMOylation is considered a safeguard of cell identity given its indispensable roles in maintaining nuclear and chromatin integrity, regulating chromatin remodeling and modulating chromatin-associated proteins associated with pluripotency and lineage commitment [235].

During reprogramming of mouse embryonic fibroblasts (MEFs), SUMOylation acts on fibroblastic enhancers to retain somatic transcription factors and impede their conversion to pluripotency (iPSCs), while in embryonic stem cells (ESCs) it inhibits the transition to totipotent-like cells (2C-like or 2-cell stage embryo). These results suggest that SUMO modification preserve somatic and pluripotent states [235, 236]. Similarly but in adult tissues, the SUMO-E3 ligase protein CBX4 acts as an inhibitor of epidermal differentiation by regulation the specific SUMOylation of transcriptional repressors related to epidermal differentiation on epidermal stem cells (epSCs) [237]. On the other hand, SUMO has also been involved in calcium-induced differentiation of keratinocytes [238]. During this process, global protein SUMOylation dynamically changes, starting with high SUMOylation levels at the basal state, followed by an abrupt decrease right after differentiation induction and a subsequent reestablishment of global SUMOylation after 48 hours of high calcium exposure. However, fluctuations in the SUMOylation of individual proteins was highly variable, and there are examples of both, SUMOylation increase and decrease as differentiation progresses [239]. In agreement with this, a study performed in human keratinocytes demonstrated that the tumor suppressor p14(ARF) targets p63 to proteasomal degradation by stimulating its SUMOylation during

differentiation. Interestingly, authors of this study observed an increase of ARF levels in differentiated keratinocytes, which is concomitant to the already mentioned increased on SUMOylation during differentiation [240].

4.3. SUMOylation, stress and disease

SUMOylation is considered a molecular glue for many stress-response complexes, which components remain together and act coordinately aided by different SUMO moieties and SIMs. These include several membranes-less organelles with well-known roles as stress sensors, such as stress granules, nuclear bodies, telomeric regions or the nucleolus [241-243]. Another good example are paraspeckles, nuclear bodies built mainly by RNA-binding and editing proteins (such as PSP1, PSF or NONO) which are related with mRNA processing and splicing [243, 244]. In addition, they also act as protective cellular barriers upon genotoxic stress by temporarily retaining incompletely processed or defective mRNA [245]. Interestingly, the SUMOylation of different paraspeckles components has been reported and seems to be essential for the assembly and integrity of these bodies [246-249].

The fact that SUMOylation levels are regulated in a delicate balanced in somatic cells, modulating stress responses, stemness and differentiation, has significant implications in disease. Indeed, a growing list of pathological conditions have been associated with either aberrant SUMOylation of specific proteins and/or dysregulation of global cellular SUMOylation [229, 250, 251]. In cancer, SUMOylation has been proved to promote or inhibit different aspects of the tumorigenic process. As instance, *Myc* oncogene is usually found highly SUMOylated in tumors, which promotes its phosphorylation and enhances its stability [252]. In this scenario, SUMOylation facilitates MYC-driven carcinogenesis. On the other hand, the tumor suppressor p53 is typically stabilized through SUMOylation, leading to an increase in its transcriptional and tumor suppressive activities [253-255]. Several studies have reported the prognostic potential of the dysregulation of global SUMOylation in cancer, although in opposite directions depending on the tumor type. The upregulation of components of the SUMO pathway were found in some cancers (including breast, melanoma, gastric and colorectal cancer) correlating with a higher histological grade or stage, metastases, therapeutic resistance, and poor prognosis [256-260]. On the contrary, high expression of the SUMO machinery has been also associated with tumor suppressive activities and better prognosis in some other cancer types, such as hepatocellular carcinoma, intestinal and bladder cancer, leukemia and osteosarcoma [261-265].

Given the mounting evidence of SUMO dysregulation in cancer, SUMO pathway is increasingly viewed as a potential therapeutic target in cancer [266, 267]. Indeed, several modulators of SUMO E1-activating, E2-conjugating or deconjugating enzymes have shown e promising results in *in vitro* and *ex vivo* models [267, 268]. One example is TAK-981, a SAE (E1 enzyme) inhibitor that is being currently evaluated in clinical trials in combination with immunotherapy drugs showing with promising results [266]. However, given the complexity of SUMOylation impact on cell physiology, its targeting remains as a risky and yet-to-be explored pharmacological strategy.

HYPOTHESIS & OBJECTIVES

Technical advances during the last decade have revealed the existence of microproteins, small bioactive proteins largely overlooked until recently and with pivotal roles in both physiological and pathological scenarios. These microproteins, commonly coded in previously assumed non-coding regions, represent an unexplored field that has opened a new layer of biological complexity that may change current paradigms, from basic research to translational and clinical settings.

The process of cell differentiation is of crucial importance for tissue homeostasis. On the contrary, the loss of cell identity through dedifferentiation is an essential initial step in epithelial tumorigenesis.

The overarching hypothesis of this thesis project is that the microproteome could be a source of novel key players involved in cell identity stablishement and tumor formation. We speculate that the discovery of these microproteins will provide valuable knowledge to understand cell development, tissue homeostasis and cancer biology.

To demonstrate this hypothesis, we proposed the following specific objectives:

1. Identify novel microproteins related with cellular identity and cancer.

2. Generate gain- and loss-of function tools to study the most interesting candidate.

3. Characterize the best microprotein candidate, including its subcellular localization, tissue distribution and the regulation of its expression.

4. Study the role of the microprotein in the regulation of epithelial identity both, in physiological contexts and in cancer.

5. Analyze the molecular mechanisms responsible for the microprotein's functions.

MATERIALS & METHODS

1. General cell culture and treatments

1.1. Cell Culture Conditions

HEK293T, U2OS, MCF7 and A549 cells were cultured in DMEM with GlutaMAX supplemented with 10% of fetal bovine serum (FBS) and 1% of Penicillin-Streptomycin (P/S). cSCC-patient derived hSCC10 cell line was obtained as described [58, 59] and cultured in DMEM-F12 supplemented with 2% B27 (Gibco) and 1% P/S. mESC V6.4 cell line was cultured in DMEM GlutaMax supplemented with 1% Sodium Pyruvate (Invitrogen), 15% FBS, 50mM β-mercaptoethanol, 100X non-essential amino acids (Invitrogen), 1% P/S and 1000 U/ml LIF (ESGRO, Chemicon). Low (basal)- or high (differentiation)-Ca^{2+} media were prepared by supplementing Ca^{2+}-free DMEM (Gibco, #21068028) with 10% Ca^{2+}-depleted FBS, 4mM L-glutamine (Invitrogen), 1% P/S, and Calcium Chloride (Sigma) to a final concentration of 0.03mM (low-Ca^{2+}, basal media) or 2.8mM (high-Ca^{2+}, differentiation media). FBS was Ca^{2+}-depleted by incubation with BT Chelex 100 resin (BIO-RAD #143-2832) according to the manufacturer's instructions. HaCaT and MCF7 cells were maintained in basal medium. Dedifferentiation was induced by culturing cells in basal medium for at least three weeks. Confluence was maintained lower than 70% to ensure basal conditions. Differentiation was induced by cultering cells in differentiation medium for at least three culture days. All cells were incubated at 37 ºC, 5% CO2, in a humidified incubator.

1.2. Cell treatments

Doxycycline (Sigma Aldrich) was added at 1µg/ml (unless otherwise indicated) to activate Tet-inducible constructs. To induce genotoxic stress and/or p53 activation, cells were incubated with doxorubicin (1µM), actinomycin-D (5nM) or nutlin-3a (10µM) for 24 hours or irradiated without medium with a UV light lamp emitting at 254nm at a rate of 50 J/m2. For protein stability assays, cells were incubated with cycloheximide (100 µg/ml) for the specified time periods, 24 hours after doxycycline-induced pTINCR expression. For CDC42-activation assay, hEGF (Invitrogen) was added at 200µg/ml for 5 minutes. For SUMOylation inhibition, ML-792 (#HY-108702, MedChemExpress) was added at 10µM for 16 hours minimum prior to additional treatments. For CDC42-GTP inhibition, CASIN and ML-141 (#HY-12874 and #HY-12755 respectively, MedChemExpress) were added at 5µM for 2 hours minimum prior to additional treatments.

2. Cloning Procedures

pTINCR human ORF was synthesized (IDT technologies) fused with a flexible linker (3xGGGGS) and an HA-Tag epitope at the C-terminal or N-terminal part of the microprotein and flanked by *EcoRI* enzyme restrictions sites at both ends. After enzymatic digestion, constructs were ligated into the pENTR1A vector (Addgene). pTINCR-SIM mut plasmid was purchased directly cloned into the pDONR201 plasmid, in frame with an HA-tag epitope in its C-terminal (Proteogenix). Subsequently, pTINCR-HA, HA-pTINCR and pTINCR-SIMmut constructs were obtained by recombining donor vectors with the lentiviral inducible system pINDUCER20 (Invitrogen) using the Gateway

Cloning Technology, following manufacturer's instructions. As specified, some experiments were reproduced using a pTINCR synthetic ORF (syORF), generated by mutating the pTINCR ORF in pTINCR-HA and pTINCR SIMmut constructs (Proteogenix) (**Table 1**).

2.1. Retro- and lenti-viral infections

HEK293T cells were used as packaging cells and transfected with 5 µg of specific plasmid and 5 µg of packaging vectors (**Table 2**) using Fugene HD (Promega) following manufacturer's instructions. Viral supernatants were collected twice a day on two consecutive days, filtered through a 0.45 µm syringe filter, supplemented with of 8 µg/ml of polybrene and used to infect cells of interest. Successfully infected cells were established by geneticin (Gibco) selection at 400-600 µg/ml (pINDUCER-20) in the case of inducible vectors or puromycin (VWR) selection at 1 mg/ml (pMSCV) in the case of constitutive vectors.

3. Generation of pTINCR-KO cells

HaCaT and MCF7 cells were transfected with pSpCas9(BB)-2A-Puro plasmid, which contains a sgRNAs (AGCCGGGCGGGCGCCATGGAGGG) design to target the first exon of *TINCR* gene loci. 24 hours after transfection, puromycin was added to select for infected cells. Isolation of pTINCR-KO single colonies was assessed by serial limiting dilution in 96-multiwell plate. Successfully CRISPR-CAS9 edited clones were screened by genotyping PCR and DNA sequencing.

4. mRNA analysis by RT-qPCR

Total RNA was extracted with Trizol (Invitrogen) following manufacturer's protocol. Genomic DNA was cleaned up and retrotranscription performed using the iScript gDNA Clear cDNA Synthesis Kit (BioRad). Gene expression was analyzed by RT-qPCR using PowerUp SYBR Green Master Mix (Thermo Fisher Scientific) in the 7900HT Fast Real-Time PCR System (Applied Biosystems). Gene-specific primers are listed in **Table 3** and **Table 4**. Cycle threshold (Ct) values were normalized to GAPDH.

5. Protein analysis by Western blot

Cells were scrapped and homogenized in medium-salt lysis buffer (150 mM NaCl, 50 mM Tris pH 8, 1% NP40 supplemented with protease (Roche) and phosphatase (SigmaAldrich) inhibitors cocktails) and concentrations determined using the Pierce™ BCA Protein Assay Kit (Thermo Fisher). Lysates were loaded in acrylamide gels for electrophoresis in Tris-Glycine SDS Running Buffer. Primary antibodies were incubated overnight at 4°C. Secondary HRP-conjugated antibodies were incubated the following day for 1 hour at room temperature, and ECL Prime Western Blotting Detection Reagent (Fisher Scientific) was used as a chemiluminescent reagent for protein detection. Antibodies and dilutions are listed in **Table 5**.

5.1. Subcellular Fractionation

Cells were homogenized in Buffer A (HEPES pH 7.8 (10mM), MgCl2 (1.5mM), KCl (10mM) and DTT). The homogenate was incubated on ice for 10 minutes and then, 10% Triton-X was added to favor cellular disruption. Samples were centrifuged at 11.000 rpm for 1 minute at 4ºC and supernatant (Cytoplasmic extract, CE) was separated from the pellet (Nuclear extract, NE). CE supernatant was washed with Buffer B (HEPES pH 7.8 (0.3mM), MgCl2 (1.4mM) and KCl (30mM)), followed by centrifugation at 15.000 g for 15 minutes at 4ºC. New supernatant was used as final cytoplasmic extract and pellet was disrupted in SDS lysis buffer (Tris-HCl pH 8.0 (50 mM), EDTA (1 mM), SDS (2%)) to obtained final cellular membranes extract. NE pellet was washed with Buffer B and disrupted using Buffer C (HEPES pH 7.8 (20mM), MgCl2 (1.5mM), NaCl (0.42mM), EDTA (0.2mM), glycerol (25%) and DTT). After 30 minutes of iced-incubation and centrifugation at 15.000 g for 15 minutes at 4ºC, supernatant was used as final nuclear extract.

6. Immunofluorescence

Cells were seeded in fibronectin-coated coverslips (Sigma Aldrich). When desired, cells were fixed in 2% paraformaldehyde for fifteen minutes and permeabilized with 0.5% Triton X-100 for 10 minutes at room temperature. Blocking step was made in 3% Bovine Serum Albumin (BSA) for 1 hour. Cells were incubated overnight at 4°C with the primary antibody diluted in blocking buffer. Next day, secondary antibodies were incubated for 1 hour at room temperature in the dark. Antibodies and dilutions are listed in **Table 5**. Finally, cells were mounted in Prolong Mounting Medium with DAPI (Invitrogen), and images were taken in a Nikon Eclipse Ti-E inverted microscope system. For quantification, at least 200 cells per staining were evaluated using ImageJ software.

7. Cell proliferation analysis

Cells were seeded into 24-well cell culture plates at density $5x10^3$ cells/well. Every two days, cells were trypsinized and counted using a Neubauer chamber. For inducible pTINCR-overexpressing models, doxycycline was added every two days.

8. BrdU incorporation and cell cycle analysis

For bromodeoxyuridine (BrdU) incorporation and cell cycle analysis, hSCC10 cells transduced with control or pTINCR-HA tag vectors were treated with doxycycline for 4 days and then label with 10 µM BrdU for 24 hours, harvested by trypsinization and fixed/permeabilized on ice using absolute ethanol. After denaturation in 2M HCl for 30 minutes at room temperature, they were neutralized with 0.5M sodium borate and stained with anti-BrdU-FITC antibody (BD Pharmingen, 1:5) diluted in PBS + 0.5% BSA and 0.5% Tween. Cells were incubated in 25 mg/L propidium iodide (PI) + 100 µg/mL RNAse A and 0.3 µM Triton X100 for 30 minutes. Flow cytometry analysis was performed on a Navios EX-Powerful, Dependable Flow Cytometer and analyzed using FlowJo™ software.

9. Senescence-associated (SA) β-galactosidase assay

A549 cells constitutively expressing the microprotein or control vector were cultured for one week, then washed with PBS and fixed with 0.2% glutaraldehyde. After washing with PBS, cells were incubated overnight at 37 °C with a staining solution containing 1 mg/mL X-Gal (Melford BioLaboratories, MB1001) prepared in dimethylformamide (DMF, Sigma, D4551) at pH 6. Cells were then washed in PBS and visualized using a Olympus CellR microscope.

10. Annexin V assay

A549 cells were cultured and treated with doxycycline for microprotein induction for at least 4 days or not treated to be used as control cells. Then, cells were washed twice with cold PBS and resuspended in Binding Buffer at a cell density of 1×10^6 cells/mL. Following manufacturer's instructions, 100 µL of cells were treated with 5 µL of Annexin V-FITC antibody (556420, BD Pharmingen™). After incubation for 15 min in the dark at room temperature, 400 µl of Binding Buffer was added and each sample was supplemented with 10 µL of propidium iodide. Flow cytometry analysis was performed on a Navios EX-Powerful, Dependable Flow Cytometer and analyzed using FlowJo™ software.

11. Cell migration assay

A549 cells were seeded and culture until confluence was reached. Then, a pipette tip was used to scratch the surface of cell monolayer, forming a wound. An Olympus CellR microscope equipped with a Hamamatsu C9100 camera was used to follow the closure of the wound up to 48 hours.

hSCC10 cells were seeded in a Corning cell culture insert (Boyden chamber). Cells were allowed to migrate through for 24 hours at 37°C. After that time, culture inserts were fixed in methanol for 5 minutes and stained using Crystal Violet during twenty minutes at room temperature. Pictures were taken using an Olympus CellR microscope. Cells were counted using ImageJ.

12. Cell invasion assay

$2x10^4$ cells were seeded in a Corning cell culture insert (Boyden chamber) coated with a layer of Matrigel (Corning). Cells were allowed to invade through the membrane for 24 hours at 37°C. After that time, culture inserts were fixed in methanol for five minutes and stained using Crystal Violet during twenty minutes at room temperature. Pictures were taken using an Olympus CellR microscope. Cells were counted using ImageJ.

13. Transcriptomic analysis

13.1. RNA sequencing

Cells were cultured and treated with doxycycline for microprotein induction during the specified experimental time periods. Total RNA was isolated with Trizol following manufacturer's protocol. RNA quantity and purity were measured with a Nanodrop spectrophotometer (Thermo Scientific) and 1 µg of total RNA was processed for sequencing analysis. RNA integrity, determined by the RNA integrity number (RIN), was determined with the 2100 Bioanalyzer (Agilent Technologies Inc., Santa Clara, CA). Cytoplasmatic and mitochondrial ribosomal RNAs were depleted using the RiboZero Magnetic Gold Kit (Illumina Inc). rRNA-depleted samples were fragmented, cDNA was synthesized and converted into sequentiable libraries using the TruSeq Stranded Total RNA kit protocol (Illumina Inc.). The size and quality of the libraries were assessed with a High Sensitivity DNA Bioanalyzer assay (Agilent Technologies Inc., Santa Clara, CA). Libraries were sequenced in a NextSeq 500™ (Illumina Inc.), with a read length of 2x76 bp. On average, 77 million paired-end reads were generated per sample. Image analysis, base calling and quality scoring of the run were processed using the manufacturer's software Real Time Analysis (RTA 1.18.64) and followed by generation of FASTQ sequence files by CASAVA.

13.2. RNAseq data analysis

Paired end reads were aligned to the hg38 human genome with STAR (version 2.5.2b) [272] and default parameters. Sambamba [273] was used to convert to bam and sort resulting sam files. All subsequent analyses were performed in the R programming environment (https://www.R-project.org) unless otherwise stated. Count matrices were generated with the Rsubread package [274] using the inbuilt annotation hg38. Genes were annotated using biomaRt [275] version GRCh38.p12. Normalization and contrasts were performed using the DESeq2 [276] R package. Time points were compared against 0 hours (no induction condition) within the control and pTINCR-overexpressing condition.

In order to find genes whose expression was significantly associated with time we applied the runImpulseDE2 function from the ImpulseDE2 [277] R package on the count matrix generated with Rsubread. The biological replicate was included as a technical confounder. Genes with adjusted p-values lower than 0.001 were selected for downstream analyses. The matrix containing the fold changes resulting from the comparisons against 0 hours was filtered for those genes with significant association with time according to ImpulseDE. The cmeans function from the e107 R package was used to find clusters of genes with similar expression patterns along time. The centers parameter was fixed at 6 after manual inspection. The ggplot2 (https://ggplot2.tidyverse.org), viridis, ggExtra and gridExtra R packages were used for plotting resulting clusters. Genes were classified in a given cluster if their membership probabilities were larger than 0.5. Functional enrichment on the resulting gene clusters was performed using a hypergeometric test against the Gene Ontology (GO) database [278], KEGG pathways [279] and Broad Hallmarks gene sets [280].

Networks of overlapping GO terms were computed using the ClueGO [281] module of Cytoscape [282] with default parameters.

The interaction between treatment and time was computed for times 0 hours and 21 days. Functional enrichment of the interaction coefficients was performed using a rotation-based methodology. The ROAST algorithm [283] as implemented in the R package limma was used to represent the null distribution. The maxmean enrichment statistic proposed in [284], under restandardization, was considered for competitive testing.

14. Interactome analysis and validation

Cells were lysed in a buffer containing 50 mM Tris-Hcl pH 7.5-8, 150 mM NaCl, 1%Triton X-100 and protease inhibitors, and homogenized for thirty minutes in a rotor wheel. Lysates (3mg) were immunoprecipitated with 5 µg of monoclonal HA-Tag antibody overnight. Immunocomplexes were collected using PureProteome™ Protein A Magnetic Beads (MERCK) and eluted by competition with a synthetic HA peptide (Sigma). Eluate was digested with trypsin and analyzed by liquid chromatography-mass spectrometry (LC-MS) on an LTQ Orbitrap Velos instrument (ThermoFisher). Progenesis ® QI for proteomics software v3.0 (Nonlinear dynamics, UK) was used for MS data analysis using default settings. The LC-MS runs were automatically aligned to an automatically selected reference sample with manual supervision Peak lists were generated from Progenesis and loaded to Proteome Discoverer v2.1 (Thermo Fisher Scientific) for protein identification. Proteins were identified using Mascot v2.5 (Matrix Science, London UK) to search the SwissProt database (2018_11, taxonomy restricted to human proteins, 20.240 sequences). Significance threshold for the identifications was set to $p<0.05$, minimum ions score of 20. Statistical analysis was performed using Progenesis software. Proteins displaying greater than 2-fold change, and $p<0.05$ (t-test) between pTINCR and control groups were considered significantly differential. When indicated, interactions were analyzed using CRAPome algorithm [285]. For validation of the results, U2OS cells were co-transfected with pTINCR-HA and CDC42-Flag or NONO-Flag plasmids and immunoprecipitation was performed as explained before. SDS-PAGE was used to visualized anti-HA pTINCR and anti-Flag proteins.

15. Non-covalent SUMO binding assay

Glutathione S-transferase (GST) pull-down experiments were performed by incubating protein extracts from U2OS cells stably expressing pTINCR-HA or pTINCR SIMmut treated with Doxycycline for 48 hours, together with GST, GST-SUMO1 or GST-SUMO2/3 bound onto glutathione-sepharose beads for 2 hours at 4ºC. Then, beads were washed, and bound proteins were eluted, separated on 14% SDS-PAGE and detected by Western blot.

16. SUMO conjugation assays

SUMO conjugation assays were performed by transfecting U2OS stably expressing pTINCR cell line with HIS6-SUMO2 protein. When required, cells were co-transfected

with CDC42 or NONO Flag-tagged plasmids. For p53 and B23, SUMOylation of the endogenous protein was assessed. After forty-eight hours of doxycycline treatment, cells were lysed with 6 M guanidinium-HCl, 0.1 M Na_2HPO_4/NaH_2PO_4 buffer. Then, lysates were mixed with Ni^{2+}-nitrilotriacetic acid-agarose beads and incubated for two hours at room temperature. After washing, beads were resuspended in Laemmli buffer and purified proteins were subjected to SDS-PAGE electrophoresis as indicated above.

17. Alternative splicing analysis

17.1. RNA isolation and library construction

A549 cells were cultured and treated with doxycycline for microprotein expression induction for 24 hours. Total RNA was isolated with Trizol following manufacturer's protocol. RNA quantity and purity were measured with the Nanodrop spectrophotometer (Thermo Scientific). RNA integrity from three replicates per condition was analyzed using a Bioanalyzer (Agilent), and the best two replicates for each condition according to RNA quality were used for RNA sequencing. Strand-specific libraries were built and sequenced at the Center for Genomic Regulation (CRG) Genomics Unit using 125 bp paired-end reads.

17.2. RNA-seq analysis

Differential gene expression. Reads alignment was performed using STAR [272] (genome annotation hg19), and differential gene expression analyses were performed using CuffDiff [286], setting the following parameters: $\log_2$ fold change ≥ 0.6 or ≤−0.6, ΔFPKM > 1, status OK, significant YES.

Gene ontology. Functional enrichment of alternatively spliced genes was performed using GOrilla (http://cbl-gorilla.cs.technion.ac.il/), comparing with a background set provided by VAST-TOOLS as a control.

Alternative splicing analysis. Splicing analyses of RNA-seq data were performed using VAST-TOOLS v2.2.2 using the hg19 library [287, 288]. For all events, a minimum read coverage of 10 actual reads per sample was required. PSI values for single replicates were quantified for all types of alternative events, including single and complex cassette exons, alternative 5′ss and 3′ss (Alt5, Alt3) and retained introns. A minimum absolute ΔPSI of 15% was required to define significant differentially spliced events (|ΔPSI| ≥ 15%), as well as a minimum range of 5% between the PSI values of the two samples. See https://github.com/vastgroup/vast-tools for details.

18. CDC42 activation assay

The activation of CDC42 (CDC42-GTP) was estimated using the CDC42 activation kit (Cytoskeleton Inc.). Briefly, after 4 days of pTINCR doxycycline induction cells were washed with cold PBS, quickly lysed using provided lysis buffer and concentrations determined using the Pierce™ BCA Protein Assay Kit (Thermo Fisher). GTP-bound CDC42 was pulled-down by GST-p21 binding domain (PBD) and detected with anti-

CDC42 antibody (**Table 5**) according to the manufacturer's instructions. Positive control was the result of treating control cells with EGF for 5 minutes prior to the procedure.

19. Human cytoskeleton phospho-array

Phosphorylation-specific antibody microarray (Fullmoon Biosystems Inc.) was used to determine the up- and down- phophorylated proteins in hSCC10 cells after 1 day of pTINCR induction. The array contains 141 site-specific antibodies against phosphorylated and unphosphorylated proteins involved in cytoskeletal pathways, each replicated six times. Actin and GAPDH are included as loading controls. The assay was performed following manufacturer's protocol.

20. Mouse experiments

20.1. Teratoma formation

For subcutaneous teratomas mESCs V6.4 expressing either pTINCR HA-tagged constructs or control vector were trypsinized and $1x10^6$ cells were subcutaneously injected into the flanks of 8-week-old immunocompromised NMRI mice. Teratoma growth was monitored every two days using the formula height x width x width x (3,1416/6) and animals were sacrificed when tumors reached 1.5 cm^3 or at the indicated time. Doxycycline was administered in the drinking water at 1mg/ml supplemented with 7.5% of sucrose every two days throughout all the experiment.

20.2. Xenograft of transformed NIH3T3 cells

hSCC10 cells transduced with pINDUCER20-pTINCR HA or its empty vector were trypsinized and $1x10^6$ cells were subcutaneously injected into the flanks of 8-week-old immunocompromised NMRI mice. Tumor growth was monitored every two days using the formula height x width x width x (3,1416/6) and animals were sacrificed when tumors reached 1.5 cm^3 or at the indicated time. Doxycycline was administered in the drinking water at 1mg/ml supplemented with 7.5% of sucrose every two days throughout all the experiment.

21. In *silico* analyses

21.1. Codon conservation

TINCR coding potential was assessed using PhyloCSF [15], a comparative genomics algorithm that analyzes a multispecies nucleotide sequence alignment and score it according to phylogenetic codon conservation.

21.2. lncRNA structure prediction

TINCR secondary RNA structure was predicted using RNAfold web server from ViennaRNA web.

21.3. Protein features prediction

The different characteristics of the microprotein were predicted using publicly available prediction softwares in different webservers. Multiple sequence alignment performed by Clustal Omega. Three dimensional protein-structure prediction was obtained with I-Tasser software [289]. SUMO-interacting motives (SIMs) were predicted by GPS-SUMO 2.0 [290].

21.4. RNA-Seq Analysis of TCGA data

Transcriptomic data were extracted from TCGA (database https://portal.gdc.cancer.gov/) in the form of count tables or from Das Mahapatra *et al*. study (cSCC cases) [291] in the form of fastq files. In this last case, paired-end reads from RNA-Seq were aligned using Tophat v.2.1.0 [292] to the human genome (hg19). Predicted transcripts from Ensembl database (release 87) were analyzed. Differentially expressed genes (DEG) were identified using HTSeq + DESeq2 v. 1.24.0 [276, 293].

21.5. LC-MS/MS identification of pTINCR microprotein

Identification of pTINCR peptides was performed using publicly available data of organotypic skin cell cultures from Elias *et al*. [294]. In brief, dataset with identifier PDX014088 was downloaded from ProteomeXchange. Raw files were processed with Proteome Discoverer 2.2 (Thermo Fisher Scientific) using the standard settings against a human protein database (UniProtKB/Swiss-Prot, July 2018, 20.373 sequences), supplemented with contaminants and the pTINCR sequences. Carbamidomethylation of cysteines was set as a fixed modification whereas methionine oxidation and N-terminal acetylation were variable protein modifications. The minimal peptide length was set to 6 amino acids and a maximum of two tryptic missed cleavages were allowed. The results were filtered at 0.01 FDR (peptide and protein level).

21.6. Generation of Kaplan-Meier Plots

Kaplan-Meier (KM) plots were generated for survival analysis using KM plotter database (http://kmplot.com/analysis), a website database based on resources from TCGA database. The final prognostic KM plots were presented with a hazard ratio (HR), 95% confidence interval (CI) and log-rank P value. P value < .05 was considered statistically significant

22. Ribosome profiling analysis

We adapted a computational approach to identify translated sORFs [16] by analyzing a public ribosome profiling dataset from mouse skin (GSE83332, [52]). In brief, read adapters were trimmed and reads mapping to annotated ribosomal and transfer RNAs were filtered out. Resulting reads were mapped to the assembled mouse genome (mm10) using Bowtie2 v2.3.4.3 with default parameters. Next, mapped reads from experimental replicates were merged and we used the ribORF algorithm [295] to predict translated sORFs in *TINCR* lncRNA with significant read uniformity and frame periodicity (score ≥ 0.7), as this feature is indicative of ORF active translation.

23. *In vitro* transcription/translation

lncRNA *TINCR* CDS was cloned in a pcDNA-3.1 vector under the control of the T7 promoter and incubated for 1 hour with rabbit reticulocyte-coupled transcription/translation system (Promega) in the presence of [35^S] methionine. After incubation, translated product was resolved by 15% SDS-polyacrylamide gel electrophoresis (PAGE) and detected by autoradiography.

24. Histology and immunohistochemistry

Immunostaining was performed on formalin-fixed paraffin-embedded (FFPE) mouse and human tissues. In brief, paraffin blocks were sliced into 5 µM thick sections, deparaffinized with xylene (Fisher Scientific, Waltham, MA, USA) and rehydrated with decreasing concentrations of ethanol in water. For immunohistochemistry, paraffine sections underwent antigenic exposure process into the Discovery Ultra (Ventana) system with CC1 buffer for 64 minutes at 95ºC. Anti-pTINCR antibody was incubated for 1 hour at room temperature (**Table 5**). Next, slides were incubated with the secondary antibody Discovery UltraMap anti-Rabbit HRP (Ventana). Hematoxylin and eosin (H&E) staining was performed on 5 µm paraffin sections in a Robust carousel tissue stainer (Slee Medical) according to common method.

25. Human cSCC samples collection and processing

A historical cohort of 51 patients diagnosed with cSCC from January 2009 to August 2010 was included in the present study provided by the Tumor Bank of the Vall d'Hebron University Hospital Biobank. For histological examination, H&E staining of FFPE tumor sample was performed. All cases were evaluated independently by two pathologists.

For immunohistochemistry studies of pTINCR and p53, whole slide FFPE tissue sections of 5 µm of selected samples were stained as previously described. All cases were evaluated independently by one expert dermatopathologist and one trained molecular biologist blinded for patient groups, considering the percentage of positive cells and intensity of the staining, which was assessed semiquantitatively. Final results were obtained utilizing the average of the two values. Whenever a major discrepancy was observed between both observers, the cases were discussed using a multi-headed microscope.

For mutational profiling, FFPE tissue sections of 10 µm were stained with H&E and examined by a pathologist to select for a minimum tumor content of 20% as a requirement for further processing.

25.1. Amplicon-seq analysis

DNA was extracted from 5x10 µm sliced FFPE sections using the Maxwell FFPE Tissue LEV DNA Purification Kit (Promega), according to the manufacturer's instructions. DNA quality and concentration were determined by fluorometric quantification using Qubit Fluorimeter and Qubit dsDNA BR Assay Kit (Life Technologies, Carlsbad, CA).

Tumor DNA was sequenced with an in-house developed amplicon-sequencing panel of over 60 genes and 1330 primer pairs targeting frequent mutations in oncogenes and several tumor suppressors. A total of 500 ng of DNA from each tissue sample were used for library preparation, according to our established protocols. An initial multiplex-PCR with a proof-reading polymerase was performed on all samples. Indexed libraries were pooled and sequenced in a MiSeq instrument (2X100) at an average coverage of 1000X. Initial alignment was performed with BWA after primer sequence clipping and variant calling was done with the GATK Unified Genotyper and VarScan2 followed by ANNOVAR annotation. Mutations were called at a minimum 3% allele frequency. SNPs were filtered out with dbSNP and 1000 genome datasets. All detected variants were manually checked.

26. Statistical analysis

Data is expressed as the mean ± standard deviation (SD) or the mean ± standard error of the mean (SEM), as specified. Differences between groups were analyzed using unpaired Student's t-test, one or two-way ANOVA, Fisher exact test, multiple comparison t-test or non-parametric Kruskal-Wallis test, as specified. Corrections for multiple comparison were performed when necessary (one and two-way ANOVA with Bonferroni method, multiple comparison t-test with Holn-Sidak method and Kruskal-Wallis test with Dunn's test). All statistical tests were two-sided and performed using GraphPad Prism (GraphPad Software Inc., San Diego, CA, USA).

27. Ethical statement

The animal studies in this work comply with the European, Spanish and Catalan Regulations for the Protection of Vertebrate Animals used for Experimental and other Scientific Purposes (Directive 2010/63; Spanish BOE RD 53/2013; Catalan DOGC 214/1997). All studies were carried out in the "Lab Animal Service Campus Vall d'Hebron (LAS-CVH)", registered and accredited at the Department de Medi Ambient i Habitatge by Generalitat of Catalonia government with register number B9900062. The experiments performed for this manuscript were linked to a project approved by Vall d'Hebron Ethics Committee, and the Commission of Animal Experimentation of Generalitat of Catalonia government.

Studies involving human samples have been performed in accordance with the Declaration of Helsinki. All patient samples studied have the approval of the Ethics Committee for Research of the Vall d'Hebron University Hospital (PR research project PR research project (AG) 191/2019)). Samples recruitment was not done specifically for our study.

28. Data availability

RNA-seq raw data will be available from the GEO database before publication (GSE175463). Mass spectrometry data will be also available from the ProteomeXchange Consortium/PRIDE repository (PXD026181). Published datasets included in this study: PDX014088, GSE83332, GSE139505, GSE58506, GSE100292. The data that support the findings of this study are available from the corresponding author upon reasonable request.

Table 1: Constructs for pTINCR exogenous expression

In capital letters, pTINCR sORF; in italics, linker sequence; in bold, HA-Tag sequence.

Name	Backbone	ORF
HA-pTINCR	pInducer20 /pMSCV	ATG**tatccttatgatgtgcctgattatgct***ggcggtggaggcagcggaggtgggggaag tggtggcggtggaagc*GAGGGGCTGCGGCGGGGGCTGTCGCGCTG GAAGCGCTACCACATCAAGGTGCACCTGGCGGACGAGGCGCT GCTGCTACCGCTGACCGTGCGGCCGCGGGACACGCTCAGCG ACCTGCGCGCCCAGCTGGTGGGCCAGGGCGTGAGCTCCTGG AAGCGCGCCTTCTACTACAACGCGCGGCGGCTGGACGACCAC CAGACGGTGCGCGACGCGCGCCTGCAGGACGGCTCGGTGCT GCTGCTCGTCAGCGACCCCAGGTGA
pTINCR-HA	pInducer20 /pMSCV	ATGGAGGGGCTGCGGCGGGGGCTGTCGCGCTGGAAGCGCTA CCACATCAAGGTGCACCTGGCGGACGAGGCGCTGCTGCTACC GCTGACCGTGCGGCCGCGGGACACGCTCAGCGACCTGCGCG CCCAGCTGGTGGGCCAGGGCGTGAGCTCCTGGAAGCGCGCC TTCTACTACAACGCGCGGCGGCTGGACGACCACCAGACGGTG CGCGACGCGCGCCTGCAGGACGGCTCGGTGCTGCTGCTCGT CAGCGACCCCAGG*ggcggtggaggcagcggaggtgggggaagtggtggcggtg gaagc***tatccttatgatgtgcctgattatgct**TGA
pTINCR-HA (syORF)	pInducer20	ATGGAGGGCCTGAGAAGAGGCCTGAGCAGATGGAAGAGATAC CACATCAAGGTGCACCTGGCCGACGAGGCCCTGCTGCTGCCA CTGACAGTGAGGCCCAGGGACACCCTGAGCGACCTGAGAGC CCAGCTGGTGGGCCAGGGCGTGAGCTCTTGGAAGAGAGCCTT CTACTACAACGCCAGGAGGCTGGACGACCACCAGACAGTGAG GGATGCCAGGCTGCAGGACGGCAGCGTGCTGCTGCTGGTGT CCGACCCCAGG*ggcggaggaggaagcggaggaggaggctccggcggaggag gttcc***taccttacgacgtgcctgactacgcc**TGA
pTINCR-SIMmut (syORF)	pInducer20	ATGGAGGGCCTGAGAAGAGGCCTGAGCAGATGGAAGAGATAC CACATCAAGGTGCACCTGGCCGACGAGGCCCTGCTGCTGCCA CTGACAGTGAGGCCCAGGGACACCCTGAGCGACCTGAGAGC CCAGCTGGTGGGCCAGGGCGTGAGCTCTTGGAAGAGAGCCTT CTACTACAACGCCAGGAGGCTGGACGACCACCAGACAGTGAG GGATGCCAGGCTGCAGGACGGCAGCGCCGCTGCTGCCGTGT CCGACCCCAGG*ggcggaggaggaagcggaggaggaggctccggcggaggag gttcc***taccttacgacgtgcctgactacgcc**TGA

Table 2: Plasmids

Retroviral packaging vectors	Retroviral vectors
pCL-Ampho	pMSCV
Lentiviral packaging vectors	Lentiviral vectors
pLP1	pINDUCER20
pLP2	pENTR1A
pLP- VSVG	pDONR201

Table 3: Mouse primers

Primer Name	Forward	Reverse
mGAPDH	TGTGTCCGTCGTGGATCTGA	TTGCTGTTGAAGTCGCAGGAG
mTINCR	TACGAACAGAACAGGAGGAC	CTCTCCACATTGTGGCTTTG

Table 4: Human primers

Primer Name	Forward	Reverse
GAPDH	GGACTCATGACCACAGTCCATGCC	TCAGGGATGACCTTGCCCACAG
lncRNA TINCR	AAGCGCTACCACATCAAGGT	CCGCGCGTTGTAGTAGAAG
EXO HA-pTINCR	CTTATGATGTGCCTGATTATGC	CACGGTCAGCGGTAGCAG
EXO pTINCR-HA	GGTGCTGCTGCTCGTCAG	TCAGGCACATCATAAGGATAGC
pTINCR-201	GGTGCTGCTGCTCGTCAG	TTCCTTCAGCCAGTACCCAG
pTINCR-204	GGTGCTGCTGCTCGTCAG	ACCATGTTGGCCAGGCTGG
CDKN1A	CCTGTCACTGTCTTGTACCCT	GCGTTTGGAGTGGTAGAAATCT
PCNA	GGCCGAAGATAACGCGGATAC	GGCATATACGTGCAAATTCACCA
Cyclin E	ATCAGCACTTTCTTGAGCAACA	TTGTGCCAAGTAAAAGGTCTCC
Cyclin B	AATGAAATTCAGGTTGTTGCAGGAG	CATGGCAGTGACACCAACCAG
KERATIN 14	GACCATTGAGGACCTGAGGA	AGACGGGCATTGTCAATCTG
KERATIN 5	GCTGCTGCGTGAGTACCAG	CTGGTCCAACTCCTTCTCCA
KERATIN 1	GGCAGTTCCAGCGTGAAGTTTGTT	TTCTCCGGTAAGGCTGGGACAAAT
KERATIN 10	GAGCAAGGAACTGACTACAG	CTCGGTTTCAGCTCGAATCT
INVOLUCRIN	TGCCTGAGCAAGAATGTGAG	TGCTCTGGGTTTTCTGCTTT
FILAGGRIN	CATGGCAGCTATGGTAGTGCAGA	ACCAAACGCACTTGCTTTACAGA
EPCAM	ATAACCTGCTCTGAGCGAGTG	TGCAGTCCGCAAACTTTTACTA
TP63	GACGTGTCCTTCCAGCAGTC	GGGGTCATCACCTTGATCTG
ESR1	CCCACTCAACAGCGTGTCTC	CGTCGATTATCTGAATTTGGCCT

HES1	TCAACACGACACCGGATAAAC	GCCGCGAGCTATCTTTCTTCA
KERATIN 18	TCGCAAATACTGTGGACAATGC	GCAGTCGTGTGATATTGGTGT
ID3	TCATCTCCAACGACAAAAGG	ACCAGGTTTAGTCTCCAGGAA
XBP1	TGGCCGGGTCTGCTGAGTCCG	ATCCATGGGGAGATGTTCTGG
EGFR	GGCACTTTTGAAGATCATTTTCTC	CTGTGTTGAGGGCAATGAG
ERBB2	TGTGACTGCCTGTCCCTACAA	CCAGACCATAGCACACTCGG
ITGB3	CGCTACAAGGGGGAGATGT	TACGCGTGGTACAGTTGCAG
VIMENTIN	GACAATGCGTCTCTGGCACGTCTT	TCCTCCGCCTCCTGCAGGTTCTT
CDH2	CTGCACAGATGTGGACAGGA	CCACAAACATCAGCACAAGG
SNAIL1	ACCACTATGCCGCGCTCTT	GGTCGTAGGGCTGCTGGAA
SLUG	TCGGACCCACACATTACCTT	ATGAGCCCTCAGATTTGACCT
ZEB1	GCACCTGAAGAGGACCAGAG	TGCATCTGGTGTTCCATTTT
ZEB2	CGCTTGACATCACTGAAGGA	CTTGCCACACTCTGTGCATT

Table 5: Antibodies

Target	Reference	Dilution WB	Dilution IF	Dilution IHC
HA-Tag	Ab9110, Abcam	1:5000	1:150	
GAPDH	AM4300, Thermo Fisher Scientific	1:10000		
pTINCR	Custom-made antibody, Abyntek	1:50	1:200	1:500
Phalloidin-TRITC (ACTIN)	P1951, Sigma-Aldrich		20µg/ml	
CDC42 (B8)	SC-8401, Santa Cruz	1:100	1:100	
Lamin B1	PA5-19468, Thermo Fisher Scientific	1:1000		
Tubulin	SC-32293, Santa Cruz	1:1000		
Histone 3	Ab18521, Abcam	1:10000		
VDAC	SC-390996, Santa Cruz	1:500		
ZO1	GTX108592, GeneTex		1:100	
β-catenin	GTX633010, GeneTex		1:100	
E-cadherin	610182, BD		1:100	
SUMO1	4930S, Cell Signaling	1:1000	1:100	
SUMO2/3	4971S, Cell Signaling	1:1000		
HIS-Tag (D3I1O) XP®	12698, Cell Signaling	1:1.000		
FLAG-Tag	F1804, Sigma	1:1000		
INVOLUCRIN (clone SY5)	MS-126 P1, Thermo Fisher Scientific	1:500		
p53 (DO-1)	SC-126, Santa Cruz	1:1000		1:500
B23	SC-5564	1:1000		
AIF	SC-13116, Santa Cruz		1:100	
Calreticulin	NBP1-47518SS, NovusBio		1:100	

RESULTS

1. Identification of pTINCR microprotein

1.1. The lncRNA *TINCR* encodes for an 87-amino acid microprotein

The first objective of this work was to identify novel microproteins with key roles in cellular identity and cancer. To start with, we followed a computational approach for microprotein identification and used PhyloCSF, a comparative genomics algorithm that analyzes codon substitution frequencies across evolution to predict coding sequences based in phylogenetic conservation [15]. Among the "assumed" non-coding regions that could encode for microproteins, we decided to focus our search in lncRNAs, since most of them are already annotated and in many cases their expression, regulation and function are characterized. This information is very helpful to identify microproteins related with a specific cellular process, given that you can look for microproteins encoded by lncRNAs deregulated in the process of interest.

In our case, we narrowed our search to lncRNAs known to be highly expressed in skin and deregulated in Cutaneous Squamous Cell Carcinoma (cSCC) [291]. By analyzing the coding potential of these lncRNAs using PhyloCSF, we found *TINCR* lncRNA as a good candidate (**Fig. 12A**). Manual *in silico* analysis of *TINCR* locus revealed a previously unrecognized sORF of 264 bp with a strong Kozak sequence (**Fig. 12B and C**), a motif that usually signals for a protein translation initiation site in most eukaryotic mRNA transcripts.

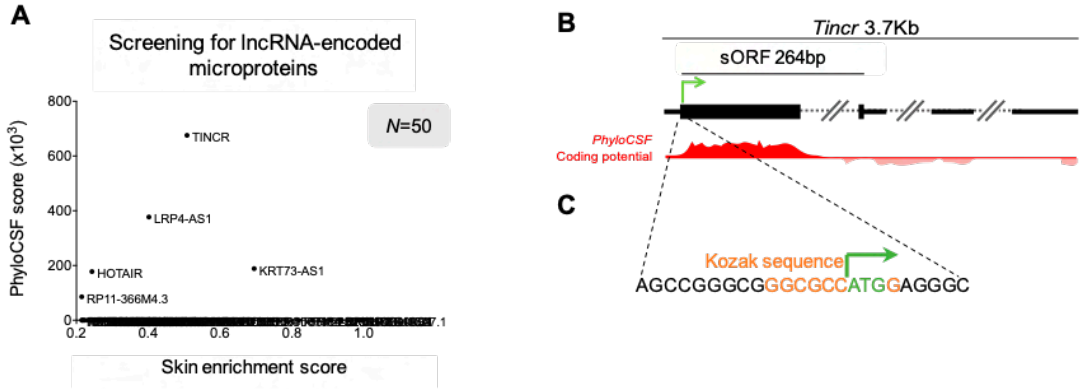


Figure 12. *TINCR* is a lncRNA with a highly conserved sORF.
A. Screening of the coding potential of lncRNAs using PhyloCSF algorithm. Graph shows lncRNA candidates whose expression is enriched in skin compared to other organs [291]. PhyloCSF score value was obtained from LNCipedia v.5.2.
B. Diagram of *TINCR* locus. The sORF is indicated with a big black square. In red, PhyloCSF score across *TINCR* lncRNA.
C. Schema of the Kozak sequence around the initiation codon of *TINCR*-encoded microprotein.

To support the coding potential of our candidate, we analyzed publicly available data of ribosome profiling experiments in mouse skin using RiboORF, a software that identifies regions of active translation based on read distribution features such as codon periodicity and uniformness [16]. This analysis revealed that *TINCR* is translated into an 87-amino acid microprotein (RiboORF score ≥ 0.7) highly conserved across the tetrapoda taxa (**Fig. 13A and B**). We have named this novel microprotein pTINCR.

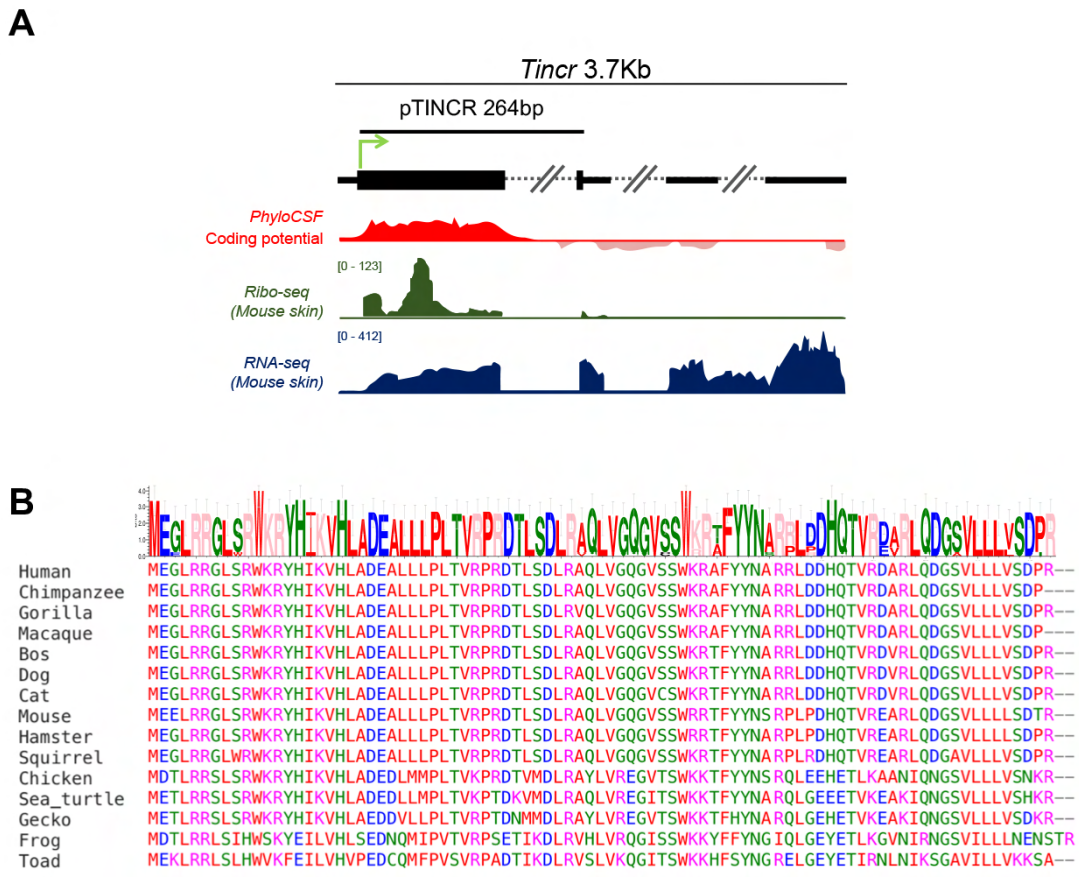


Figure 13. pTINCR is an evolutionarily conserved microprotein potentially translated in mouse skin.
A. Diagram of *TINCR* locus. The sORF encoding pTINCR is indicated with a big black square. In red, PhyloCSF score across *TINCR* lncRNA. In green, Ribo-seq analysis of *Tincr* transcript in mouse skin. In blue, expression of *Tincr* lncRNA in mouse skin by RNAseq.
B. pTINCR amino acid conservation across the tetrapoda taxa depicted by multiple sequence alignment performed by Clustal Omega. The colors of the amino acid residues indicate their properties (pink, positive charge; blue, negative charge; red, hydrophobic; green, hydrophilic). The symbols below the alignment represent the biochemical similarity of aligned amino acids, with asterisks indicating identical conservation, colons representing high similarity, and periods showing somewhat similar conservation.

Next, we looked for further experimental evidence of pTINCR translation. First, we performed *in vitro* translation using the full-length *TINCR* lncRNA (ENSEMBL: ENST00000448587.5) in the presence of 35^{S}-methionine and obtained a peptide product of ~12 kDa (**Fig. 14A and B**), confirming that pTINCR is translated into a stable microprotein. Additionally, by mutating the first ATG codon of *TINCR* sORF we abolished the translation of any detectable product, hence validating the starting codon of pTINCR (**Fig. 14A and B**). Second, we analyzed public mass spectrometry data from human organotypic skin cultures [294] and identified three different tryptic peptides corresponding to pTINCR, providing strong experimental evidence of its translation in skin (**Fig. 14C**).

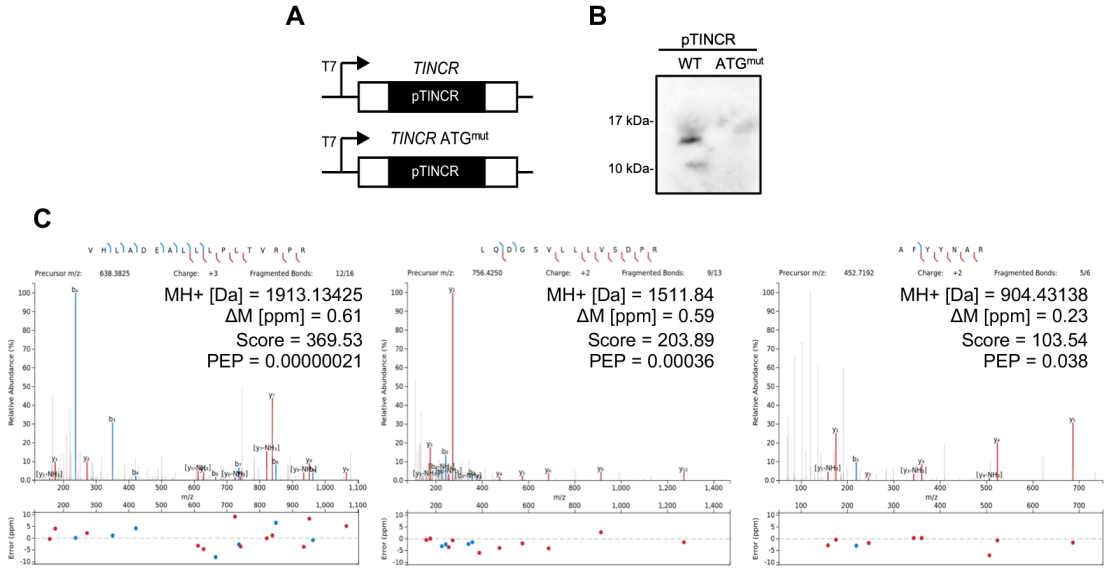


Figure 14. pTINCR is actively translated.
A. Diagram of the constructs used for *in vitro* translation of pTINCR microprotein.
B. *In vitro* translation of the WT and ATGmut full-length *TINCR* lncRNA in the presence of radiolabeled 35^{S}-methionine, visualized by SDS-PAGE.
C. MS/MS spectrum of unique peptides derived from pTINCR microprotein identified by mass spectrometry in human skin organotypic cultures.

Altogether, we demonstrated that *TINCR* is a transcript misannotated as a lncRNA that actually codes for pTINCR microprotein.

1.2. Analysis of *TINCR* locus: expression, regulation, and variants

1.2.1. *TINCR* is expressed in epithelial tissues

TINCR is a 3.7-kilobase lncRNA previously described to be expressed in skin but also in other epithelial tissues (Kretz et al., 2013, Omote et al., 2021; Zhuang et al., 2020). We compared the expression pattern of *TINCR* gene in mouse and human tissues. For human tissues, we reanalyzed GTEX data on gene expression. For mouse tissues, we evaluated *Tincr* expression on a wide panel of organs by RT-qPCR. Our analysis corroborated that *TINCR* is highly expressed in skin, but also in other epithelia (**Fig. 15A**). We also screened *TINCR* expression in different established cell lines. We confirmed the high expression of *TINCR* in some epithelial-derived cell lines (**Fig. 15B**), such as the human keratinocyte cell line HaCaT and the luminal breast cancer cell line MCF7, both documented to express high levels of *TINCR* lncRNA according to the Human Protein Atlas project.

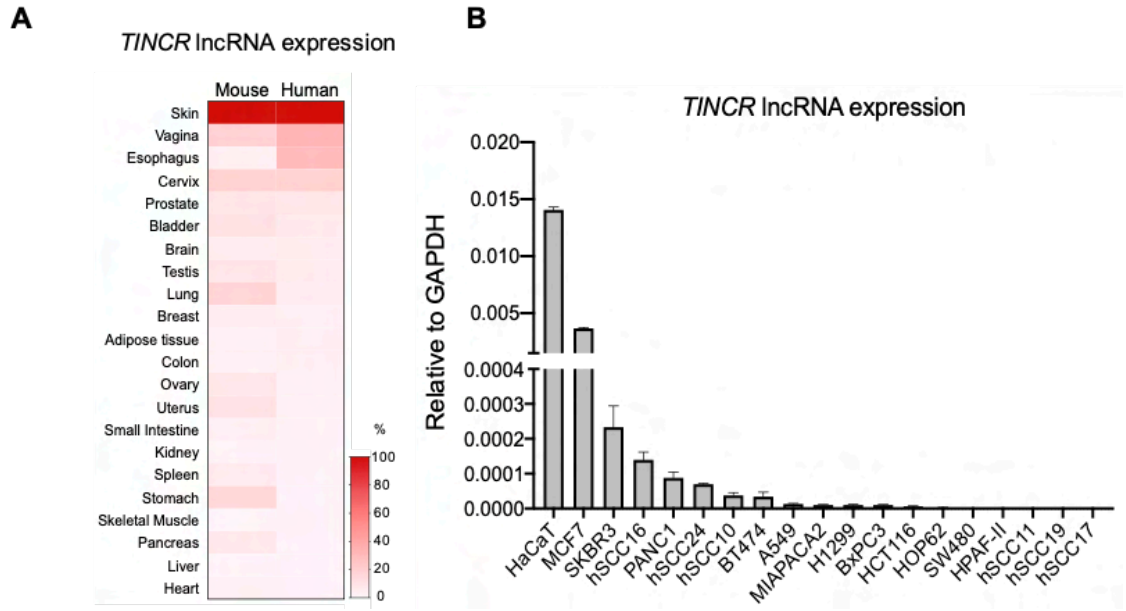


Figure 15. *TINCR* lncRNA is mainly expressed in skin and other epithelia.
A. Heat map of *TINCR* lncRNA expression in human and mouse tissues. Expression in human was obtained from GTEX. Expression in mouse was assessed by RT-qPCR (at least 3 mice per organ). Values represent the percentage of *TINCR* expression in each organ normalized to the expression in the skin.
B. *TINCR* lncRNA expression in different cell lines analyzed by RT-qPCR. Error bars represent the mean ± SD.

1.2.2. *TINCR* is upregulated during epithelial differentiation

Transcriptional upregulation of *TINCR* locus has been defined as a typical event during differentiation of keratinocytes [296] and other epithelial tissues [297, 298]. To confirm this, we followed a known stablished culture protocol in which epithelial differentiation is triggered by modulating the Ca^{2+} concentration in the culture medium [299-301]. Consistent with previous reports, we observed that *TINCR* mRNA expression was higher in differentiated (cultured under high Ca^{2+} concentration) compared to non-differentiated cells (cultured under low Ca^{2+} concentration) in both HaCaT and MCF7 cell lines (**Fig. 16**).

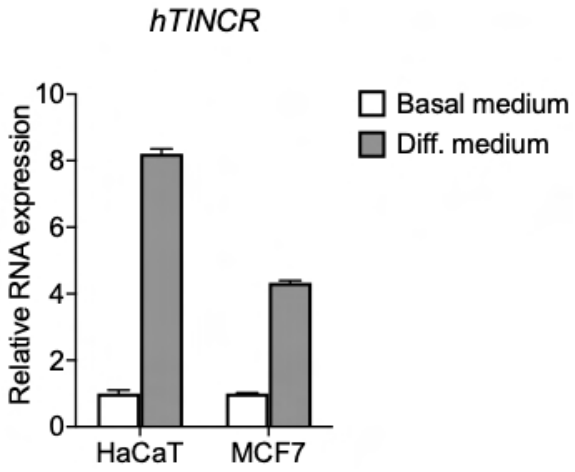


Figure 16. *TINCR* lncRNA is upregulated upon calcium-induced epithelial differentiation.
TINCR lncRNA expression in HaCaT and MCF7 cells under basal and differentiation culture conditions measured by RT-qPCR. mRNA expression is normalized to GAPDH expression and relativized to the basal condition. Error bars represent the mean ± SD.

1.2.3. *TINCR* is upregulated upon cellular stress in a p53 dependent manner

Previous studies in cardiomyocytes have reported that *TINCR* locus is upregulated by p53 upon damage [302]. To corroborate this, we treated several cell lines with different p53 mutational status (WT p53, mutant p53 or knock-out p53) with p53 inducers or stabilizers and analyzed *TINCR* expression levels by RT-qPCR. Specifically, we used doxorubicin and actinomycin-D (two chemotherapeutic drugs that induce genotoxic stress) or nutlin-3a (an inhibitor of MDM2, a E3 ubiquitin-ligase that promotes p53 degradation) (**Fig. 17A**). Our results showed that only cells with functional p53 upregulate *TINCR,* while those with mutant p53 KO p53 did not show the same response (**Fig. 17B**). Therefore, *TINCR* is upregulated by p53 activation upon cellular stress.

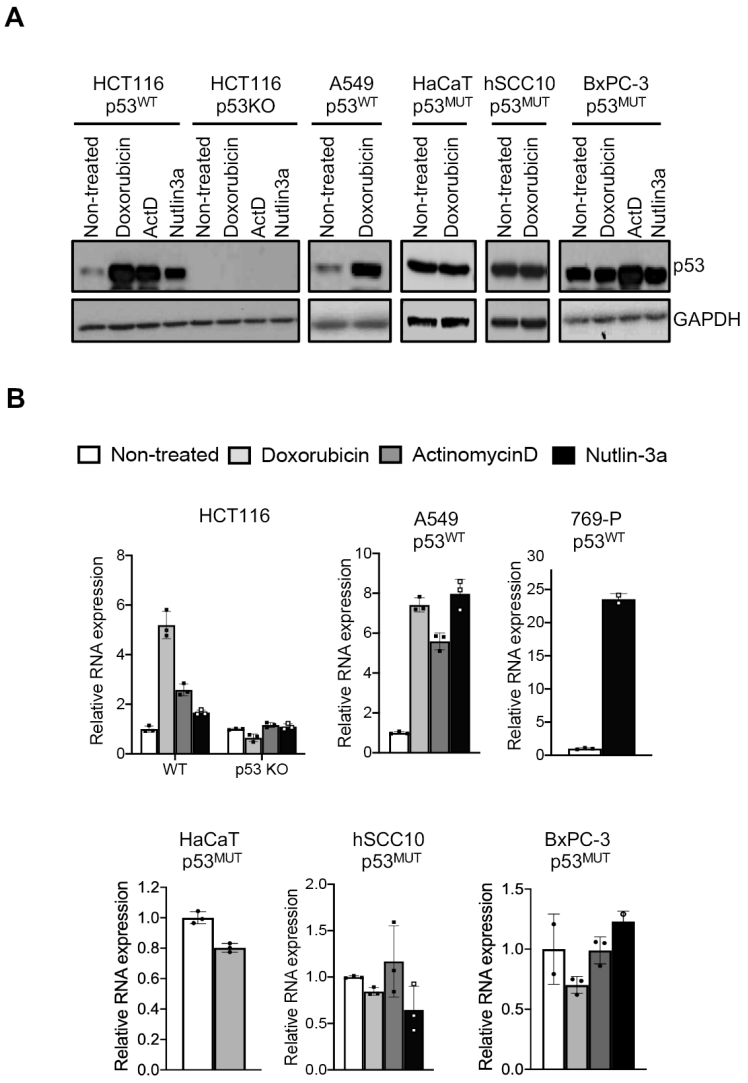


Figure 17. *TINCR* mRNA expression is regulated by p53.
A. Western blot of p53 after 24 hours of treatment with doxorubicin (1 μm), actinomycin-D (5 nm) or nutlin-3a (10 μm) in the indicated cell lines.
B. *TINCR* lncRNA expression levels analyzed by RT-qPCR after 24 hours of treatment with doxorubicin (1 μm), actinomycin D (5 nm) or nutlin-3a (10 μm) in the indicated cell lines. mRNA expression is normalized to GAPDH in each case and relative to the control. Error bars represent the mean ± SD.

1.2.4. *TINCR-201* variant is specifically upregulated during differentiation and upon stress

When this project was started, a unique transcript of *TINCR* lncRNA was annotated in the mouse (*Gm1920*) and human genome (*TINCR*-201). Later, the annotation of human *TINCR* was updated and four new splice variants were described, three of them predicted to encode for different microproteins (**Fig. 18A**). Our *in silico* and experimental studies had already provided strong evidence about the translation of pTINCR sORF within the mouse transcript and the human *TINCR*-201 variant. *TINCR*-202 variant is annotated with a 5' truncation (ENSEMBL: ENST00000587632.1), harboring a shorter form of pTINCR microprotein potentially beginning in a non-canonical starting codon. However, the indicated 5' truncation could imply that the annotation of the real start of the CDS is missing. For this reason, we considered that this transcript could be encoding a microprotein identical from pTINCR and we did not take it into consideration. *TINCR*-204 is predicted to encode for a 120-amino acid protein which shares the N-terminus with pTINCR, but it has and extended C-terminus (**Fig. 18A and B**). Taking advantage of this distinctive feature, we reanalyzed the same mass spectrometry data from human organotypic skin cultures and looked for unique peptides corresponding to putative *TINCR*-204-encoded microprotein. We were able to detect one peptide derived from the C-terminal part of *TINCR*-204 sORF (**Fig. 18C**), which confirmed its translation in skin.

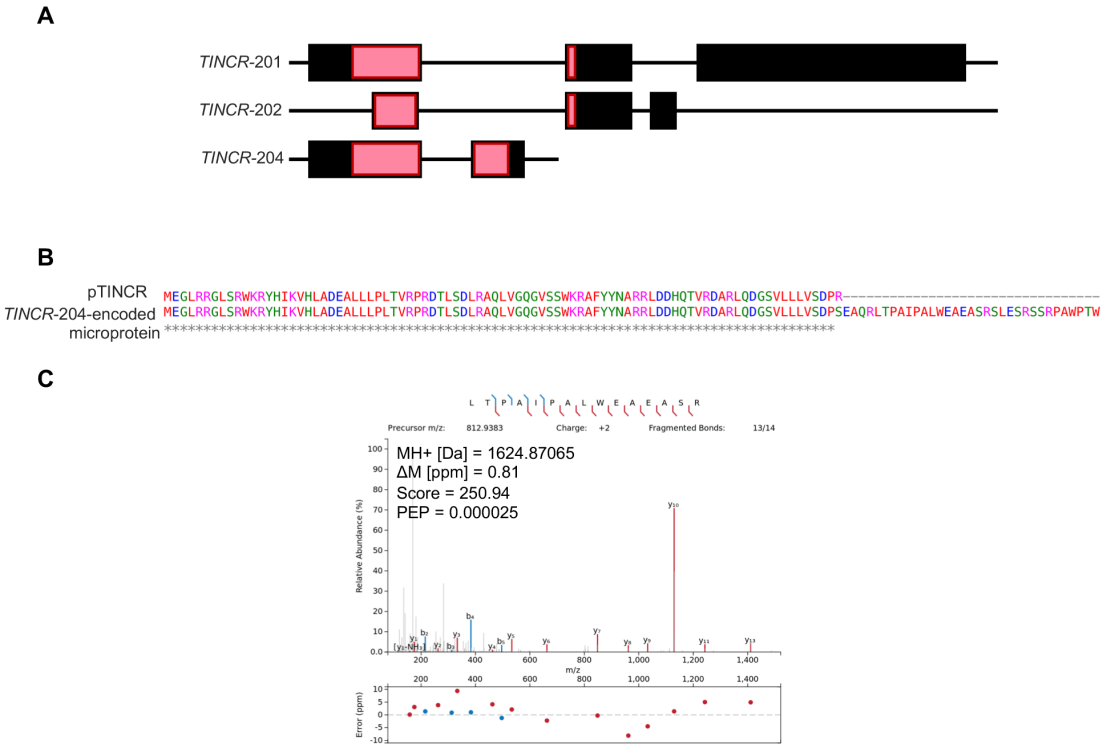


Figure 18. *TINCR* lncRNA transcripts and their encoded microproteins.
A. Diagram of *TINCR* splice variants predicted to encode for microproteins. The different sORFs are indicated with a pink square.
B. Comparison of pTINCR and *TINCR*-204-encoded microprotein amino acid sequence by Clustal Omega. The colors of the amino acid residues indicate their properties (pink, positive charge; blue, negative charge; red, hydrophobic; green, hydrophilic). The symbols below the alignment represent the biochemical similarity of aligned amino acids, with asterisks indicating identical conservation.
C. MS/MS spectrum of a unique peptide derived from *TINCR*-204-encoded microprotein identified by mass spectrometry in human skin organotypic cultures.

To better understand the regulation of *TINCR*-204 variant we measured its expression levels upon calcium-induced differentiation and genotoxic stress and compared it with *TINCR*-201 regulation. RT-qPCR results showed that although basal levels of *TINCR*-204 are higher, only *TINCR*-201 transcript was upregulated upon both stimuli (**Fig. 19A and B**). Several microproteins have been described to be regulated upon cellular stress and are thought to play an important role in coping with damage and adaptation [36-38]. Therefore, we found more interesting to select pTINCR microprotein to further functional and molecular characterization.

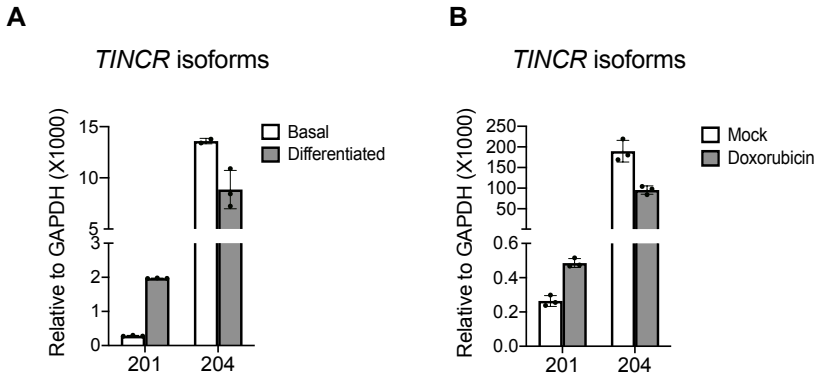


Figure 19. *TINCR*-201, but not *TINCR*-204, responses to calcium-induced differentiation and damage.
A. *TINCR* splice variants expression levels in HaCaT cells under basal and differentiation culture conditions analyzed by RT-qPCR. mRNA expression is normalized to GAPDH in each case. Error bars represent the mean ± SD.
B. *TINCR* splice variants expression levels analyzed by RT-qPCR after 24 hours of treatment with doxorubicin (1 μm) in A549 cell line. mRNA expression is normalized to GAPDH in each case. Error bars represent the mean ± SD.

1.3. Generation of different tools to study pTINCR microprotein

1.3.1. Gain-of-function tools

To start the characterization of pTINCR microprotein, we cloned pTINCR sORF tagged with an HA epitope into the pMSCV constitutive retroviral vector or the pINDUCER20 doxycycline-inducible lentiviral vector. As represented in **Figure 20A**, we generated two different constructs placing an HA-tag either in the C-terminal (pTINCR-HA) or the N-terminal (HA-pTINCR) part of the microprotein. To minimize the possible effect of the tag on pTINCR, we introduced a flexible linker of low interacting amino acids (3xGGGGS) between the tag and pTINCR sORF. Moreover, to uncouple the function of *TINCR* lncRNA and pTINCR microprotein, we also developed a synthetic ORF (syORF) by mutating ~20% of pTINCR-HA nucleotide sequence, significantly changing the secondary structure of the RNA while producing the same protein (**Fig. 20B**).

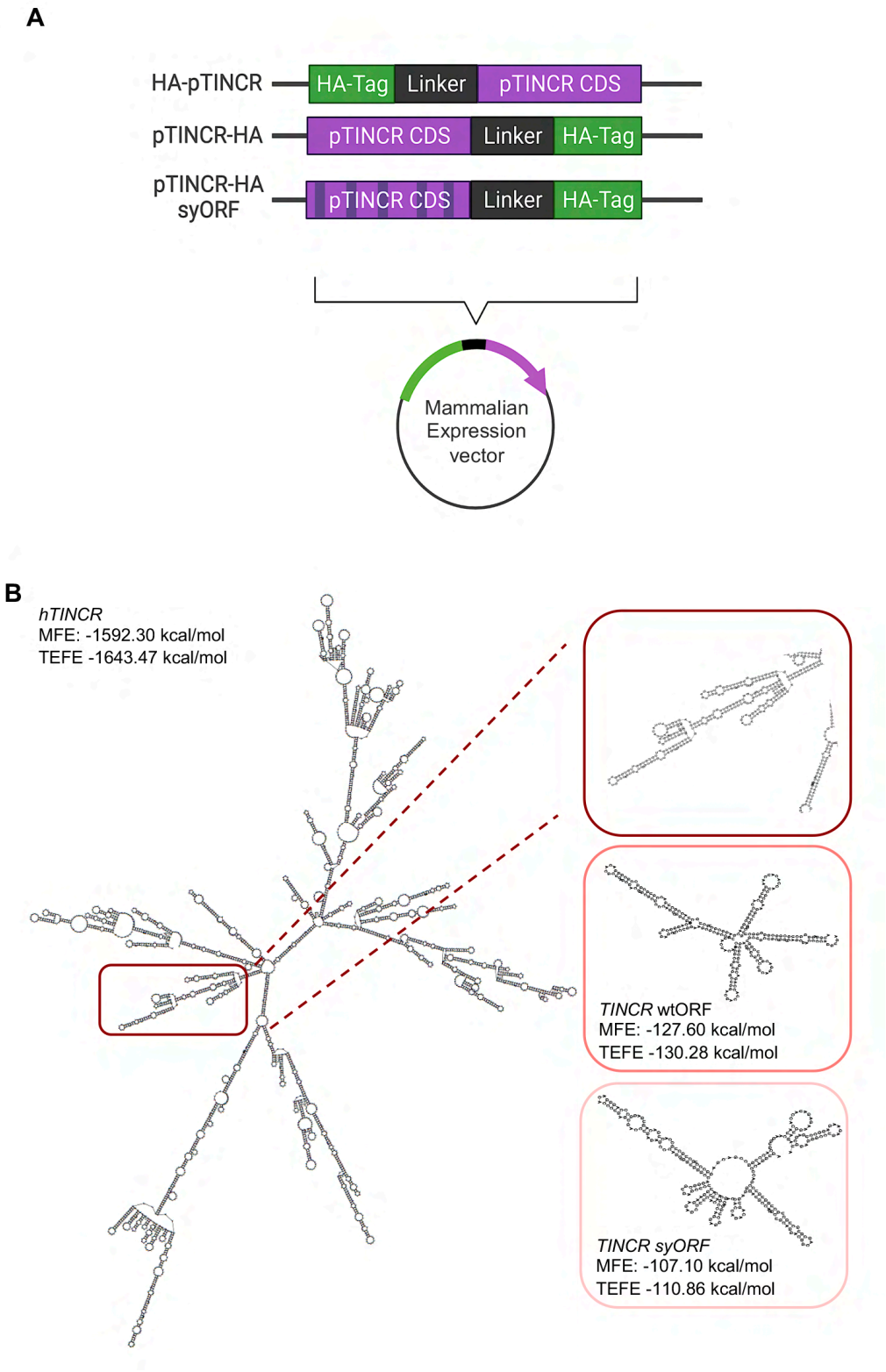


Figure 20. Generation of gain-of-function tools to assess pTINCR function.
A. Diagram representing the HA-tagged pTINCR constructs cloned in mammalian expression vectors.
B. Minimum Free Energy (MFE) structure of the human *TINCR* lncRNA, *TINCR* wtORF and *TINCR* syORF, using RNAfold web server from ViennaRNA web.

We have used these constructs to transduce all different cell lines used across this thesis and generate stable cell lines overexpressing HA-Tagged pTINCR. Although overexpression of both constructs is detected by RT-qPCR **(Fig. 21A)**, N-terminal HA-tagged pTINCR was not always detected by Western blot. We needed high levels of HA-pTINCR (rendered by transient expression) or high stringent protein lysis buffers to properly detect this construct variant by Western blot (**Fig. 21B and C**). For this reason, we always corroborated the overexpression of both constructs by RT-qPCR and by Western blot. Worthy to mentioned, C-terminal and N-terminal pTINCR presented different mobilities in an SDS-PAGE gel. These particularities will be further discussed in the discussion section.

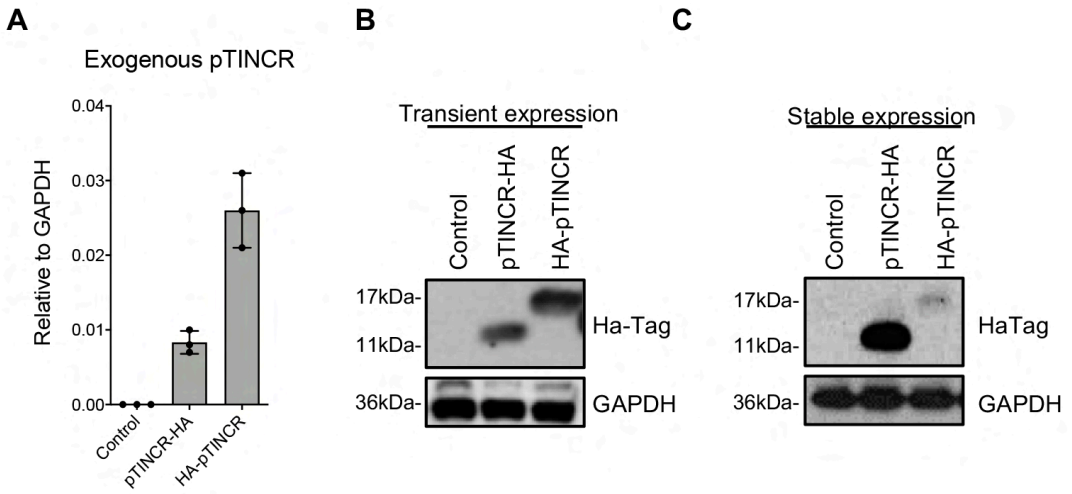


Figure 21. Detection of exogenous pTINCR expression.
A. Exogenous pTINCR expression in U2OS cells transduced with C-terminal (pTINCR-HA) and N-terminal (HA-pTINCR) HA-tagged microprotein by RT-qPCR. mRNA expression is normalized to GAPDH in each case and relative to the control. Error bars represent the mean ± SD.
B and C. Exogenous pTINCR protein levels in U2OS cells analyzed by Western blot using an anti-HA antibody in cells transfected (**B**) or transduced (**C**) with C-terminal (pTINCR-HA) and N-terminal (HA-pTINCR) HA-tagged microprotein.

1.3.1.1. Subcellular localization of pTINCR microprotein

Using our gain-of-function tools, we investigated pTINCR subcellular localization and performed subcellular fractionation (**Fig. 22A**) and immunofluorescence (**Fig. 22B-D**) on HA-tagged pTINCR overexpressing cells. We observed that both constructs displayed similar distributions, being detected mainly in the nucleus and at the cellular membrane. Curiously, we also observed co-localization with mitochondria and partially with the endoplasmic reticulum (ER), although this pattern was only confirmed in some cell lines.

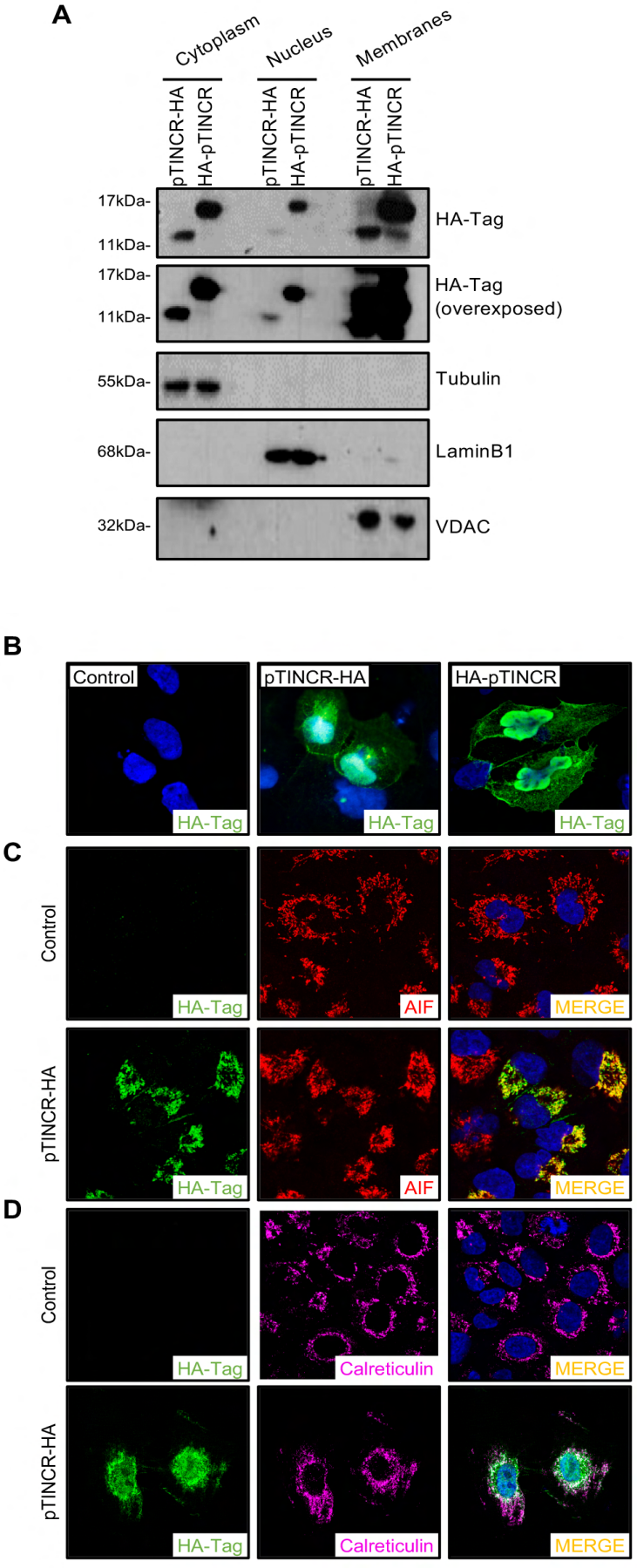


Figure 22. pTINCR is detected in the nucleus, cell-to-cell contacts, mitochondria and partially in the ER.
A. Detection of exogenous pTINCR by Western blot in HaCaT cells after subcellular fractionation. Enrichment for cytosol, nucleus and cell membranes was verified by Western blot analysis using Tubulin, LaminB1 and VDAC, respectively
B. Immunostaining images showing HA-Tagged pTINCR in U2OS cells transiently transfected with C-terminal (pTINCR-HA) and N-terminal (HA-pTINCR) HA-tagged microprotein constructs using an anti-HA antibody.
C and D. Immunostaining images of A549 cells transduced with C-terminal (pTINCR-HA) microprotein showing the co-localization with AIF (**C**) and Calreticulin (**D**) proteins.

1.3.2. Loss-of-function tools

To assess pTINCR loss of function, we generated pTINCR-deficient cells using CRISPR-Cas9 technology in two of our epithelial models, HaCaT and MCF7 cells. It is known that lncRNAs function strongly relies on their secondary and tertiary structure [303, 304]. Therefore, it was crucial to design a strategy that would ensure the lncRNA integrity while disrupting pTINCR translation. We were able to generate pTINCR-KO cells containing a single-nucleotide insertion that disrupts the ATG start codon, which abolish the translation of pTINCR (**Fig. 23A and B**). Importantly, this modification is predicted to have no effect on the secondary structure of *TINCR* transcript (**Fig. 23C**).

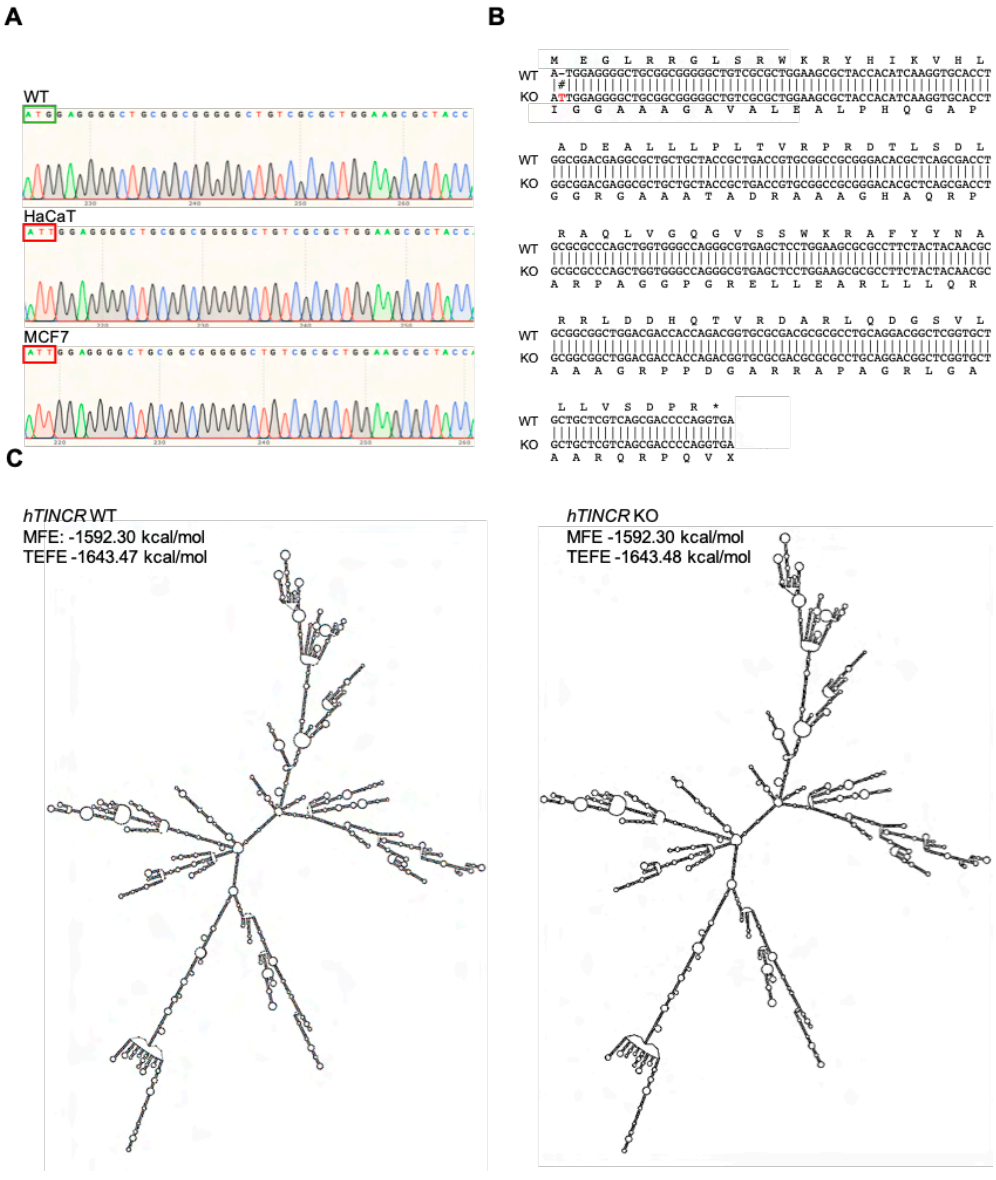


Figure 23. Generation of pTINCR-KO cells.
A. Sequencing of WT and pTINCR-KO cell lines. The editing of the start codon is highlighted.
B. Alignment of WT and pTINCR-KO nucleotide sequence. Hashtags mark differences between the two sequences.
C. Minimum Free Energy (MFE) structure of *TINCR* WT and *TINCR* KO, using RNAfold web server from ViennaRNA web.

1.3.3. Generation of an anti-pTINCR antibody

To further characterized pTINCR microprotein and to validate our pTINCR-KO models, we generated a custom rabbit polyclonal antibody. We subcontracted the generation of the antibody to the company Abyntek, which immunized the rabbits using a peptide located at the C-terminal part of pTINCR microprotein. Of note, this peptide shares an 85% identity (11 out of 13 residues) with mouse pTINCR, which makes it also suitable for its application in mouse samples. We were provided with a purified antibody, which we validated by Western blot and immunohistochemistry (**Fig. 24A and B**). Interestingly, these results also confirmed that pTINCR, as well as *TINCR* lncRNA, is upregulated upon calcium-induced differentiation.

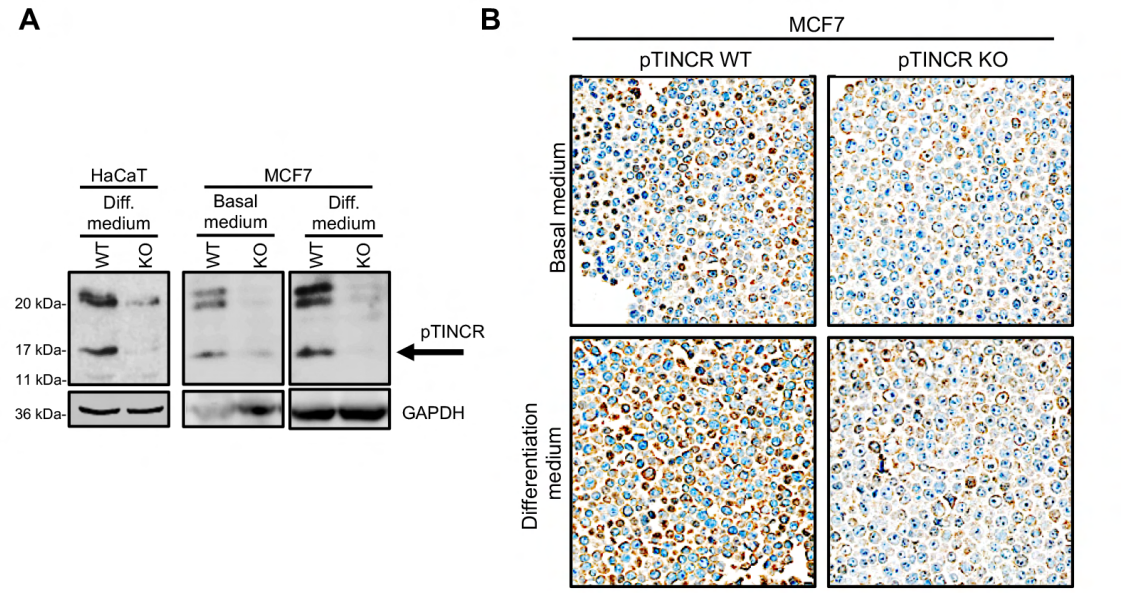


Figure 24. Validation of pTINCR-KO cell lines.
A. Western blot of endogenous pTINCR using a polyclonal custom-made antibody in WT and pTINCR-KO HaCaT and MCF7 cells.
B. Representative IHC staining of endogenous pTINCR using a polyclonal custom-made antibody in of WT and pTINCR-KO MCF7 cells, in basal or differentiation media.

Of note, apart from the band with the expected size (indicated by the arrow), pTINCR antibody also detected additional bands that disappear in pTINCR-KO cells. On one hand, we hypothesized that those bands may correspond to posttranslational modifications of pTINCR. On the other hand, we could not discard neither the possible detection of *TINCR*-204-encoded microprotein. The peptide used for the generation of the antibody is also present in this microprotein variant and, since both microproteins share the same start codon, our genetically engineered cells will be deficient for both. However, we have shown that *TINCR*-204 transcript is not upregulated upon calcium-induced differentiation (**Fig. 8A**). Therefore, we could expect the same for its derived microprotein. These questions will be further addressed in the discussion section.

2. Functional characterization of pTINCR microprotein

2.1. pTINCR subcellular localization and tissue expression pattern

Using our custom-made antibody, we assessed endogenous pTINCR subcellular localization in differentiated HaCaT and MCF7 cells by immunofluorescence experiments. Endogenous pTINCR microprotein was detected in these cell lines localizing mainly in the nucleus and at the cell-to-cell junctions, where it co-localizes with cortical F-actin (**Fig. 25**).

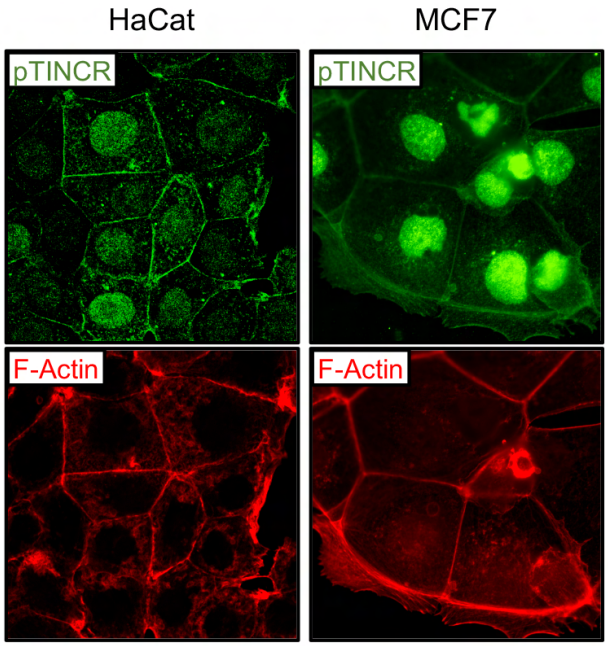


Figure 25. Endogenous pTINCR is detected in the nucleus and in the cell membrane.
Representative immunofluorescence images of endogenous pTINCR and F-actin in HaCaT and MCF7 cells using a pTINCR custom-made antibody (green) and Phalloidin-TRITC (red), respectively.

We further used pTINCR antibody to study pTINCR tissue distribution. We performed immunostaining of histological sections from both, mouse and human tissues. In agreement with *TINCR* lncRNA expression, we detected pTINCR microprotein in mouse and human epidermis (**Fig. 26A and D**). Moreover, we also detected pTINCR in other mouse stratified epithelia (tongue, palate, esophagus, bladder, cervix and mammary glands) (**Fig. 26B and D)**, as well as in other epithelial tissues (stomach, uterus, sebaceous and sweat glands and lung) (**Fig. 26B and C**).

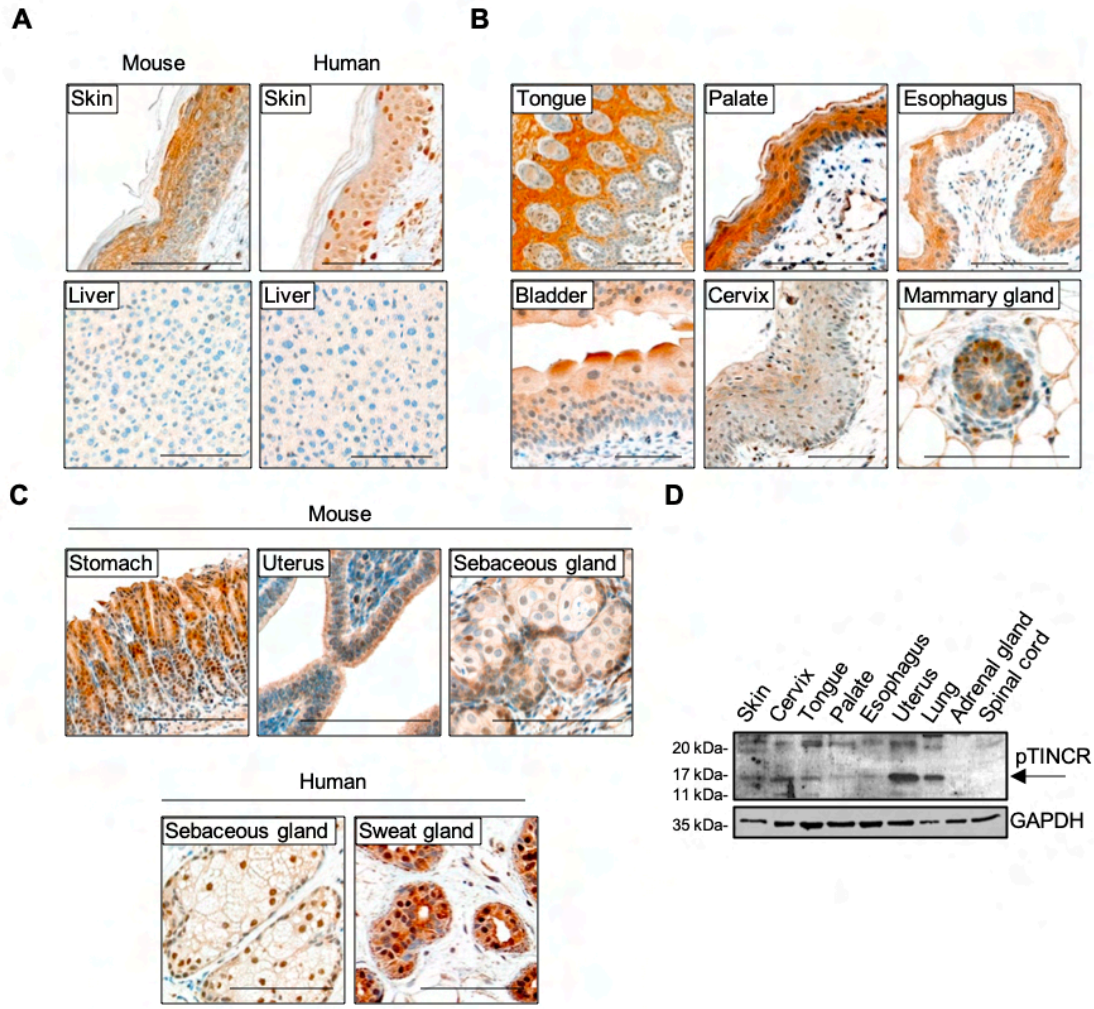


Figure 26. pTINCR is expressed in skin and other epithelia.
A-C. Representative immunostaining images of mouse and human pTINCR in skin and liver (**A**), a panel of mouse stratified epithelia (**B**) and a panel of mouse and human single epithelial (**C**) using a polyclonal custom-made antibody. Size bar is 100 μm.
D. Western blot analysis of endogenous pTINCR in a panel of mouse epithelial tissues using a polyclonal pTINCR antibody. Adrenal gland and spinal cord were added as negative controls.

2.2. pTINCR is upregulated upon cellular stress in a p53 dependent manner

Considering *TINCR* upregulation upon stress, we wanted to study whether the expression of pTINCR microprotein was also regulated by p53 activation. We treated different cells with doxorubicin to trigger DNA damage and p53 activation as explained above and analyzed pTINCR expression by immunofluorescence using our custom-made antibody. We observed that only cells with functional p53 upregulated pTINCR microprotein, whereas p53-KO cells or with mutant p53 did not (**Fig. 27**). Therefore, p53 activation leads to pTINCR upregulation at the mRNA and protein level.

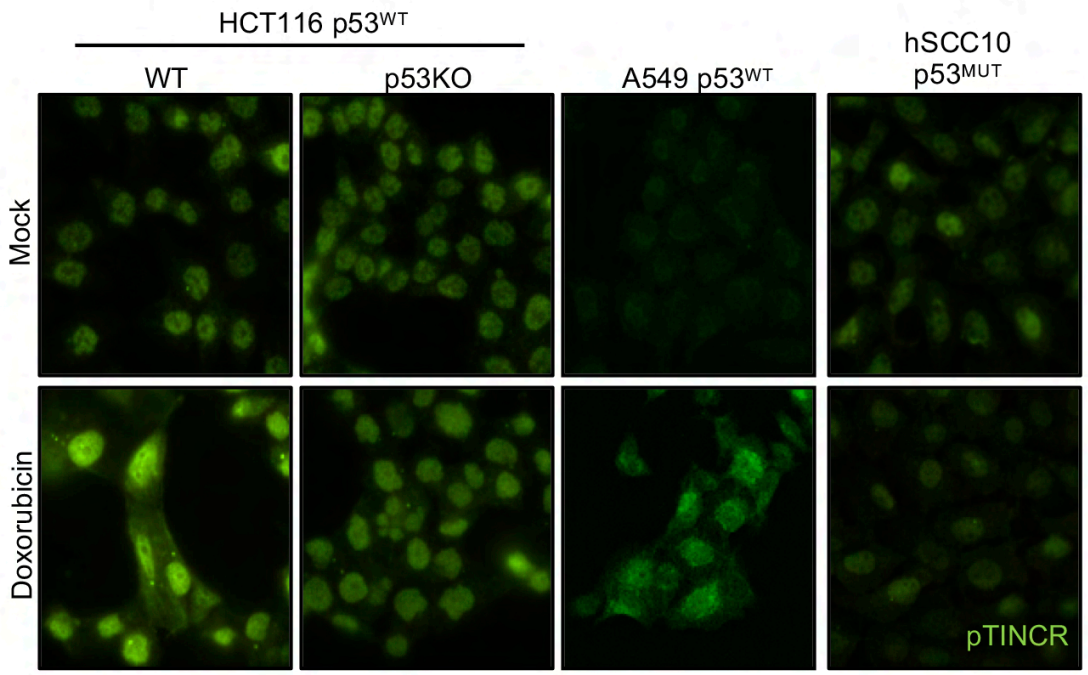


Figure 27. pTINCR is upregulated by p53 activation.
Endogenous pTINCR expression analyzed by immunofluorescence using a custom-made antibody after 24 hours of treatment with doxorubicin (1 μm), in the indicated cell lines.

Given these results, we considered of interest to test whether pTINCR overexpression was related with some of the most common outcomes of p53 activation, such as cell cycle arrest, apoptosis, senescence, and cell differentiation.

2.3. pTINCR overexpression reduces cell proliferation

As a first approach to study the functions of pTINCR microprotein, we studied the effect of pTINCR overexpression in cell proliferation. We transduced different cancer cell lines with constitutive (BxPC-3) or inducible (A549, hSCC10 and MCF7) HA-Tagged pTINCR vectors and measured cell proliferation *in vitro.* Given that doxycycline is known to affect cell proliferation and viability by itself, we performed a titration to determine the minimum doxycycline concentration needed to induce pTINCR without side-effects. Importantly, 0.5 μg/ml of doxycycline was enough to induce pTINCR overexpression (**Fig. 28A**) but did not significantly affect cell viability (**Fig. 28B**), as measured by Cell Titer Glo in cells treated or not with doxycycline for 4 days.

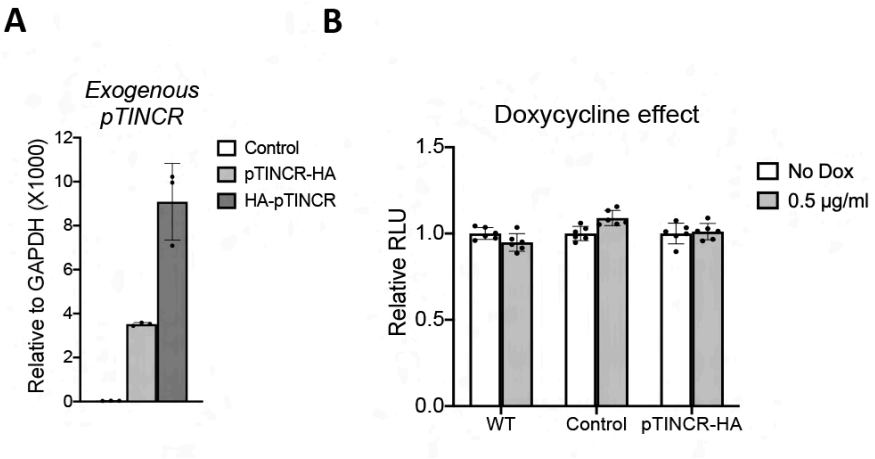


Figure 28. Low concentrations of doxycycline does not affect cell viability.
A. Exogenous pTINCR expression in cells transduced with C-terminal (pTINCR-HA) and N-terminal (HA-pTINCR) HA-tagged microprotein by RT-qPCR. mRNA expression is normalized to GAPDH in each case and relative to the control. Error bars represent the mean ± SD.
B. Cell viability measured by Cell Titer Glo assay in WT cells, cells transduced with empty vector (control) or cells transduced with pTINCR HA-Tagged vector (pTINCR-HA) treated or not with doxycycline at the indicated concentration. Relative Light Units (RLU) are relative to the untreated condition.

To characterize the effect of pTINCR in the cell cycle, we measured the expression level of genes related with cell cycle progression and performed a bromodeoxyuridine (BrdU) incorporation assay coupled with propidium iodide (PI) staining in the $p53^{WT}$ A549 cell line. For these experiments, we used cells transduced with inducible HA-Tagged pTINCR, treated or not (control) with 0.5 µg/ml of doxycycline. RT-qPCR results showed that pTINCR-overexpressing cells presented a moderate increase in the expression of cyclin-dependent kinase inhibitor p21 (**Fig. 29A**), a known target of p53. This protein functions as a regulator of cell cycle progression at G1 and S phase by inhibiting the activity of cyclin-CDK1, -CDK2, and -CDK4/6 complexes. Moreover, we also observed reduced levels of the proliferating cell nuclear antigen (PCNA) (**Fig. 29A**), a DNA polymerase accessory factor essential for DNA replication, as well as a downregulation of Cyclin B (necessary for the progression throughout M phase) and Cyclin E (necessary for the G1-S phase transition) (**Fig. 29A**). Accordingly, we observed that pTINCR-overexpressing cells progressed more slowly through the cell cycle, presenting a clear accumulation of cells in the G1 phase and a reduction in DNA-replicating cells in S phase (**Fig. 29B**).

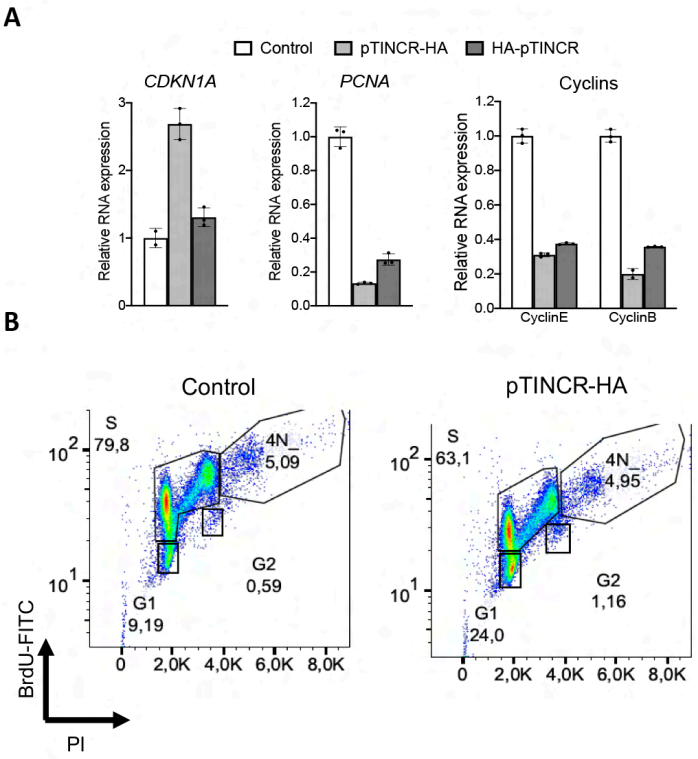


Figure 29. pTINCR overexpression downregulates cell cycle-related genes and arrests cells in G1 phase.
A. Expression of the indicated genes by RT-qPCR after 4 days of pTINCR induction with doxycycline. mRNA expression is normalized to GAPDH in each case and relative to the control. Error bars represent the mean ± SD. **B.** BrdU incorporation assay coupled to PI staining by Flow Cytometry in control or pTINCR-overexpressing cells after 24 hours of BrdU pulse.

Accordingly, our results showed that cells overexpressing pTINCR displayed a significant reduction in their proliferation rate *in vitro* (**Fig. 30**).

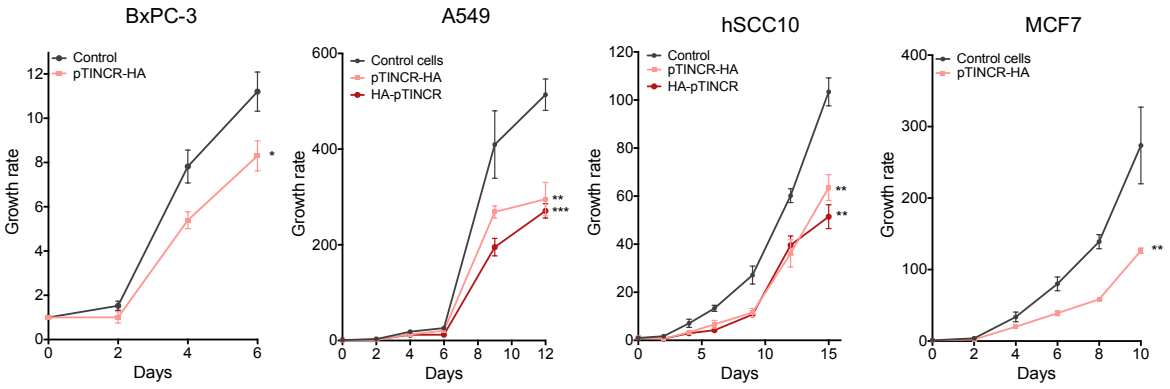


Figure 30. pTINCR overexpression reduces cell growth.
Growth curves of BxPC-3, A549, hSCC10 and MCF7 cells expressing pTINCR or not. Growth rate indicates cell number in millions at each corresponding time point relative to the starting number of cells.

2.4. pTINCR overexpression does not induce apoptosis or senescence

We transduced A549 cells with constitutive or inducible HA-Tagged pTINCR vectors and measured apoptosis and senescence induction. Cells constitutively expressing pTINCR for one week did not present any difference in the percentage of senescent cells, as measured by SA β-Gal staining (**Fig. 31A**). Then, we measured apoptosis by performing Annexin V assay in cells treated or not with 0.5 µg/ml of doxycycline to induce pTINCR overexpression. As positive control, we used Staurosporine to induce cell death (**Fig. 31B**). We observed that pTINCR did not induce apoptosis in A549 cells.

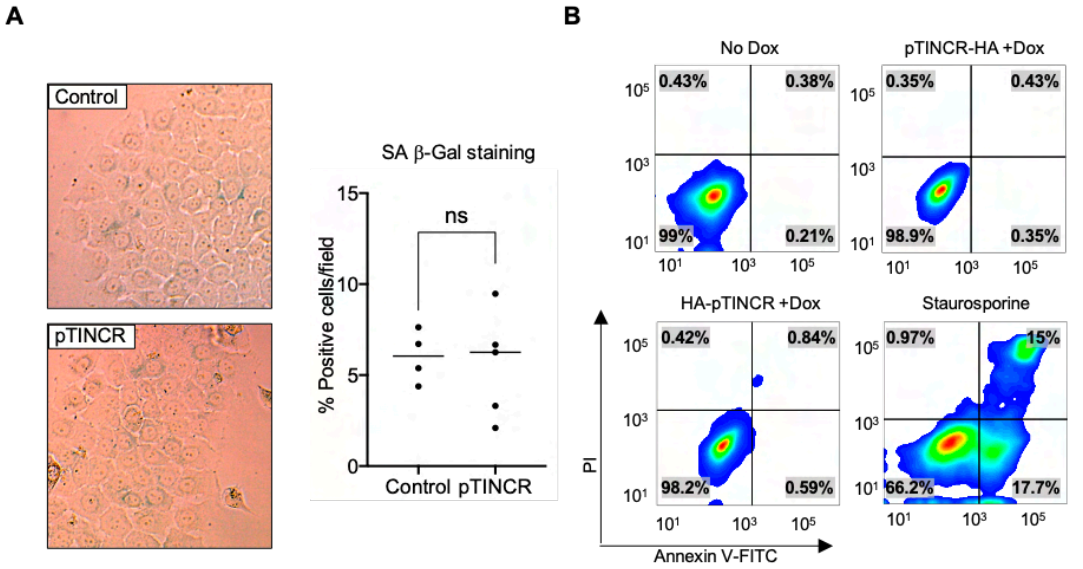


Figure 31. pTINCR overexpression does not trigger cellular senescence or apoptosis.
A. On the left, representative images of SA-β-Gal staining in BXPC-3 control or overexpressing pTINCR cells for one week. On the right, number of senescent cells represented as relative to the total number of cells. Each dot represents a different field used for SA-β-Gal quantification.
B. Representative flow cytometry plots using Annexin V-FITC/PI staining to measure apoptosis in A549 cells treated or not with doxycycline to induce pTINCR overexpression. Staurosporine condition was used as a positive control for apoptosis induction.

2.5. pTINCR overexpression promotes epithelial differentiation

The high expression of *TINCR* lncRNA and pTINCR microprotein in skin and other epithelial tissues prompted us to evaluate the role of pTINCR specifically in this context, in which p53 has been described to play a role during differentiation [305-308].

2.5.1. pTINCR promotes epithelial differentiation *in vitro*

First, we used HaCaT keratinocytes, a model that allows the recapitulation of typical events occurring during skin differentiation by modulating the Ca^{2+} concentration in the culture medium [309, 310]. We transduced our doxycycline-inducible constructs into HaCaT cells and induced pTINCR overexpression for 4 days in basal or high calcium

(differentiation) conditions (**Fig. 32A**). Then, we measured the effect of pTINCR on several differentiation traits. Our results showed that pTINCR overexpression during HaCaT differentiation led to an increased upregulation of genes related with epidermal differentiation compared to the differentiated control cells, indicating that pTINCR improves keratinocyte differentiation (**Fig. 32B, differentiated panels**). Remarkably, pTINCR overexpression induced an increase in the expression of some differentiation markers even in low Ca^{2+} conditions, suggesting that pTINCR is sufficient to trigger differentiation events (**Fig. 32B, basal panels**).

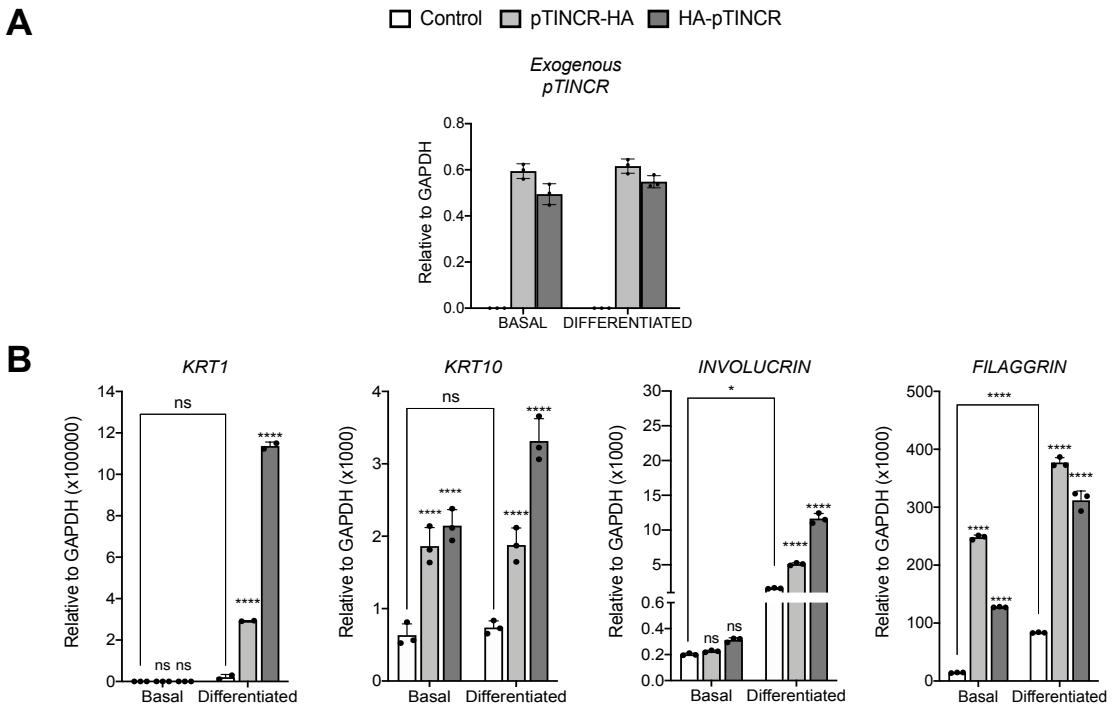


Figure 32. pTINCR overexpression induces the upregulation of epidermal differentiation markers.
A. Exogenous pTINCR expression in cells transduced with C-terminal (pTINCR-HA) and N-terminal (HA-pTINCR) HA-tagged microprotein measured by RT-qPCR in HaCaT cells cultured in basal or differentiation conditions. mRNA expression is normalized to GAPDH in each case. Error bars represent the mean ± SD.
B. Expression of the indicated differentiation markers upon the overexpression of C-terminal syORF (pTINCR-HA) and N-terminal (HA-pTINCR) HA-tagged microprotein in HaCaT cells measured by RT-qPCR. mRNA expression was measured 4 days after doxycycline induction in cells cultured in basal medium or differentiation medium. Error bars represent the mean ± SD. Statistical analysis was performed using a two-way ANOVA test with multiple comparison. We show statistical differences between pTINCR-overexpressing cells and control cells in each experimental condition (basal or differentiation medium) and between basal and differentiated control cells (shown in brackets). *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001, (n = 3 technical replicates).

Second, we measured the effect of pTINCR on actin dynamics and observed that pTINCR overexpression *per se* (without changing the Ca^{2+} concentration) switched actin cytoskeleton organization from a stress fiber disposition to a cortical pattern (**Fig. 33, middle panels**), resembling the keratinocyte differentiation process (**Fig. 33, differentiated panels**).

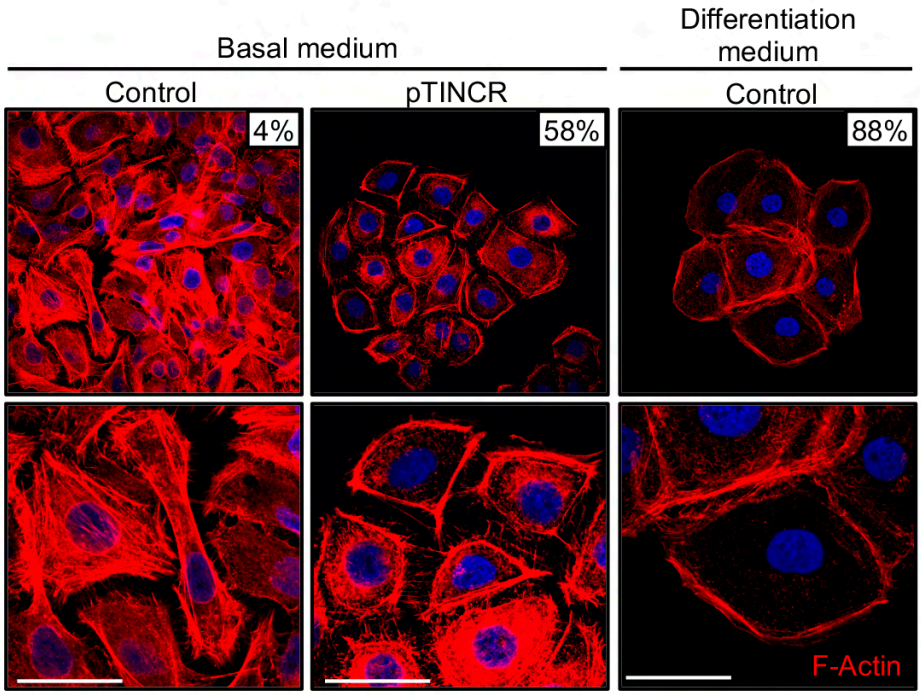


Figure 33. pTINCR triggers actin cytoskeleton remodeling towards a cortical pattern.
Immunostaining images showing actin organization using Phalloidin-TRITC 24 hours after doxycycline induction of syORF pTINCR-HA in HaCaT cells. Control cells cultured in differentiation medium for 24 hours are shown as positive control for the differentiation process. Size bar is 100 µm. The percentage of cells showing a cortical actin disposition is indicated in each image.

Third, pTINCR overexpression also induced the establishment of adherens and tight junctions similarly to high calcium conditions, as seen by the increase in E-cadherin and β-catenin (markers of adherens junctions) (**Fig. 34A and B**) and ZO-1 (marker of tight junctions) (**Fig. 34C**) at the cell membrane. Altogether, our results show that pTINCR promotes keratinocytes differentiation *in vitro*.

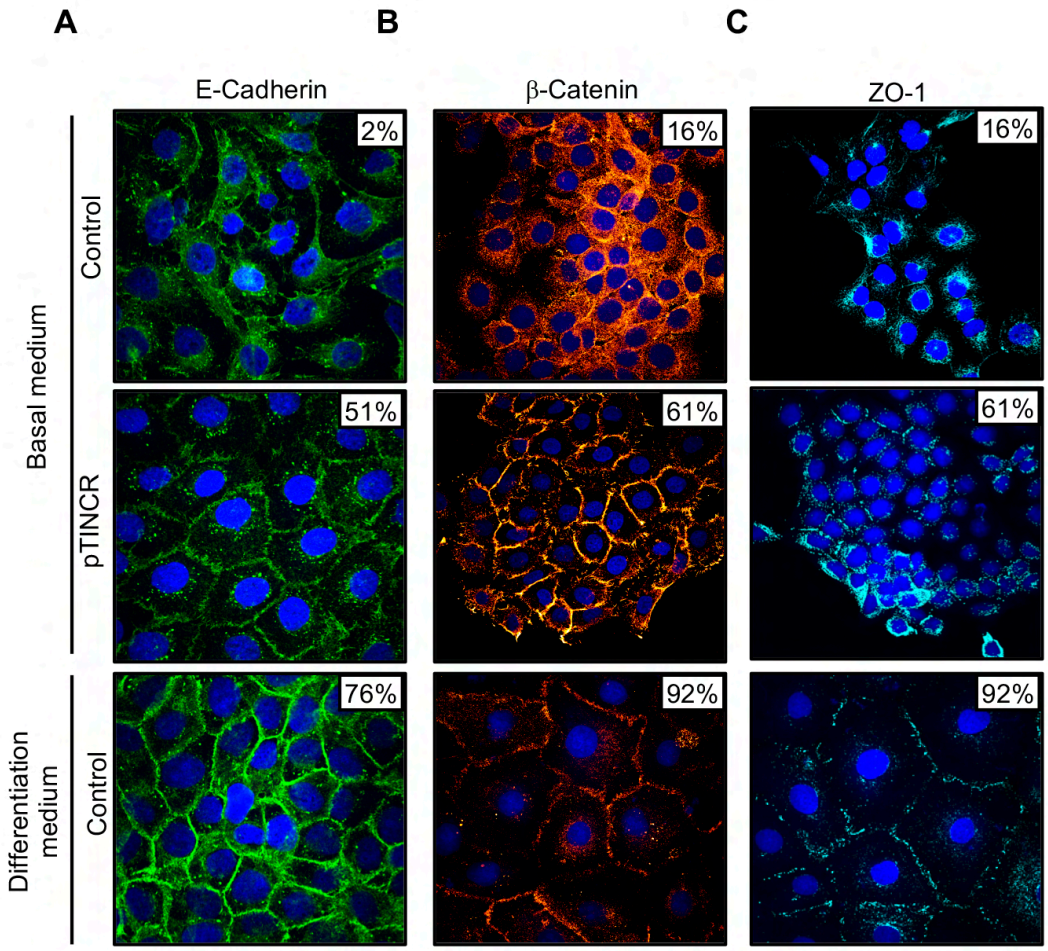


Figure 34. pTINCR enhances cell-to-cell contacts.
A-C. Immunostainings showing E-cadherin (**A**), β-catenin (**B**) and ZO-1 (**C**) after 24 hours (E-cadherin) or 4 days (β-catenin and ZO-1) of doxycycline induction of C-terminal syORF (pTINCR-HA) in HaCaT cells. Control cells cultured in differentiation medium for 24 hours are shown as positive control for the differentiation process. The percentage of cells showing a membranous staining is indicated in each image.

Then, we also wanted to address the function of pTINCR in other epithelial contexts: a patient-derived cSCC cell line (hSCC10 [311, 312]), the A549 lung adenocarcinoma cell line, and the MCF7 luminal breast cancer cell line. We induced pTINCR overexpression in these models for several days and measured different traits of epithelial differentiation.

In hSCC10 and A549 cells, pTINCR overexpression was associated with an upregulation of several genes related to epithelial differentiation and a downregulation of basal epithelial markers (**Fig. 35A and B**). In agreement, pTINCR-overexpressing MCF7 cells upregulated luminal markers and downregulated basal cell markers (**Fig. 35C**).

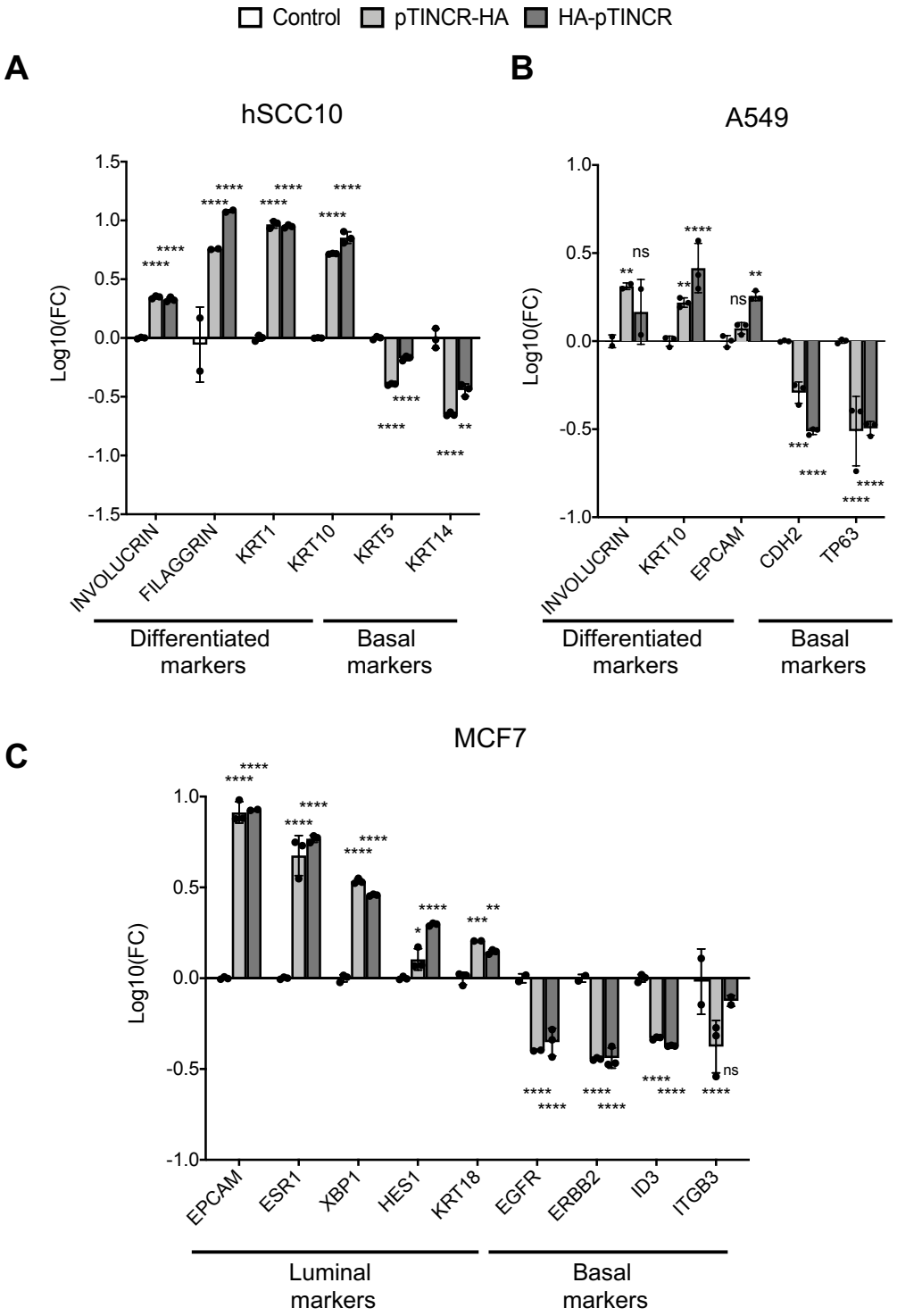


Figure 35. pTINCR overexpression induces the upregulation of epithelial differentiation markers and the downregulation of basal markers.
A-C. Expression of the indicated differentiation and basal markers upon the overexpression of HA-tagged pTINCR in hSCC10 (**A**) and A549 (**B**) or syORF (pTINCR-HA) in MCF7 cells (**C**) measured by RT-qPCR. mRNA expression was measured at 21 days (A549 and hSCC10) or 4 days (MCF7) after doxycycline induction. mRNA expression is normalized to GAPDH and relative to the control in each case. Error bars represent the mean ± SD. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001, using two-way ANOVA test (n = 3 technical replicates).

As previously observed in HaCaT cells, pTINCR expression triggered a cortical actin cytoskeleton disposition (**Fig. 36A-C, first columns**) and reinforced cell-to-cell contacts in these models, as seen by an increased accumulation of β-catenin (**Fig. 36A and B, second columns**), ZO-1 (**Fig. 36A and B, third columns**) and E-cadherin at the cell membrane (**Fig. 36C, second column**).

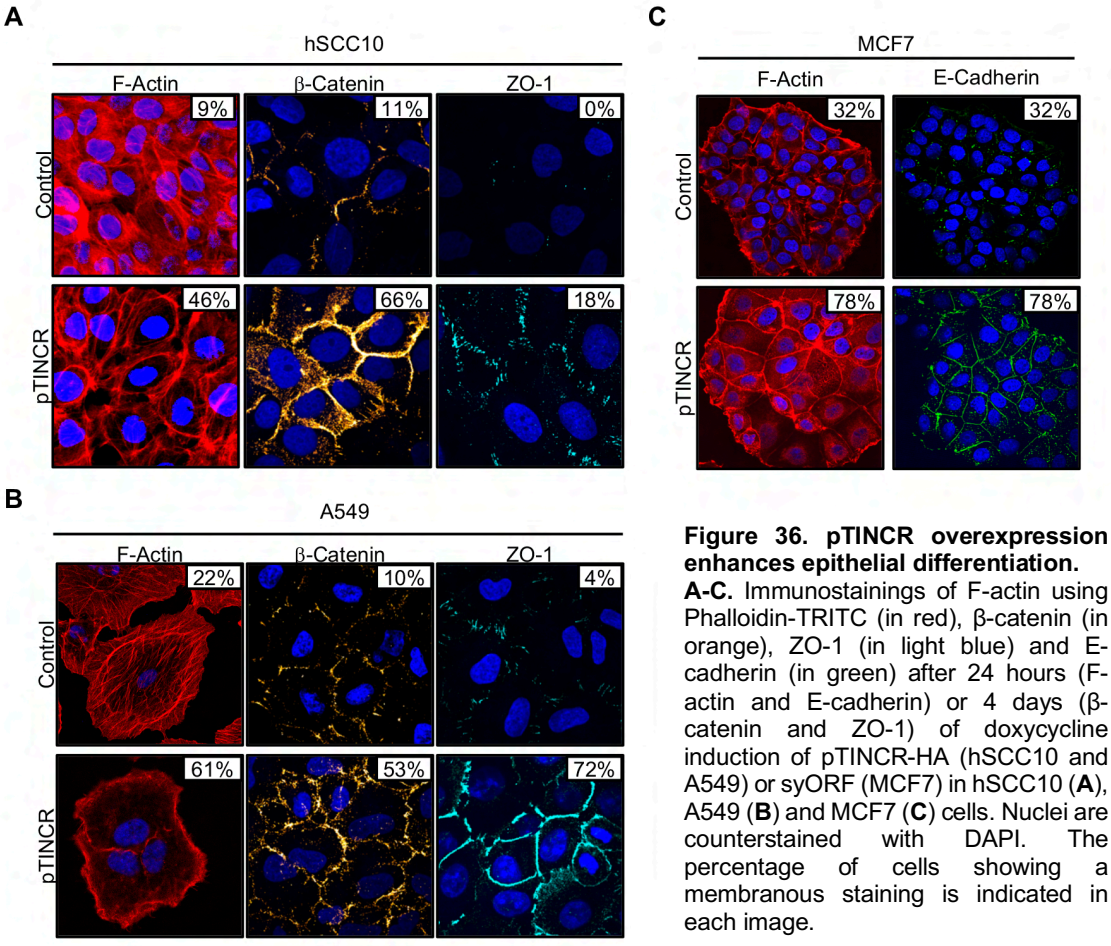


Figure 36. pTINCR overexpression enhances epithelial differentiation.
A-C. Immunostainings of F-actin using Phalloidin-TRITC (in red), β-catenin (in orange), ZO-1 (in light blue) and E-cadherin (in green) after 24 hours (F-actin and E-cadherin) or 4 days (β-catenin and ZO-1) of doxycycline induction of pTINCR-HA (hSCC10 and A549) or syORF (MCF7) in hSCC10 (**A**), A549 (**B**) and MCF7 (**C**) cells. Nuclei are counterstained with DAPI. The percentage of cells showing a membranous staining is indicated in each image.

2.5.2. pTINCR favors epidermal differentiation *in vivo*

To further explore the role of pTINCR as a driver of epithelial differentiation, we performed an *in vivo* proof-of-concept experiment using teratoma formation assay, which allows to study pluripotent stem cell differentiation in an unbiased manner.

First, we transduced mouse Embryonic Stem Cells (mESCs) with constitutive HA-tagged pTINCR and confirmed pTINCR overexpression (**Fig. 37A**). Interestingly, after one week of cell culture pTINCR-overexpressing mESCs colonies presented a flatter morphology with less defined edges (**Fig. 37B**), which are typical characteristics of mESCs differentiation. Indeed, we observed a downregulation of the classical pluripotency markers SOX2, NANOG and OCT4 by RT-qPCR (**Fig. 37C**).

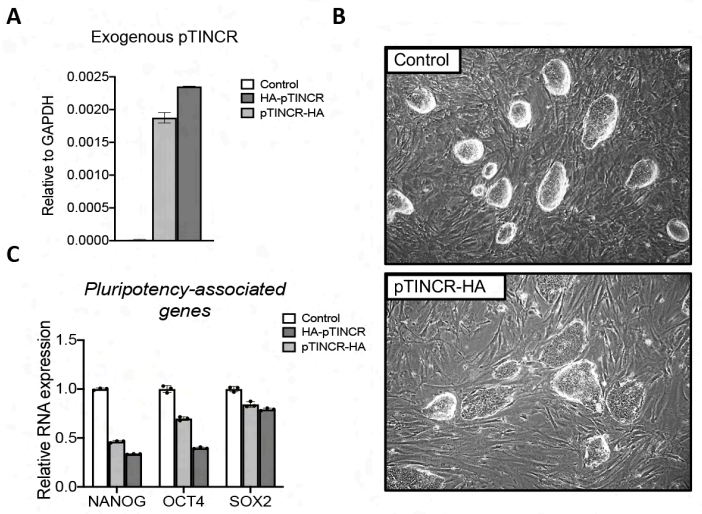


Figure 37. mESCs overexpressing pTINCR present decreased pluripotency.
A. Exogenous pTINCR expression in mESCs transduced with C-terminal (pTINCR-HA) and N-terminal (HA-pTINCR) HA-tagged microprotein by RT-qPCR. mRNA expression is normalized to GAPDH in each case. Error bars represent the mean ± SD.
B. Representative phase contrast images of Control and pTINCR-HA mESCs.
C. Expression of the indicated pluripotency markers upon in Cterminal (pTINCR-HA) and N-terminal (HA-pTINCR) HA-tagged microprotein overexpressing mESCs measured by RT-qPCR. mRNA expression was measured after one week of pTINCR overexpression. Error bars represent the mean ± SD.

Second, we inoculated the mESCs modified to express pTINCR or not in the flanks of immunodeficient mice and monitored teratoma growth. We observed that pTINCR expression led to reduced teratoma growth, which is consistent with more differentiated teratomas (**Fig. 38A**). Once the control group reached the ethical human endpoint, we sacrificed the mice and collected teratomas for histopathological examination. Results revealed a significant increase in skin areas within pTINCR-overexpressing teratomas, measured by the increase in keratin deposition (**Fig. 38B and C**).

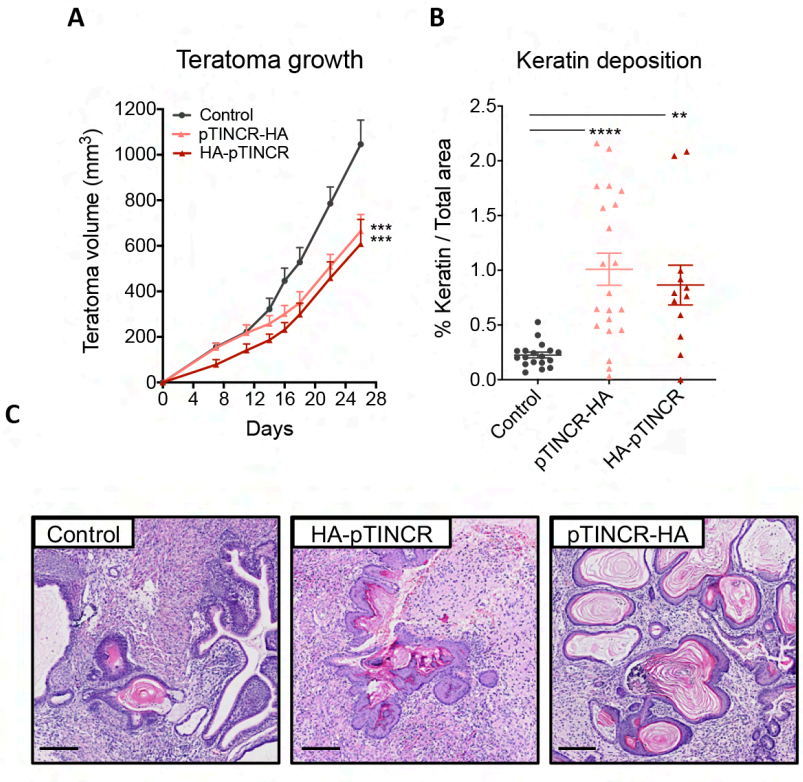


Figure 38. pTINCR overexpression favors epidermal formation *in vivo*.
A. Effect of pTINCR on teratoma growth. Teratomas were generated by the subcutaneous injection of pTINCR-expressing mESCs or control mESCs in both flanks of NMRI nude mice. Teratoma growth was monitored by measuring the volume at the indicated time points. Error bars represent the mean ± SEM. N= 18/20 teratomas in control group; N= 21/22 teratomas in pTINCR-HA group; N= 12/22 teratomas in HA-pTINCR group. ***p < 0.001, using one-way ANOVA.
B. Quantification of keratin deposition in differentiated teratomas using QuPath. Dots represent the percentage of area with keratin deposition relative to the total teratoma area. Error bars represent the mean ± SEM. N= 18/20 teratomas in control group; N= 21/22 teratomas in pTINCR-HA group; N= 12/22 teratomas in HA-pTINCR group.**p < 0.01, ****p < 0.0001, using non-parametric Kruskal-Wallis test.
C. Representative images of H&E stainings of teratoma areas with skin development, as seen by the formation of keratin pearls. Size bar is 200 µm.

Further immunostaining experiments in control teratomas (with no pTINCR overexpression) showed endogenous pTINCR expression in skin (**Fig. 39A**). Moreover, some teratomas presented transitions from single to stratified epithelia and, remarkably, pTINCR expression notably increased in the stratified areas compared to the non-stratified ones (**Fig. 39A and B**), further suggesting a possible role of pTINCR in the differentiation of stratified epithelia.

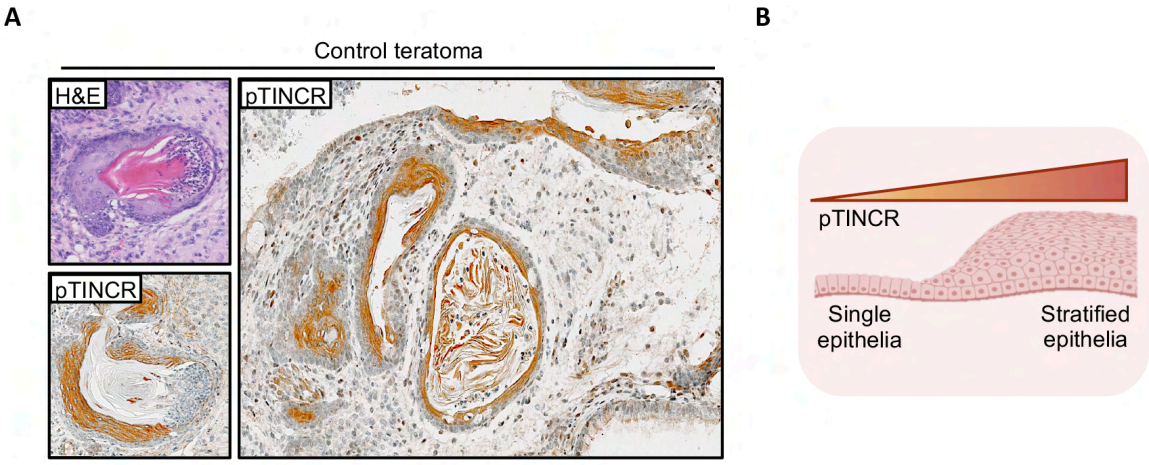


Fig. 39. Endogenous pTINCR is detected in keratin positive regions and increased in stratified epithelial areas compared to the non-stratified ones.
A. Representative images of H&E and IHC stainings of control teratomas (generated with WT mESCs) using a custom-made antibody against endogenous pTINCR. Pictures show areas of single and stratified epithelia with keratin deposition.
B. Diagram illustrating the upregulation of endogenous pTINCR in the transition from single to stratified epithelia. Image was created with BioRender.com.

2.5.3. pTINCR is required for *in vitro* epithelial differentiation

To investigate the impact of pTINCR deficiency on cellular differentiation, we subjected our HaCaT and MCF7 pTINCR-deficient cells to high calcium conditions and studied differentiation markers. Importantly, we observed that *TINCR* lncRNA was upregulated upon calcium-induced differentiation to the same extent in WT and pTINCR-KO cells (**Fig. 40A**), indicating that the regulation of the lncRNA is not affected in pTINCR-KO

cells. However, pTINCR-deficient cells failed to upregulate differentiation markers to the same extent as WT cells (**Fig. 40B and C**).

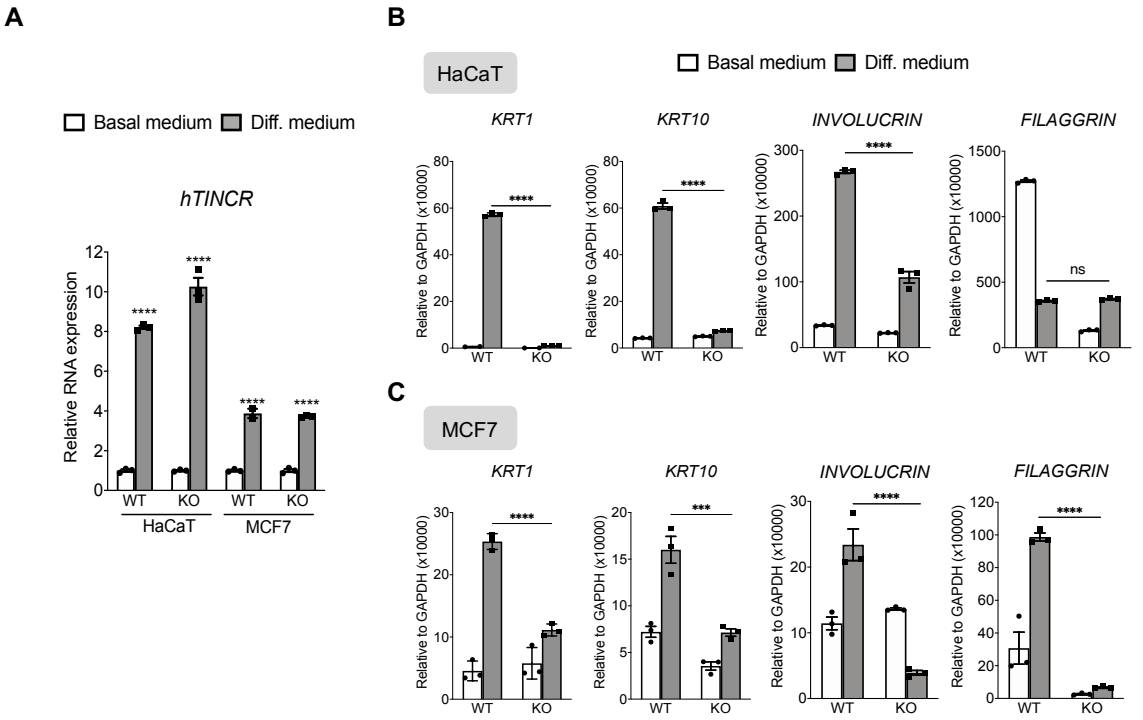


Figure 40. pTINCR-KO cells do not upregulate differentiation markers upon calcium-induced differentiation.
A. *TINCR* lncRNA expression in WT and pTINCR-KO HaCaT and MCF7 cells, under basal and differentiation culture conditions measured by RT-qPCR. mRNA expression is normalized to GAPDH expression and relative to the basal condition. Error bars represent the mean ± SD. ****p < 0.0001, using two-way ANOVA test (n = 3 technical replicates).
B-C. Expression of the indicated differentiation markers in WT and pTINCR-KO HaCaT (**B**) and MCF7 (**C**) cells under basal and differentiation culture conditions measured by RT-qPCR. mRNA expression is normalized to GAPDH expression. Error bars represent the mean ± SD. ***p < 0.001, ****p < 0.0001, n.s: not significant using two-way ANOVA test (n = 3 technical replicates).

In addition, pTINCR-KO cells did not acquire epithelial morphology (**Fig. 41A and C**) nor remodeled their actin cytoskeleton towards a cortical disposition (**Fig. 41B and D**) upon high calcium exposure, demonstrating that pTINCR is required for *in vitro* epithelial differentiation.

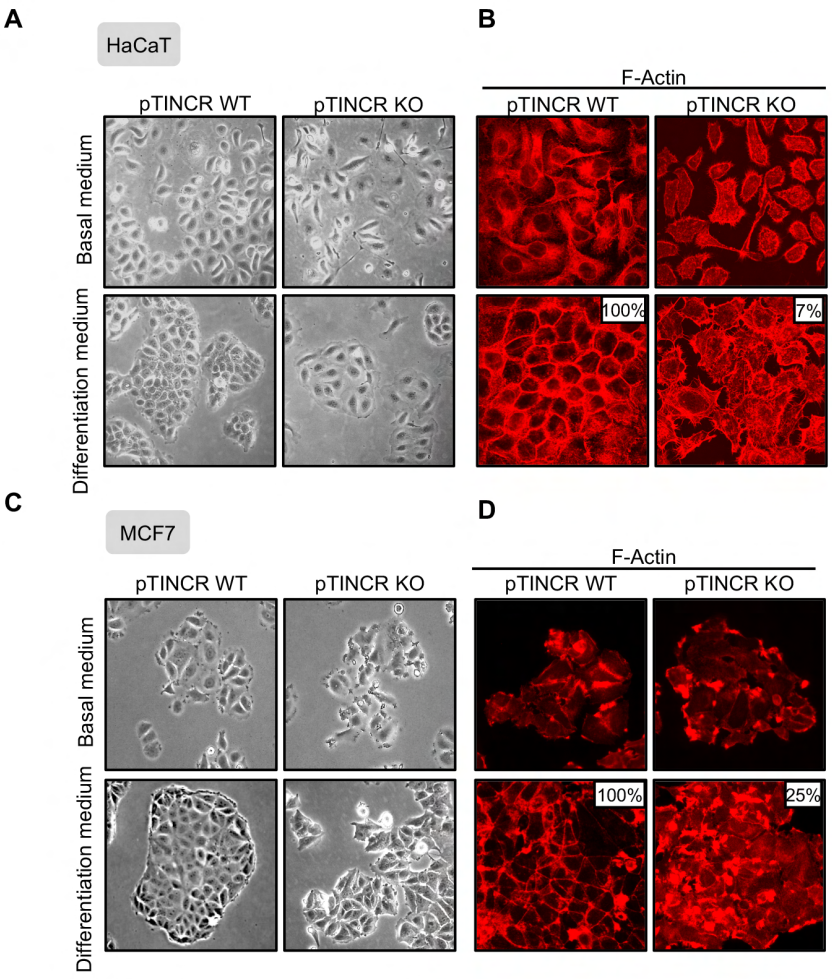


Figure 41. pTINCR-KO cells fail to acquire epithelial morphology and remodel actin cytoskeleton.
A and C. Representative phase contrast images of WT and pTINCR-KO HaCaT (**A**) and MCF7 (**C**) cells cultured in basal and differentiation conditions.
B and D. Representative immunofluorescence images of F-actin in WT and pTINCR-KO HaCaTs (**B**) and MCF7 (**D**) cells cultured in basal and differentiation conditions using Phalloidin-TRITC. Total percentage of cells displaying cortical actin is indicated in each image.

2.6. pTINCR triggers an epithelial differentiation transcriptional program

To further strength the pro-differentiation function of pTINCR, we analyzed the transcriptional profile induced by pTINCR overexpression. We performed an extensive RNA-seq analysis in hSCC10 cell line upon the overexpression of pTINCR to assess both, short- (6, 12 and 24 hours) and long-term transcriptomic changes (4, 7, 14 and 21 days), the latter probably reflecting changes in cell identity. First, we studied transcriptional dynamics driven by pTINCR overexpression using impulseDE algorithm, a framework for longitudinal sequencing experiments that reveals differential gene expression associated with time [277]. We detected 6 different groups of genes that clustered together based on their dynamic expression (**Fig. 42A**) and performed gene ontology enrichment analysis on each cluster (**Fig. 42B**). Interestingly, two of them (cluster 1 and cluster 2) were dynamically opposite clusters, both enriched in gene ontology terms related to cytoskeleton. Another cluster with significant relevance in the analysis was cluster 3, enriched in gene ontology terms related with mitochondria and metabolism.

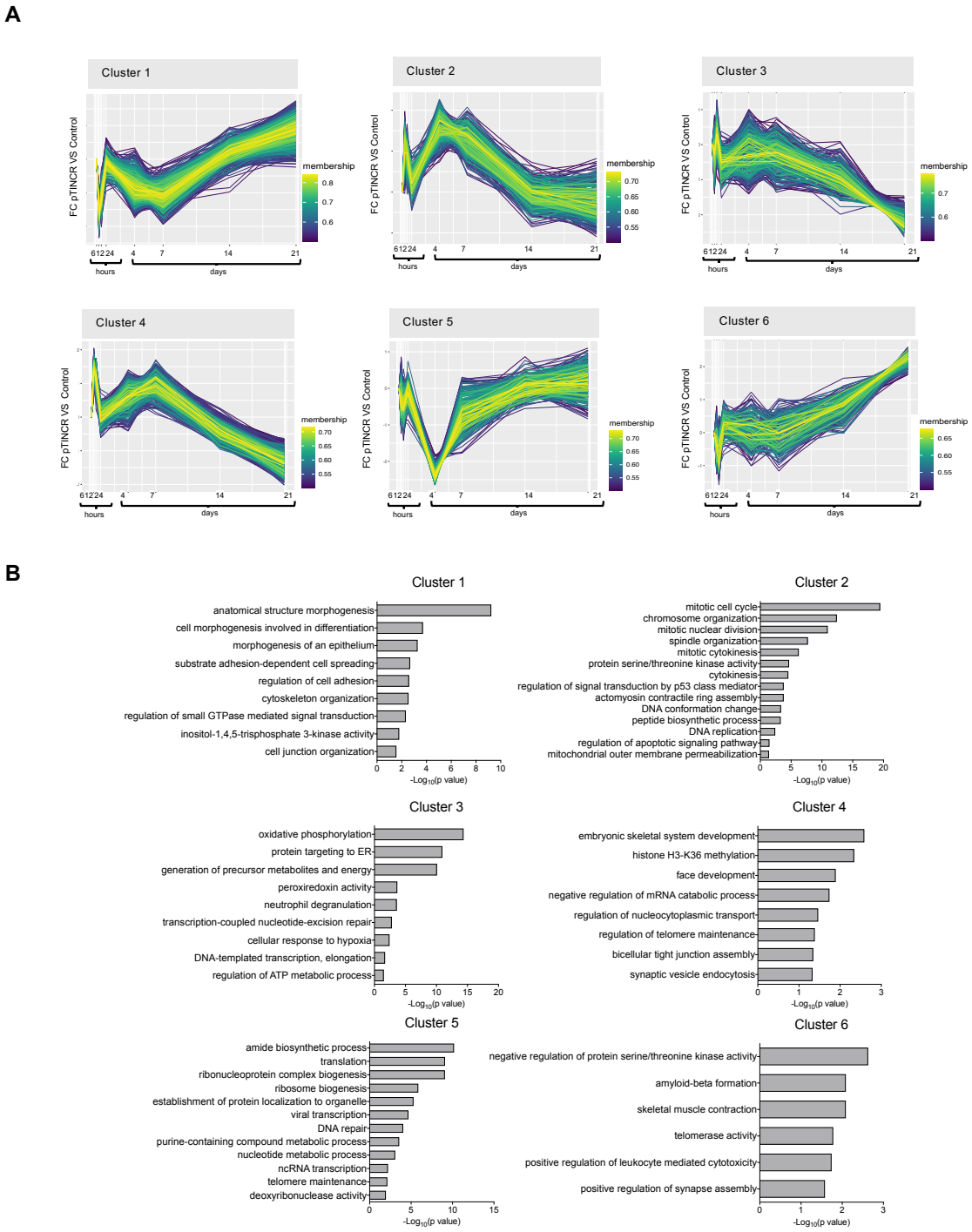


Figure 42. pTINCR-induced transcriptional program analyzed by RNA-seq.
RNA-seq analysis was performed in hSCC10 cell line upon the overexpression of syORF pTINCR to assess both early (6, 12 and 24 hours) and late transcriptomic changes (4, 7, 14 and 21 days).
A. Transcriptional dynamics were analyzed using impulseDE, and genes were classified in clusters according to their similar expression dynamics.
B. GO terms enrichment analysis using ClueGO software indicating the functional term enrichments of theidentified genes in each cluster. Plots show significant GO term enrichment (pV ≤ 0.05, Bonferroni correction).

Given the phenotype observed in actin remodeling and cell cycle arrest, we were particularly interested in cluster 1 and 2. We used the ClueGo tool to functionally organize gene ontology networks within each cluster (**Fig. 43**) and observed that cluster 1 is enriched in genes connected with cell morphogenesis, cell-to-cell junctions, and adhesion; while cluster 2 is formed by cytoskeleton organization-related genes associated with cell division and chromosome segregation.

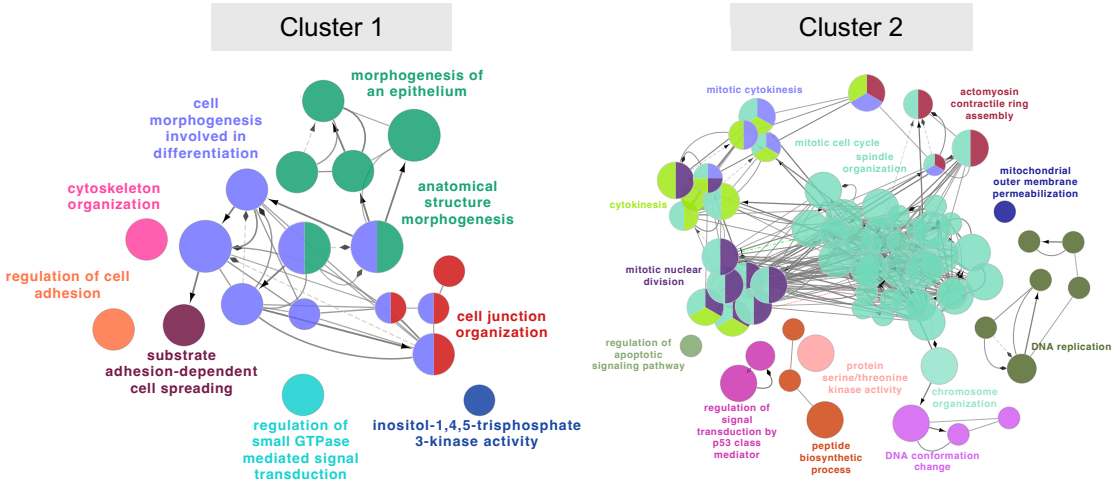


Figure 43. pTINCR triggers transcriptional changes associated with actin cytoskeleton, epithelial morphogenesis and cell cycle.
GO enrichment analysis of genes from cluster 1 and 2 using ClueGO 2.5.1. Each node represents a different GO term and is proportional to the Bonferroni-corrected p-value (p-V ≤ 0.05 were considered significant). An edge between two nodes indicates shared common genes between the two GO terms. Significant GO terms are grouped based on their similarity and labeled in the same color. Most significant GO terms in each group are labeled.

Next, we ran DESeq2 analysis [276] in samples collected at 0 hours and at 21 days and looked for genes whose expression changes differently in pTINCR-overexpressing cells compared to the control. A rotation-based functional enrichment analysis showed that gene-sets positively correlated with pTINCR expression were associated with actin organization and cell polarity, such as “actin filament organization” or “establishment or maintenance of cell polarity” (**Fig. 44A**). In addition, other epithelia-specific gene sets (“Epithelial cell differentiation, “Apical surface”, “Adherens junctions” or “Tight junctions”) were also enriched after 21 days of pTINCR overexpression (**Fig. 44A**). On the other hand, genes sets that negatively correlates with pTINCR expression were enriched in pathways associated with cell cycle, cell metabolism (“Oxidative phosphorylation”, “Cellular fatty acid metabolism”) and protein processing (“Protein targeting to the membrane”) among others (**Fig. 44B**). Interestingly, we found that pTINCR-downregulated genes were also enriched in the signature of “Myc targets”, an oncogenic pathway closely connected with undifferentiated grade and bad prognosis in cSCC [291, 313, 314].

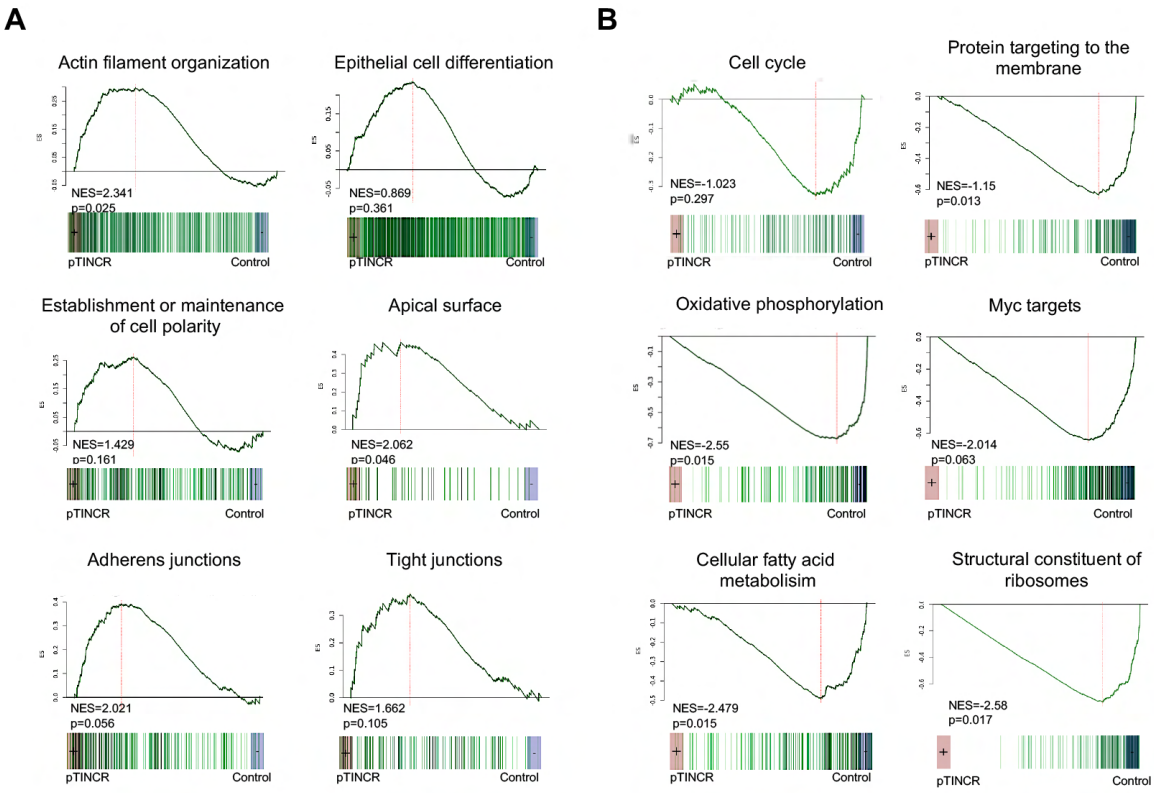


Figure 44. pTINCR-induced transcriptome in long-term overexpressing cells.
A. Positive correlation of Gene Set Enrichment Analysis of pTINCR-induced genes versus the indicated gene signatures.
B. Negative correlation of Gene Set Enrichment Analysis of pTINCR-induced genes versus the indicated gene signatures.

3. Molecular characterization of pTINCR microprotein

3.1. pTINCR is a new ubiquitin-like protein that interacts with SUMO and modulates SUMOylation

3.1.1. pTINCR structure resembles ubiquitin protein

To understand the molecular mechanisms behind the function of our microprotein, we did structural modeling using I-TASSER [315], which revealed that pTINCR is predicted to be a ubiquitin-like protein (**Fig. 45A**). Ubiquitin-like proteins (UBLs) are a family of proteins with structural similarity to ubiquitin, despite presenting low similarity in their amino acid sequences [226]. UBLs are classified into two types (see introduction section) and given the absence of a di-Gly motif, pTINCR arised as a new Type II UBL (**Fig. 45B**). Moreover, subsequent analysis of the pTINCR amino acid sequence using GPS-SUMO software [290, 316] highlighted two well-conserved and overlapping SUMO-interacting Motifs (SIM) in the C-terminal part of the microprotein (**Fig. 45B**).

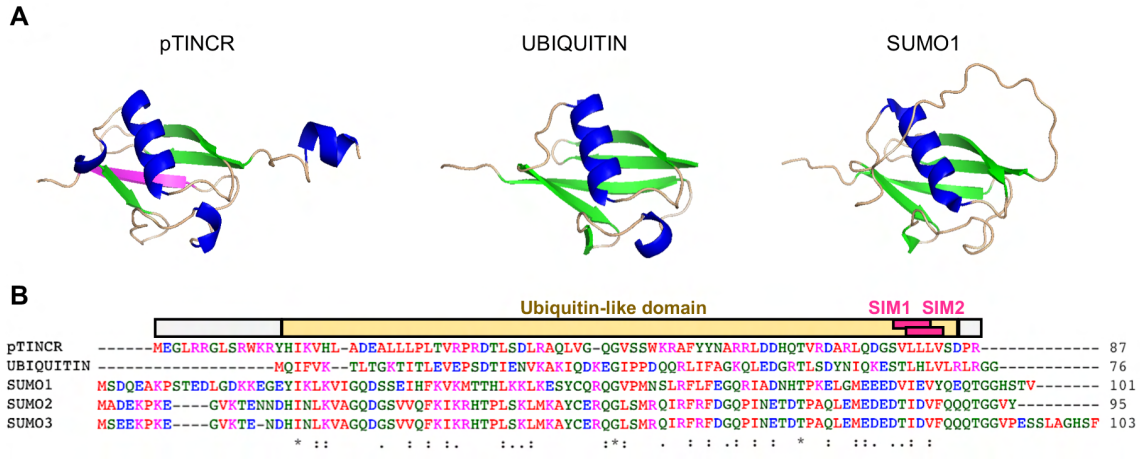


Figure 45. pTINCR is a novel type II UBL protein.
A. Image depicting the predicted pTINCR ubiquitin-like structure using I-TASSER software, and Ubiquitin and SUMO structures using PyMOL. Color code indicates the following features: dark blue, α-helix; green, β–sheet; purple, SUMO-interacting motives (SIMs) (predicted by GPS-SUMO 2.0).
B. Upper image, scheme of the predicted pTINCR domains and its amino acid sequence. Lower image, sequence alignment of human pTINCR with Ubiquitin and other Ubiquitin-like proteins by multiple sequence alignment performed by Clustal Omega. The colors of the amino acid residues indicate their properties (pink, positive charge; blue, negative charge; red, hydrophobic; green, hydrophilic).

3.1.2. pTINCR binds to SUMO in a non-covalent manner through its SIM domains

SIM domains are well-known hydrophobic motifs that allow proteins that harbor them to interact non-covalently with SUMO proteins. To test the possible interaction of pTINCR with SUMO, we generated a SIM mutant version of pTINCR-HA (pTINCR-SIMmut) in which most of the amino acids constituting the SIM domains were replaced by alanines (VLLLV > AAAAV). Then, we performed GST-pull down assays of SUMO1 and SUMO2 in cells overexpressing pTINCR-HA or pTINCR-SIMmut. Results confirmed that pTINCR binds to SUMO1 and SUMO2/3 in a non-covalent manner and that this interaction is lost in the pTINCR-SIMmut (**Fig. 46A**). Interestingly, SIM-mutant pTINCR resulted in a less stable protein, as shown by cycloheximide treatment experiments, pointing to an important role of SUMO in stabilizing pTINCR (**Fig. 46B**).

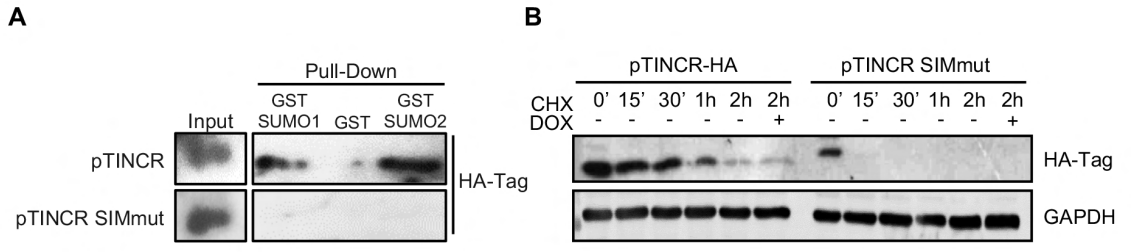


Figure 46. pTINCR binds non-covalently to SUMO proteins and this interaction is crucial for its stability.
A. pTINCR and pTINCR-SIMmut were *in vitro*-translated using radiolabeled 35^S-methionine followed by GST pull-down with GST, GST-SUMO1 or GST-SUMO2. pTINCR-SUMO interaction was visualized by SDS-PAGE.
B. Analysis of pTINCR-HA and pTINCR-SIMmut stability in U2OS cell line. Cells were treated with CHX for the indicated time points. An additional timepoint including simultaneous CHX treatment and doxycycline induction was included as an internal control.

3.1.3. Protein SUMOylation is modulated by pTINCR

To understand the meaning of pTINCR-SUMO interaction we analyzed global protein SUMOylation in different cell lines by Western blot. We observed that pTINCR overexpression modifies SUMO1 and SUMO2/3 conjugation patterns in a cell-dependent manner, increasing the intensity of some SUMOylated bands and decreasing the intensity of others, an effect that seemed to be partially reverted in pTINCR-SIMmut cells (**Fig. 47A and B**). Interestingly, levels of SUMOylated proteins are reduced in pTINCR-deficient cells (**Fig. 47A, MCF7 panel**).

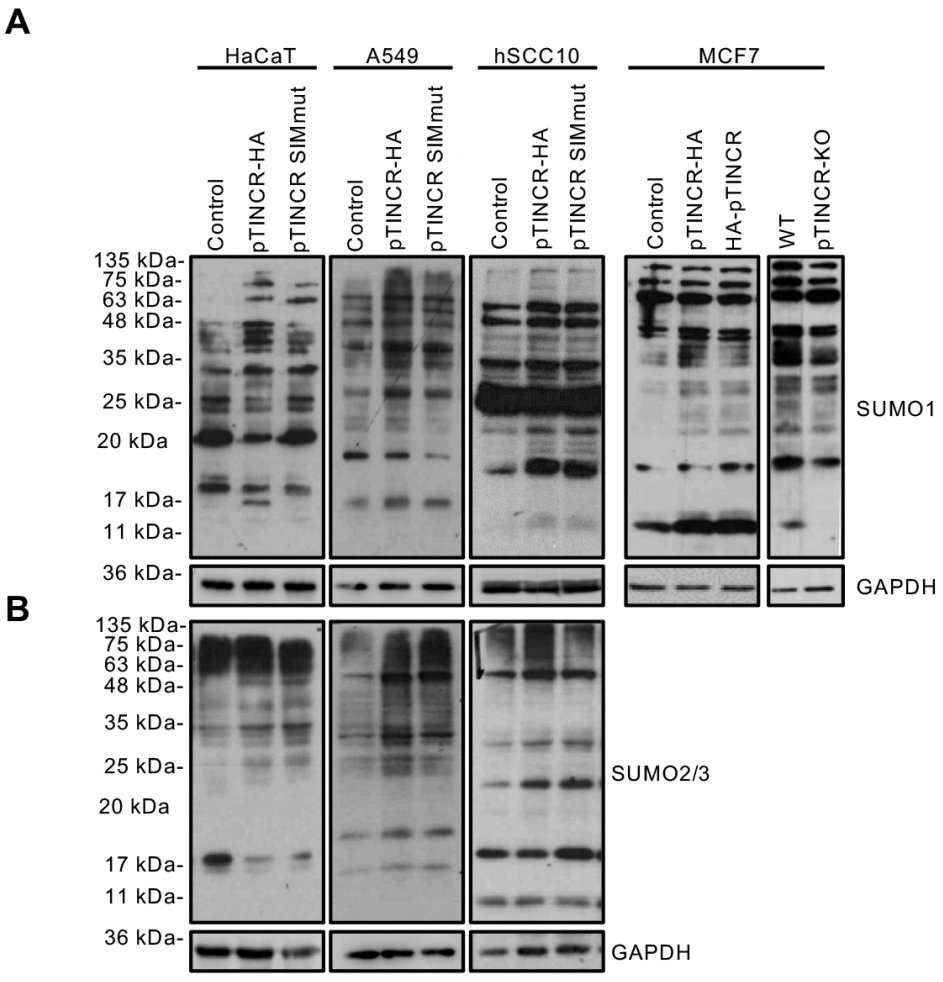


Figure 47. pTINCR modulates cell SUMOylation.
A and B. Effect of pTINCR overexpression on total SUMO1 (**A**) and SUMO2/3 (**B**) levels by Western blot in the indicated cell lines. In HaCaT, we used the syORF construct for pTINCR overexpression.

A bigger question for us was to determine whether pTINCR modulation of SUMOylation was a global or a protein-specific effect. By performing immunoprecipitation assays of SUMO and SUMOylated proteins (using HIS-Tagged SUMO), we measured the SUMOylation levels of two well-known SUMO targets, p53 and B23. We did not observe changes on the SUMOylation of these proteins after pTINCR overexpression (**Fig. 48**), which could suggest that the effect of pTINCR on SUMOylation is protein-specific.

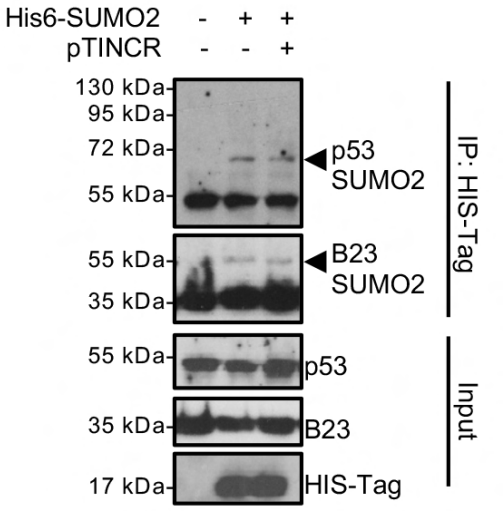


Figure 48. pTINCR overexpression does not affect the SUMOylation levels of some common SUMO targets.
Representative SUMOylation assay of endogenous p53 and B23 proteins in U2OS cells transfected with His6-SUMO2 and pTINCR-HA. 36h after transfection, anti-HIS pull-down products were analyzed by Western blot with anti-HIS, anti-HA, anti-p53 and anti-B23 antibodies.

3.2. Study of pTINCR interactome

The fact that pTINCR binds to SUMO implies that it potentially interacts with many other proteins. Therefore, we wanted to uncover its interactome. We overexpressed pTINCR in A549 cells and performed pTINCR-HA immunoprecipitation followed by mass spectrometry. We analyzed the results obtained following different criteria of quality, probability and significance. First analysis (**Fig. 49A**) ranked the candidates according to their abundance in pTINCR condition compared to the control (fold change (FC)) and the confidence score obtained from the spectra, which increases with the number of peptides identified for each hit. Following this approach, we obtained 35 candidates with a FC > 2 and a confidence score > 25 (detection reliability limit). NONO protein emerged as one of the best ranked candidates in this analysis.

A second and more stringent analysis (**Fig. 49B**) was performed applying CRAPome software to our data. This pipeline incorporates the SAINT parameter to filtering out high probable false positive interactions. Then, we ranked the obtained cleaner outcome according to their FC in the pTINCR condition and their SAINT probability. From this analysis, CDC42 protein emerged as the best ranked candidate.

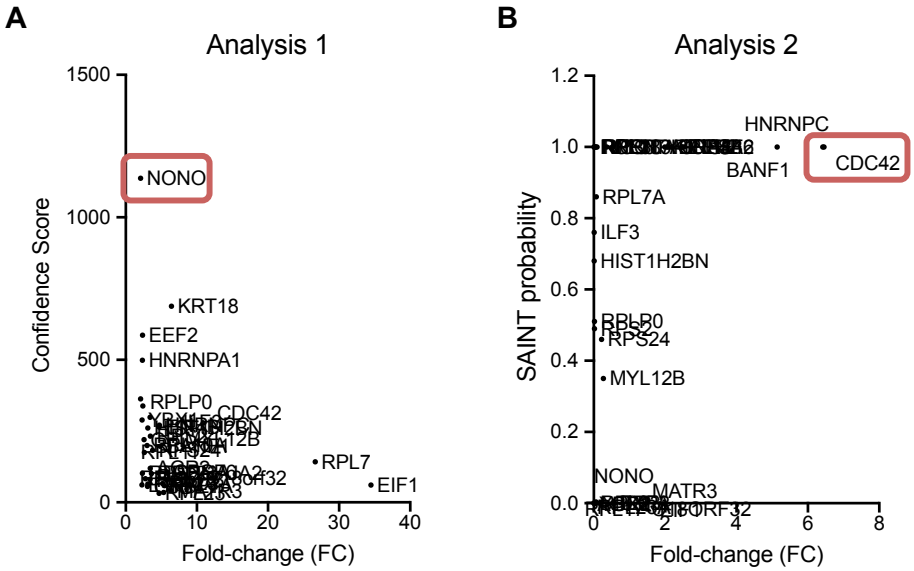


Figure 49. pTINCR interactome candidates.
A. Graphical representation of the interaction candidates identified by mass spectrometry. Graph plots interactors according to their FC and its confidence score.
B. Graphical representation of the interaction candidates identified by mass spectrometry using CRAPome algorithm to remove false positives and background proteins. Graph plots interactors according to their SAINT probability and FC score.

3.2.1. Non-POU domain-containing octamer-binding protein (NONO)

NONO is a nuclear protein belonging to the *Drosophila* behavior/human splicing (DBHS) protein family. It is implicated in several processes, such as RNA splicing, processing and transcriptional regulation, DNA repair upon damage and paraspeckles formation. Paraspeckles are a type of liquid-liquid phase separation (LLPS) condensates found in the nucleus of mammalian cells and. Although its function is not fully understood, they are thought to regulate gene expression by nuclear retention of mRNAs. Interestingly for this thesis, paraspeckles formation is associated with loss of pluripotency in ESC and connected to the process of differentiation [317]. Therefore, we wanted to further explore the relationship between pTINCR and NONO.

3.2.1.1. pTINCR binds to NONO and enhances its SUMOylation

First, we performed co-immunoprecipitation assays using pTINCR-HA and Flag-tagged NONO and validated the interaction of both proteins (**Fig. 50A**). Given the essential role of SUMO in the assembly of LLPS bodies [241, 242, 318], we wanted to assess the interplay between pTINCR, NONO and SUMO. We performed immunoprecipitation assays of SUMO and SUMOylated proteins (using HIS-tagged SUMO) in cells that were transfected with Flag-tagged NONO, with or without pTINCR-HA. Our preliminary results confirmed that NONO is SUMOylated, and that this modification was enhanced by pTINCR overexpression (**Fig. 50B**).

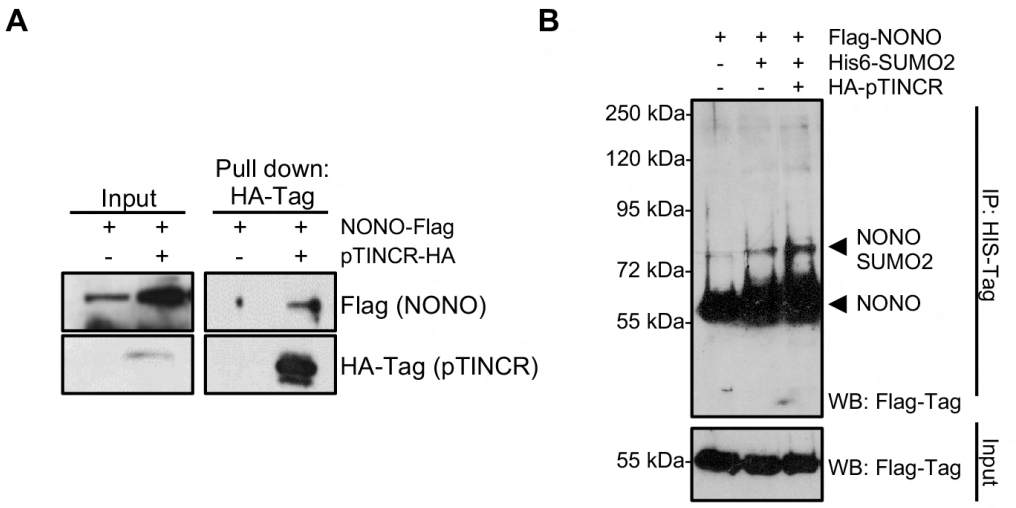


Figure 50. pTINCR interacts with NONO and increases its SUMOylation.
A. Validation of the interaction of pTINCR with NONO by co-immunoprecipitation using an anti-HA antibody in cells overexpressing exogenous pTINCR-HA and Flag-NONO, followed by Western blot.
B. Representative SUMOylation assay in cells transfected with pTINCR-HA and Flag-tagged NONO together with His6-SUMO2. 36h after transfection, anti-HIS pull-down products were analyzed by Western blot with anti-HIS, anti-Flag and anti-HA antibodies. SUMOylated NONO is indicated with arrowheads.

Although we did not have the chance to further explore the meaning of pTINCR enhancement of NONO SUMOylation, we find it very interesting and it will be further considered in the discussion section. Nevertheless, this result supports the idea that pTINCR modulates SUMOylation not globally but in a substrate-specific manner, suggesting that it is probably restricted to pTINCR interactors.

3.2.1.2. pTINCR induces an alternative splicing related with epithelial cell identity

One of the main functions of NONO protein is to participate in the regulation of splicing. Therefore, we wanted to explore whether pTINCR interaction with NONO could affect cellular spliceome. RNA-seq was performed on A549 cells overexpressing pTINCR for 24 hours and analyzed using VAST-TOOLS software, a toolset for profiling, identifying and comparing alternative splicing events [287, 288]. This analysis revealed 473 alternative splicing events directed by pTINCR overexpression (**Fig. 51A**), the vast majority of them affecting exons (~60%). GO analysis of genes whose splicing was altered by pTINCR overexpression revealed an enrichment in key pathways related to epithelial homeostasis, p53 and cytoskeleton (**Fig. 51B**).

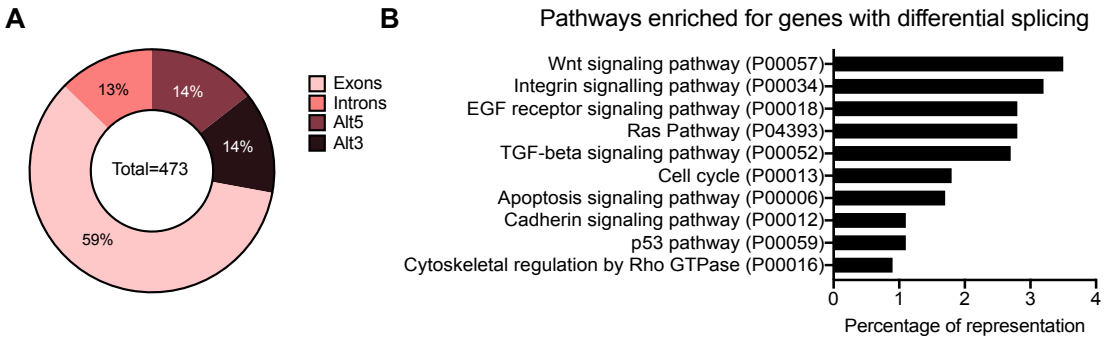


Figure 51. pTINCR overexpression modifies cellular splicing.
A. Summary of alternative splicing changes upon pTINCR overexpression in A549 cells.
B. GO analysis of genes affected by alternative splicing in pTINCR-overexpressing cells using PantherGO tool.

Alternative splicing (AS) events are divided into 4 main types: inclusion or exclusion of exons, retention of introns and alternative used of 5′ (donor) or 3′ (acceptor) splice sites. Compared to the control, cells overexpressing pTINCR are more prone to exon inclusion events and less prone to intron retention, with a preferential use of alternative 3′ splice sites rather than 5’ splice sites (**Fig. 52A**). Moreover, pTINCR triggered more splicing events predicted to preserve the coding potential of their transcripts than events predicted to disrupt the ORF (**Fig. 52B**). This result, together with the observation of decreased intron retention (which usually leads to ORF disruption), suggested that pTINCR-induced AS acts as a modulator of protein isoforms, rather than an on/off regulator of protein translation. Comparison of pTINCR-induced AS with public data of the AS of NONO-KO cells revealed no major similarities between both signatures (**Fig. 52C**), suggesting that pTINCR effect on NONO would not be related with an impairment in its function.

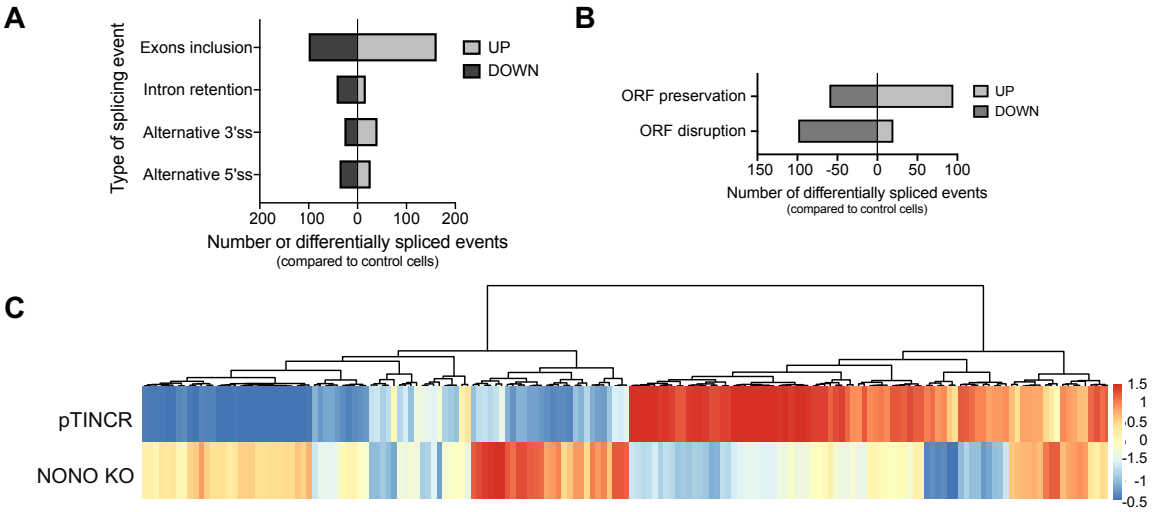


Figure 52. pTINCR-induced alternative splicing acts as a modulator of protein isoforms.
A. Bar plot representing the number of events differentially spliced between pTINCR-overexpressing cells and control cells. The x-axis represents the number of differentially spliced events in pTINCR cells compared to the control. The y-axis represents the type of AS event.
B. Bar plot representing the number of events differentially spliced between pTINCR-overexpressing cells and control cells classified according to the following categories: preserving the transcript ORF or disrupting the transcript ORF.
C. Hierarchical clustering heatmap analysis of AS events in pTINCR overexpressing cells and NONO knock-out cells.

The sequencing results were validated by RT-PCR and capillary electrophoresis and the biological effect of each splicing event was reviewed using VastDB bioportal (https://vastdb.crg.eu/wiki/Main_Page), an atlas of alternative splicing profiles. Among validated events, we found some interesting ones that could be related with pTINCR phenotype (**Fig. 53A**). For example, pTINCR overexpression causes an AS of ANAPC1 protein (key component on the APC/C complex that controls the progression of mitosis and G1 phase) that results in ORF disruption. Also, SUN1 protein (linker between nucleoskeleton and cytoskeleton with a role in DNA repair) suffers a splicing event resulting in an alternative isoform. However, the most interesting result for us was the splice event on CD44 protein. CD44 is a receptor for hyaluronic acid that mediates cell-cell and cell-matrix interactions and participates in cell migration, tumor growth and progression. Our results revealed that pTINCR overexpressing cells have higher amounts of the CD44 splicing variant CD44.V4, that involves the retention of exon 9. This splicing variant has been reported to be specific of epithelial tissues (**Fig. 53B**). We validated the increase in CD44.V4 variant in our epithelial cellular models by RT-qPCR, using primers specifics for the retained exon (**Fig. 53C**).

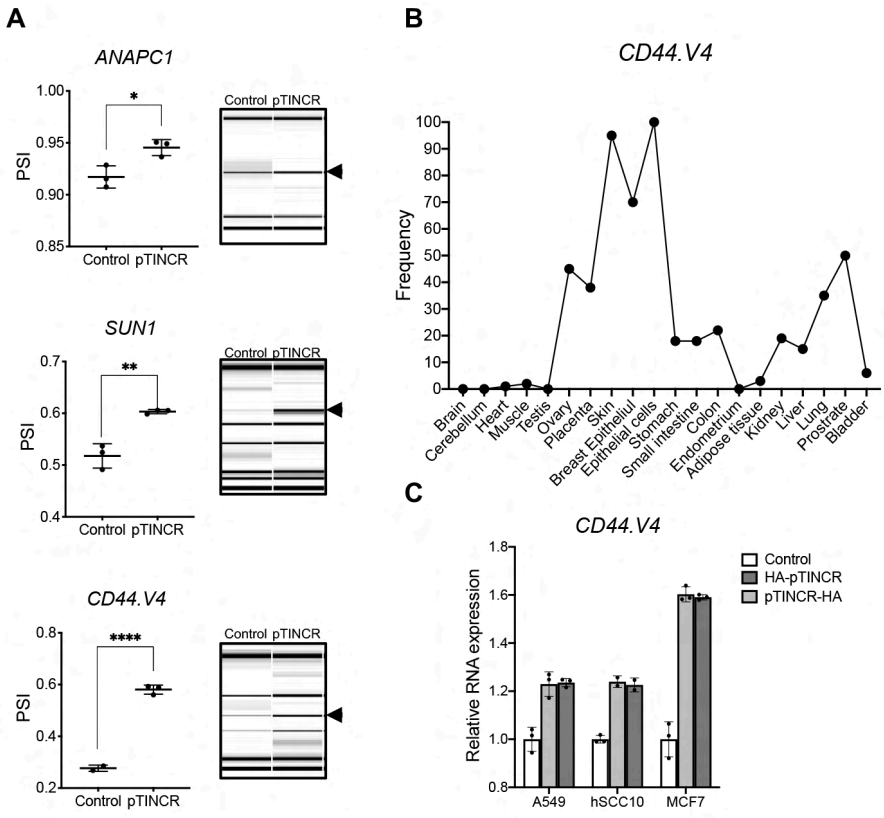


Figure 53. pTINCR modifies CD44 splicing towards an epithelial-specific isoform.
A. RT-PCR validation of selected splicing events inferred from RNA-seq analyses. Selected candidates were analyzed by semi-quantitative RT-PCR and quantified by capillary electrophoresis. Each panel includes a picture of the gel (right) showing the exon inclusion event (of one of the replicates) and the corresponding quantification (left) of the triplicates (PSI = molarity of inclusion product / molarity of inclusion + skipping products).
B. Diagram representing the frequency of CD44.V4 variant across tissues. Data was obtained from the VastDB bioportal.
C. Levels of CD44.V4 variant in A549, hSCC10 and MCF7 cells after pTINCR overexpression for 24 hours measured by RT-qPCR. mRNA expression is normalized to GAPDH expression and relative to the control.

3.2.2. Cell division control protein 42 homolog (CDC42)

Another good candidate in the pTINCR interactome was CDC42, a small GTPase of the Rho family with an important function in actin cytoskeleton regulation and epithelial morphogenesis [73, 101, 109, 119]. Interestingly, our RNAseq analysis pointed to a statistically significant positive correlation between pTINCR expression and the gene sets “GTPase binding” and “GTPase activator activity” (**Fig. 54**).

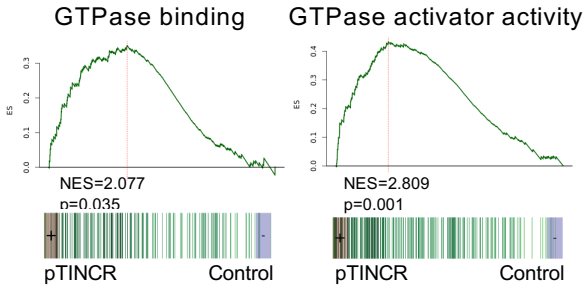


Figure 54. pTINCR-overexpressing cells present an enrichment in gene sets related with GTPase proteins.
Gene Set Enrichment Analysis of pTINCR-induced genes in hSSC10 cells versus the indicated gene signatures.

3.2.2.1. pTINCR binds CDC42 and promotes its activation

First, we validated the interaction between pTINCR and CDC42 by co-immunoprecipitation assays (**Fig. 55A**). Next, we assessed the effect of pTINCR on the activation status of CDC42 by pull-down experiments and observed that pTINCR, but not pTINCR-SIMmut, increases GTP-CDC42 levels (**Fig. 55B**). Of note, we also observed reduced levels of co-immunoprecipitated pTINCR-SIMmut compared to WT pTINCR-HA. This could suggest that pTINCR needs to bind SUMO for its interaction with CDC42 and its consequent activation, although we could not discard the reduced stability of pTINCR-SIMmut **(Fig. 46B)** as an alternative explanation.

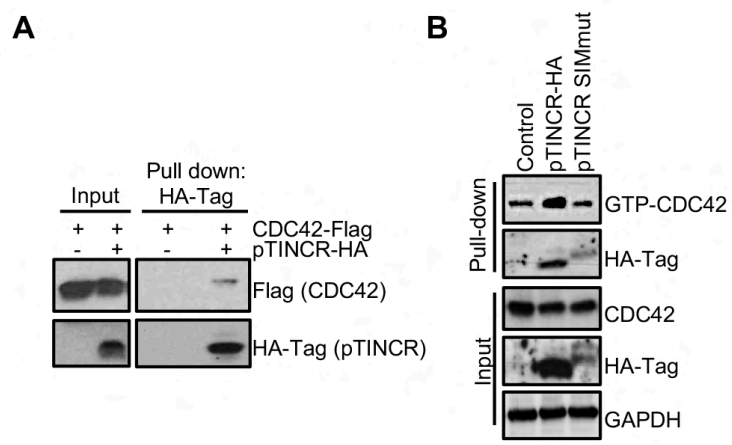


Figure 55. pTINCR interacts and activates CDC42.
A. Validation of the interaction of pTINCR with CDC42 by coimmunoprecipitation using an anti-HA antibody in cells overexpressing exogenous pTINCR-HA or pTINCR-SIMmut and Flag-CDC42, followed by Western blot.
B. CDC42 activation assay performed by pulling down GTP-CDC42 followed by Western blot using an anti-CDC42 antibody in hSCC10 cells after 4 days pTINCR overexpression. syORF was used for C-terminal and pTINCR-SIMmut conditions.

3.2.2.2. pTINCR promotes CDC42 activation by enhancing its SUMOylation

Our previous results with NONO drove us to investigate whether pTINCR modulation of SUMOylated proteins was also applicable to CDC42. To our knowledge, the SUMOylation of CDC42 has not been described before, so we first studied whether CDC42 was a target of SUMO and if so, how pTINCR affected this modification. Importantly, we confirmed that CDC42 is SUMOylated and that this modification is enhanced when we overexpressed pTINCR (**Fig. 56A and B**). These results, together with the effect on NONO SUMOylation, strongly support that pTINCR modulates the SUMOylation of its interactors.

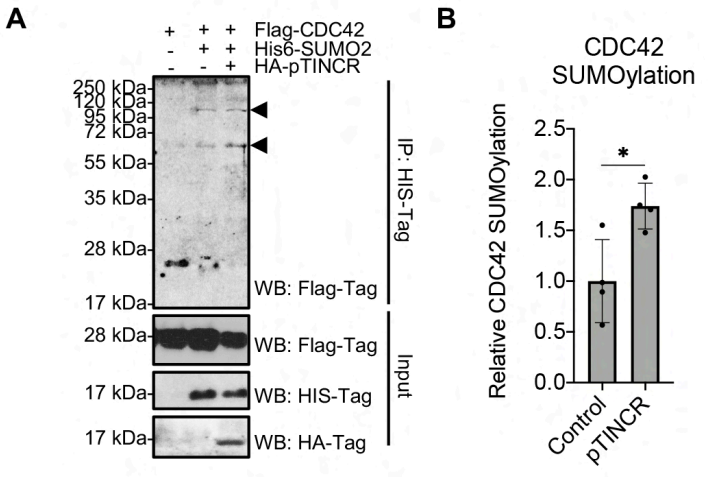


Figure 56. pTINCR increases CDC42 SUMOylation.
A. Representative SUMOylation assay in U2OS transduced with pTINCR-HA and Flag-tagged CDC42 together with His6-SUMO2. 36h after transfection, anti-HIS pull-down products were analyzed by Western blot with anti-HIS, anti-Flag and anti-HA antibodies. SUMOylated CDC42 is indicated with arrowheads.
B. Quantification of 4 independent CDC42 SUMOylation assays. Values of CDC42 SUMOylation are normalized to total CDC42 and relative to the control. *p < 0.05, using Multiple t-test.

To understand the biological significance of CDC42 SUMOylation, we used the SUMOylation inhibitor ML-792 (**Fig. 57A**) and analyzed CDC42 protein localization, stability and activation. Immunofluorescence staining showed that CDC42 SUMOylation did not significantly modify protein localization (**Fig. 57B**) or protein stability (**Fig. 57C**), given that we only saw effects beyond 24 hours of cycloheximide.

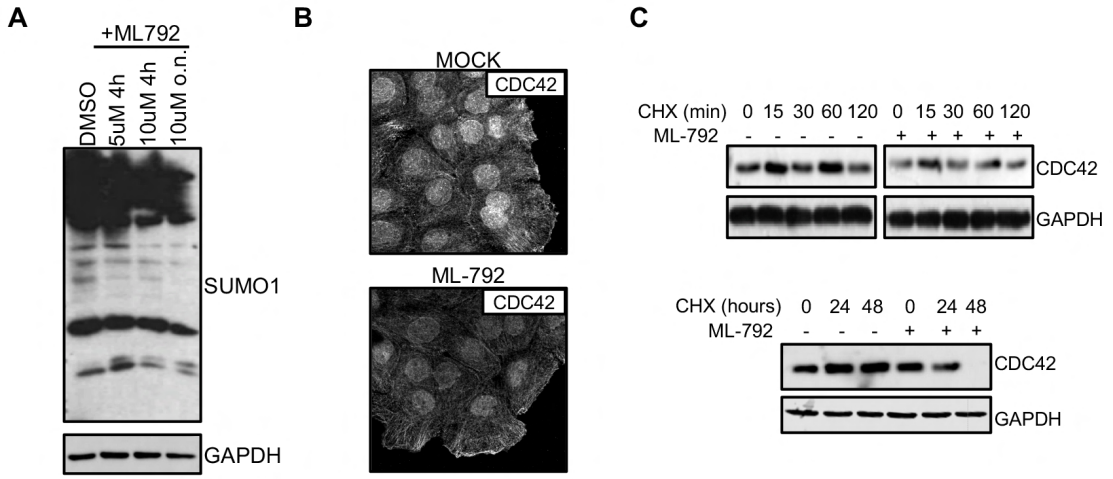


Figure 57. Inhibition of CDC42 SUMOylation does affect its localization or stability.
A. Western blot of SUMO1 in HaCaT cells treated or not with ML-792 at the indicated conditions.
B. Immunostaining images showing CDC42 in HaCaT cells treated or not with ML-792 for 16 hours.
C. Analysis of CDC42 stability in HaCaT cells upon SUMOylation inhibition. Cells were treated with ML-792 for 16 hours prior to CHX treatment for the indicated time points and protein levels were analyzed by Western blot analysis.

However, the inhibition of SUMOylation did decrease the levels of GTP-CDC42 in control cells and partially reverted pTINCR-induced CDC42 activation (**Fig. 58**), suggesting that CDC42 SUMOylation is needed for its activation.

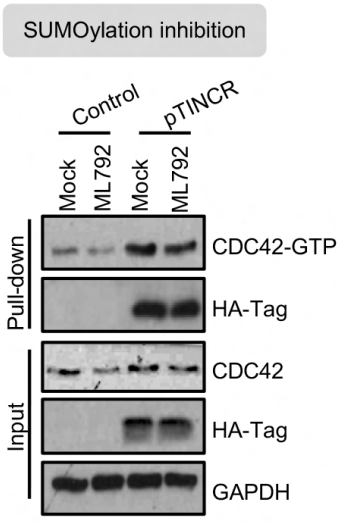
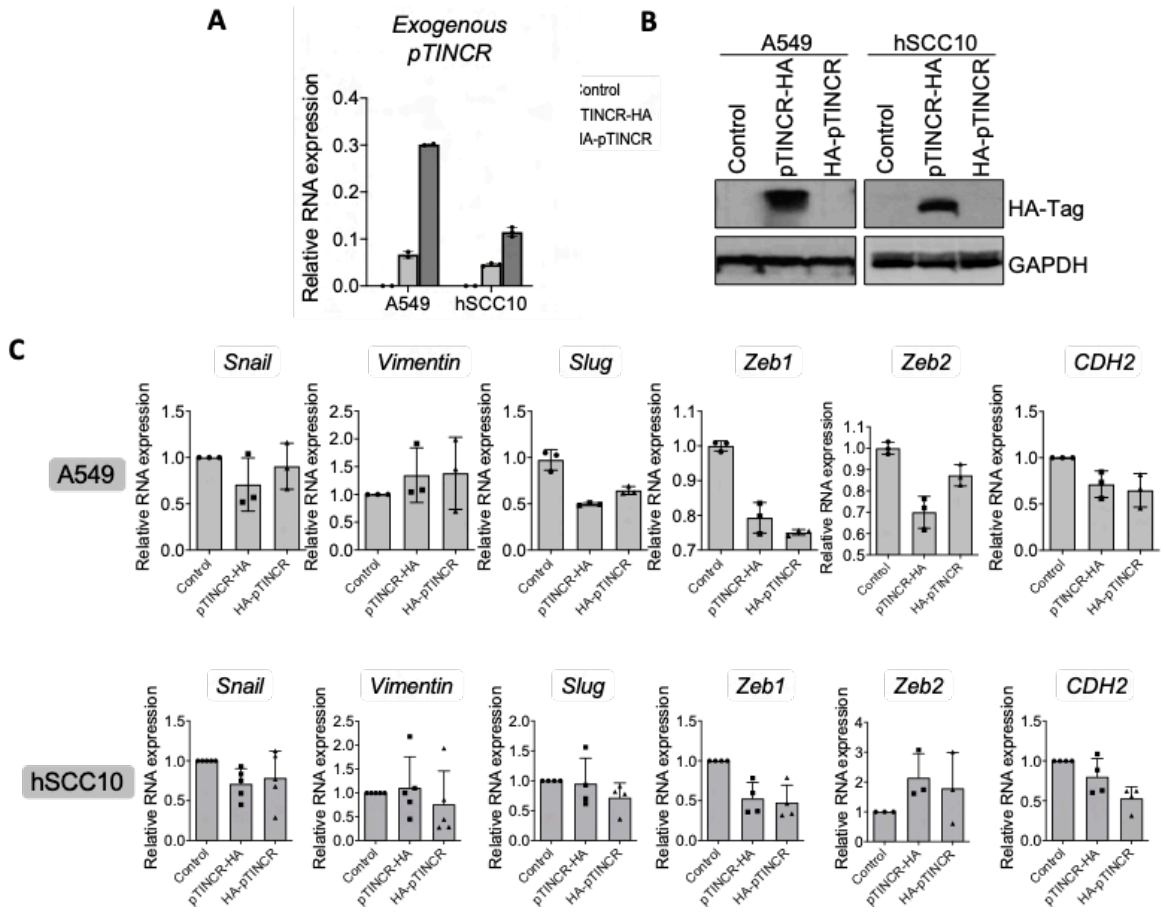


Figure 58. pTINCR-enhanced CDC42 SUMOylation is crucial for its activation.
CDC42 activation assay performed by pulling down GTP-CDC42 followed by Western blot anti-CDC42 in HaCaT cells after 16h of ML-792 treatment and 24 hours of doxycycline induction. syORF was used.

3.2.2.3. pTINCR activation of CDC42 does not promote EMT

Activation of CDC42 culminates with the stimulation of a variety of signaling cascades regulating cellular processes such as cytoskeletal remodeling and establishment of cell polarity, but also promotion of cell migration, invasion and cell proliferation. Therefore, we wanted to analyze the effect of our microprotein on epithelial-to-mesenchymal transition (EMT). We used doxycycline-inducible pTINCR-overexpressing cells and measured the expression of EMT genes, as well as cell migration and invasion after 4 days of pTINCR induction (**Fig. 59A and B**). We observed a slight downregulation in the expression of some EMT markers, although these results were not consistent across cell lines (**Fig. 59C**). Moreover, no relevant effects were observed in cell migration nor invasion after pTINCR overexpression (**Fig. 59D-G**). Altogether, we conclude that pTINCR does not induce an EMT program.



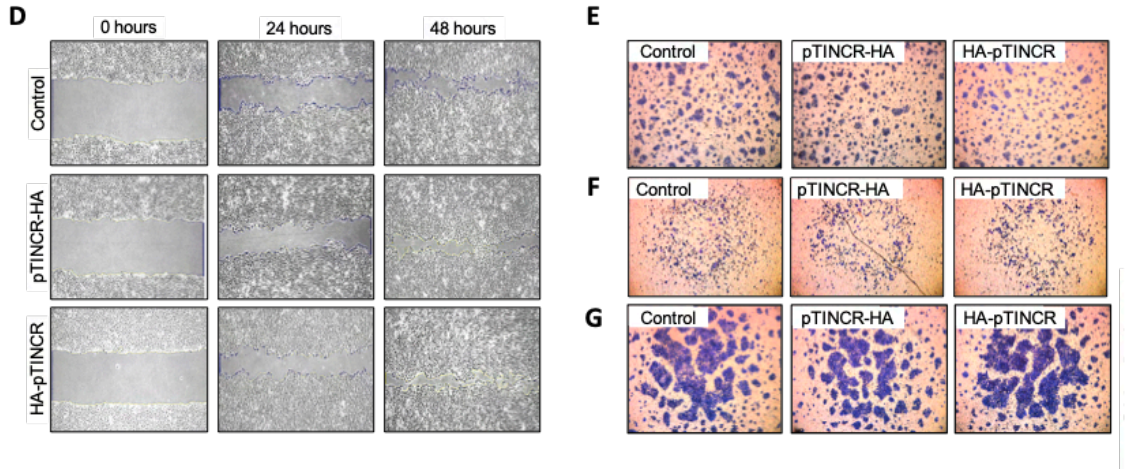


Figure 59. pTINCR does not promote mesenchymal features.
A. Exogenous pTINCR expression in cells transduced with C-terminal (pTINCR-HA) and N-terminal (HA-pTINCR)
HA-tagged microprotein by RT-qPCR. mRNA expression is normalized to GAPDH in each case and relative to the control. Error bars represent the mean ± SD.
B. Exogenous pTINCR protein levels by Western blot using an anti-HA antibody in cells transduced with C-terminal (pTINCR-HA) and N-terminal (HA-pTINCR) HA-tagged microprotein.
C. Expression of the indicated EMT markers in A549 and hSCC10 cells after 4 days of pTINCR induction with doxycycline by RT-qPCR. mRNA expression is normalized to GAPDH and relative to the control in each case. Error bars represent the mean ± SD (n ≥ 2 independent experiments).
D. Cell migration was assessed by wound healing assay in A549. Cells were treated with doxycycline for 4 days to induce pTINCR expression before performing the scratch on the monolayer. Figure shows representative pictures of migrating cells 24- and 48-hours post-scratch.
E. Cell migration in hSCC10 was assessed by transwell assay. Cells were treated with doxycycline for 4 days to induce pTINCR expression before seeded in transwells. Figure shows representative pictures of migrating cells 24 hours post-seeding.
F and G. Cell invasion was assessed using Matrigel-covered transwells in A549 (**F**) and hSCC10 (**G**) cell lines. Cells were treated with doxycycline for 4 days to induce pTINCR expression before seeded in transwells. Figure shows representative pictures of invading cells 24 hours post-seeding.

3.2.2.4. pTINCR promotes epithelial differentiation through the activation of CDC42

Given the direct relationship of CDC42 with epithelial morphogenesis and differentiation, we wanted to assess whether pTINCR pro-differentiation phenotype was mediated by its ability to bind and activate CDC42. First, we analyzed the activation of CDC42 during calcium-induced differentiation in HaCaT WT, pTINCR-KO and pTINCR-KO transduced with pTINCR syORF (pTINCR-KO+OE). We observed that CDC42 is activated during calcium-induced differentiation in WT cells (**Fig. 60, WT cells**). This result agrees with previous reports showing that regulation of CDC42 activity is critical for processes associated with epithelial morphogenesis, such as cell junction maturation [96, 101, 319]. By contrast, pTINCR-KO cells failed to activate CDC42 upon exposure to differentiation conditions (**Fig. 60, pTINCR-KO cells**). Remarkably, this effect was reverted by pTINCR exogenous re-expression (**Fig. 60, pTINCR-KO+OE cells**), demonstrating that pTINCR microprotein -and not *TINCR* lncRNA- mediates CDC42 activation during differentiation.

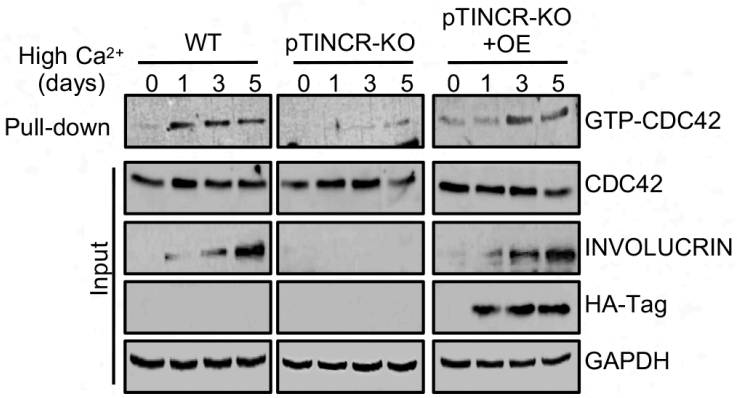


Figure 60. pTINCR is essential for CDC42 activation during calcium-induced epithelial differentiation.
CDC42 activation assay performed by GTP-CDC42 pull-down followed by Western blot using an anti-CDC42 antibody in WT, pTINCR-KO or pTINCR-KO re-expressing syORF pTINCR-HA (pTINCR-KO + OE) HaCaT cells during calcium-induced differentiation for the indicated timepoints. Western blot of INVOLUCRIN was performed as a marker of cell differentiation.

To further confirm that pTINCR-enhanced differentiation was driven by CDC42 activation, we used two chemical inhibitors of CDC42: CASIN and ML-141, which block GTP loading onto CDC42 (**Fig. 61**).

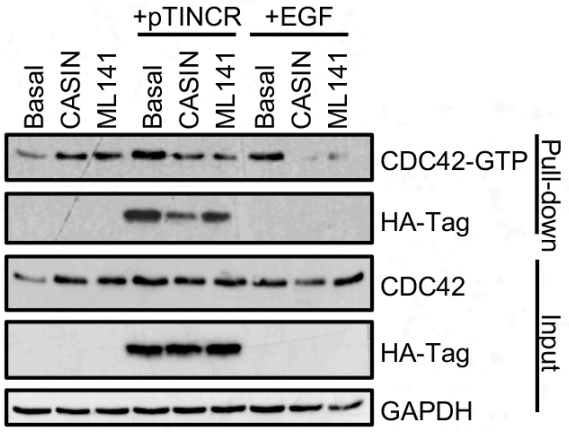


Figure 61. Inhibition of CDC42 activation by CASIN and ML-141.
CDC42 activation assay in HaCaT cells performed by pulling down GTP-CDC42 followed by Western blot using an anti-CDC42 antibody. Cells were treated with the CDC42 inhibitors CASIN or ML-141 for 16 hours and then CDC42 activation was induced by doxycycline treatment for 24 hours (pTINCR overexpression) or EGF treatment 5 minutes before protein collection.

Of relevance, treatment with CASIN or ML-141 inhibits both calcium- and pTINCR-induced differentiation, as seen by an impaired upregulation of differentiation markers (**Fig. 62**).

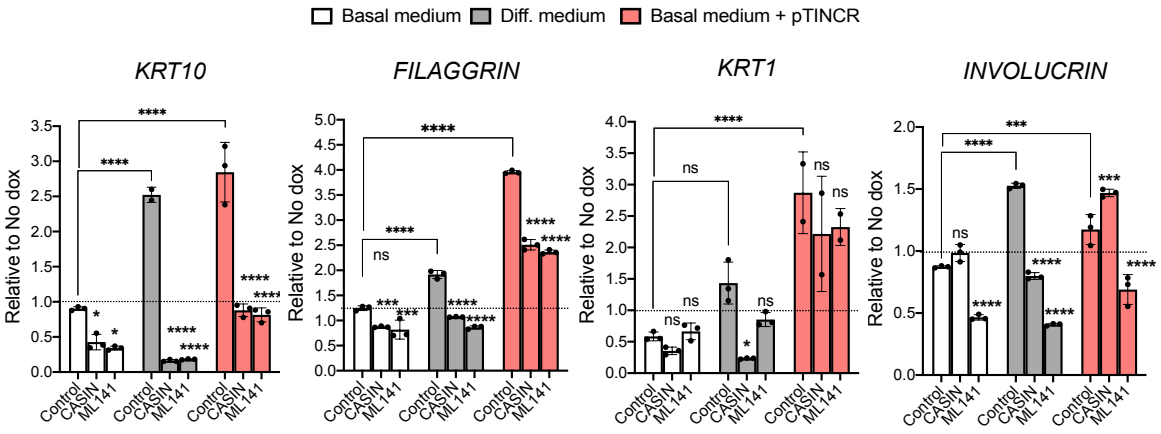


Figure 62. Inhibition of pTINCR-induced CDC42 activation impairs the upregulation of epithelial differentiation markers.
Expression of the indicated differentiation markers upon calcium-induced differentiation or the overexpression of syORF pTINCR-HA in HaCaT cells treated or not with CASIN or ML-141. mRNA expression was measured by RT-qPCR 24 hours after doxycycline induction in cells cultured in basal medium (control cells in white, pTINCR-overexpressing cells in pink) or differentiation medium (control cells in grey). Error bars represent the mean ± SD. Statistical analysis was performed using a two-way ANOVA test with multiple comparison. We show statistical differences within each experimental condition and between basal control cells and differentiated control cells or pTINCR-overexpressing cells (shown in brackets). *p < 0.05, ***p < 0.001, ****p < 0.0001, n.s: not significant using two-way ANOVA test (n = 3 technical replicates).

Moreover, CDC42 inhibition also impaired the remodeling of acting cytoskeleton and the establishment of cell-to-cell junctions induced by pTINCR (**Fig. 63A**) as well as by calcium-induced differentiation (**Fig. 63B**).

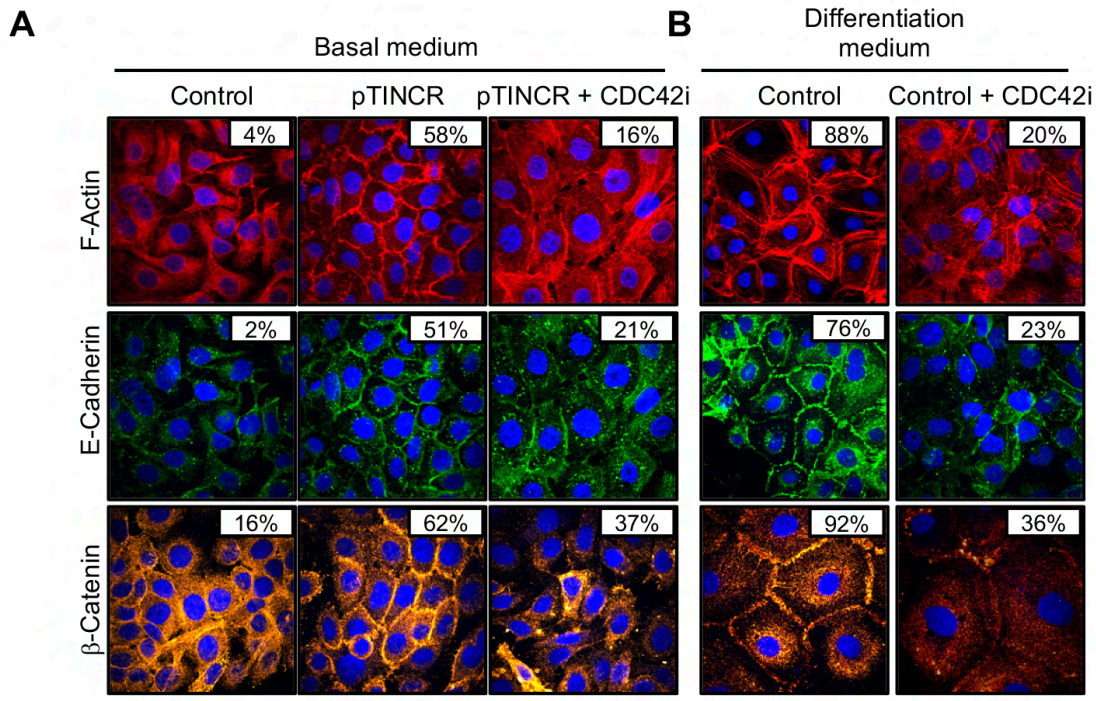


Figure 63. Inhibition of pTINCR-induced CDC42 activation disrupts cell-to-cell junction establishment during epithelial differentiation.
A. Immunostainings using Phalloidin-TRITC (in red), E-cadherin (in green) and β-catenin (in orange) after 24 hours (F-actin and E-cadherin) or 4 days (β-catenin) of doxycycline induction of syORF pTINCR-HA in basal HaCaT cells treated or not with CASIN (F-actin and E-cadherin) or ML-141 (β-catenin). Nuclei are counterstained with DAPI. Total percentage of cells showing a membranous staining is indicated in each image.
B. Immunostainings using Phalloidin-TRITC (in red), E-cadherin (in green) and β-catenin (in orange) after calcium-induced differentiation (24 hours for F-actin and E-cadherin and 4 days for β-catenin) in HaCaT cells treated or not with CASIN (F-actin and E-cadherin) or ML-141 (β-catenin). Nuclei are counterstained with DAPI. Total percentage of cells showing a membranous staining is indicated in each image.

To better define the molecular mechanisms behind the pro-differentiation function of pTINCR, we performed an antibody-based array to analyze the phosphorylation status of 141 proteins associated with actin dynamics (**Fig. 64A**). Among the proteins with increased phosphorylation levels, we found an enrichment in members of the PKC, PLC and the phosphatidylinositol lipid (PtdIns) families, such as phospho-Ser307 PIP5K or phospho-Thr638 PKC α/β II (**Fig. 64B**). This group of proteins and their effectors have been widely related to epithelial morphogenesis and to CDC42 protein activation [115, 319, 320]. In contrast, several members of the MAPK family, whose activation is usually connected with pro-oncogenic and pro-proliferative outcomes, showed decreased phosphorylation levels (**Fig. 64C**). It is also worthy to mention that with this approach we detected a striking increase in F-actin levels in pTINCR-overexpressing cells (**Fig. 64D**).

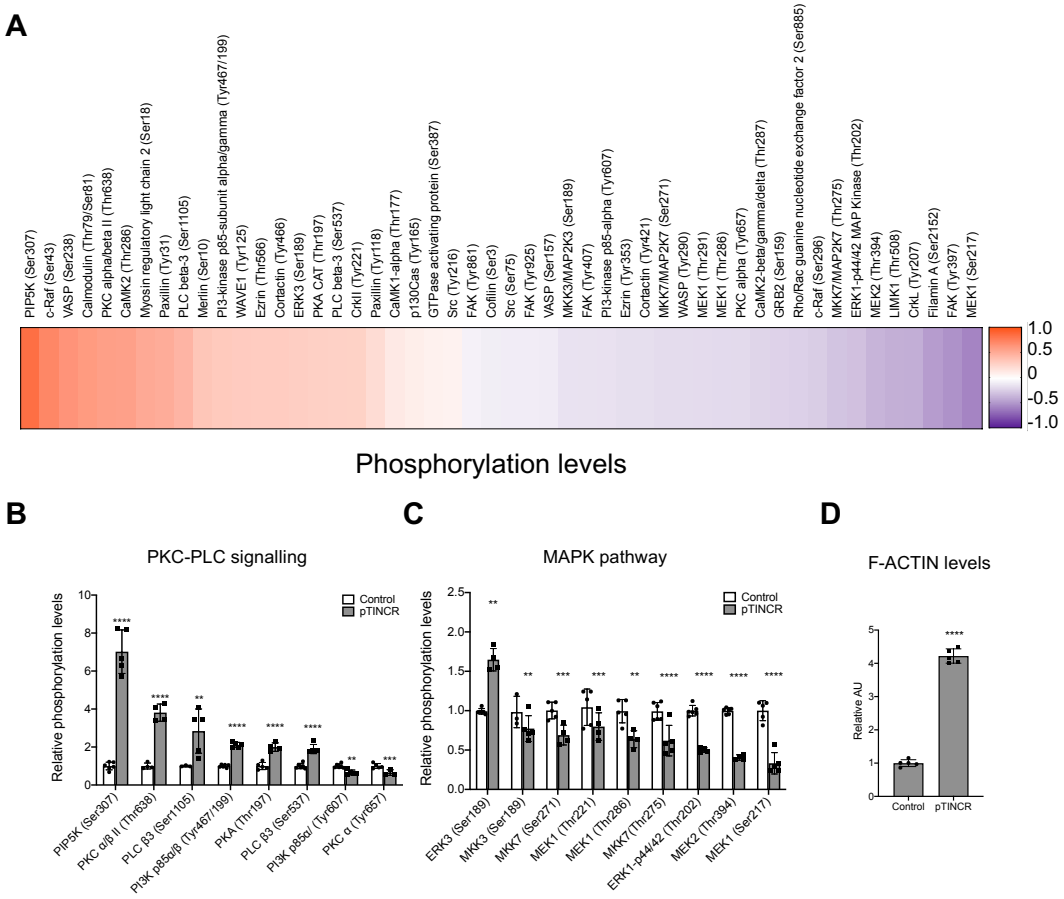


Figure 64. pTINCR overexpression modifies the phosphorylation levels of several proteins in signaling pathways related with epithelial differentiation.
A. Heatmap representing phosphorylation levels of proteins associated with actin dynamics upon pTINCR overexpression in hSCC10 after 4 days of doxycycline induction. Values are normalized first to the loading control and second to the total amount of each protein.
B and C. Phosphorylation levels of PKC-PLC (**B**) and MAPK (**C**) pathway-related proteins upon pTINCR overexpression in hSCC10 after 4 days of doxycycline induction. Data is normalized first to the loading control and second to the total amount of each protein. Values are represented relativized to the control.
D. Relative protein levels of total F-actin upon pTINCR overexpression in hSCC10 after 4 days of doxycycline induction. Data is normalized to the loading control. Represented values are relativized by fold change to the control. ****p < 0.0001, using Student's t-test (n=5 technical replicates).

4. pTINCR in cancer

4.1. pTINCR acts downstream the activation of p53 and leads to damage-induced differentiation

It is known that upon oncogenic insults, cells can trigger a damage-induced differentiation program as a strategy to prevent damaged cells from becoming cancer cells. Therefore, we studied the effect of pTINCR deficiency in p53-dependent DNA-damage-induced differentiation using the pTINCR-KO MCF7 cell line. As showed above in other models, p53-proficient MCF7 (both WT and pTINCR-KO cells) upregulated pTINCR at the mRNA and protein level upon exposure to UV light (a DNA damaging agent), an effect that was abolished in p53-deficient cells (transduced with E6 protein of the oncogenic HPV-16 virus, which induces p53 degradation) (**Fig. 65A-C**). By measuring differentiation markers by RT-qPCR, we corroborated that UV damage induced differentiation in WT but not E6-MCF7 cells (**Fig. 65D**). Interestingly, pTINCR-KO MCF7 do not differentiate upon damage (**Fig. 65D**), even if they upregulated *TINCR* lncRNA to the same extent as WT cells (**Fig. 65B**), suggesting that this effect is triggered by the microprotein rather than the lncRNA. In agreement, this UV-induced differentiation was partially restored when re-expressing exogenous pTINCR in pTINCR-KO cells (**Fig. 65D**). Thus, pTINCR acts downstream p53 to induce differentiation upon damage.

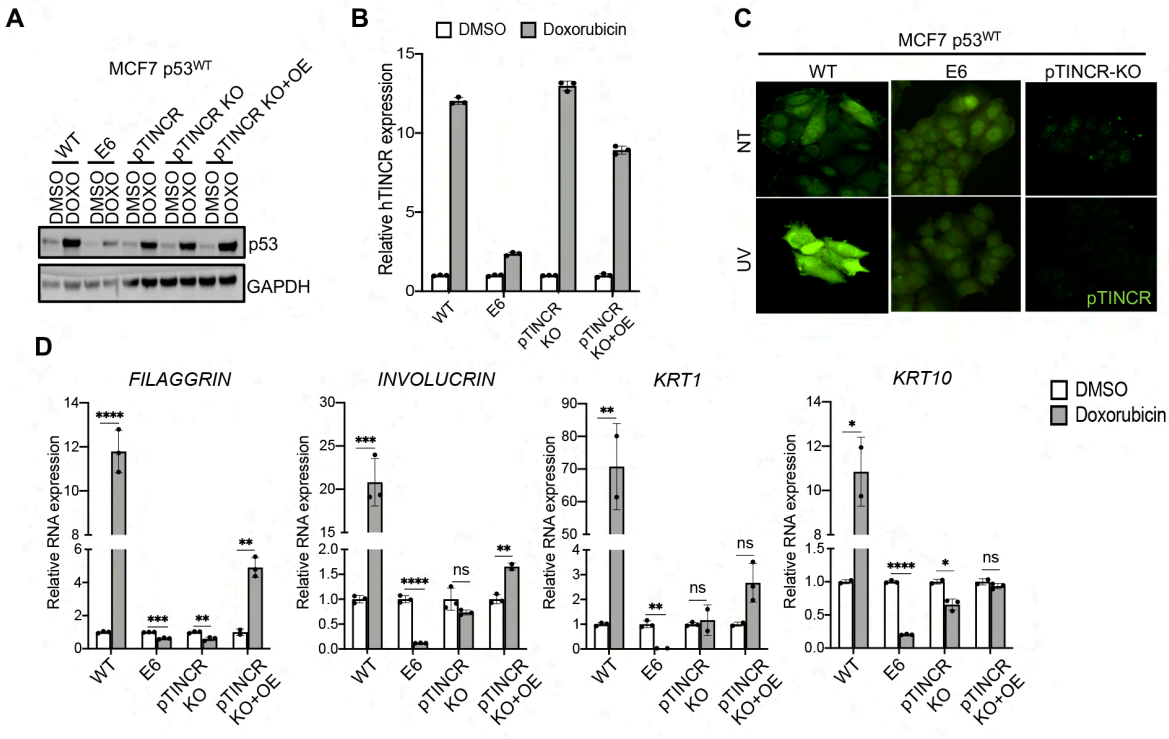


Figure 65. pTINCR upregulation induced by p53 activation is required for damage-induced differentiation.
A. Western blot of p53 after 24 hours of treatment with 1 µm of doxorubicin (DOXO) in the indicated cell lines.
B. *TINCR* lncRNA expression levels analyzed by RT-qPCR after 24 hours of doxorubicin (1 µm) in WT, E6-transduced, pTINCR-KO or pTINCR-KO + OE MCF7 cells. mRNA expression is normalized to GAPDH in each case and relative to the control. Error bars represent the mean ± SD. *p < 0.05, ****p < 0.0001, using Multiple t-test (n = 3 technical replicates).
C. Endogenous pTINCR expression analyzed by immunofluorescence using a custom-made antibody after 24 hours of exposure to UV light in WT, E6-transduced or pTINCR-KO MCF7 cells. NT: non-treated cells.
D. Expression of the indicated differentiation markers by RT-qPCR after 24 hours of treatment with doxorubicin (1 µm) in WT, E6-transduced, pTINCR-KO or pTINCR-KO re-expressing pTINCR-HA (pTINCR-KO + OE) MCF7 cells. mRNA expression is normalized to GAPDH and relative to the control in each case. Error bars represent the mean ± SD. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.001; n.s: not significant, using Multiple t-test (n = 3 technical replicates).

4.2. pTINCR has tumor suppressor activity in epithelial tumors

TINCR lncRNA has been previously documented to be downregulated in human cSCC [296]. Of note, inactivating mutations in *TP53* are early events and well-known prevalent risk factors during the development of cSCC ([321], http://p53.free.fr). Our own analysis of public transcriptomic data corroborated that *TINCR* expression is significantly decreased in cSCC compared to healthy skin (**Fig. 66A**). Moreover, we screened a panel of different human patient-derived cSCC cell lines (previously characterized by their epithelial or mesenchymal features [311, 312]) for *TINCR* expression by RT-qPCR and confirmed its downregulation when compared to HaCaT keratinocytes (**Fig. 66B**).

Notably, epithelial-like cancer cell lines expressed higher levels of *TINCR* lncRNA compared to mesenchymal-like cancer cell lines. Moreover, immunohistochemical analyses in a wide cohort of human cSCC tumors (N=51) using our custom-made antibody revealed that pTINCR expression is decreased in 50 out of 51 samples, where 40% of tumors have a more than 50% reduction below normal levels (**Fig. 66C and D**), supporting a tumor suppressive role of pTINCR in cSCC. We evaluated the *TP53* mutational profile of this cohort by DNA sequencing analysis and immunohistochemistry (**Fig. 66E**) and found that the loss of pTINCR expression did not correlate with the mutational state of p53 (**Fig. 66F**), possibly meaning that tumors benefit from losing pTINCR expression independently of p53 mutation status.

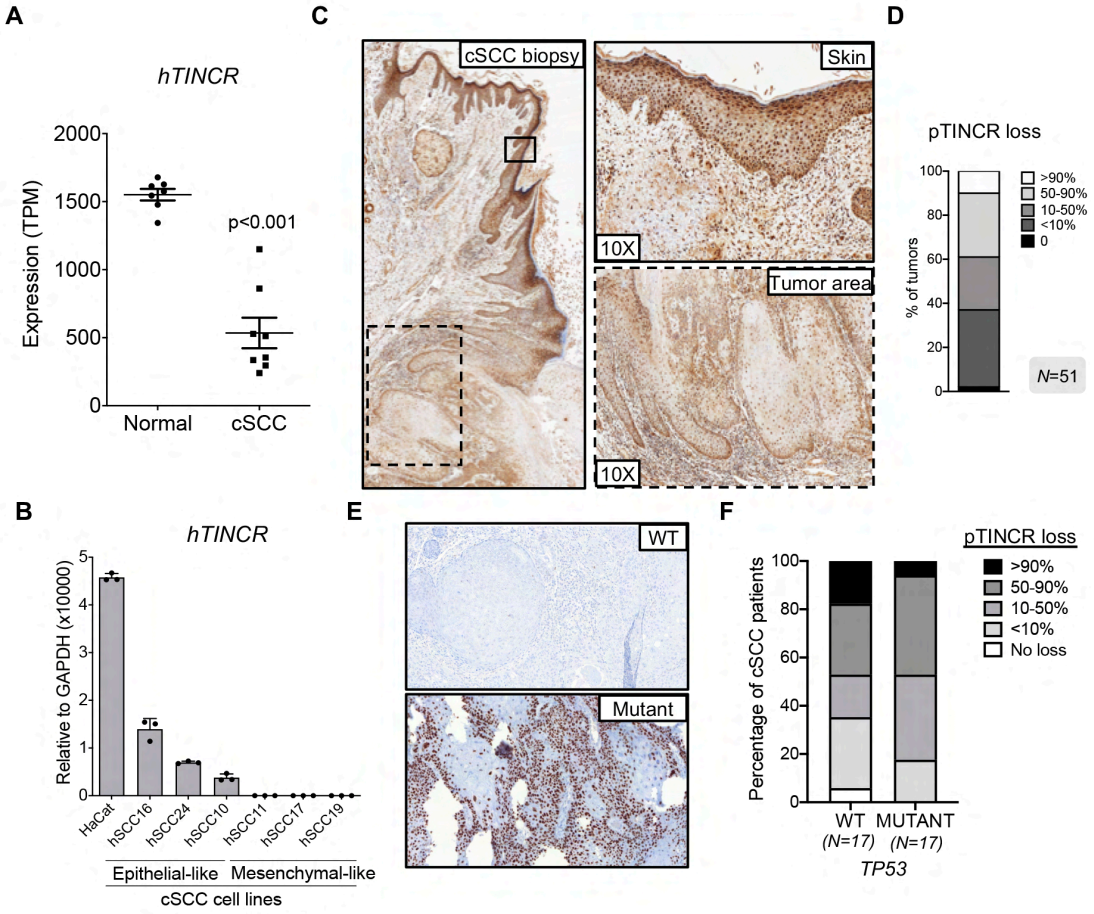


Figure 66. cSCC tumors loss pTINCR expression independently of their TP53 status.
A. Expression of *TINCR* lncRNA in cSCC tumors compared to healthy skins (data subtracted from TCGA data).
B. RT-qPCR analysis of *TINCR* lncRNA expression in a panel of patient-derived cSCC cell lines and in HaCaT keratinocytes. Cancer cell lines are grouped according to their epithelial/mesenchymal traits. mRNA expression is normalized to GAPDH in each case. Error bars represent the mean ± SD.
C. pTINCR expression was analyzed by IHC in a cSCC patient cohort (N=51). Representative image shows the loss of pTINCR expression from healthy epidermis to cSCC (left) and magnification of healthy skin and tumor areas (right).
D. Quantification of the percentage of pTINCR loss in the cSCC patient cohort (N=51).
E. Representative images of IHC against p53 in a WT and a p53-mutant tumor (N=41), showing the typical positive staining associated with p53 mutations.
F. Distribution of pTINCR loss in cSCC tumors (N=34) with wild-type (WT) or mutant TP53 gene. WT; tumors without p53 mutations. Mutant; tumors with a p53 mutation.

Consistent with these results, pTINCR overexpression significantly decreased tumor growth in PDXs generated with hSCC10 cells (**Fig. 67A**). Interestingly, pTINCR-overexpressing tumors presented a significantly higher deposition of hyaline matrix (**Fig. 67B**), a degenerative phenomenon associated with regression, DNA damage, autophagy, apoptosis and cell death [322].

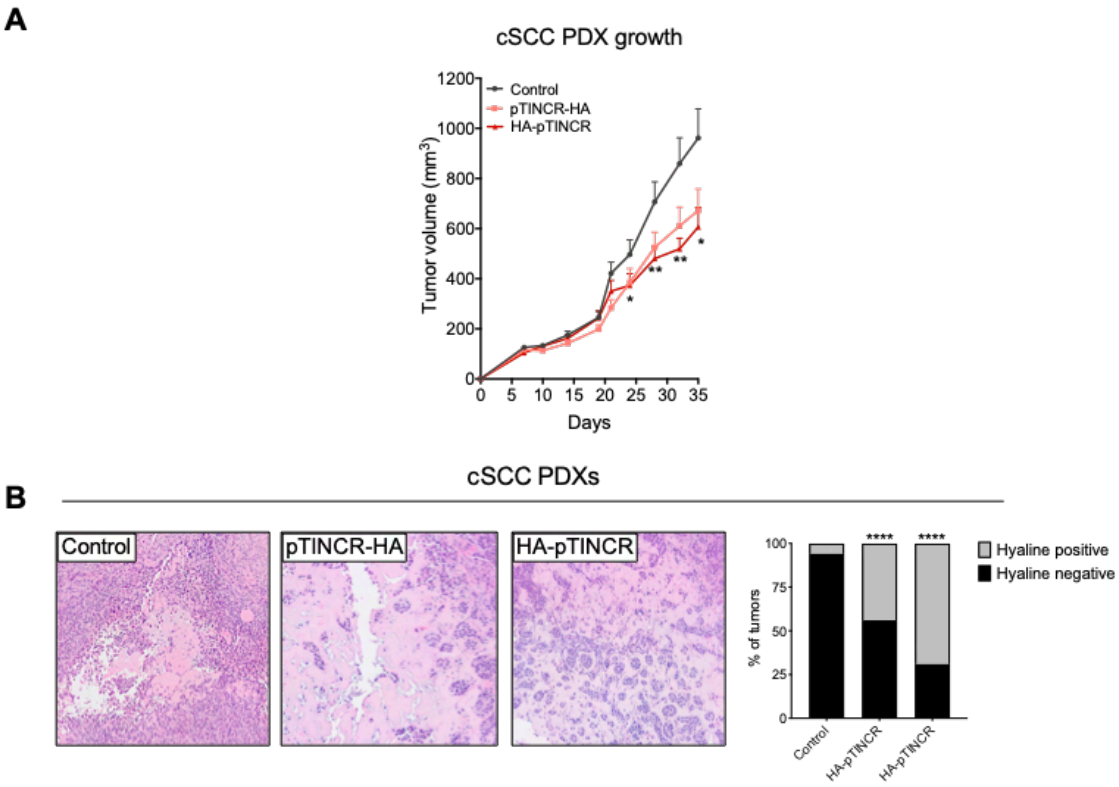


Figure 67. pTINCR overexpression reduces tumor growth in cSCC-PDXs.
A. Effect of pTINCR on cSCC PDXs tumor growth. PDXs were generated by the subcutaneous injection of pTINCR-expressing hSCC10 or control hSCC10 in both flanks of NSG mice. Tumor growth was monitored by measuring the volume at the indicated time points. Error bars represent the mean ± SEM (N=8 mice per group).
B. Representative images of H&E stainings showing areas with hyaline deposition in control and pTINCR-overexpressing cSCC PDXs. Graph shows the quantification of the number of tumors with hyaline deposits. ****p < 0.0001, using Fisher exact test.
C. Kaplan-Meier curves showing the correlation between TINCR expression and patient survival in the indicated tumor types.

Finally, as pTINCR is expressed in several epithelia, we studied the prognostic value of pTINCR expression in other epithelial cancers and observed that high expression of *TINCR* lncRNA significantly correlates with a better overall survival (OS) of patients with bladder carcinoma (Log rank test p = 0.00034, HR = 0.59), PDAC (Log rank test p = 0.0024, HR = 0.51), stomach adenocarcinoma (Log rank test p = 0.022, HR = 0.63), head and neck squamous cell carcinoma (Log rank test p = 0.0022, HR = 0.66) and lung adenocarcinoma (Log rank test p = 0.021, HR = 0.7) (**Fig. 68**).

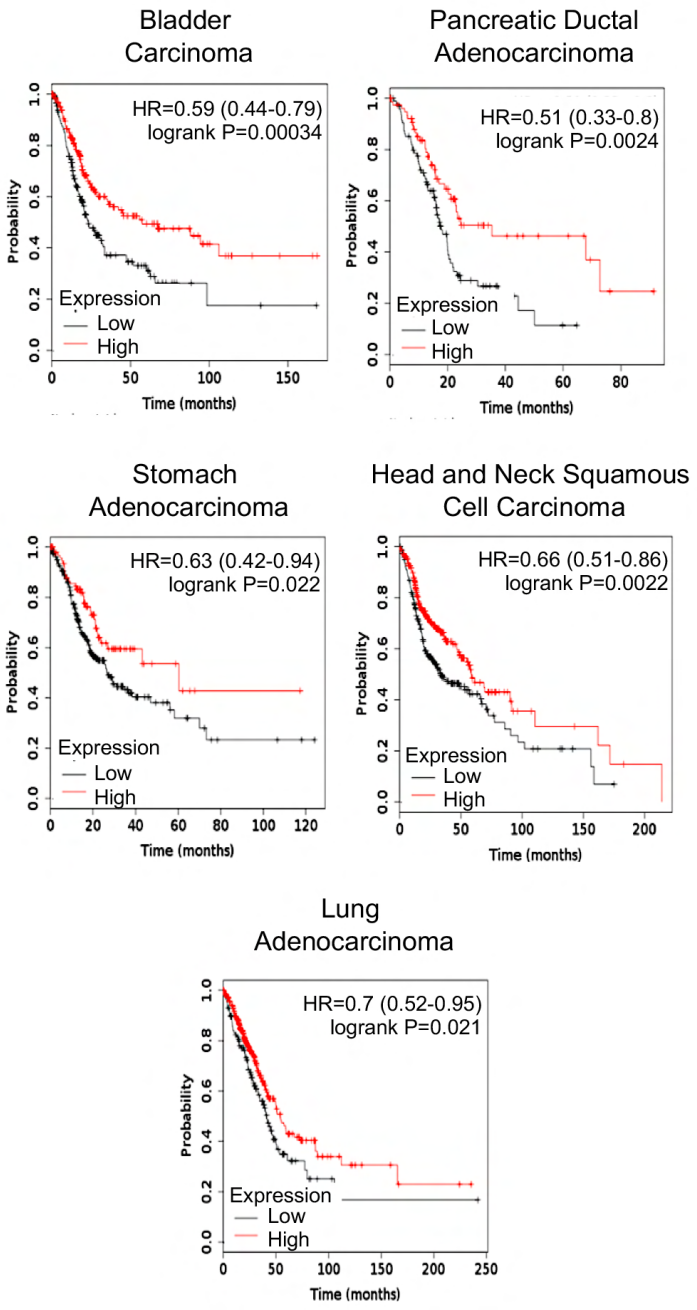


Figure 68. *TINCR* expression correlates with better prognosis in several epithelial tumors.
Kaplan-Meier curves showing the correlation between *TINCR* expression and patient overall survival in the indicated tumor types.

Altogether, these results highly support the role of pTINCR as a tumor suppressor in cSCC and potentially in other epithelial tumors.

DISCUSSION

The microproteome has been largely overlooked until recently, when large-scale bioinformatics and OMICS analyses started to demonstrate that many types of non-coding RNA molecules, such as lncRNAs, do actually code for small functional proteins. Today, their functions are just beginning to be uncovered, highlighting crucial roles of microproteins in many important cellular processes. However, much more research is needed to ascertain the extent of their biological relevance.

In this thesis, we have provided pioneer evidence on the role of microproteins as regulators of cell identity. We have identified and characterized pTINCR, a new ubiquitin-like protein that regulates epithelial differentiation and acts as a tumor suppressor in epithelial cancers.

1. The microprotein behind the lncRNA

1.1. The discovery of pTINCR

In this work we have uncovered pTINCR, a new 87-amino acid microprotein encoded by *TINCR*, a gene previously annotated as a lncRNA. Using PhyloCSF, we computationally identified pTINCR sORF **(Fig. 12B)**, which starts in a methionine located in the first exon of *TINCR*, then splices and continues into the second exon to assemble the final stop codon. The coding potential of this highly evolutionary conserved sORF was further supported by the presence of a strong Kozak sequence and by validating its translation *in vitro* **(Fig. 12C & Fig. 14A and B)**. Expression of *TINCR* lncRNA was previously reported mainly in skin, but also in other stratified and simple epithelia. Accordingly, we have obtained compelling evidence of pTINCR translation in this type of tissue by the performance of different experimental techniques. First, Ribo-seq analysis revealed active translation of pTINCR sORF in mouse skin **(Fig. 13A)**. Second, our own analysis of public mass spectrometry data of organotypic cultures of human skin successfully identified several different tryptic peptides of pTINCR microprotein **(Fig. 14C)**. And third, using our custom-made antibody against pTINCR, we have detected high levels of the endogenous microprotein in several epithelial tissues, such as skin, tongue, palate, esophagus, bladder, cervix, mammary glands, stomach, uterus and glands **(Fig. 26A-C)**. Additionally, we have also detected pTINCR in HaCaT and MCF7 epithelial cell lines, reported to highly express the lncRNA **(Fig. 25)**. Altogether, we concluded that *TINCR* was miss-annotated as a lncRNA and codes for a novel 87-amino acid microprotein, which we have named pTINCR.

In agreement with our results and parallel to this work, peptides derived from pTINCR were detected by proteomics in the human stratum corneum [323], and another study shown that pTINCR is associated with keratinocyte proliferation [324]. However, the molecular mechanisms beyond pTINCR expression were not further explored.

1.2. Analysis of *TINCR* locus: expression, regulation and variants

The transcriptional upregulation of *TINCR* locus has been reported as a key event during differentiation of keratinocytes and other epithelial tissues. We have corroborated this data in our models of epithelial differentiation **(Fig. 16)** and, in addition, we have observed an increase in the expresssion of pTINCR microprotein in differentiated

epithelial cells using our custom-made antibody **(Fig. 24A and B)**. Previous studies also described a p53-dependent upregulation of *TINCR* lncRNA upon damage [302]. We have confirmed this evidence in several models using a panel of different DNA damaging agents and added also evidence of pTINCR microprotein upregulation upon genotoxic stress in a p53-dependent manner **(Fig. 17B & Fig. 27)**. p53, also known as the guardian of the genome, is a very well-studied transcription factor that senses cellular stress and triggers different cellular responses such as cell cycle arrest, apoptosis, senesence or differentation. The p53-dependent regulation of pTINCR suggests a potential role of the microprotein in these processes, acting as a tumor suppressor.

At the time this thesis project began, the human and mouse *TINCR* gene only had one annotated transcript variant (*TINCR*-201 in humans). Our initial analysis identified and validated a sORF within this transcript and our studies focused in the microprotein it encodes, pTINCR. Later, human genome annotation was updated and 4 additional transcripts were identified in *TINCR* locus, two of them also containing potential sORFs **(Fig. 18A)**. *TINCR*-202 was annotated as a 5' truncated transcript which potentially harbors a shorter form of pTINCR microprotein beginning in a non-canonical starting codon. Giving the dubious origin of this transcript variant, together with the evidence provided by our Ribo-seq analysis highlighting a highly conserved starting methionine, we did not take into consideration this new variant for further exploration. *TINCR*-204 transcript arise from a different splicing event involving a previously unknown exon that could give as a result a longer microprotein **(Fig. 18B)**. By re-analyzing the previously used mass spectrometry data from human organotypic skin cultures we were able to identify one peptide derived from the specific C-terminal part of *TINCR*-204-encoded microprotein, confirming its translation in skin **(Fig. 18C)**. These results evidenced the existence of two isoforms of the same microprotein. However, due to some difficulties inherent to the methodology followed in the MS study that we reanalyzed, it was not possible to extract any quantitative conclusion regarding the relative abundance of both microproteins in organotypic skin cultures.

Our custom-made antibody can potentially detect pTINCR-204 microprotein. Indeed, a band that could fit the pTINCR-204 size (~15 kDa) is usually detected by Western blot **(Fig. 24A)**, and it disappears in pTINCR-KO cells (both microproteins share the starting codon that we modified by CRISPR). Intriguingly, while the levels of the higher band increase upon differentiation **(Fig. 24A, MCF7 panel)**, *TINCR*-204 mRNA levels do not **(Fig. 19A)**. mRNA and protein levels do not always correlate with each other, which could be a possible explanation, but it is also possible that the higher band observed is not pTINCR-204 but pTINCR posttranslationally modified (see also section 2.1). Evolutionary speaking, *TINCR*-204 isoform has been described only in human and some primates, but not beyond. This could suggest that *TINCR*-204 is a younger transcript, and therefore maybe with less functional relevance compared with *TINCR*-201. However, this could be simply a matter of gen annotation deficiency in other species.

In any case, our results suggest that *TINCR*-201 is specifically upregulated upon damage and during cell differentiation **(Fig. 19A and B, Fig. 16A and B & Fig. 17)**, which we found very interesting. Therefore, this thesis project focused on characterizing the 87-amino acid microprotein encoded by *TINCR*-201 transcript.

2. Functional characterization of pTINCR

TINCR was first described as a lncRNA involved in epidermal differentiation [296-298]. In this work, we have demonstrated that *TINCR*-encoded microprotein pTINCR has its own role promoting epithelial differentiation, which in turn allows pTINCR act as a tumor suppressor in epithelial tumors.

2.1. Generation of gain-of-function tools

With the purpose of functionally characterize our microprotein, we developed gain-of-functions tools based on cloning pTINCR sORF tagged with an HA epitope in different viral vectors. It has been described that molecular tags can affect proteins in many aspects [325-327]. To minimize all the possible effects of the tag on pTINCR, we generated two constructs placing the HA epitope in either N-terminal or C-terminal part of the microprotein and added a flexible linker of low interacting amino acids between the microprotein and the tag **(Fig. 20A)**. Surprisingly, when detecting the microprotein overexpression by Western blot, pTINCR-HA constructs migrated differentially **(Fig. 21B and C)**. Given that the amino acid composition of both constructs is the same, one possible explanation is that the place of the linker-HA tag affects the SDS binding distribution throughout the protein when performing Western blot, altering the overall net charge of the construct and resulting in different mobilities. However, it is also possible that both constructs suffer different posttranslational modifications or with different efficiency. Indeed, we consistently observed in specific experimental settings that overexpression of HA-tagged pTINCR shows multiple banding. For example, in subcellular fractionation experiments, both C- and N-terminal versions of HA-Tagged pTINCR display a common pattern of three bands in the “membrane-enriched fraction” **(Fig. 22A)**. This pattern is also observed when immunoprecipitating pTINCR-HA (data not shown), a condition in which we have enriched amounts of pTINCR and possibly favoring the detection of less abundant forms of the microprotein. These observations could fit with the hypothesis of a differential efficiency on a given posttranslational modification between constructs. Logically, we considered pTINCR-SUMO as a possible explanation, but it is rather unlikely since SUMO conjugation results in a ~17 kDa shift, which is much bigger that the one we observed. In any case, our results show that both constructs are functional and have the same effect when overexpressed in all the cell lines tested, meaning that this difference is not having a significant impact on their performance.

Interestingly, endogenous pTINCR detection also showed a pattern of three different bands by Western blot, all of them disappearing in pTINCR-KO models **(Fig. 24A)**. We have previously discussed the hypothesis of pTINCR-204 corresponding to the upper band but given the pattern of exogenous pTINCR, it could also be the detection of a putative posttranslationally modified pTINCR. However, whether the bands observed with exogenous pTINCR correspond to the bands observed with the endogenous pTINCR is difficult to confirm by Western blot, given the limited resolution of the technique in that low molecular weight range.

2.2. Uncoupling lncRNA and microprotein functions

The concept of microproteins encoded by lncRNAs usually awakes scientific skepticisms, mainly regarding the possible overlap between the function of the lncRNA and the microprotein. The secondary and tertiary structures of lncRNAs are crucial for their function as modulators of the activity of other proteins, DNA or RNAs [303, 304]. A question that often araises is whether the biological function attributed to the microprotein could be simply the effect of overexpressing an important section of an already functional biomolecule, the lncRNA. In this work, we have been extremely conscious about this concern and took preventive measures to address it to the extent possible. First, we performed our gain-of-function assays using only the pTINCR sORF (264bp) and not the full-length *TINCR* lncRNA (3.7kb). This was an important measure given that the predicted secondary structure of pTINCR sORF do not resembles the lncRNA folding **(Fig. 20B)**. Additionally, our key observations have been confirmed using a syORF that differs by ~20% in its nucleotide sequence from the wild type pTINCR sORF while producing the same protein, measure that exacerbated the differences among secondary structures (**Fig. 20B)**. In addition, our pTINCR-KO cells were engineered by introducing a single nucleotide change in the start codon of pTINCR sORF **(Fig. 23A and B)**. This change is predicted to have a minor effect on the folding of *TINCR* lncRNA and did not affect its expression or transcriptional regulation **(Fig. 23C & Fig. 40A)**, strongly supporting that the results observed in these models were solely explained by the deficiency of the microprotein. In agreement, some of the phenotypes described in pTINCR-KO cells were rescued by overexpressing exogenous pTINCR **(Fig. 60 & Fig. 65D)**.

Altogether, we are confident that our observations are mediated by the microprotein and not by the lncRNA.

2.3. pTINCR role in mediating p53-triggered processes

The fact that transcription and translation of pTINCR was enhanced by activation of p53 prompted us to evaluate the possible role of the microprotein in cellular processes regulated by the tumor suppressor. No significant evidence was observed regarding the participation of pTINCR in cellular senescence or apoptosis **(Fig. 21A and B)**. However, pTINCR overexpression did affect cell proliferation and altered cell cycle **(Fig. 29 & 30)**. We have shown that cells overexpressing pTINCR displayed a significant reduction in their proliferation rate **(Fig. 30)** and in the progression of the cell cycle, presenting an arrest in G1 phase and a reduction in DNA-replicating cells during the S phase **(Fig. 29B)**. These results were further supported by the increase in p21, a known target of p53, and the downregulation of the proliferating cell nuclear antigen PCNA, Cyclin E and Cyclin B, cyclins needed for the proper progression of G1-S phase and M phase, respectively (**Fig. 29A**). The observation that pTINCR is upregulated downstream p53 activation to reduce proliferation pointed to a possible role of the microprotein as a promoter of tumor suppressive processes.

To further understand the p53-*TINCR* axis we analyzed two published p53 Chip-seq experiments performed in cells treated with DNA-damaging agents [328, 329] and we

did not observe any specific binding of p53 to *TINCR* locus (data not shown) suggesting that, although dependent, the activation of pTINCR by p53 is not direct. Nevertheless, the fact that this is not direct does not affect the relevance of the finding, given the importance of p53 pathway in cancer and in differentiation.

2.4. pTINCR role in epithelial differentiation

By using our gain-of-functions tools in several models and through the performance of a wide variety of experimental approaches we have demonstrated that pTINCR favors epithelial differentiation. First, we have shown that HaCaT keratinocytes and three different cancer cell lines of epithelial origin upregulate epithelial differentiation markers upon pTINCR overexpression **(Fig. 32B and Fig. 35)**. In agreement, we reported that pTINCR reinforces cell-to-cell junctions **(Fig. 34 & Fig. 36)** and triggers actin cytoskeleton reorganization towards the formation of a cortical actin network **(Fig. 33 & Fig. 36)**, all crucial processes for cell polarity establishment and epithelial differentiation. The effect of pTINCR decreasing cell proliferation can be reasonably linked to its pro-differentiation role. Furthermore, our results in teratoma formation assays strongly supported the enhancement of epithelial differentiation by pTINCR also *in vivo* **(Fig. 38)**.

Next, we performed loss-of-function experiments with pTINCR-KO cells that, as previously discussed, have been generated by disturbing the microprotein translation but not the lncRNA structure and regulation. Using these models, we have demonstrated that pTINCR deficiency is enough to impair the differentiation process **(Fig. 40 & Fig. 41)**. To be noted, the impairment of differentiation in pTINCR deficient cells is not complete **(Fig. 40B),** which could be explained by the upregulation (and subsequent function) of the lncRNA during the process of differentiation *per se* **(Fig. 16)**. These results raise interesting questions, such as the possibility that some of the functions previously ascribed to the lncRNA could be in fact mediated by pTINCR. But also, they nurture the idea of a possible additional level of regulation involving the coordinated action of the lncRNA and its microprotein. Exploring and understanding this interplay could uncover new mechanisms critical for the regulation of epithelial differentiation, in both physiological and pathological contexts.

Lastly, we have performed an extensive transcriptomic analysis by RNA-seq that covered early and late transcriptomic changes governed by pTINCR **(Fig. 42, Fig. 43 & Fig. 44)**. With this experimental setting, we aimed to capture the initial and direct effects of pTINCR in gene expression, but also to determine long-term and stable transcriptomic changes related to cell identity. In this regard, it is worth noticing that in addition to its localization at cell-to-cell contacts, pTINCR was also found in the nucleus **(Fig. 22 & Fig. 25)**, leaving open the possibility that it may have a direct role as a transcriptional regulator. The results obtained supported our previous findings regarding the pro-differentiation role of pTINCR, showing two predominant group of genes related to cytoskeleton, morphogenesis and cell division whose changes in expression cluster together **(Fig. 43).** Moreover, we found a positive correlation between pTINCR expression and gene signatures associated with actin organization, cell-to-cell junctions, cell polarity or epithelial/epidermal differentiation **(Fig. 44A)**. In agreement, genes sets that negatively correlates with pTINCR expression were related with cell cycle or the Myc

signature **(Fig. 44B)**, an oncogenic pathway closely linked with the prognosis and progression of cSCCs [291, 313, 314], pointing once more to the possible relevance of the microprotein also in cancer.

RNA-seq is a high-throughput technique that provides tremendous amount of information, sometimes beyond the expected. Surprisingly, we observed a cluster of genes (cluster 3) related to mitochondria dynamics and protein targeting, which are downregulated after long-term expression of pTINCR **(Fig. 42)**, as well as a negative correlation with "oxidative phosphorylation" and "protein targeting to the membrane" signatures **(Fig. 44B)**. Although this work has not addressed the role of pTINCR in these cellular processes, we found these results really interesting given the co-localization of HA-Tagged pTINCR in mitochondria and partially in the ER **(Fig. 22C and D)**.

As mentioned in the introduction, protein targeting and vesicle sorting is a key process for the establishment of epithelial cell polarity that ensures the proper sorting of membrane-related proteins to their target localization: apical, lateral o basal membrane domains. Given the relationship of pTINCR with actin cytoskeleton (the cellular roads for vesicle transport) and CDC42 (known to localized in the Golgi apparatus regulating polarized vesicular trafficking), it is tentative to speculate a possible role of pTINCR in this process.

It has been described that during the cornification of the skin (the last step of epidermal differentiation) the keratinocytes programmed to die and become part of the epidermal barrier undergo several internal degradative processes. Mitochondria are known to depolarized, become fragmented and then degraded by mitophagy [330, 331]. This depolarization causes loss of mitochondria membrane potential, which is essential for the oxidative phosphorylation process [332]. Accordingly, fragmented mitochondria are energetically less efficient. Of note, pTINCR has been previously detected by mass spectrometry in the cornified layer of keratinocytes [323]. All together, these observations may point to a functional role of pTINCR during the last step of keratinocytes differentiation related with processes of mitophagy. Although this project did not go any further into this aspect, it represents a potential and amusing future work to be exploited.

It is perhaps noticeable that some of the functional enrichment analysis did not show statistical significance. As described in the methods section, we have used Roast-GSA [283] to compute enrichment scores and their significance. This method, in contrast to the more widely used GSEA [333], is more conservative when computing significance given that it penalizes those gene sets whose genes are expressed in a highly-coordinated manner. Repeating the analysis using the classical pre-ranked methodology implemented by the Broad Institute shown that all selected signatures were highly significant (data not shown), which confirms the stringency of our methodology. Moreover, functional enrichment analysis was computed over an interaction (the statistical term referring to the assessment of differences of differences), which by construction reduces the actual sample size by half. This implies that the power to detect significant results is much lower than when directly comparing the difference between two conditions. With all this in mind, we have rather preferred to follow this stringent methodology despite some p-values displayed larger than the commonly used 0.05

threshold, and we strongly believe that the evidence in our analyses is consistent and in line with the hypothesis being developed in this study.

2.5. pTINCR role in epithelial tumors

Up to now, our results have evidenced a role of pTINCR in epithelial differentiation, but they have also clearly pointed to a potential participation in tumorigenesis. *TINCR* lncRNA was previously documented as deregulated in various tumor types, although its role as an oncogene or as a tumor suppressor seems to be tumor specific [334-342]. In esophageal SCC cells, siRNA-mediated silencing of *TINCR* represses cell proliferation and EMT features [343]. However, in other studies performed in cell lines derived from cervix, head and neck, and lung SCC, its silencing enhanced cell growth and migration [344]. In cSCC, *TINCR* lncRNA has been consistently reported to have a tumor suppressor activity [344, 345]. We have corroborated the downregulation of *TINCR* lncRNA in cSCC **(Fig. 66A and B)** and, moreover, we have demonstrated that pTINCR protein is also lost in human cSCC compared to healthy epidermis **(Fig. 66C and D)**. Of note, despite the p53-dependent regulation of *TINCR* consistently observed *in vitro*, we did not find a correlation between the percentage of pTINCR loss and the mutational status of p53 in these tumors **(Fig. 66E and F)**, suggesting that tumors could benefit from losing pTINCR expression independently of their p53 mutational state. Indeed, its reexpression in cSCC PDXs reduced tumor growth *in vivo* **(Fig. 67A)**, which agrees with previous results on cell proliferation *in vitro,* and further strength the tumor suppressive role of pTINCR. Of note, a marked deposition of hyaline matrix was evident in tumors expressing pTINCR **(Fig. 67B).** Hyaline globules have been reported in numerous types of neoplasms as a rare event usually related to benign lesions [322, 346-351]. Interestingly, the presence of hyaline increases during differentiation (keratohyalin deposits observed in the granular layer) but also in tissues after injury or in tumors after therapy [322, 352-356]. Although there is no clear consensus about its origin, this phenomenon is generally regarded as a hallmark of apoptosis and degeneration [322].

In cSCC, p53 mutations are reported as high-risk factor events and oncogenic drivers of the tumorigenic process. Recently, it has been also described in these tumors that damage-induced differentiation triggered by p53 represents a novel mechanism to prevent malignant progression while preserving epithelial identity [161, 166]. We corroborated in the epithelial breast cancer cell line MCF7 that p53 activation upon damage induces *TINCR* upregulation and epithelial differentiation markers **(Fig. 65)**. Strikingly, pTINCR-deficient cells did not show the same response upon damage, suggesting the requirement of the microprotein for this process **(Fig. 65D)**.

Finally, to expand the relevance of pTINCR in other cancer types, we performed correlation studies using public data and observed that high *TINCR* expression is associated with a favorable prognosis in patients with different tumors of epithelial origin, such as bladder, head & neck, pancreatic, stomach and lung adenocarcinomas **(Fig. 68)**.

Altogether, our results compellingly show that pTINCR acts as a tumor suppressor in cSCC through its ability to assure epithelial cell identity.

3. Molecular characterization of pTINCR microprotein

3.1. pTINCR is a new UBL protein that interacts with SUMO and modulates SUMOylation

By *in silico* analysis of pTINCR structure we observed that it belongs to the ubiquitin-like protein family **(Fig. 45)**. Interestingly, pTINCR is a Type II UBL, lacking the C-terminal GG residues that Type I UBLs typically use for their conjugation, as in the case of UFM, Atg12, Atg8 and UBL5 [226]. Of note, pTINCR has two overlapping SIMs and we have demonstrated that pTINCR interacts non-covalently with SUMO1 and SUMO2/3, an interaction that is necessary for pTINCR stability **(Fig. 45B & Fig. 46)**. The analysis of global SUMOylation pointed to a modulatory role of pTINCR on protein SUMOylation, which extent seems to depend on the cell type **(Fig. 47)**. In line with these results, previous evidences have associated changes in cellular SUMOylation with embryonic development and cell fate stabilization [235, 357], keratinocyte differentiation [237-239] and tumor suppression [251, 261, 358]. The nature of this modulation was an additional interesting point to address. SUMOylation is a complex and dynamic process widely affecting the proteome to fine-tune cell behavior, hence, understanding whether pTINCR effect on SUMOylation was a global effect on all SUMO targets or specific to some substrates was important for the field. By analyzing the SUMOylation levels of common SUMOylated proteins (p53 and B23) we did not observe any changes upon pTINCR overexpression **(Fig. 48)**. However, when assessing the SUMOylation levels of two bona fide interactors of pTINCR (CDC42 and NONO) we did observe that the microprotein increases their SUMOylation **(Fig. 50B & Fig. 56)**, suggesting that the effect of pTINCR in SUMOylation is not global but rather protein-specific.

We recognize that the effect of pTINCR enhancing NONO/CDC42-SUMO2 conjugation is rather modest. However, the stoichiometry of SUMO is typically very low, and the fact that small effects on SUMOylation could have a big biological impact (known as the “SUMO paradox” [359]) is widely accepted in the field. In fact, reverting the increase in CDC42-SUMOylation by treating with a SUMOylation inhibitor was enough to impair the ability of pTINCR to activate CDC42 **(Fig. 58D)**, suggesting that CDC42 SUMOylation is required for its activation. These observations are consistent with other small GTPases, such as RAC, that are also activated by SUMOylation [360].

SUMO conjugation is a very complex field and further molecular studies are needed to fully understand the interplay between pTINCR and SUMO. Nevertheless, we propose a molecular model in which pTINCR acts as a scaffold to enhance or stabilize the SUMOylation of its interactors, modulating their activities **(Fig. 69).**

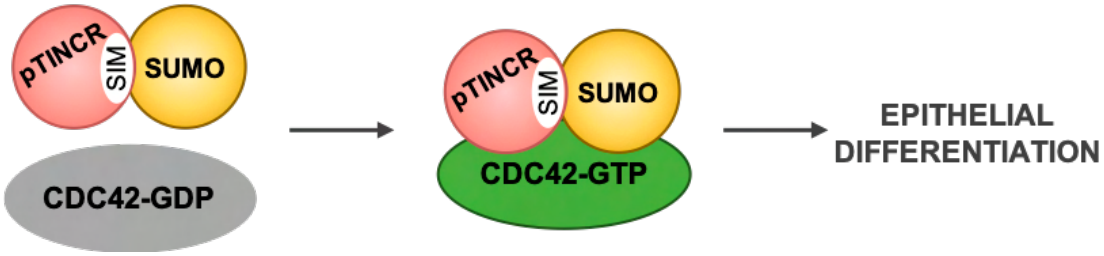


Figure 69. Proposed working model of pTINCR acting as a scaffold to modulate SUMOylation. Model shows the particular case of the pTINCR interactor CDC42. pTINCR is a novel UBL microprotein that interacts with SUMO through its SIM domain and modulates protein SUMOylation patterns. In the image, pTINCR increases CDC42 SUMOylation and promotes its activation, leading to epithelial differentiation and to tumor suppression.

3.2. pTINCR interactome: is pTINCR behaving as a posttranslational modification?

The nature of pTINCR as an UBL is suggestive of pTINCR acting as a posttranslational modification (PTM), although the definition of PTM involves enzyme-dependent covalent binding to its substrates. As explained in the results section (on page 114), pTINCR lacks the GG domain present in Type I UBLs for covalent binding, and we do not have evidence of covalent binding by other residues. Instead, we have shown that pTINCR has two overlapping SIM domains for binding non-covalently SUMO1 and SUMO2/3. The ability to bind SUMO allows pTINCR to potentially bind many SUMOylated proteins and to hypothesize a possible PTM-like role. In this study, we have uncovered two proteins which pTINCR binds to and modulates its activity: NONO and CDC42 **(Fig. 49, Fig. 50A & Fig. 55A)**.

3.2.1. pTINCR and NONO

One of the pTINCR interactors unveiled in this work was NONO, whose SUMOylation is enhanced by pTINCR overexpression **(Fig. 50B)**. Although we did not deeply explore the nature and the impact of this interaction, NONO represents and interesting target for further studies given its involvement in transcriptional regulation, pre-mRNA processing and splicing, and nuclear retention of defective RNAs. The crucial role of NONO in cellular stress and during DNA double strand breaks repair is particularly interesting for us [361-365]. It has been reported that after damage, NONO knock-down impairs DNA damage recovery by delaying the non-homologous end joining pathway [366, 367] and the resolution of γH2AX foci, increasing chromosomal aberrations [361]. The relationship of NONO with damage response is directly linked with its involvement in paraspeckles formation, a class of membrane-less condensates known to promote the retention and sequestration of defective RNA upon DNA damage. Interestingly, the formation of this liquid-liquid phase structures has been reported to be promoted directly by p53 activation [368, 369], which is of great interest in this study given the p53-dependent upregulation of pTINCR. Moreover, SUMO is well-known to be essential for the assemble, maintenance and functionality of paraspeckles. We hypothesize that in response to

cellular damage, p53 activation triggers pTINCR expression, which in turn enhances NONO SUMOylation and promotes paraspeckles formation **(Fig. 70)**. However, much more studies should be performed to fully elucidate the biological meaning of pTINCR-NONO interaction.

mRNA splicing is a critical step in posttranscriptional gene regulation. NONO is known to interact with several components of the spliceosome machinery [370] to promote pre-mRNA alternative splicing [370-372]. We therefore study the possible role of pTINCR in this scenario and observed that it triggers a significant amount of alternative splicing events upon its overexpression **(Fig. 51 & Fig. 52)**. Whether this effect is dependent of the interaction of pTINCR with NONO remains to be clarified. Interestingly, among the pTINCR-induced alternative spliceome, we identified a particular event specific for epithelial tissues, the CD44.v4 variant **(Fig. 53B)**. CD44 is a cell surface protein that modulates cellular signaling. In humans, CD44 encodes nine constitutive exons and nine variable exons, which combination produces an extensive family of protein isoforms. Alternative splicing of this gene generates two major families: CD44 standard isoforms (CD44s, containing only constitutive exons) and CD44 variants isoforms (CD44v, containing at least one of the variable exons). While CD44s isoforms are ubiquitously expressed in many tissue types, CD44v are specifically expressed in epithelial cells [373]. Indeed, it has been well-documented that switching from CD44v to CD44s is essential for epithelial cancer cell to lose their epithelial identity and undergo EMT [374-377].

Tissue-specific alternative splicing patterns have been identified in several studies [378, 379], playing an important role in determining cell identity. Although little is known on how this specific splicing profiles are established, it has been suggested a possible differential and tissue-specific fine-tuning of splicing factors expression [380]. The fact that pTINCR promotes an epithelial-specific alternative splicing event, presumably through its interaction with NONO, leads us to hypothesize that pTINCR could be an epithelial-specific splicing daterminant important for epithelial cell identity adquisition. This represents an exciting future perspective yet-to-be-explored.

3.2.2. pTINCR and CDC42

CDC42 is a master regulator of cell polarity and actin dynamics [91]. During epithelial differentiation, CDC42 is involved in the establishment of the apico-basal cell polarity through the control of cell-to-cell junction and asymmetrical cell division. CDC42 knock-out mice die soon after birth with severe defects in the formation of the epidermal barrier, indicating the essential role of CDC42 in epidermal differentiation [137]. In cancer, CDC42 activation has been commonly associated with oncogenic traits such as increased cell migration and invasion [214]. However, several publications support its tumor suppressor role in certain contexts given its pro-differentiation role [381]. Recent studies in SCC showed that activation of the RHO and RAC GTPases, but not CDC42, is common in most types of this cancer [382]. In addition, high levels of VAV2 (a guanine nucleotide exchange factor of the Rho family) in both cutaneous and head and neck SCC favors tumorigenesis through activation of RAC and RHO proteins, but not through CDC42 [383].

Our results show that pTINCR-induced CDC42 activation does not lead to lamellipodia or filopodia formation **(Fig. 33 & Fig. 36)**, or an enhancement of mesenchymal traits **(Fig. 59)**. Instead, it leads to cytoskeleton remodeling towards a cortical deposition of actin filaments resembling a more differentiated state. In agreement, we found pTINCR located in the cell membrane **(Fig. 22 & Fig. 25)**, where CDC42 is required to stabilized cell junctions. By using two different CDC42 inhibitors (ML-141 and CASIN), we have demonstrated that pTINCR needs active CDC42 to promote HaCaT keratinocyte differentiation **(Fig. 62 & Fig. 63)**. Moreover, we showed that pTINCR-KO HaCaT cells fail to activate CDC42 under high calcium-differentiation medium, a defect that is rescued by the overexpression of exogenous pTINCR **(Fig. 60).** With these results, we concluded that the pro-differentiation role of pTINCR is mediated, at least in part, by the activation of CDC42.

The outcome of CDC42 activation relies on its downstream effectors. By performing a phosphorylation array of cytoskeleton-related proteins we have observed the pTINCR induces a substantial increase in the phosphorylation status of proteins that have been widely associated with epithelial differentiation and stratification, such as Merlin [121, 384-387], the p85-subunit α/γ of PI3K [388-393] or several members of the Par complex [394-396] **(Fig. 64B)**. These last two families of proteins are direct downstream effectors of CDC42 and extremely important for the establishment of the apico-basal polarity of epithelial tissues, as well as for the control of asymmetric cell division, spindle orientation, vesicular trafficking, and stem cell differentiation [27, 319, 320]. On the other side, we observed a significant decrease in the phosphorylation status of many proteins of the MAPK/ERK pathway **(Fig. 64C)**, reported to be inhibited during keratinocyte differentiation [397, 398]. The impact of most of these specific phosphorylations has not been reported and hence we cannot anticipate its biological significance, but our results encourage to further characterize them as potential drivers of epithelial differentiation.

4. Final considerations and working models

In this thesis project we have performed an in-depth mechanistic analysis to characterize pTINCR, an 87-amino acid microprotein regulated in a p53-dependent manner, expressed in epithelial tissues and downregulated in cSCC. We have demonstrated that pTINCR is a critical regulator of epithelial differentiation and that it has tumor suppressor activity.

At the molecular level, we have described a new member of the UBL family that binds SUMO through its SIM domain and changes cell SUMOylation patterns. Moreover, we have identified two targets of pTINCR, NONO and CDC42, and have shown that pTINCR promotes their SUMOylation. NONO is a protein belonging to the paraspeckles family with a known role in damage response, RNA processing and splicing. Interestingly, we have observed that pTINCR overexpression changes the cellular spliceosome and identified at least one specific alternative splicing event related to epithelial cell identity. We proposed a working model in which pTINCR enhances NONO SUMOylation to 1) trigger an epithelial splicing program and 2) promote paraspeckles formation and defective RNA retention in response to cellular damage and p53 activation **(Fig. 70)**. On the other hand, we described the interplay between pTINCR and CDC42, demonstrating

that pTINCR-induced CDC42 activation is required for the pro-differentiation phenotype of the microprotein that we described **(Fig. 71)**.

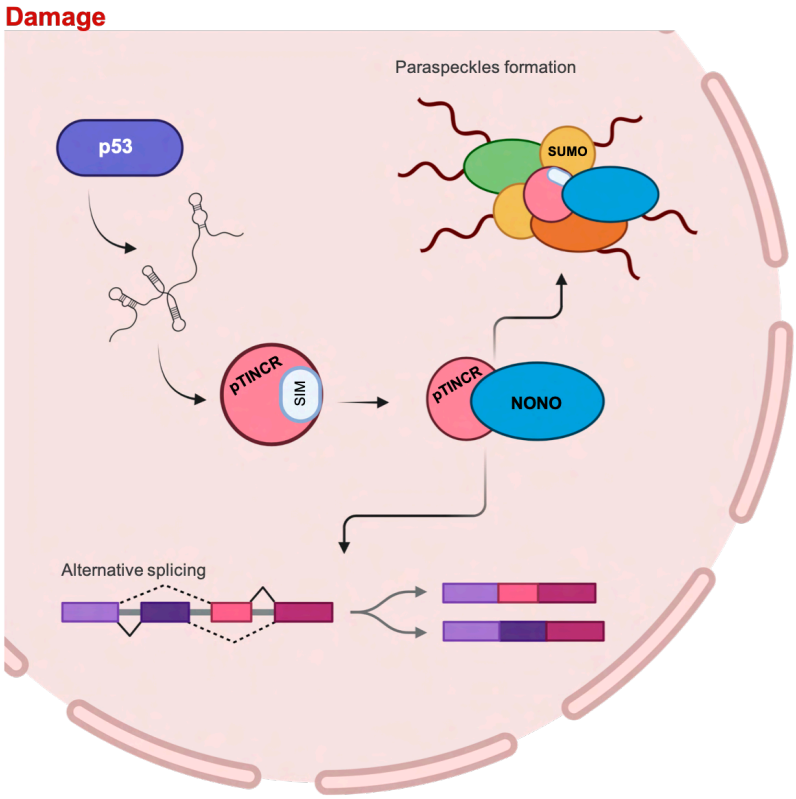


Figure 70. Proposed working model for pTINCR-NONO interaction.
pTINCR interaction with NONO enhances its SUMOylation. This event may trigger an epithelial splicing program and promote paraspeckles formation for the retention of defective RNA in response to cellular damage and p53 activation.

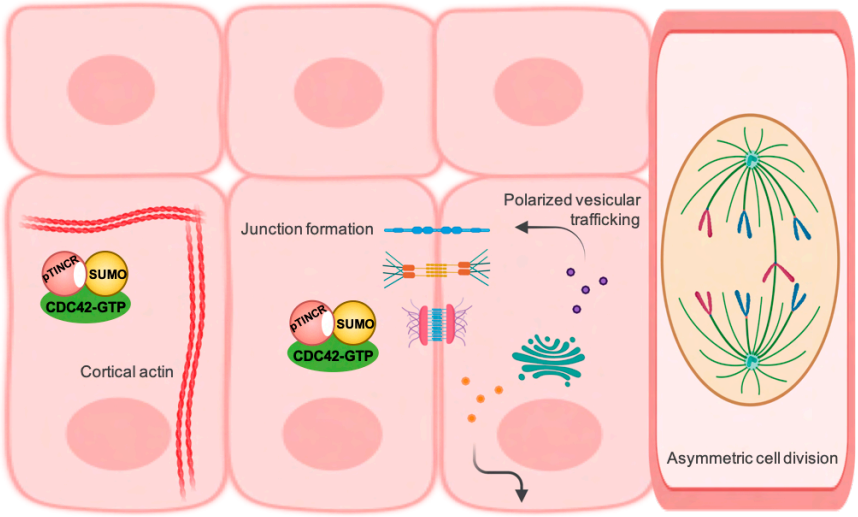


Figure 71. Proposed working model for pTINCR-CDC42 interaction.
pTINCR enhances CDC42 SUMOylation and promotes its activation, leading to a pro-differentiation cascade involving cortical actin assembly and intercellular junction formation and, hypothetically, polarized vesicular trafficking and mitotic spindle positioning during cell division.

Altogether, we hypothesize that pTINCR, as a new UBL, may bind and modulate the activity of several SUMOylated proteins to regulate epithelial differentiation and tumor suppression (**Fig. 72**). Thus, we believe that there are many more pTINCR substrates to be discovered. Indeed, the ubiquitous cellular localization of pTINCR reinforces the idea of a wide spectrum of partners and processes which pTINCR may regulate. As an instance, mitochondria dynamics or protein sorting. This work adds novel and valuable knowledge to the field of ubiquitin and ubiquitin-like proteins and suggests the possibility of specific UBLs playing roles in specific contexts such as pTINCR, which regulates multiple processes during epithelial differentiation.

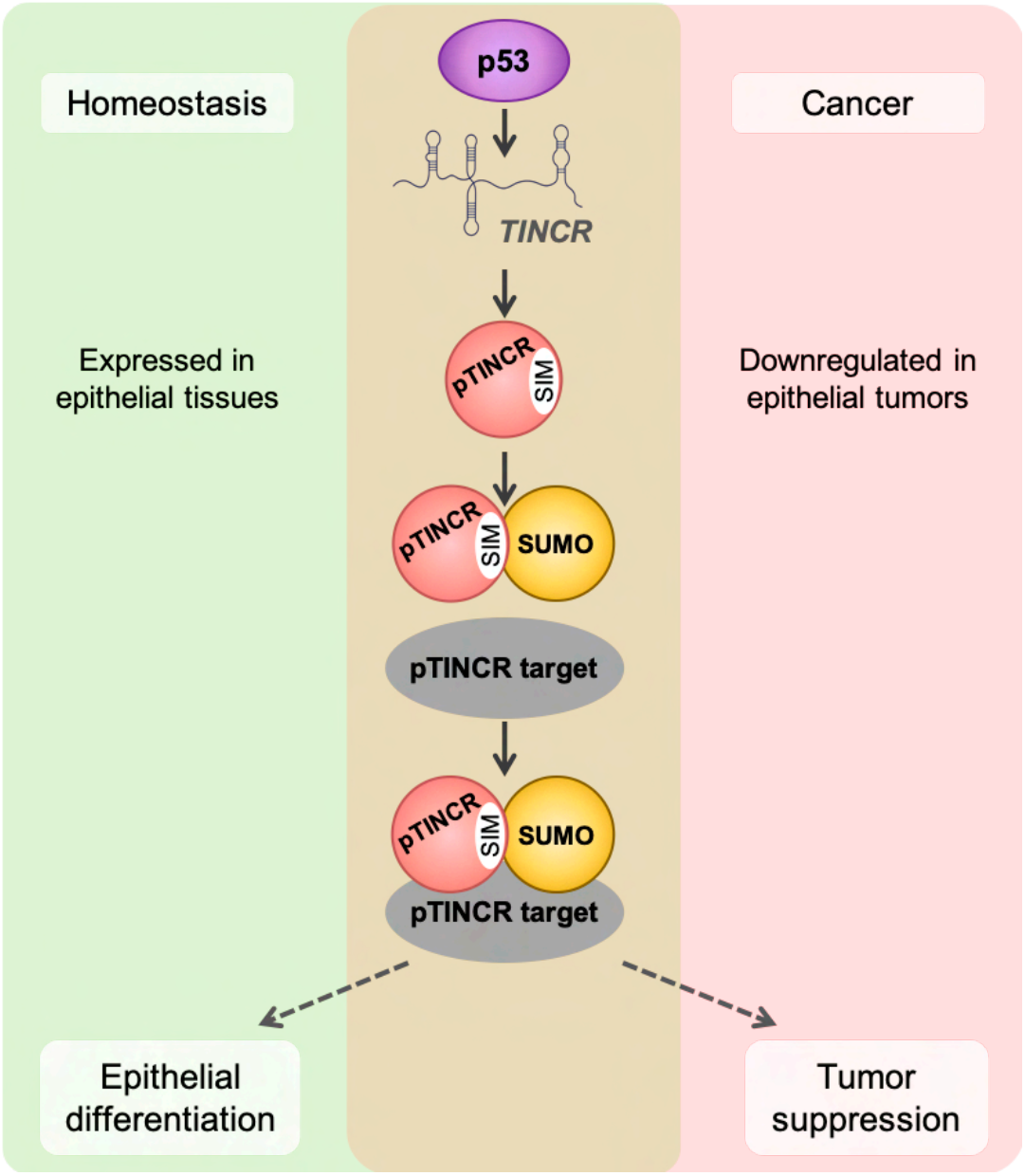


Figure 70. Proposed working model for pTINCR function.
pTINCR is a novel UBL microprotein regulated by p53 which is expressed in epithelial tissues and downregulated in cSCC. pTINCR interacts with SUMO through its SIM domain and modulates protein SUMOylation patterns. In our model, pTINCR modulates the SUMOylation of its interactors, triggering different molecular events leading to epithelial differentiation and to tumor suppression.

In conclusion, microproteins have emerged as a novel class of molecular regulators with functions in many important biological processes, in physiological and pathological contexts. With this study, we have uncovered the role of microproteins as regulators of cell identity. We believe that further exploring the microproteome will provide novel players critical for this process, bringing potential therapeutic perspectives in regenerative medicine and cancer.

CONCLUSIONS

1. *TINCR* lncRNA contains a sORF that encodes a highly conserved 87-amino acid microprotein that we have named pTINCR.

2. pTINCR is naturally expressed in human and mouse skin and in several other simple and stratified epithelia. At the subcellular level, pTINCR is found ubiquitously in the cell but enriched in the nucleus and in the cell membrane.

3. pTINCR expression is upregulated during epithelial differentiation and also upon genotoxic stress in a p53 dependent manner.

4. pTINCR overexpression promotes epithelial differentiation *in vitro* and epidermal differentiation *in vivo*.

5. pTINCR is essential for epithelial differentiation *in vitro.*

6. pTINCR has tumor suppressor activity in cSCC: it is downregulated in human cSCCs and its overexpression reduces the growth and malignant features of cSCC-PDXs.

7. *TINCR* expression correlates with good prognosis in several epithelial cancers, including bladder carcinoma, PDAC, stomach carcinoma, head and neck squamous cell carcinoma and lung adenocarcinoma.

8. pTINCR is a new Type II Ubiquitin-like protein that binds to SUMO through its SIM domain and modulates SUMO conjugation in a protein-specific manner.

9. pTINCR interacts with NONO, promotes its SUMOylation and triggers a specific splicing program associated with epithelial cell identity.

10. pTINCR interacts with CDC42, promotes its SUMOylation and its activation, leading to cytoskeleton remodeling and cell junctions formation, critical events of epithelial cell polarity and differentiation.

BIBLIOGRAPHY

1. Crick, F.H., *On protein synthesis.* Symp Soc Exp Biol, 1958. **12**: p. 138-63.

2. Baltimore, D., *Central dogma reversed.* Nature, 1970. **226**(5252): p. 1198-9.

3. Prusiner, S.B., *Novel proteinaceous infectious particles cause scrapie.* Science, 1982. **216**(4542): p. 136-44.

4. Consortium, E.P., *An integrated encyclopedia of DNA elements in the human genome.* Nature, 2012. **489**(7414): p. 57-74.

5. van de Sande, K., et al., *Modification of phytohormone response by a peptide encoded by ENOD40 of legumes and a nonlegume.* Science, 1996. **273**(5273): p. 370-3.

6. Basrai, M.A., P. Hieter, and J.D. Boeke, *Small open reading frames: beautiful needles in the haystack.* Genome Res, 1997. **7**(8): p. 768-71.

7. Fickett, J.W., *ORFs and genes: how strong a connection?* J Comput Biol, 1995. **2**(1): p. 117-23.

8. Andrews, S.J. and J.A. Rothnagel, *Emerging evidence for functional peptides encoded by short open reading frames.* Nat Rev Genet, 2014. **15**(3): p. 193-204.

9. Brunet, M.A., S. Leblanc, and X. Roucou, *Reconsidering proteomic diversity with functional investigation of small ORFs and alternative ORFs.* Exp Cell Res, 2020. **393**(1): p. 112057.

10. Makarewich, C.A. and E.N. Olson, *Mining for Micropeptides.* Trends Cell Biol, 2017. **27**(9): p. 685-696.

11. Patraquim, P., et al., *Developmental regulation of canonical and small ORF translation from mRNAs.* Genome Biol, 2020. **21**(1): p. 128.

12. Merino-Valverde, I., E. Greco, and M. Abad, *The microproteome of cancer: From invisibility to relevance.* Exp Cell Res, 2020. **392**(1): p. 111997.

13. Renz, P.F., F. Valdivia-Francia, and A. Sendoel, *Some like it translated: small ORFs in the 5'UTR.* Exp Cell Res, 2020. **396**(1): p. 112229.

14. Samandi, S., et al., *Deep transcriptome annotation enables the discovery and functional characterization of cryptic small proteins.* Elife, 2017. **6**.

15. Lin, M.F., I. Jungreis, and M. Kellis, *PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions.* Bioinformatics, 2011. **27**(13): p. i275-82.

16. Ruiz-Orera, J., et al., *Translation of neutrally evolving peptides provides a basis for de novo gene evolution.* Nat Ecol Evol, 2018. **2**(5): p. 890-896.

17. Cao, X. and S.A. Slavoff, *Non-AUG start codons: Expanding and regulating the small and alternative ORFeome.* Exp Cell Res, 2020. **391**(1): p. 111973.

18. Dunn, J.G., et al., *Ribosome profiling reveals pervasive and regulated stop codon readthrough in Drosophila melanogaster.* Elife, 2013. **2**: p. e01179.

19. Peeters, M.K.R. and G. Menschaert, *The hunt for sORFs: A multidisciplinary strategy.* Exp Cell Res, 2020. **391**(1): p. 111923.

20. Zhu, Y., et al., *Discovery of coding regions in the human genome by integrated proteogenomics analysis workflow.* Nat Commun, 2018. **9**(1): p. 903.

21. Zhang, Z., et al., *Proteomics-driven identification of short open reading frame-encoded peptides.* Proteomics, 2022: p. e2100312.

22. Slavoff, S.A., et al., *A human short open reading frame (sORF)-encoded polypeptide that stimulates DNA end joining.* J Biol Chem, 2014. **289**(16): p. 10950-10957.

23. Arnoult, N., et al., *Regulation of DNA repair pathway choice in S and G2 phases by the NHEJ inhibitor CYREN.* Nature, 2017. **549**(7673): p. 548-552.

24. D'Lima, N.G., et al., *A human microprotein that interacts with the mRNA decapping complex.* Nat Chem Biol, 2017. **13**(2): p. 174-180.

25. Huang, J.Z., et al., *A Peptide Encoded by a Putative lncRNA HOXB-AS3 Suppresses Colon Cancer Growth.* Mol Cell, 2017. **68**(1): p. 171-184 e6.

26. Anderson, D.M., et al., *A micropeptide encoded by a putative long noncoding RNA regulates muscle performance.* Cell, 2015. **160**(4): p. 595-606.

27. Nelson, B.R., et al., *A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle.* Science, 2016. **351**(6270): p. 271-5.

28. Anderson, D.M., et al., *Widespread control of calcium signaling by a family of SERCA-inhibiting micropeptides.* Sci Signal, 2016. **9**(457): p. ra119.

29. Senis, E., et al., *TUNAR lncRNA Encodes a Microprotein that Regulates Neural Differentiation and Neurite Formation by Modulating Calcium Dynamics.* Front Cell Dev Biol, 2021. **9**: p. 747667.

30. Stein, C.S., et al., *Mitoregulin: A lncRNA-Encoded Microprotein that Supports Mitochondrial Supercomplexes and Respiratory Efficiency.* Cell Rep, 2018. **23**(13): p. 3710-3720 e8.

31. Makarewich, C.A., et al., *MOXI Is a Mitochondrial Micropeptide That Enhances Fatty Acid beta-Oxidation.* Cell Rep, 2018. **23**(13): p. 3701-3709.

32. Friesen, M., et al., *Mitoregulin Controls beta-Oxidation in Human and Mouse Adipocytes.* Stem Cell Reports, 2020. **14**(4): p. 590-602.

33. Cloutier, P., et al., *Upstream ORF-Encoded ASDURF Is a Novel Prefoldin-like Subunit of the PAQosome.* J Proteome Res, 2020. **19**(1): p. 18-27.

34. Mise, S., et al., *Kastor and Polluks polypeptides encoded by a single gene locus cooperatively regulate VDAC and spermatogenesis.* Nat Commun, 2022. **13**(1): p. 1071.

35. Modzelewski, A.J., et al., *A mouse-specific retrotransposon drives a conserved Cdk2ap1 isoform essential for development.* Cell, 2021. **184**(22): p. 5541-5558 e22.

36. Khitun, A., T.J. Ness, and S.A. Slavoff, *Small open reading frames and cellular stress responses.* Mol Omics, 2019. **15**(2): p. 108-116.

37. Moro, S.G., et al., *Impact of uORFs in mediating regulation of translation in stress conditions.* BMC Mol Cell Biol, 2021. **22**(1): p. 29.

38. Starck, S.R., et al., *Translation from the 5' untranslated region shapes the integrated stress response.* Science, 2016. **351**(6272): p. aad3867.

39. Matsumoto, A., et al., *mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide.* Nature, 2017. **541**(7636): p. 228-232.

40. Chu, Q., et al., *Regulation of the ER stress response by a mitochondrial microprotein.* Nat Commun, 2019. **10**(1): p. 4883.

41. Hashimoto, Y., et al., *Mechanisms of neuroprotection by a novel rescue factor humanin from Swedish mutant amyloid precursor protein.* Biochem Biophys Res Commun, 2001. **283**(2): p. 460-8.

42. Zapala, B., et al., *Humanins, the neuroprotective and cytoprotective peptides with antiapoptotic and anti-inflammatory properties.* Pharmacol Rep, 2010. **62**(5): p. 767-77.

43. Yen, K., et al., *The mitochondrial derived peptide humanin is a regulator of lifespan and healthspan.* Aging (Albany NY), 2020. **12**(12): p. 11185-11199.

44. Bachar, A.R., et al., *Humanin is expressed in human vascular walls and has a cytoprotective effect against oxidized LDL-induced oxidative stress.* Cardiovasc Res, 2010. **88**(2): p. 360-6.

45. Matsuoka, M., *Protective effects of Humanin and calmodulin-like skin protein in Alzheimer's disease and broad range of abnormalities.* Mol Neurobiol, 2015. **51**(3): p. 1232-9.

46. Wek, R.C., *Role of eIF2alpha Kinases in Translational Control and Adaptation to Cellular Stress.* Cold Spring Harb Perspect Biol, 2018. **10**(7).

47. Andreev, D.E., et al., *Oxygen and glucose deprivation induces widespread alterations in mRNA translation within 20 minutes.* Genome Biol, 2015. **16**: p. 90.

48. Advani, V.M. and P. Ivanov, *Translational Control under Stress: Reshaping the Translatome.* Bioessays, 2019. **41**(5): p. e1900009.

49. Walters, B. and S.R. Thompson, *Cap-Independent Translational Control of Carcinogenesis.* Front Oncol, 2016. **6**: p. 128.

50. Razooky, B.S., et al., *Viral Infection Identifies Micropeptides Differentially Regulated in smORF-Containing lncRNAs.* Genes (Basel), 2017. **8**(8).

51. Jackson, R., et al., *The translation of non-canonical open reading frames controls mucosal immunity.* Nature, 2018. **564**(7736): p. 434-438.

52. Sendoel, A., et al., *Translation from unconventional 5' start sites drives tumour initiation.* Nature, 2017. **541**(7638): p. 494-499.

53. Zhu, S., et al., *Peptides/Proteins Encoded by Non-coding RNA: A Novel Resource Bank for Drug Targets and Biomarkers.* Front Pharmacol, 2018. **9**: p. 1295.

54. Hashimoto, Y., et al., *Humanin inhibits neuronal cell death by interacting with a cytokine receptor complex or complexes involving CNTF receptor alpha/WSX-1/gp130.* Mol Biol Cell, 2009. **20**(12): p. 2864-73.

55. Lightfoot, J.W., et al., *Small peptide-mediated self-recognition prevents cannibalism in predatory nematodes.* Science, 2019. **364**(6435): p. 86-89.

56. Yang, P., et al., *Elabela/Toddler Is an Endogenous Agonist of the Apelin APJ Receptor in the Adult Cardiovascular System, and Exogenous Administration of the Peptide Compensates for the Downregulation of Its Expression in Pulmonary Arterial Hypertension.* Circulation, 2017. **135**(12): p. 1160-1173.

57. Rheinbay, E., et al., *Analyses of non-coding somatic drivers in 2,658 cancer whole genomes.* Nature, 2020. **578**(7793): p. 102-111.

58. Hu, X., et al., *The role of long noncoding RNAs in cancer: the dark matter matters.* Curr Opin Genet Dev, 2018. **48**: p. 8-15.

59. Carbonnelle, D., et al., *The melanoma antigens MELOE-1 and MELOE-2 are translated from a bona fide polycistronic mRNA containing functional IRES sequences.* PLoS One, 2013. **8**(9): p. e75233.

60. Laumont, C.M., et al., *Noncoding regions are the main source of targetable tumor-specific antigens.* Sci Transl Med, 2018. **10**(470).

61. Chong, C., et al., *Integrated proteogenomic deep sequencing and analytics accurately identify non-canonical peptides in tumor immunopeptidomes.* Nat Commun, 2020. **11**(1): p. 1293.

62. Piroli, M.E., J.O. Blanchette, and E. Jabbarzadeh, *Polarity as a physiological modulator of cell function.* Front Biosci (Landmark Ed), 2019. **24**(3): p. 451-462.

63. Giepmans, B.N. and S.C. van Ijzendoorn, *Epithelial cell-cell junctions and plasma membrane domains.* Biochim Biophys Acta, 2009. **1788**(4): p. 820-31.

64. Garcia, M.A., W.J. Nelson, and N. Chavez, *Cell-Cell Junctions Organize Structural and Signaling Networks.* Cold Spring Harb Perspect Biol, 2018. **10**(4).

65. Otani, T. and M. Furuse, *Tight Junction Structure and Function Revisited.* Trends Cell Biol, 2020. **30**(10): p. 805-817.

66. Zihni, C., et al., *Tight junctions: from simple barriers to multifunctional molecular gates.* Nat Rev Mol Cell Biol, 2016. **17**(9): p. 564-80.

67. Hartsock, A. and W.J. Nelson, *Adherens and tight junctions: structure, function and connections to the actin cytoskeleton.* Biochim Biophys Acta, 2008. **1778**(3): p. 660-9.

68. Epifano, C. and M. Perez-Moreno, *Crossroads of integrins and cadherins in epithelia and stroma remodeling.* Cell Adh Migr, 2012. **6**(3): p. 261-73.

69. Garrod, D. and M. Chidgey, *Desmosome structure, composition and function.* Biochim Biophys Acta, 2008. **1778**(3): p. 572-87.

70. Burdick, D., A.H. Le Gall, and E. Rodriguez-Boulan, *Vesicular transport: implications for cell polarity.* Biocell, 1996. **20**(3): p. 343-53.

71. Ravichandran, Y., B. Goud, and J.B. Manneville, *The Golgi apparatus and cell polarity: Roles of the cytoskeleton, the Golgi matrix, and Golgi membranes.* Curr Opin Cell Biol, 2020. **62**: p. 104-113.

72. Raman, R., C.S. Pinto, and M. Sonawane, *Polarized Organization of the Cytoskeleton: Regulation by Cell Polarity Proteins.* J Mol Biol, 2018. **430**(19): p. 3565-3584.

73. Vaezi, A., et al., *Actin cable dynamics and Rho/Rock orchestrate a polarized cytoskeletal architecture in the early steps of assembling a stratified epithelium.* Dev Cell, 2002. **3**(3): p. 367-81.

74. Nigam, S.K., E. Rodriguez-Boulan, and R.B. Silver, *Changes in intracellular calcium during the development of epithelial polarity and junctions.* Proc Natl Acad Sci U S A, 1992. **89**(13): p. 6162-6.

75. Perez-Moreno, M., et al., *Vinculin but not alpha-actinin is a target of PKC phosphorylation during junctional assembly induced by calcium.* J Cell Sci, 1998. **111 (Pt 23)**: p. 3563-71.

76. Brown, R.C. and T.P. Davis, *Calcium modulation of adherens and tight junction function: a potential mechanism for blood-brain barrier disruption after stroke.* Stroke, 2002. **33**(6): p. 1706-11.

77. Bandyopadhyay, B.C., et al., *Apical localization of a functional TRPC3/TRPC6-Ca2+-signaling complex in polarized epithelial cells. Role in apical Ca2+ influx.* J Biol Chem, 2005. **280**(13): p. 12908-16.

78. Panousopoulou, E. and J.B. Green, *Spindle orientation processes in epithelial growth and organisation.* Semin Cell Dev Biol, 2014. **34**: p. 124-32.

79. Bergstralh, D.T. and D. St Johnston, *Spindle orientation: what if it goes wrong?* Semin Cell Dev Biol, 2014. **34**: p. 140-5.

80. Bergstralh, D.T., T. Haack, and D. St Johnston, *Epithelial polarity and spindle orientation: intersecting pathways.* Philos Trans R Soc Lond B Biol Sci, 2013. **368**(1629): p. 20130291.

81. Seldin, L. and I. Macara, *Epithelial spindle orientation diversities and uncertainties: recent developments and lingering questions.* F1000Res, 2017. **6**: p. 984.

82. Hodge, R.G. and A.J. Ridley, *Regulating Rho GTPases and their regulators.* Nat Rev Mol Cell Biol, 2016. **17**(8): p. 496-510.

83. Melendez, J., M. Grogg, and Y. Zheng, *Signaling role of Cdc42 in regulating mammalian physiology.* J Biol Chem, 2011. **286**(4): p. 2375-81.

84. Bishop, A.L. and A. Hall, *Rho GTPases and their effector proteins.* Biochem J, 2000. **348 Pt 2**: p. 241-55.

85. Watson, J.R., D. Owen, and H.R. Mott, *Cdc42 in actin dynamics: An ordered pathway governed by complex equilibria and directional effector handover.* Small GTPases, 2017. **8**(4): p. 237-244.

86. Ma, L., R. Rohatgi, and M.W. Kirschner, *The Arp2/3 complex mediates actin polymerization induced by the small GTP-binding protein Cdc42.* Proc Natl Acad Sci U S A, 1998. **95**(26): p. 15362-7.

87. Rohatgi, R., et al., *The interaction between N-WASP and the Arp2/3 complex links Cdc42-dependent signals to actin assembly.* Cell, 1999. **97**(2): p. 221-31.

88. Kim, S.H., Z. Li, and D.B. Sacks, *E-cadherin-mediated cell-cell attachment activates Cdc42.* J Biol Chem, 2000. **275**(47): p. 36999-7005.

89. Kim, S.K., *Cell polarity: new PARtners for Cdc42 and Rac.* Nat Cell Biol, 2000. **2**(8): p. E143-5.

90. Nunes de Almeida, F., et al., *Cdc42 defines apical identity and regulates epithelial morphogenesis by promoting apical recruitment of Par6-aPKC and Crumbs.* Development, 2019. **146**(15).

91. Etienne-Manneville, S., *Cdc42--the centre of polarity.* J Cell Sci, 2004. **117**(Pt 8): p. 1291-300.

92. Baum, B. and M. Georgiou, *Dynamics of adherens junctions in epithelial establishment, maintenance, and remodeling.* J Cell Biol, 2011. **192**(6): p. 907-17.

93. Gopalakrishnan, S., et al., *aPKC-PAR complex dysfunction and tight junction disassembly in renal epithelial cells during ATP depletion.* Am J Physiol Cell Physiol, 2007. **292**(3): p. C1094-102.

94. Assemat, E., et al., *Polarity complex proteins.* Biochim Biophys Acta, 2008. **1778**(3): p. 614-30.

95. Firth, N.A. and P.C. Reade, *Comparison of eosinophil densities in oral mucosal lichen planus and lichenoid drug reactions.* J Oral Pathol Med, 1990. **19**(2): p. 86-8.

96. Wallace, S.W., et al., *Cdc42 regulates apical junction formation in human bronchial epithelial cells through PAK4 and Par6B.* Mol Biol Cell, 2010. **21**(17): p. 2996-3006.

97. Fukuhara, T., et al., *Activation of Cdc42 by trans interactions of the cell adhesion molecules nectins through c-Src and Cdc42-GEF FRG.* J Cell Biol, 2004. **166**(3): p. 393-405.

98. Otani, T., et al., *Cdc42 GEF Tuba regulates the junctional configuration of simple epithelial cells.* J Cell Biol, 2006. **175**(1): p. 135-46.

99. Broman, M.T., et al., *Cdc42 regulates adherens junction stability and endothelial permeability by inducing alpha-catenin interaction with the vascular endothelial cadherin complex.* Circ Res, 2006. **98**(1): p. 73-80.

100. Wu, X., et al., *Cdc42 is crucial for the establishment of epithelial polarity during early mammalian development.* Dev Dyn, 2007. **236**(10): p. 2767-78.

101. Elbediwy, A., et al., *Epithelial junction formation requires confinement of Cdc42 activity by a novel SH3BP1 complex.* J Cell Biol, 2012. **198**(4): p. 677-93.

102. Rojas, R., et al., *Cdc42-dependent modulation of tight junctions and membrane protein traffic in polarized Madin-Darby canine kidney cells.* Mol Biol Cell, 2001. **12**(8): p. 2257-74.

103. Bruewer, M., et al., *RhoA, Rac1, and Cdc42 exert distinct effects on epithelial barrier via selective structural and biochemical modulation of junctional proteins and F-actin.* Am J Physiol Cell Physiol, 2004. **287**(2): p. C327-35.

104. Cerione, R.A., *Cdc42: new roads to travel.* Trends Cell Biol, 2004. **14**(3): p. 127-32.

105. Goldstein, B. and I.G. Macara, *The PAR proteins: fundamental players in animal cell polarization.* Dev Cell, 2007. **13**(5): p. 609-622.

106. Suzuki, A. and S. Ohno, *The PAR-aPKC system: lessons in polarity.* J Cell Sci, 2006. **119**(Pt 6): p. 979-87.

107. Musch, A., et al., *cdc42 regulates the exit of apical and basolateral proteins from the trans-Golgi network.* EMBO J, 2001. **20**(9): p. 2171-9.

108. Kroschewski, R., A. Hall, and I. Mellman, *Cdc42 controls secretory and endocytic transport to the basolateral plasma membrane of MDCK cells.* Nat Cell Biol, 1999. **1**(1): p. 8-13.

109. Cohen, D., A. Musch, and E. Rodriguez-Boulan, *Selective control of basolateral membrane protein polarity by cdc42.* Traffic, 2001. **2**(8): p. 556-64.

110. Trifaro, J.M., S. Gasman, and L.M. Gutierrez, *Cytoskeletal control of vesicle transport and exocytosis in chromaffin cells.* Acta Physiol (Oxf), 2008. **192**(2): p. 165-72.

111. Matas, O.B., J.A. Martinez-Menarguez, and G. Egea, *Association of Cdc42/N-WASP/Arp2/3 signaling pathway with Golgi membranes.* Traffic, 2004. **5**(11): p. 838-46.

112. Luna, A., et al., *Regulation of protein transport from the Golgi complex to the endoplasmic reticulum by CDC42 and N-WASP.* Mol Biol Cell, 2002. **13**(3): p. 866-79.

113. Fucini, R.V., et al., *Golgi vesicle proteins are linked to the assembly of an actin complex defined by mAbp1.* Mol Biol Cell, 2002. **13**(2): p. 621-31.

114. Yasuda, S., et al., *Cdc42 and mDia3 regulate microtubule attachment to kinetochores.* Nature, 2004. **428**(6984): p. 767-71.

115. Mitsushima, M., F. Toyoshima, and E. Nishida, *Dual role of Cdc42 in spindle orientation control of adherent cells.* Mol Cell Biol, 2009. **29**(10): p. 2816-27.

116. Gotta, M., M.C. Abraham, and J. Ahringer, *CDC-42 controls early cell polarity and spindle orientation in C. elegans.* Curr Biol, 2001. **11**(7): p. 482-8.

117. Ma, C., et al., *Cdc42 activation couples spindle positioning to first polar body formation in oocyte maturation.* Curr Biol, 2006. **16**(2): p. 214-20.

118. Na, J. and M. Zernicka-Goetz, *Asymmetric positioning and organization of the meiotic spindle of mouse oocytes requires CDC42 function.* Curr Biol, 2006. **16**(12): p. 1249-54.

119. Jaffe, A.B., et al., *Cdc42 controls spindle orientation to position the apical surface during epithelial morphogenesis.* J Cell Biol, 2008. **183**(4): p. 625-33.

120. Durgan, J., et al., *Par6B and atypical PKC regulate mitotic spindle orientation during epithelial morphogenesis.* J Biol Chem, 2011. **286**(14): p. 12461-74.

121. Simpson, C.L., D.M. Patel, and K.J. Green, *Deconstructing the skin: cytoarchitectural determinants of epidermal morphogenesis.* Nat Rev Mol Cell Biol, 2011. **12**(9): p. 565-80.

122. Gutowska-Owsiak, D., et al., *Orchestrated control of filaggrin-actin scaffolds underpins cornification.* Cell Death Dis, 2018. **9**(4): p. 412.

123. Gdula, M.R., et al., *Remodeling of three-dimensional organization of the nucleus during terminal keratinocyte differentiation in the epidermis.* J Invest Dermatol, 2013. **133**(9): p. 2191-201.

124. Rogerson, C., D. Bergamaschi, and R.F.L. O'Shaughnessy, *Uncovering mechanisms of nuclear degradation in keratinocytes: A paradigm for nuclear degradation in other tissues.* Nucleus, 2018. **9**(1): p. 56-64.

125. Ipponjima, S., et al., *Live imaging of alterations in cellular morphology and organelles during cornification using an epidermal equivalent model.* Sci Rep, 2020. **10**(1): p. 5515.

126. Debrand-Passard, A., et al., *Thrombogenicity of dialyzer membranes as assessed by residual blood volume and surface morphology at different heparin dosages.* Contrib Nephrol, 1989. **74**: p. 2-9.

127. Kubler, M.D., et al., *Changes in the abundance and distribution of actin and associated proteins during terminal differentiation of human epidermal keratinocytes.* J Cell Sci, 1991. **100 (Pt 1)**: p. 153-65.

128. Perez-Moreno, M., C. Jamora, and E. Fuchs, *Sticky business: orchestrating cellular signals at adherens junctions.* Cell, 2003. **112**(4): p. 535-48.

129. Forni, M.F., M. Trombetta-Lima, and M.C. Sogayar, *Stem cells in embryonic skin development.* Biol Res, 2012. **45**(3): p. 215-22.

130. Fuchs, E. and H. Green, *Changes in keratin gene expression during terminal differentiation of the keratinocyte.* Cell, 1980. **19**(4): p. 1033-42.

131. Eckert, R.L. and E.A. Rorke, *Molecular biology of keratinocyte differentiation.* Environ Health Perspect, 1989. **80**: p. 109-16.

132. Castilho, R.M., et al., *Requirement of Rac1 distinguishes follicular from interfollicular epithelial stem cells.* Oncogene, 2007. **26**(35): p. 5078-85.

133. McMullan, R., et al., *Keratinocyte differentiation is regulated by the Rho and ROCK signaling pathway.* Curr Biol, 2003. **13**(24): p. 2185-9.

134. Perez-Moreno, M., et al., *Loss of p120 catenin and links to mitotic alterations, inflammation, and skin cancer.* Proc Natl Acad Sci U S A, 2008. **105**(40): p. 15399-404.

135. Wu, X., et al., *Cdc42 controls progenitor cell differentiation and beta-catenin turnover in skin.* Genes Dev, 2006. **20**(5): p. 571-85.

136. Wu, X., F. Quondamatteo, and C. Brakebusch, *Cdc42 expression in keratinocytes is required for the maintenance of the basement membrane in skin.* Matrix Biol, 2006. **25**(8): p. 466-74.

137. Zhang, M., et al., *Cdc42 Deficiency Leads To Epidermal Barrier Dysfunction by Regulating Intercellular Junctions and Keratinization of Epidermal Cells during Mouse Skin Development.* Theranostics, 2019. **9**(17): p. 5065-5084.

138. Page, A., et al., *Protective role of p53 in skin cancer: Carcinogenesis studies in mice lacking epidermal p53.* Oncotarget, 2016. **7**(15): p. 20902-18.

139. Shetty, S.K., et al., *Regulation of airway and alveolar epithelial cell apoptosis by p53-Induced plasminogen activator inhibitor-1 during cigarette smoke exposure injury.* Am J Respir Cell Mol Biol, 2012. **47**(4): p. 474-83.

140. Siganaki, M., et al., *Deregulation of apoptosis mediators' p53 and bcl2 in lung tissue of COPD patients.* Respir Res, 2010. **11**: p. 46.

141. Merritt, A.J., et al., *The role of p53 in spontaneous and radiation-induced apoptosis in the gastrointestinal tract of normal and p53-deficient mice.* Cancer Res, 1994. **54**(3): p. 614-7.

142. Clarke, A.R., et al., *p53 dependence of early apoptotic and proliferative responses within the mouse intestinal epithelium following gamma-irradiation.* Oncogene, 1994. **9**(6): p. 1767-73.

143. Einspahr, J.G., et al., *Relationship of p53 mutations to epidermal cell proliferation and apoptosis in human UV-induced skin carcinogenesis.* Neoplasia, 1999. **1**(5): p. 468-75.

144. Loureiro, J.B., et al., *P53 in skin cancer: From a master player to a privileged target for prevention and therapy.* Biochim Biophys Acta Rev Cancer, 2020. **1874**(2): p. 188438.

145. Campisi, J., *Senescent cells, tumor suppression, and organismal aging: good citizens, bad neighbors.* Cell, 2005. **120**(4): p. 513-22.

146. Burton, D.G. and V. Krizhanovsky, *Physiological and pathological consequences of cellular senescence.* Cell Mol Life Sci, 2014. **71**(22): p. 4373-86.

147. Campisi, J., *Aging, cellular senescence, and cancer.* Annu Rev Physiol, 2013. **75**: p. 685-705.

148. Munoz-Espin, D. and M. Serrano, *Cellular senescence: from physiology to pathology.* Nat Rev Mol Cell Biol, 2014. **15**(7): p. 482-96.

149. Naylor, R.M., D.J. Baker, and J.M. van Deursen, *Senescent cells: a novel therapeutic target for aging and age-related diseases.* Clin Pharmacol Ther, 2013. **93**(1): p. 105-16.

150. Acosta, J.C. and J. Gil, *Senescence: a new weapon for cancer therapy.* Trends Cell Biol, 2012. **22**(4): p. 211-9.

151. Campisi, J. and F. d'Adda di Fagagna, *Cellular senescence: when bad things happen to good cells.* Nat Rev Mol Cell Biol, 2007. **8**(9): p. 729-40.

152. Demaria, M., et al., *An essential role for senescent cells in optimal wound healing through secretion of PDGF-AA.* Dev Cell, 2014. **31**(6): p. 722-33.

153. Takeuchi, S., et al., *Intrinsic cooperation between p16INK4a and p21Waf1/Cip1 in the onset of cellular senescence and tumor suppression in vivo.* Cancer Res, 2010. **70**(22): p. 9381-90.

154. Waaijer, M.E., et al., *The number of p16INK4a positive cells in human skin reflects biological age.* Aging Cell, 2012. **11**(4): p. 722-5.

155. Adamus, J., et al., *p16INK4A influences the aging phenotype in the living skin equivalent.* J Invest Dermatol, 2014. **134**(4): p. 1131-1133.

156. Santos, M.A., et al., *DNA-damage-induced differentiation of leukaemic cells as an anti-cancer barrier.* Nature, 2014. **514**(7520): p. 107-11.

157. Mandal, P.K. and D.J. Rossi, *DNA-damage-induced differentiation in hematopoietic stem cells.* Cell, 2012. **148**(5): p. 847-8.

158. Wang, J., et al., *A differentiation checkpoint limits hematopoietic stem cell self-renewal in response to DNA damage.* Cell, 2012. **148**(5): p. 1001-14.

159. Sherman, M.H., C.H. Bassing, and M.A. Teitell, *Regulation of cell differentiation by the DNA damage response.* Trends Cell Biol, 2011. **21**(5): p. 312-9.

160. Puri, P.L., et al., *A myogenic differentiation checkpoint activated by genotoxic stress.* Nat Genet, 2002. **32**(4): p. 585-93.

161. Ying, Z., M. Sandoval, and S. Beronja, *Oncogenic activation of PI3K induces progenitor cell differentiation to suppress epidermal growth.* Nat Cell Biol, 2018. **20**(11): p. 1256-1266.

162. de Pedro, I., et al., *Sublethal UV irradiation induces squamous differentiation via a p53-independent, DNA damage-mitosis checkpoint.* Cell Death Dis, 2018. **9**(11): p. 1094.

163. Molinuevo, R., et al., *The DNA damage response links human squamous proliferation with differentiation.* J Cell Biol, 2020. **219**(11).

164. Sanz-Gomez, N., et al., *Squamous differentiation requires G2/mitosis slippage to avoid apoptosis.* Cell Death Differ, 2020. **27**(8): p. 2451-2467.

165. Gandarillas, A., *The mysterious human epidermal cell cycle, or an oncogene-induced differentiation checkpoint.* Cell Cycle, 2012. **11**(24): p. 4507-16.

166. Guinea-Viniegra, J., et al., *Differentiation-induced skin cancer suppression by FOS, p53, and TACE/ADAM17.* J Clin Invest, 2012. **122**(8): p. 2898-910.

167. Feinstein, E., et al., *Expression of the normal p53 gene induces differentiation of K562 cells.* Oncogene, 1992. **7**(9): p. 1853-7.

168. Campos, M.A., J.M. Lopes, and P. Soares, *The genetics of cutaneous squamous cell carcinogenesis.* Eur J Dermatol, 2018. **28**(5): p. 597-605.

169. Fania, L., et al., *Cutaneous Squamous Cell Carcinoma: From Pathophysiology to Novel Therapeutic Approaches.* Biomedicines, 2021. **9**(2).

170. Jemal, A., et al., *Cancer statistics, 2008.* CA Cancer J Clin, 2008. **58**(2): p. 71-96.

171. Combalia, A. and C. Carrera, *Squamous Cell Carcinoma: An Update on Diagnosis and Treatment.* Dermatol Pract Concept, 2020. **10**(3): p. e2020066.

172. Daley, G.Q., *Common themes of dedifferentiation in somatic cell reprogramming and cancer.* Cold Spring Harb Symp Quant Biol, 2008. **73**: p. 171-4.

173. Friedmann-Morvinski, D. and I.M. Verma, *Dedifferentiation and reprogramming: origins of cancer stem cells.* EMBO Rep, 2014. **15**(3): p. 244-53.

174. Yamada, Y., H. Haga, and Y. Yamada, *Concise review: dedifferentiation meets cancer development: proof of concept for epigenetic cancer.* Stem Cells Transl Med, 2014. **3**(10): p. 1182-7.

175. Li, L. and W. Li, *Epithelial-mesenchymal transition in human cancer: comprehensive reprogramming of metabolism, epigenetics, and differentiation.* Pharmacol Ther, 2015. **150**: p. 33-46.

176. Cao, Y., *Tumorigenesis as a process of gradual loss of original cell identity and gain of properties of neural precursor/progenitor cells.* Cell Biosci, 2017. **7**: p. 61.

177. Hanahan, D., *Hallmarks of Cancer: New Dimensions.* Cancer Discov, 2022. **12**(1): p. 31-46.

178. Bruner, H.C. and P.W.B. Derksen, *Loss of E-Cadherin-Dependent Cell-Cell Adhesion and the Development and Progression of Cancer.* Cold Spring Harb Perspect Biol, 2018. **10**(3).

179. Martin-Belmonte, F. and M. Perez-Moreno, *Epithelial cell polarity, stem cells and cancer.* Nat Rev Cancer, 2011. **12**(1): p. 23-38.

180. Cicalese, A., et al., *The tumor suppressor p53 regulates polarity of self-renewing divisions in mammary stem cells.* Cell, 2009. **138**(6): p. 1083-95.

181. Tao, L., et al., *Repression of mammary stem/progenitor cells by p53 is mediated by Notch and separable from apoptotic activity.* Stem Cells, 2011. **29**(1): p. 119-27.

182. Blanpain, C., et al., *Canonical notch signaling functions as a commitment switch in the epidermal lineage.* Genes Dev, 2006. **20**(21): p. 3022-35.

183. Rangarajan, A., et al., *Notch signaling is a direct determinant of keratinocyte growth arrest and entry into differentiation.* EMBO J, 2001. **20**(13): p. 3427-36.

184. Pickering, C.R., et al., *Mutational landscape of aggressive cutaneous squamous cell carcinoma.* Clin Cancer Res, 2014. **20**(24): p. 6582-92.

185. South, A.P., et al., *NOTCH1 mutations occur early during cutaneous squamous cell carcinogenesis.* J Invest Dermatol, 2014. **134**(10): p. 2630-2638.

186. Wang, N.J., et al., *Loss-of-function mutations in Notch receptors in cutaneous and lung squamous cell carcinoma.* Proc Natl Acad Sci U S A, 2011. **108**(43): p. 17761-6.

187. Nicolas, M., et al., *Notch1 functions as a tumor suppressor in mouse skin.* Nat Genet, 2003. **33**(3): p. 416-21.

188. Melino, G., et al., *Maintaining epithelial stemness with p63.* Sci Signal, 2015. **8**(387): p. re9.

189. Blanpain, C. and E. Fuchs, *p63: revving up epithelial stem-cell potential.* Nat Cell Biol, 2007. **9**(7): p. 731-3.

190. Dotto, G.P., *Notch tumor suppressor function.* Oncogene, 2008. **27**(38): p. 5115-23.

191. Chen, Y. and G.L. Sen, *SOX2 expression inhibits terminal epidermal differentiation.* Exp Dermatol, 2015. **24**(12): p. 974-6.

192. Royer, C. and X. Lu, *Epithelial cell polarity: a major gatekeeper against cancer?* Cell Death Differ, 2011. **18**(9): p. 1470-7.

193. Lee, M. and V. Vasioukhin, *Cell polarity and cancer--cell and tissue polarity as a non-canonical tumor suppressor.* J Cell Sci, 2008. **121**(Pt 8): p. 1141-50.

194. Muthuswamy, S.K. and B. Xue, *Cell polarity as a regulator of cancer cell behavior plasticity.* Annu Rev Cell Dev Biol, 2012. **28**: p. 599-625.

195. Coradini, D., C. Casarsa, and S. Oriana, *Epithelial cell polarity and tumorigenesis: new perspectives for cancer detection and treatment.* Acta Pharmacol Sin, 2011. **32**(5): p. 552-64.

196. Brembeck, F.H., M. Rosario, and W. Birchmeier, *Balancing cell adhesion and Wnt signaling, the key role of beta-catenin.* Curr Opin Genet Dev, 2006. **16**(1): p. 51-9.

197. Huber, M.A., N. Kraut, and H. Beug, *Molecular requirements for epithelial-mesenchymal transition during tumor progression.* Curr Opin Cell Biol, 2005. **17**(5): p. 548-58.

198. Marshall, T.W., et al., *The tumor suppressor adenomatous polyposis coli controls the direction in which a cell extrudes from an epithelium.* Mol Biol Cell, 2011. **22**(21): p. 3962-70.

199. Hogan, C., et al., *Characterization of the interface between normal and transformed epithelial cells.* Nat Cell Biol, 2009. **11**(4): p. 460-7.

200. Tanos, B. and E. Rodriguez-Boulan, *The epithelial polarity program: machineries involved and their hijacking by cancer.* Oncogene, 2008. **27**(55): p. 6939-57.

201. Desai, R.A., et al., *Cell polarity triggered by cell-cell adhesion via E-cadherin.* J Cell Sci, 2009. **122**(Pt 7): p. 905-11.

202. Zhou, P.J., et al., *Loss of Par3 promotes prostatic tumorigenesis by enhancing cell growth and changing cell division modes.* Oncogene, 2019. **38**(12): p. 2192-2205.

203. Saito, Y., R.R. Desai, and S.K. Muthuswamy, *Reinterpreting polarity and cancer: The changing landscape from tumor suppression to tumor promotion.* Biochim Biophys Acta Rev Cancer, 2018. **1869**(2): p. 103-116.

204. Mescher, M., et al., *The epidermal polarity protein Par3 is a non-cell autonomous suppressor of malignant melanoma.* J Exp Med, 2017. **214**(2): p. 339-358.

205. Archibald, A., et al., *Atypical protein kinase C induces cell transformation by disrupting Hippo/Yap signaling.* Mol Biol Cell, 2015. **26**(20): p. 3578-95.

206. Paul, A., et al., *PKCzeta Promotes Breast Cancer Invasion by Regulating Expression of E-cadherin and Zonula Occludens-1 (ZO-1) via NFkappaB-p65.* Sci Rep, 2015. **5**: p. 12520.

207. Xue, B., et al., *Loss of Par3 promotes breast cancer metastasis by compromising cell-cell cohesion.* Nat Cell Biol, 2013. **15**(2): p. 189-200.

208. Haga, R.B. and A.J. Ridley, *Rho GTPases: Regulation and roles in cancer cell biology.* Small GTPases, 2016. **7**(4): p. 207-221.

209. Zheng, Y., S. Bagrodia, and R.A. Cerione, *Activation of phosphoinositide 3-kinase activity by Cdc42Hs binding to p85.* J Biol Chem, 1994. **269**(29): p. 18727-30.

210. Mizukawa, B., et al., *The cell polarity determinant CDC42 controls division symmetry to block leukemia cell differentiation.* Blood, 2017. **130**(11): p. 1336-1346.

211. Zhang, Z., et al., *CDC42 controlled apical-basal polarity regulates intestinal stem cell to transit amplifying cell fate transition via YAP-EGF-mTOR signaling.* Cell Rep, 2022. **38**(2): p. 110009.

212. Li, W., H. Chong, and K.L. Guan, *Function of the Rho family GTPases in Ras-stimulated Raf activation.* J Biol Chem, 2001. **276**(37): p. 34728-37.

213. Ozdamar, B., et al., *Regulation of the polarity protein Par6 by TGFbeta receptors controls epithelial cell plasticity.* Science, 2005. **307**(5715): p. 1603-9.

214. Maldonado, M.D.M. and S. Dharmawardhane, *Targeting Rac and Cdc42 GTPases in Cancer.* Cancer Res, 2018. **78**(12): p. 3101-3111.

215. Ye, H., et al., *Cdc42 expression in cervical cancer and its effects on cervical tumor invasion and migration.* Int J Oncol, 2015. **46**(2): p. 757-63.

216. Chen, Q.Y., et al., *Expression analysis of Cdc42 in lung cancer and modulation of its expression by curcumin in lung cancer cell lines.* Int J Oncol, 2012. **40**(5): p. 1561-8.

217. Kamai, T., et al., *Overexpression of RhoA, Rac1, and Cdc42 GTPases is associated with progression in testicular cancer.* Clin Cancer Res, 2004. **10**(14): p. 4799-805.

218. Pichot, C.S., et al., *Cdc42-interacting protein 4 promotes breast cancer cell invasion and formation of invadopodia through activation of N-WASp.* Cancer Res, 2010. **70**(21): p. 8347-56.

219. Sakurai-Yageta, M., et al., *The interaction of IQGAP1 with the exocyst complex is required for tumor cell invasion downstream of Cdc42 and RhoA.* J Cell Biol, 2008. **181**(6): p. 985-98.

220. Wilkinson, S., H.F. Paterson, and C.J. Marshall, *Cdc42-MRCK and Rho-ROCK signalling cooperate in myosin phosphorylation and cell invasion.* Nat Cell Biol, 2005. **7**(3): p. 255-61.

221. Bouzahzah, B., et al., *Rho family GTPases regulate mammary epithelium cell growth and metastasis through distinguishable pathways.* Mol Med, 2001. **7**(12): p. 816-30.

222. Stengel, K.R. and Y. Zheng, *Essential role of Cdc42 in Ras-induced transformation revealed by gene targeting.* PLoS One, 2012. **7**(6): p. e37317.

223. Wang, J.B., W.J. Wu, and R.A. Cerione, *Cdc42 and Ras cooperate to mediate cellular transformation by intersectin-L.* J Biol Chem, 2005. **280**(24): p. 22883-91.

224. Qiu, R.G., et al., *Cdc42 regulates anchorage-independent growth and is necessary for Ras transformation.* Mol Cell Biol, 1997. **17**(6): p. 3449-58.

225. Walters, K.J., et al., *Ubiquitin family proteins and their relationship to the proteasome: a structural perspective.* Biochim Biophys Acta, 2004. **1695**(1-3): p. 73-87.

226. Cappadocia, L. and C.D. Lima, *Ubiquitin-like Protein Conjugation: Structures, Chemistry, and Mechanism.* Chem Rev, 2018. **118**(3): p. 889-918.

227. Hochstrasser, M., *Origin and function of ubiquitin-like proteins.* Nature, 2009. **458**(7237): p. 422-9.

228. Kerscher, O., R. Felberbaum, and M. Hochstrasser, *Modification of proteins by ubiquitin and ubiquitin-like proteins.* Annu Rev Cell Dev Biol, 2006. **22**: p. 159-80.

229. Sahin, U., H. de The, and V. Lallemand-Breitenbach, *Sumoylation in Physiology, Pathology and Therapy.* Cells, 2022. **11**(5).

230. Celen, A.B. and U. Sahin, *Sumoylation on its 25th anniversary: mechanisms, pathology, and emerging concepts.* FEBS J, 2020. **287**(15): p. 3110-3140.

231. Vertegaal, A.C., et al., *A proteomic study of SUMO-2 target proteins.* J Biol Chem, 2004. **279**(32): p. 33791-8.

232. Rosas-Acosta, G., et al., *A universal strategy for proteomic studies of SUMO and other ubiquitin-like modifiers.* Mol Cell Proteomics, 2005. **4**(1): p. 56-72.

233. Zhu, J., et al., *Small ubiquitin-related modifier (SUMO) binding determines substrate recognition and paralog-selective SUMO modification.* J Biol Chem, 2008. **283**(43): p. 29405-15.

234. Meulmeester, E., et al., *Mechanism and consequences for paralog-specific sumoylation of ubiquitin-specific protease 25.* Mol Cell, 2008. **30**(5): p. 610-9.

235. Cossec, J.C., et al., *SUMO Safeguards Somatic and Pluripotent Cell Identities by Enforcing Distinct Chromatin States.* Cell Stem Cell, 2018. **23**(5): p. 742-757 e8.

236. Theurillat, I., et al., *Extensive SUMO Modification of Repressive Chromatin Factors Distinguishes Pluripotent from Somatic Cells.* Cell Rep, 2020. **32**(11): p. 108146.

237. Luis, N.M., et al., *Regulation of human epidermal stem cell proliferation and senescence requires polycomb- dependent and -independent functions of Cbx4.* Cell Stem Cell, 2011. **9**(3): p. 233-46.

238. Deyrieux, A.F., et al., *Sumoylation dynamics during keratinocyte differentiation.* J Cell Sci, 2007. **120**(Pt 1): p. 125-36.

239. Heaton, P.R., et al., *Analysis of global sumoylation changes occurring during keratinocyte differentiation.* PLoS One, 2012. **7**(1): p. e30165.

240. Vivo, M., et al., *Downregulation of DeltaNp63alpha in keratinocytes by p14ARF-mediated SUMO-conjugation and degradation.* Cell Cycle, 2009. **8**(21): p. 3545-51.

241. Keiten-Schmitz, J., et al., *SUMO: Glue or Solvent for Phase-Separated Ribonucleoprotein Complexes and Molecular Condensates?* Front Mol Biosci, 2021. **8**: p. 673038.

242. Banani, S.F., et al., *Compositional Control of Phase-Separated Cellular Bodies.* Cell, 2016. **166**(3): p. 651-663.

243. Matunis, M.J., E. Coutavas, and G. Blobel, *A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex.* J Cell Biol, 1996. **135**(6 Pt 1): p. 1457-70.

244. Girard, C., et al., *Post-transcriptional spliceosomes are retained in nuclear speckles until splicing completion.* Nat Commun, 2012. **3**: p. 994.

245. Hochberg-Laufer, H., et al., *Uncoupling of nucleo-cytoplasmic RNA export and localization during stress.* Nucleic Acids Res, 2019. **47**(9): p. 4778-4797.

246. Pozzi, B., et al., *When SUMO met splicing.* RNA Biol, 2018. **15**(6): p. 689-695.

247. Pawellek, A., et al., *Characterisation of the biflavonoid hinokiflavone as a pre-mRNA splicing modulator that inhibits SENP.* Elife, 2017. **6**.

248. Carvalho, T., et al., *Pharmacological inhibition of the spliceosome subunit SF3b triggers exon junction complex-independent nonsense-mediated decay.* J Cell Sci, 2017. **130**(9): p. 1519-1531.

249. Richard, P., V. Vethantham, and J.L. Manley, *Roles of Sumoylation in mRNA Processing and Metabolism.* Adv Exp Med Biol, 2017. **963**: p. 15-33.

250. Chen, X., et al., *The function of SUMOylation and its crucial roles in the development of neurological diseases.* FASEB J, 2021. **35**(4): p. e21510.

251. Gonzalez-Santamaria, J., et al., *Regulation of the tumor suppressor PTEN by SUMO.* Cell Death Dis, 2012. **3**: p. e393.

252. Gonzalez-Prieto, R., et al., *c-Myc is targeted to the proteasome for degradation in a SUMOylation-dependent manner, regulated by PIAS1, SENP7 and RNF4.* Cell Cycle, 2015. **14**(12): p. 1859-72.

253. Rodriguez, M.S., et al., *SUMO-1 modification activates the transcriptional response of p53.* EMBO J, 1999. **18**(22): p. 6455-61.

254. Ivanschitz, L., et al., *PML IV/ARF interaction enhances p53 SUMO-1 conjugation, activation, and senescence.* Proc Natl Acad Sci U S A, 2015. **112**(46): p. 14278-83.

255. Gostissa, M., et al., *Activation of p53 by conjugation to the ubiquitin-like protein SUMO-1.* EMBO J, 1999. **18**(22): p. 6462-71.

256. Chen, S.F., et al., *Ubc9 expression predicts chemoresistance in breast cancer.* Chin J Cancer, 2011. **30**(9): p. 638-44.

257. Moschos, S.J., et al., *SAGE and antibody array analysis of melanoma-infiltrated lymph nodes: identification of Ubc9 as an important molecule in advanced-stage melanomas.* Oncogene, 2007. **26**(29): p. 4216-25.

258. Shao, D.F., et al., *High-level SAE2 promotes malignant phenotype and predicts outcome in gastric cancer.* Am J Cancer Res, 2015. **5**(1): p. 140-54.

259. Zhang, H., et al., *Over-expression of small ubiquitin-related modifier-1 and sumoylated p53 in colon cancer.* Cell Biochem Biophys, 2013. **67**(3): p. 1081-7.

260. Eifler, K. and A.C.O. Vertegaal, *SUMOylation-Mediated Regulation of Cell Cycle Progression and Cancer.* Trends Biochem Sci, 2015. **40**(12): p. 779-793.

261. Lopez, I., et al., *An unanticipated tumor-suppressive role of the SUMO pathway in the intestine unveiled by Ubc9 haploinsufficiency.* Oncogene, 2020. **39**(43): p. 6692-6703.

262. Shen, H.J., et al., *SENP2 regulates hepatocellular carcinoma cell growth by modulating the stability of beta-catenin.* Asian Pac J Cancer Prev, 2012. **13**(8): p. 3583-7.

263. Tan, M.Y., et al., *SUMO-specific protease 2 suppresses cell migration and invasion through inhibiting the expression of MMP13 in bladder cancer cells.* Cell Physiol Biochem, 2013. **32**(3): p. 542-8.

264. Chen, X.L., et al., *SENP2 exerts an antitumor effect on chronic lymphocytic leukemia cells through the inhibition of the Notch and NFkappaB signaling pathways.* Int J Oncol, 2019. **54**(2): p. 455-466.

265. Pei, H., et al., *SUMO-specific protease 2 (SENP2) functions as a tumor suppressor in osteosarcoma via SOX9 degradation.* Exp Ther Med, 2018. **16**(6): p. 5359-5365.

266. Kukkula, A., et al., *Therapeutic Potential of Targeting the SUMO Pathway in Cancer.* Cancers (Basel), 2021. **13**(17).

267. Kroonen, J.S. and A.C.O. Vertegaal, *Targeting SUMO Signaling to Wrestle Cancer.* Trends Cancer, 2021. **7**(6): p. 496-510.

268. Seeler, J.S. and A. Dejean, *SUMO and the robustness of cancer.* Nat Rev Cancer, 2017. **17**(3): p. 184-197.

269. Gandalovicova, A., et al., *Cell polarity signaling in the plasticity of cancer cell invasiveness.* Oncotarget, 2016. **7**(18): p. 25022-49.

270. Thiery, J.P., *Epithelial-mesenchymal transitions in tumour progression.* Nat Rev Cancer, 2002. **2**(6): p. 442-54.

271. Yuan, S., R.J. Norgard, and B.Z. Stanger, *Cellular Plasticity in Cancer.* Cancer Discov, 2019. **9**(7): p. 837-851.

272. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.

273. Tarasov, A., et al., *Sambamba: fast processing of NGS alignment formats.* Bioinformatics, 2015. **31**(12): p. 2032-4.

274. Liao, Y., G.K. Smyth, and W. Shi, *The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads.* Nucleic Acids Res, 2019. **47**(8): p. e47.

275. Durinck, S., et al., *BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis.* Bioinformatics, 2005. **21**(16): p. 3439-40.

276. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biol, 2014. **15**(12): p. 550.

277. Fischer, D.S., F.J. Theis, and N. Yosef, *Impulse model-based differential expression analysis of time course sequencing data.* Nucleic Acids Res, 2018. **46**(20): p. e119.

278. Gene Ontology, C., *The Gene Ontology resource: enriching a GOld mine.* Nucleic Acids Res, 2021. **49**(D1): p. D325-D334.

279. Ogata, H., et al., *KEGG: Kyoto Encyclopedia of Genes and Genomes.* Nucleic Acids Res, 1999. **27**(1): p. 29-34.

280. Liberzon, A., et al., *Molecular signatures database (MSigDB) 3.0.* Bioinformatics, 2011. **27**(12): p. 1739-40.

281. Bindea, G., et al., *ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks.* Bioinformatics, 2009. **25**(8): p. 1091-3.

282. Shannon, P., et al., *Cytoscape: a software environment for integrated models of biomolecular interaction networks.* Genome Res, 2003. **13**(11): p. 2498-504.

283. Wu, D., et al., *ROAST: rotation gene set tests for complex microarray experiments.* Bioinformatics, 2010. **26**(17): p. 2176-82.

284. Bradley Efron, R.T., *On testing the significance of sets ofgenes.* Ann. Appl. Stat., 2007. **1**(1): p. 107-129.

285. Mellacheruvu, D., et al., *The CRAPome: a contaminant repository for affinity purification-mass spectrometry data.* Nat Methods, 2013. **10**(8): p. 730-6.

286. Trapnell, C., et al., *Differential analysis of gene regulation at transcript resolution with RNA-seq.* Nat Biotechnol, 2013. **31**(1): p. 46-53.

287. Irimia, M., et al., *A highly conserved program of neuronal microexons is misregulated in autistic brains.* Cell, 2014. **159**(7): p. 1511-23.

288. Tapial, J., et al., *An atlas of alternative splicing profiles and functional associations reveals new regulatory programs and genes that simultaneously express multiple major isoforms.* Genome Res, 2017. **27**(10): p. 1759-1768.

289. Yang, J., et al., *The I-TASSER Suite: protein structure and function prediction.* Nat Methods, 2015. **12**(1): p. 7-8.

290. Zhao, Q., et al., *GPS-SUMO: a tool for the prediction of sumoylation sites and SUMO-interaction motifs.* Nucleic Acids Res, 2014. **42**(Web Server issue): p. W325-30.

291. Das Mahapatra, K., et al., *A comprehensive analysis of coding and non-coding transcriptomic changes in cutaneous squamous cell carcinoma.* Sci Rep, 2020. **10**(1): p. 3637.

292. Kim, D., et al., *TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions.* Genome Biol, 2013. **14**(4): p. R36.

293. Anders, S. and W. Huber, *Differential expression analysis for sequence count data.* Genome Biol, 2010. **11**(10): p. R106.

294. Elias, M.S., et al., *EMSY expression affects multiple components of the skin barrier with relevance to atopic dermatitis.* J Allergy Clin Immunol, 2019. **144**(2): p. 470-481.

295. Ji, Z., et al., *Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins.* Elife, 2015. **4**: p. e08890.

296. Kretz, M., et al., *Control of somatic tissue differentiation by the long non-coding RNA TINCR.* Nature, 2013. **493**(7431): p. 231-5.

297. Omote, N., et al., *Long noncoding RNA TINCR is a novel regulator of human bronchial epithelial cell differentiation state.* Physiol Rep, 2021. **9**(3): p. e14727.

298. Zhuang, Z., et al., *Down-Regulation of Long Non-Coding RNA TINCR Induces Cell Dedifferentiation and Predicts Progression in Oral Squamous Cell Carcinoma.* Front Oncol, 2020. **10**: p. 624752.

299. Hennings, H. and K.A. Holbrook, *Calcium regulation of cell-cell contact and differentiation of epidermal cells in culture. An ultrastructural study.* Exp Cell Res, 1983. **143**(1): p. 127-42.

300. Hennings, H., K.A. Holbrook, and S.H. Yuspa, *Factors influencing calcium-induced terminal differentiation in cultured mouse epidermal cells.* J Cell Physiol, 1983. **116**(3): p. 265-81.

301. Vandewalle, B., et al., *Intracellular calcium and breast-cancer cell-growth and differentiation.* Int J Oncol, 1993. **2**(4): p. 613-20.

302. Meng, L., et al., *Doxorubicin induces cardiomyocyte pyroptosis via the TINCR-mediated posttranscriptional stabilization of NLR family pyrin domain containing 3.* J Mol Cell Cardiol, 2019. **136**: p. 15-26.

303. Johnsson, P., et al., *Evolutionary conservation of long non-coding RNAs; sequence, structure, function.* Biochim Biophys Acta, 2014. **1840**(3): p. 1063-71.

304. Zampetaki, A., A. Albrecht, and K. Steinhofel, *Long Non-coding RNA Structure and Function: Is There a Link?* Front Physiol, 2018. **9**: p. 1201.

305. Charruyer, A., et al., *Decreased p53 is associated with a decline in asymmetric stem cell self-renewal in aged human epidermis.* Aging Cell, 2021. **20**(2): p. e13310.

306. McConnell, A.M., et al., *p53 Regulates Progenitor Cell Quiescence and Differentiation in the Airway.* Cell Rep, 2016. **17**(9): p. 2173-2182.

307. Saifudeen, Z., S. Dipp, and S.S. El-Dahr, *A role for p53 in terminal epithelial cell differentiation.* J Clin Invest, 2002. **109**(8): p. 1021-30.

308. Yugawa, T., et al., *Regulation of Notch1 gene expression by p53 in epithelial cells.* Mol Cell Biol, 2007. **27**(10): p. 3732-42.

309. Colombo, I., et al., *HaCaT Cells as a Reliable In Vitro Differentiation Model to Dissect the Inflammatory/Repair Response of Human Keratinocytes.* Mediators Inflamm, 2017. **2017**: p. 7435621.

310. Wilson, V.G., *Growth and differentiation of HaCaT keratinocytes.* Methods Mol Biol, 2014. **1195**: p. 33-41.

311. Bernat-Peguera, A., et al., *FGFR Inhibition Overcomes Resistance to EGFR-targeted Therapy in Epithelial-like Cutaneous Carcinoma.* Clin Cancer Res, 2021. **27**(5): p. 1491-1504.

312. Bernat-Peguera, A., et al., *PDGFR-induced autocrine SDF-1 signaling in cancer cells promotes metastasis in advanced skin carcinoma.* Oncogene, 2019. **38**(25): p. 5021-5037.

313. Zheng, Y., et al., *Expressions of oncogenes c-fos and c-myc in skin lesion of cutaneous squamous cell carcinoma.* Asian Pac J Trop Med, 2014. **7**(10): p. 761-4.

314. Zhang, Y., et al., *MALAT1-KTN1-EGFR regulatory axis promotes the development of cutaneous squamous cell carcinoma.* Cell Death Differ, 2019. **26**(10): p. 2061-2073.

315. Yang, J. and Y. Zhang, *I-TASSER server: new development for protein structure and function predictions.* Nucleic Acids Res, 2015. **43**(W1): p. W174-81.

316. Ren, J., et al., *Systematic study of protein sumoylation: Development of a site-specific predictor of SUMOsp 2.0.* Proteomics, 2009. **9**(12): p. 3409-3412.

317. Chen, L.L. and G.G. Carmichael, *Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA.* Mol Cell, 2009. **35**(4): p. 467-78.

318. Zhao, X., *SUMO-Mediated Regulation of Nuclear Functions and Signaling Processes.* Mol Cell, 2018. **71**(3): p. 409-418.

319. Pichaud, F., R.F. Walther, and F. Nunes de Almeida, *Regulation of Cdc42 and its effectors in epithelial morphogenesis.* J Cell Sci, 2019. **132**(10).

320. van den Bout, I. and N. Divecha, *PIP5K-driven PtdIns(4,5)P2 synthesis: regulation and cellular functions.* J Cell Sci, 2009. **122**(Pt 21): p. 3837-50.

321. Brash, D.E., *Roles of the transcription factor p53 in keratinocyte carcinomas.* Br J Dermatol, 2006. **154 Suppl 1**: p. 8-10.

322. Papadimitriou, J.C., et al., *"Thanatosomes": a unifying morphogenetic concept for tumor hyaline globules related to apoptosis.* Hum Pathol, 2000. **31**(12): p. 1455-65.

323. Eckhart, L., et al., *TINCR is not a non-coding RNA but encodes a protein component of cornified epidermal keratinocytes.* Exp Dermatol, 2020. **29**(4): p. 376-379.

324. Nita, A., et al., *A ubiquitin-like protein encoded by the "noncoding" RNA TINCR promotes keratinocyte proliferation and wound healing.* PLoS Genet, 2021. **17**(8): p. e1009686.

325. Araujo, A.P., et al., *Influence of the histidine tail on the structure and activity of recombinant chlorocatechol 1,2-dioxygenase.* Biochem Biophys Res Commun, 2000. **272**(2): p. 480-4.

326. Mohanty, A.K. and M.C. Wiener, *Membrane protein expression and production: effects of polyhistidine tag length and position.* Protein Expr Purif, 2004. **33**(2): p. 311-25.

327. Ledent, P., et al., *Unexpected influence of a C-terminal-fused His-tag on the processing of an enzyme and on the kinetic and folding parameters.* FEBS Lett, 1997. **413**(2): p. 194-6.

328. Sanchez, Y., et al., *Genome-wide analysis of the human p53 transcriptional network unveils a lncRNA tumour suppressor signature.* Nat Commun, 2014. **5**: p. 5812.

329. Hafner, A., et al., *Identification of universal and cell-type specific p53 DNA binding.* BMC Mol Cell Biol, 2020. **21**(1): p. 5.

330. Simpson, C.L., et al., *NIX initiates mitochondrial fragmentation via DRP1 to drive epidermal differentiation.* Cell Rep, 2021. **34**(5): p. 108689.

331. Mellem, D., et al., *Fragmentation of the mitochondrial network in skin in vivo.* PLoS One, 2017. **12**(6): p. e0174469.

332. Liberman, E.A., et al., *Mechanism of coupling of oxidative phosphorylation and the membrane potential of mitochondria.* Nature, 1969. **222**(5198): p. 1076-8.

333. Subramanian, A., et al., *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.* Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15545-50.

334. Ghafouri-Fard, S., et al., *TINCR: An lncRNA with dual functions in the carcinogenesis process.* Noncoding RNA Res, 2020. **5**(3): p. 109-115.

335. Xu, T.P., et al., *E2F1 induces TINCR transcriptional activity and accelerates gastric cancer progression via activation of TINCR/STAU1/CDKN2B signaling axis.* Cell Death Dis, 2017. **8**(6): p. e2837.

336. Zhu, Z.J. and J.K. He, *TINCR facilitates non-small cell lung cancer progression through BRAF-activated MAPK pathway.* Biochem Biophys Res Commun, 2018. **497**(4): p. 971-977.

337. Zhang, Z.Y., et al., *Loss of TINCR expression promotes proliferation, metastasis through activating EpCAM cleavage in colorectal cancer.* Oncotarget, 2016. **7**(16): p. 22639-49.

338. Hu, Y.W., et al., *LncRNA PLAC2 down-regulates RPL36 expression and blocks cell cycle progression in glioma through a mechanism involving STAT1.* J Cell Mol Med, 2018. **22**(1): p. 497-510.

339. Dong, L., et al., *LncRNA TINCR is associated with clinical progression and serves as tumor suppressive role in prostate cancer.* Cancer Manag Res, 2018. **10**: p. 2799-2807.

340. Liu, X., et al., *TINCR suppresses proliferation and invasion through regulating miR-544a/FBXW7 axis in lung cancer.* Biomed Pharmacother, 2018. **99**: p. 9-17.

341. Xia, H., et al., *LncRNA PLAC 2 downregulated miR-21 in non-small cell lung cancer and predicted survival.* BMC Pulm Med, 2019. **19**(1): p. 172.

342. Tian, F., et al., *TINCR expression is associated with unfavorable prognosis in patients with hepatocellular carcinoma.* Biosci Rep, 2017. **37**(4).

343. Xu, Y., et al., *Long noncoding RNA, tissue differentiation-inducing nonprotein coding RNA is upregulated and promotes development of esophageal squamous cell carcinoma.* Dis Esophagus, 2016. **29**(8): p. 950-958.

344. Hazawa, M., et al., *ZNF750 is a lineage-specific tumour suppressor in squamous cell carcinoma.* Oncogene, 2017. **36**(16): p. 2243-2254.

345. Zhou, W., et al., *lncRNA TINCR participates in ALA-PDT-induced apoptosis and autophagy in cutaneous squamous cell carcinoma.* J Cell Biochem, 2019. **120**(8): p. 13893-13902.

346. Ohata, C., *Hyaline Cell-Rich Apocrine Mixed Tumor with Cytologic Atypia.* Dermatopathology (Basel), 2018. **5**(3): p. 108-112.

347. Dikov, D.I., et al., *Hyaline globules (thanatosomes) in gastrointestinal epithelium: pathophysiologic correlations.* Am J Clin Pathol, 2007. **127**(5): p. 792-9.

348. Meriden, Z., et al., *Hyaline globules in neuroendocrine and solid-pseudopapillary neoplasms of the pancreas: a clue to the diagnosis.* Am J Surg Pathol, 2011. **35**(7): p. 981-8.

349. Gatalica, Z., et al., *Hyaline globules in renal cell carcinomas and oncocytomas.* Hum Pathol, 1997. **28**(4): p. 400-3.

350. Rubio, C.A., et al., *Intraepithelial bodies in colorectal adenomas: Leuchtenberger bodies revisited.* Dis Colon Rectum, 1991. **34**(1): p. 47-50.

351. Michal, M., et al., *Ovarian fibromas with heavy deposition of hyaline globules: a diagnostic pitfall.* Int J Gynecol Pathol, 2009. **28**(4): p. 356-61.

352. Dikov, D., et al., *Hyaline globules (thanatosomes) in prostate disease.* Am J Surg Pathol, 2003. **27**(5): p. 700-2.

353. Kerr, J.F., *An electron microscopic study of giant cytosegresomes in acute liver injury due to heliotrine.* Pathology, 1969. **1**(2): p. 83-94.

354. Anderson, P.J., S. Cohen, and T. Barka, *Hepatic injury. A histochemical study of intracytoplasmic globules occurring in liver injury.* Arch Pathol, 1961. **71**: p. 89-95.

355. Aishima, S., et al., *p62+ Hyaline inclusions in intrahepatic cholangiocarcinoma associated with viral hepatitis or alcoholic liver disease.* Am J Clin Pathol, 2010. **134**(3): p. 457-65.

356. Rekhi, B., et al., *Raspberry bodies and hyaline globules with positive napsin A immunoexpression are useful features in diagnosing clear cell carcinoma of the female genital tract in cytology samples.* Cytopathology, 2018. **29**(6): p. 600-602.

357. Theurillat, I., et al., *Extensive SUMO Modification of Repressive Chromatin Factors Distinguishes Pluripotent from Somatic Cells.* Cell Rep, 2020. **33**(1): p. 108251.

358. den Besten, W., et al., *Ubiquitination of, and sumoylation by, the Arf tumor suppressor.* Isr Med Assoc J, 2006. **8**(4): p. 249-51.

359. Hay, R.T., *SUMO: a history of modification.* Mol Cell, 2005. **18**(1): p. 1-12.

360. Castillo-Lluva, S., et al., *SUMOylation of the GTPase Rac1 is required for optimal cell migration.* Nat Cell Biol, 2010. **12**(11): p. 1078-85.

361. Salton, M., et al., *Involvement of Matrin 3 and SFPQ/NONO in the DNA damage response.* Cell Cycle, 2010. **9**(8): p. 1568-76.

362. Alfano, L., et al., *NONO regulates the intra-S-phase checkpoint in response to UV radiation.* Oncogene, 2016. **35**(5): p. 567-76.

363. Deshar, R., et al., *RNF8 mediates NONO degradation following UV-induced DNA damage to properly terminate ATR-CHK1 checkpoint signaling.* Nucleic Acids Res, 2019. **47**(2): p. 762-778.

364. Bladen, C.L., et al., *Identification of the polypyrimidine tract binding protein-associated splicing factor.p54(nrb) complex as a candidate DNA double-strand break rejoining factor.* J Biol Chem, 2005. **280**(7): p. 5205-10.

365. Jaafar, L., et al., *SFPQ*NONO and XLF function separately and together to promote DNA double-strand break repair via canonical nonhomologous end joining.* Nucleic Acids Res, 2017. **45**(4): p. 1848-1859.

366. Kowalska, E., et al., *NONO couples the circadian clock to the cell cycle.* Proc Natl Acad Sci U S A, 2013. **110**(5): p. 1592-9.

367. Maier, B. and A. Kramer, *A NONO-gate times the cell cycle.* Proc Natl Acad Sci U S A, 2013. **110**(5): p. 1565-6.

368. Taiana, E., et al., *LncRNA NEAT1 in Paraspeckles: A Structural Scaffold for Cellular DNA Damage Response Systems?* Noncoding RNA, 2020. **6**(3).

369. Adriaens, C., et al., *p53 induces formation of NEAT1 lncRNA-containing paraspeckles that modulate replication stress response and chemosensitivity.* Nat Med, 2016. **22**(8): p. 861-8.

370. Zhang, W.J. and J.Y. Wu, *Functional properties of p54, a novel SR protein active in constitutive and alternative splicing.* Mol Cell Biol, 1996. **16**(10): p. 5400-8.

371. Kameoka, S., P. Duque, and M.M. Konarska, *p54(nrb) associates with the 5' splice site within large transcription/splicing complexes.* EMBO J, 2004. **23**(8): p. 1782-91.

372. Emili, A., et al., *Splicing and transcription-associated proteins PSF and p54nrb/nonO bind to the RNA polymerase II CTD.* RNA, 2002. **8**(9): p. 1102-11.

373. Ponta, H., L. Sherman, and P.A. Herrlich, *CD44: from adhesion molecules to signalling regulators.* Nat Rev Mol Cell Biol, 2003. **4**(1): p. 33-45.

374. Brown, R.L., et al., *CD44 splice isoform switching in human and mouse epithelium is essential for epithelial-mesenchymal transition and breast cancer progression.* J Clin Invest, 2011. **121**(3): p. 1064-74.

375. Mima, K., et al., *High CD44s expression is associated with the EMT expression profile and intrahepatic dissemination of hepatocellular carcinoma after local ablation therapy.* J Hepatobiliary Pancreat Sci, 2013. **20**(4): p. 429-34.

376. Bhattacharya, R., et al., *Mesenchymal splice isoform of CD44 (CD44s) promotes EMT/invasion and imparts stem-like properties to ovarian cancer cells.* J Cell Biochem, 2018. **119**(4): p. 3373-3383.

377. Chen, Q., et al., *TGF-beta1 promotes epithelial-to-mesenchymal transition and stemness of prostate cancer cells by inducing PCBP1 degradation and alternative splicing of CD44.* Cell Mol Life Sci, 2021. **78**(3): p. 949-962.

378. Rodriguez, J.M., et al., *An analysis of tissue-specific alternative splicing at the protein level.* PLoS Comput Biol, 2020. **16**(10): p. e1008287.

379. Yeo, G., et al., *Variation in alternative splicing across human tissues.* Genome Biol, 2004. **5**(10): p. R74.

380. Matlin, A.J., F. Clark, and C.W. Smith, *Understanding alternative splicing: towards a cellular code.* Nat Rev Mol Cell Biol, 2005. **6**(5): p. 386-98.

381. Qadir, M.I., A. Parveen, and M. Ali, *Cdc42: Role in Cancer Management.* Chem Biol Drug Des, 2015. **86**(4): p. 432-9.

382. Hoadley, K.A., et al., *Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin.* Cell, 2014. **158**(4): p. 929-944.

383. Lorenzo-Martin, L.F., et al., *VAV2 signaling promotes regenerative proliferation in both cutaneous and head and neck squamous cell carcinoma.* Nat Commun, 2020. **11**(1): p. 4788.

384. Morrow, K.A. and L.A. Shevde, *Merlin: the wizard requires protein stability to function as a tumor suppressor.* Biochim Biophys Acta, 2012. **1826**(2): p. 400-6.

385. Gladden, A.B., et al., *The NF2 tumor suppressor, Merlin, regulates epidermal development through the establishment of a junctional polarity complex.* Dev Cell, 2010. **19**(5): p. 727-39.

386. Xiao, G.H., et al., *p21-activated kinase links Rac/Cdc42 signaling to merlin.* J Biol Chem, 2002. **277**(2): p. 883-6.

387. Laulajainen, M., et al., *Protein kinase A-mediated phosphorylation of the NF2 tumor suppressor protein merlin at serine 10 affects the actin cytoskeleton.* Oncogene, 2008. **27**(23): p. 3233-43.

388. Thrash, B.R., et al., *AKT1 provides an essential survival signal required for differentiation and stratification of primary human keratinocytes.* J Biol Chem, 2006. **281**(17): p. 12155-62.

389. Dainichi, T., et al., *PDK1 Is a Regulator of Epidermal Differentiation that Activates and Organizes Asymmetric Cell Division.* Cell Rep, 2016. **15**(8): p. 1615-23.

390. Sayama, K., et al., *Phosphatidylinositol 3-kinase is a key regulator of early phase differentiation in keratinocytes.* J Biol Chem, 2002. **277**(43): p. 40390-6.

391. Calautti, E., et al., *Phosphoinositide 3-kinase signaling to Akt promotes keratinocyte differentiation versus death.* J Biol Chem, 2005. **280**(38): p. 32856-65.

392. Janes, S.M., et al., *PI3-kinase-dependent activation of apoptotic machinery occurs on commitment of epidermal keratinocytes to terminal differentiation.* Cell Res, 2009. **19**(3): p. 328-39.

393. Pankow, S., et al., *Regulation of epidermal homeostasis and repair by phosphoinositide 3-kinase.* J Cell Sci, 2006. **119**(Pt 19): p. 4033-46.

394. Horikoshi, Y., et al., *Interaction between PAR-3 and the aPKC-PAR-6 complex is indispensable for apical domain development of epithelial cells.* J Cell Sci, 2009. **122**(Pt 10): p. 1595-606.

395. Hirose, T., et al., *Involvement of ASIP/PAR-3 in the promotion of epithelial tight junction formation.* J Cell Sci, 2002. **115**(Pt 12): p. 2485-95.

396. Hutterer, A., et al., *Sequential roles of Cdc42, Par-6, aPKC, and Lgl in the establishment of epithelial polarity during Drosophila embryogenesis.* Dev Cell, 2004. **6**(6): p. 845-54.

397. Rohani, M.G., et al., *Cdc42 inhibits ERK-mediated collagenase-1 (MMP-1) expression in collagen-activated human keratinocytes.* J Invest Dermatol, 2014. **134**(5): p. 1230-1237.

398. Yang, G.N., Z. Kopecki, and A.J. Cowin, *Role of Actin Cytoskeleton in the Regulation of Epithelial Cutaneous Stem Cells.* Stem Cells Dev, 2016. **25**(10): p. 749-59.