Universitat Autònoma de Barcelona

MASALA: a novel regulator of cellular plasticity encoded by a lncRNA

Iñaki Merino Valverde

Universidad Autónoma de Barcelona

Departamento de Bioquímica

MASALA: a novel regulator of cellular plasticity encoded by a lncRNA

Iñaki Merino Valverde

Barcelona, 2022
Doctoral Thesis

Departamento de Bioquímica
Universidad Autónoma de Barcelona

MASALA: a novel regulator of cellular plasticity encoded by a lncRNA

Iñaki Merino Valverde
Graduate in Biology
Master in Biotechnology

Thesis director: María Abad, PhD

The work presented in this thesis has been carried out in the Cellular Plasticity & Cancer Group at the Vall d'Hebron Institute of Oncology in Barcelona under the direction and supervision of Dr. María Abad Méndez

To those who wonder

"Confía en el tiempo, que suele dar dulces salidas a muchas amargas dificultades"

Miguel de Cervantes.

Acknowledgements

El trabajo que se presenta en esta tesis ha formado parte de mi vida durante los últimos cinco años y pico. A lo largo de este tiempo he aprendido muchísimo tanto sobre el tema del que trata esta tesis como sobre mi mismo, de una manera que jamás hubiera imaginado cuando lo comencé. Tanto este trabajo, como todo mi crecimiento personal hasta el día de hoy ha sido gracias a la gente que me ha hecho llegar hasta aquí y ser lo que soy hoy. Este trabajo es tan vuestro como mío.

En primer lugar y de manera evidente, este trabajo no existiría si no fuera por ti, María. Gracias por contar conmigo y confiar en mi para esta aventura en la cual hemos creado algo impresionante de la nada. Desde aquella entrevista por Skype en la que hoy hablar por primera vez de los micropeptidos, hasta nuestro último Zoom hace dos dias han pasado muchas cosas, tantas que darían para un libro a parte. Gracias por todo lo que me has enseñado, por saber ser jefa y compañera, por transmitir tu ilusión por este trabajo. y por cuidar de nosotros, tus pollos. Siempre llevaremos algo de ti alla donde vayamos (minimo, el toc de las figuras) ¡Muchas gracias!

Por supuesto, este trabajo empezó en un laboratorio vacío y fue gracias a mis compañeras pioneras que este trabajo saliera adelante. Empezando por Elena, mi hermana mayor en el laboratorio, muchas gracias por cada conversación, cada consejo de Post-doc, cada turno de tarde que compartimos, gran parte de lo que es este trabajo y lo que soy yo hoy es gracias a ti. A Emanuela, muchas gracias por tu apoyo incondicional, por cada receta de cocina que he aprendido, por cada idea en un labmeeting y por todos esos momentos que hemos compartido desde el día que pensabas que era una niña japonesa hasta hoy. A Olga, mi otra hermana de laboratorio y compañera original de bench, se que tu te has vuelto mas Vietnam y yo quizás mas Canadá (no), pero muchas gracias por tu ayuda siempre, por cada conversación sobre metafísica que hemos tenido, por cuidar de la familia gatuna y por todo lo que me has aportado como persona estos años, compartiendo alegrías y penas.

Aunque los inicios fueron duros, las nuevas generaciones llegaron pisando fuerte y no puedo dejar agradecer a la F1 del labo todo lo que me han aportado. A Alba, no tengo palabras para agradecerte lo suficiente todo lo que me has ayudado, gracias por entenderme tanto, por cada escapada al monte o a escalar, por estar ahí para sacarme del agujero, por cada pomodoro interrumpido y todas las risas. Aun nos quedan mas picos, mas canciones que descubrir y mas agobios que superar! A Marion, gracias por sacar siempre lo bueno de mi, ayudarme a pasar mis rachas malas siempre con algún coti entre manos, por anotar cada evento importante en nuestro calendario y ser mi compañera oficial de geogessr. Gracias por estar siempre dispuesta a ayudar y por tu sinceridad siempre! También tengo que dar las gracias a Lluís, aunque has llegado hace menos, llegaste en el mejor momento, gracias por la calma que inspiras.

Tampoco puedo olvidarme por la gente que ha pasado por el laboratorio y ha dejado su huella en este trabajo y en mi. Marta, muchísimas gracias por saber ver siempre lo bueno de nosotros, por tu paciencia infinita y por tu alegría en el laboratorio. A Alejandro gracias por formar parte de este proyecto y tu ayuda en cada experimento. A Camilla gracias por cada momento que compartimos en el turno de tarde de una época rara.

Aunque mis compañeras de laboratorio han sido y son mi familia en Barcelona. Llegar hasta aquí no habría sido posible si no fuera por mi familia. A mis padres, gracias por confiar y creer siempre en mi, por apoyarme en cada decisión, por ayudar corregirme las malas y por estar siempre ahí. Por inculcarme el esfuerzo cada dia, ellos serán siempre mi ejemplo a seguir. A mi hermano Javier, gracias por compartir la carga del doctorando y cada canción y momento que hemos compartido. A Aitor gracias por las risas, los masajes y las visitas a Barcelona. A mis tías Ana, Amaia y Rosi por preocuparse siempre por la tesis y por mi; y mis primos Andoni y Ana por estar ahí en cada visita y ayudarme a desconectar cuando era necesario. A Rafa y a Tere por

preocuparse siempre por mi como si fuera un hijo mas. Y por último a mis abuelos, ellos forman parte de mi y por tanto de este trabajo.

No puedo pasar por alto agradecer a mis amigos que me han apoyado cual hooligans desde el primer momento, desde el primer "Iñaki no te vayas" de Drugo, Sebas y Showman hasta cada momento, cada abrazo, cada cerveza durante estos años. Especialmente gracias a Pako, mi hermano de otra madre, que siempre está, en las buenas y las no tan buenas y con el que llevo mucha vida vivida. Y por supuesto a Nacho por enseñarme el autentico espíritu de lucha y valorarme siempre mas de lo que yo lo hago de mi mismo.

Por supuesto gracias a mis compañeros de facultad Carlos, Nacho, Sergio, Aitor, Pepemi, Juanjo y Pachón, desde la primera clase de la facultad hasta hoy siempre un apoyo. Sin olvidarme de Helen, cada visita, cada Descent, cada conversación, ¡gracias! Tampoco me olvido de Cris, Paloma Bea y Yol gracias.

Durante estos años he coincido en el VHIO con gente maravillosa la cual me ha marcado y ha hecho de mi y del VHIO algo mejor! Empezando por nuestros vecinos los ITAG Anna, Judit, y Noe y Albert, por salvarnos siempre la papeleta con cualquier cosa. Especial gracias a María Lozano nuestra compi de inicio de laboratorio, gracias por cada calçotada y cada cerveza en Julio. Andrea gracias por cada conversación en cultis, cada consejo y perdón por cada desahogo que has pillao por mi parte.
También gracias a Alfonso, Sara, Queralt, Andrea, Faiz, Ale, Alex, Tony, Oriol, Marie-Eve, Sandra y un montón de gente que se me olvida, pero que han aportado el valor humano de este centro y un montón de momentos.

Como la ciencia es colaborativa, no puedo dejar de agradecer tampoco a toda la gente que me ha ayudado desde otras instituciones. En el IRB, el laboratorio de Manuel Serrano ha sido de gran ayuda en este proyecto, empezando por Manuel Serrano, siempre acertado en sus consejos e ideas, gracias. Gracias también a Isabel por ayudarme siempre, aguantar miles de mails pidiendo cosas y siempre con una sonrisa. A Vanessa por su ayuda con las mito cosas y a todas las personas con las que he compartido momentos dentro y fura del laboratorio.

Es ahora momento de acordarme también de la gente que ha marcado pequeñas decisiones que me han llevado hasta aquí, empezando por mi profesora de biología en el instituto, gracias Tere. A Jorge, desde aquella experiencia en el CNIO hasta aquí, siempre confiaste en mi y no puedo si no agradecer que marcases mi camino hasta aquí.
También la gente del Judo, gracias Enrique por enseñarme el esfuerzo y el espíritu de lucha tan necesarios en un doctorado, asi como al resto de mis compañeros (David, Dani Montero, Alvaro, Mariano, Jorge…) Gracias!!

Especial mención para mi gata Jazz, la única de esta lista que no sabe leer ni hablar, pero siempre sabe como tranquilizarme.
Para terminar, a la que ha tenido que aguantar la peor parte de este trabajo, mi cara mas fea en todos los sentidos, la que ha tenido que sufrir cada momento de agobio, cada vez que se me cerraba el estomago, cada día que no podía dormir, cada enfado... Gracias Laura, por acompañarme incondicionalmente en esta aventura, por cada momento que nos ha traído hasta aquí desde aquel Agosto de 2007 hasta hoy, por intentar siempre hacerme sentir mejor aunque ni yo supiera como, por hacerme querer ser mejor persona. Gracias

Gracias a todo el que ha hecho esto posible

Resumen

Después de la secuenciación de genomas durante los ‘90 y 2000s se asumió que alrededor de un 99% del genoma no tenía función, llegando incluso a considerase “ADN basura”. A lo largo de los años, ésta y otras suposiciones sobre el genoma han ido cambiando. La secuenciación de transcriptomas a gran escala reveló que en realidad la mayoría del ADN se transcribía, resultando en la anotación de un gran número de transcritos como no codificantes. Sorprendentemente, en los últimos años se ha demostrado que muchas regiones que hasta ahora se consideraban no codificantes, como los lncRNA, en realidad codifican pequeñas proteínas menores de 100 amino ácidos que se han llamado microproteínas y que, debido a su pequeño tamaño, fueron pasadas por alto. Aunque por ahora sólo se han caracterizado una pequeña fracción de ellas, sabemos que son proteínas funcionales con papeles importantes en una gran variedad de procesos celulares. Estos describrimientos han abierto todo un campo de investigación, en el cual miles de microproteínas esperan a ser caracterizadas.

La plasticidad celular es la capacidad de una célula para cambiar su identidad y transitar entre diferentes estados celulares, incluyendo las transiciones entre estados diferenciados y desdiferenciados. Esta propiedad es crucial para muchos procesos, incluyendo el desarrollo embrionario y la regeneración de tejidos. De manera importante, en los últimos años se ha demostrado que la plasticidad celular es también crucial para las células tumorales, ofreciéndoles una ventaja sin parangón a la hora de adaptarse.

En esta tesis doctoral nos marcamos como objetivo identificar y caracterizar nuevas microproteinas que jueguen un papel en los procesos de plasticidad celular y cáncer. Hemos identificado a MASALA, una microproteina de 78 aminoácidos codificada por un gen erróneamente anotado como no codificante. MASALA se localiza en la membrana externa mitocondrial, y hemos observado que en condiciones fisiológicas este gen se expresa principalmente en cerebro, glándula adrenal y testículo. Además, hemos descubierto que su expresión se induce por daño citotóxico de manera dependiente de p53. Asimismo, la expresión de esta proteína correlaciona con un mejor pronóstico en distintos tipos de cáncer y, de hecho, su expresión se encuentra reprimida en varios tipos tumorales, lo que sugiere una posible función de MASALA como supresor tumoral. Apoyando esta hipótesis, hemos observado que la sobreexpresión de esta microproteína impide la transformación oncogénica de fibroblastos primarios murinos, así como su conversión en células pluripotentes inducidas (iPSCs) mediante reprogramación celular.

Para determinar los mecanismos moleculares responsables de la función de MASALA, hemos realizado un análisis de las diferentes funciones mitocondriales en fibroblastos primarios murinos. Hemos observado que la sobreexpresión continuada de MASALA durante produce alteraciones en la dinámica mitocondrial, mostrando una mayor frecuencia de mitocondrias fragmentadas junto con signos de estrés mitocondrial como hinchazón de la matriz mitocondrial y formación de esferoides mitocondriales. De acuerdo con estas observaciones, la sobreexpresión de MASALA es capaz de producir una reducción de la tasa de respiración mitocondrial, así como de su capacidad respiratoria de reserva. Finalmente, mediante espectrometría de masas hemos podido identificar un potencial interactor: MARCH5, una E3 ubiquitina-ligasa a través de la cual MASALA podría estar modificando la actividad de DRP1 y en consecuencia la dinámica mitocondrial.

En resumen, en este trabajo hemos identificado una nueva microproteína, hasta ahora desconocida, codificada por un gen anotado erróneamente como no codificante. MASALA es un nuevo regulador de la dinámica mitocondrial que actúa como barrera en los procesos de reprogramación celular y transformación oncogénica. Nuestros resultados aumentan nuestro conocimiento sobre el todavía inexplorado campo de las microproteínas, y pone de manifiesto que el microproteoma esconde proteínas relevantes para la preservación de la identidad celular y, por tanto, con potencial para el desarrollo de nuevas terapias en medicina regenerativa y cáncer.

Summary

After the sequencing of the genomes of different model organisms during the 90's and 2000's it was assumed that 99% of the genome did not have a function and was even considered "junk DNA". Along these years, this, and other assumptions about the genome have been challenged. Large-scale transcriptome revealed revealed that in fact most of the DNA was transcribed, leaving many transcripts annotated as non-coding. Surprisingly, more recently it has been demonstrated that many assumed non-coding regions, such as lncRNAs, can actually code for small proteins of less than 100 amino acids, called microproteins, which due to their small size, had been systematically overlooked. Although only a small subset of them have been characterized so far, we know that they are functional proteins with important roles in a wide variety of cellular processes. These discoveries have opened a new field of study, with thousands of microproteins waiting to be characterized.

Cellular plasticity is the ability of a cell to change its cellular identity and transit between different cellular states, including the transitions between differentiated and dedifferentiated states. It is a fundamental property for many processes like the embryonic development or tissue regeneration. Importantly, in the last years it has been demonstrated that cellular plasticity is also crucial for cancer cells, increasing their adaptative capacity.

In this doctoral thesis, we have aimed to find and characterize novel microproteins with a role in the processes of cellular plasticity and cancer. We have identified MASALA, a 78-amino acid microprotein encoded in a gene missanotated as a long non-coding RNA (lncRNA). MASALA localizes in the outer mitochondrial membrane and we have observed that in physiological conditions this gene is expressed in brain, adrenal gland and testis. Aditionally, we have observed that its expression is upregulated upon cytotoxic damage in a p53 dependent manner. Moreover, its expression correlates with a better prognosis in different cancer types and, in fact, it is downregulated in several tumor types, suggesting that MASALA has a tumor suppressor role. In agreement with this hypothesis, we have observed that overexpression of MASALA impairs the oncogenic transformation of mouse primary fibroblasts, as well as their transformation to induced pluripotent stem cells (iPSCs) through cellular reprogramming.

To characterize the molecular mechanism responsible for MASALA's function we performed an analysis of the different mitochondrial functions in murine embryonic fibroblasts. This revealed that continued MASALA overexpression produces mitochondrial dynamics alterations, displaying a more fragmented mitochondrial phenotype, together with different signs of mitochondrial stress, such as matrix swelling and mitochondrial spheroids formation. Accordingly, MASALA overexpression is able to reduce mitochondrial respiration as well as the spare respiratory capacity. Finally, from the analysis of MASALA interactome we have selected a potential interactor: MARCH5, a E3-ubiquitin ligase which could be modifying DRP1 activity through MASALA interaction and, thus, the mitochondrial dynamics.

Summarizing, in this work we have identified a previously unknown microprotein that is encoded in a gene missanotated as lncRNA. MASALA is a novel regulator of the mitochondrial dynamics that acts as a molecular barrier in the processes of cellular reprogramming and cancer. Our results further expand our knowledge of the yet unexplored field of microproteins and reveal that the microproteome may hide numerous relevant proteins for cellular identity and, therefore, with potential for the development of novel therapies for regenerative medicine and cancer.

Index

Introduction

1. Small ORF-encoded Microproteins

1.1. The hidden microproteome

The sequencing of the genome of different model organisms during the previous decades opened a new world to both, basic and applied science. In a relatively short period of time, scientists gathered an extraordinary amount of genomic data that had to be annotated. As reviewed by some authors back in the mid 1990s, the great challenge was to annotate meaningful open reading frames (ORFs) within the genome, filtering out those that can be randomly generated but are not translated. While analyzing the genome of *Saccharomyces cerevisiae,* they realized that setting a cut-off of 100-codons was sufficient to drastically reduce the artefactual noise of small ORFs (sORFs), that were considered biologically meaningless at the time (Basrai et al., 1997; Fickett, 1995). This, however, had a main drawback: the *ad hoc* cut-off of 100-codons would systematically hide sORFs that do code for functional proteins. ORF annotation was subjected to other assumptions, such as that eukaryotic genes are monocistronic, and that the translation start codon is an ATG.

Surprisingly, recent advances in bioinformatics, ribosome-profiling and peptidomics have revealed that many assumed non-coding regions, such as lncRNAs, miRNAs, and UTRs contain non-canonical ORFs that indeed code for small active proteins (Andrews and Rothnagel, 2014; Brunet et al., 2020; Makarewich and Olson, 2017; Patraquim et al., 2020). These proteins have received different names, such as sORFs-encoded polypeptides (SEPs), micropeptides or microproteins. These surprising observations have opened a new level of complexity, bearing enormous implications from basic research to the clinical setting.

Based on their genomic location, non-canonical ORFs can be classified as 1) small ORFs (sORFs), when they are located in previously assumed non-coding transcripts, such as lncRNAs, miRNAs and circRNAs, and generally are smaller than 100 codons; or 2) alternative ORFs (alt-ORFs), located in canonical protein-coding transcripts, but different from the main annotated ORF. Alt-ORFs can overlap the main ORF (in frame or out of frame) or can be located in the UTRs, named upstream ORFs (uORFs) when they are in the 5'UTR or downstream ORFs (dORFs) when they are in the 3'UTR **(Fig 1)** (Merino-Valverde et al., 2020) (2020). Importantly, many non-canonical ORFs starts with a codon different from ATG (Cao and Slavoff, 2020).

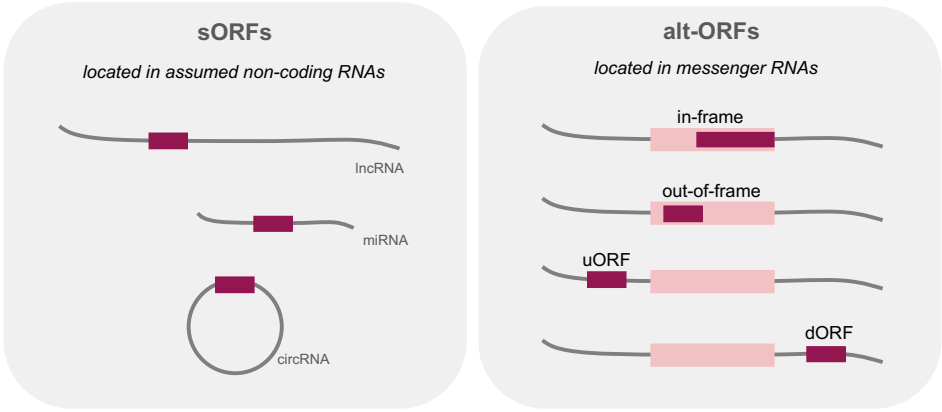


Fig 1. Classification of non-canonical ORFs. sORFs are smaller than 100 codons and are encoded by non-coding RNAs, such as lncRNAs, miRNAs or circRNAs (right panel); alt-ORFs are encoded by canonical protein-coding transcripts. They can overlap the main-ORF (pink), in-frame or out-of-frame, or they can be located in the UTRs, named upstream ORFs (uORFs) when they are in the 5′ UTR or downstream ORFs (dORFs) when they are in the 3′ UTR.

There are different methods to find novel microproteins. They can be separated into experimental methods, when the technique provides direct experimental evidence of the microprotein, or computational approaches, when microprotein existence is based on bioinformatic predictions. One of the experimental method is based on mass-spectrometry (MS). For MS analysis, the proteins of a biological sample are digested into small peptides that are analyzed to determine their specific amino acidic spectra. Then, the spectra are aligned to database containing the spectra of already known proteins, normally using available databases such as UniProt **(Fig 2)**. Since microproteins are not annotated in normal protein databases, common MS analysis overlook microproteins. To overcome this, it is needed to create a customized database that include microproteins. One way is using transcriptomic data and generate *in silico* all the microproteins that can be potentially translated. Although there has been great technical advance in the technique, it is still possible that many microproteins get lost during the identification. Their small size limits the number of tryptic peptides that are unique for the microprotein and thus, indistinguishable from other proteins. Additionally, their small size is accompanied by low abundance compared to other bigger proteins, probably due to the fact that classic proteomic approaches select the top mass-intensities in the analysis resulting in a bias towards highly abundant peptides. (Makarewich and Olson, 2017; Peeters and Menschaert, 2020; Slavoff et al., 2013; Zhu et al., 2018).

Another experimental approach is the identification of actively translated regions of the genome by Ribo-seq. In this technique, isolated ribosomes with their transcripts are subjected to nuclease treatment in order to eliminate those regions that are not protected by the ribosomes, and then sequence what remains **(Fig 2)**. This Ribo-Seq data is then evaluated for translation probability using algorithms such as RibORF, that scores ribosome pausing in the ORF based on codon frequency and periodicity (Ruiz-Orera et al., 2018).

On the other hand, computational approaches can be used for microprotein discovery. A method that has been followed in several reports is looking for potential ORFs with high phylogenetic codon conservation located in regions annotated as non-coding, which are generally non-conserved (as lncRNAs). This is performed using bioinformatic algorithms, such as PhyloCSF **(Fig 2)**. Predicted microproteins can be evaluated for additional features of meaningful ORFs, such as the presence of a Kozak sequence, microprotein homology with other known proteins or conserved protein domains as well as predicted N-terminal sequences. Although codon conservation is highly suggestive of biological relevance, this approach needs to be completed with further experimental evidences.

The combination of these experimental and computational approaches has allowed the identification of many novel microproteins. Nevertheless, these methods do not evaluate the biological relevance of the microproteins and a functional characterization needs to be performed case by case. To date, although only a small fraction of them have been functionally characterized, from these functional studies we have learnt that their small size provides them with unique characteristics, critical for different biological functions.

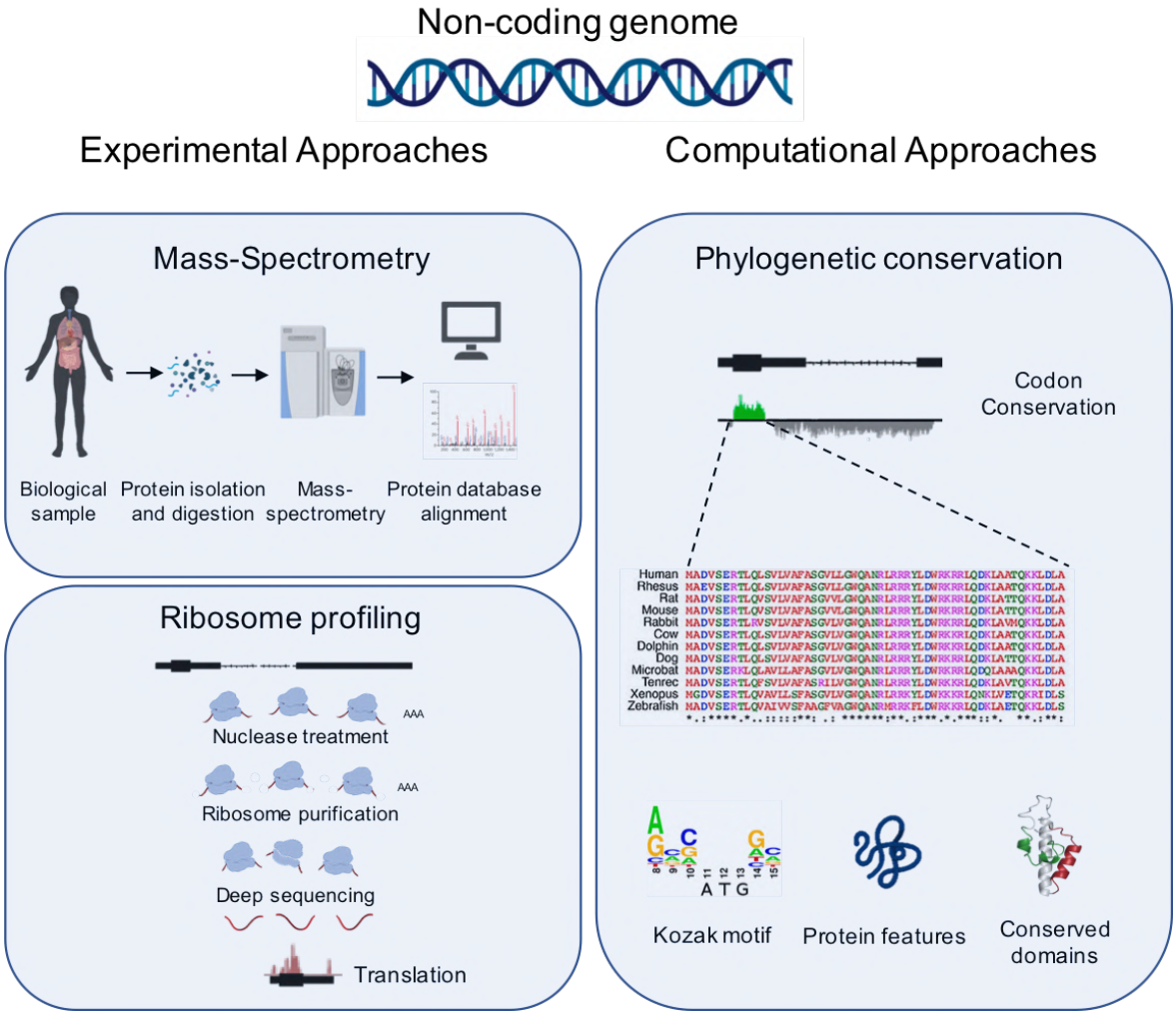


Figure 2. Microprotein discovery strategies. Methods for microprotein discovery can be divided into experimental or computational approaches. Experimental approaches give direct evidence for protein existence, by direct detection and identification of the microprotein spectrum by Mass-Spectrometry analysis or by direct evaluation of the translation of the transcript using Ribo-Seq. Computational approaches evaluate phylogenetic codon conservation and analyzes predicted protein characteristics.

1.2. Microproteins' functions

1.2.1. Microproteins in physiological processes

Despite their small size, microproteins have been demonstrated to display critical functions in different cellular processes. Unlike other known classical active peptides that come from proteolytic processing of much larger proteins, such as neuropeptides or some hormones, microproteins are translated from small ORFs directly as functional products. Given their size, microproteins have functional limitations such as those that implicate enzymatic activities, however, it makes them ideal for other roles. Many of the microproteins characterized so far have been shown to bind and regulate bigger protein complexes that participate in many different cellular processes. That is the case of the regulin family of microproteins. Myoregulin, Phospholamban, Sarcolipin, Endoregulin, Another-Regulin and DWORF, that through its interaction with the Calcium-ATPase SERCA they have been demonstrated to regulate intracellular calcium dynamics in a tissue specific manner (Anderson et al., 2015; Anderson et al., 2016; Nelson et al., 2016). Other examples include Mitoregulin, that regulates mitochondrial respiration stabilizing the formation of mitochondrial respiratory supercomplexes or SPAR microprotein, that regulates amino acid metabolism by regulating mTORC1 and lysosomal v-ATPase interaction (Makarewich et al., 2018; Matsumoto et al., 2017; Stein et al., 2018). Also, microproteins have been shown to be implicated in the process of DNA-repair; the MRI-2 microprotein has been shown to stimulate non-homologous end joining (NHEJ) by interaction with Ku heterodimeric proteins, while CYREN microprotein exhibits the opposite function (Arnoult et al., 2017; Slavoff et al., 2014). Additionally, NoBody microprotein participates in mRNA processing regulating the mRNA stability and the formation of P-bodies(D'Lima et al., 2017). Other cellular processes regulated by microproteins include, for example, splicing or fatty acid-oxidation (Makarewich et al., 2018; Polycarpou-Schwarz et al., 2018). Importantly, there is evidence of at least one microprotein that helps to regulate organelle interactions. PIGBOS microprotein localizes in the outer mitochondrial membrane and regulates the ER-Mitochondrial contacts. The interaction of PIGBOS with ER help the cells to cope with ER stress and resist to apoptosis (Chu et al., 2019).

Summarizing, microproteins have been observed in different cellular processes acting as fine-tuner regulators of major protein complexes. Interestingly, many of the new discovered microproteins, including most of the above mentioned, have hydrophobic regions in their structure and localize in membrane regions where they display their regulatory functions at ion channeling, protein recruitment or enzymatic regulation (Makarewich, 2020).

1.2.2. Microproteins in pathological processes

As many regions of the genome have been shown to encode novel microproteins, the interest on their possible implication in disease has grown over the last years. In fact, many microproteins have been linked to pathological contexts. BRAWNIN microprotein, together with the previously mentioned Mitoregulin, can be related with metabolic disorders as these proteins have been shown to regulate fatty acid-oxidation and mitochondrial respiration (Makarewich et al., 2018; Stein et al., 2018; Zhang et al., 2020). Another example is Humanin microprotein, which expression is downregulated in Alzheimer's disease and has already been linked to neuroprotective effects (Chai et al., 2014; Hashimoto et al., 2001). Also, microproteins have been found to regulate immune response. That is the case of Aw112010, that has been related with mucosal immunity, protecting against infection through the regulation of proinflammatory pathways (Chu and

Saghatelian, 2019). However, most of the microproteins related with pathological processes have been identified in the context of cancer.

1.2.2.1. Microproteins in cancer

Cancer is a complex, multistep disease in which cells acquire replicative immortality and invasion capacities to colonize adjacent or distant tissues. They do so by deregulating and hijacking many cellular processes to their advantage in order to survive and progress. The finding of novel microproteins in the context of cancer contribute to advance our understanding of cancer disease as well as providing a novel source of clinical targets.

Different microproteins have been already identified in specific cancer processes. The first one was HOXB-AS3; this microprotein has been shown to be downregulated in colorectal cancer and its overexpression in cancer cells attenuates their oncogenic capabilities by altering their glycolytic capacity (Huang et al., 2017). Other examples of tumor suppressor microproteins include PINT-87aa, that is downregulated in glioblastoma and acts as a transcriptional repressor of several oncogenes (Zhang et al., 2018) and YY1BM, a microprotein that is downregulated in Esophageal Squamous Cell Carcinoma (ESCC) and sensitizes cells to apoptosis upon nutrient deprivation (Wu et al., 2020). Additionally, MIAC microprotein has been shown to inhibit Head and Neck Squamous Cell Carcinoma (HNSCC) progression and metastasis by inactivating the actin cytoskeleton through interaction with Aquaporin-2. CIP2A-BP microprotein inhibits triple negative breast cancer metastasis by competitive binding with CIP2A protein, which results in downregulation of key proteins for metastasis such as Snail or different matrix metallopeptidases (Guo et al., 2020). Finally, another recently described tumor suppressor microprotein is MP31, which displays its tumor suppressor activity in Glioblastoma impairing the conversion of lactate to pyruvate by interacting and modifying the activity of the mitochondrial Lactate Dehydrogenase (Huang et al., 2021).

On the contrary, several microproteins have been shown to display oncogenic activity. CASIMO1 microprotein has been shown to activate the mitogen-activated protein kinases (MAPK) cascade in breast cancer (Polycarpou-Schwarz et al., 2018). CircPPP1R12A- 73aa also has been proposed to act as an oncogene, increasing the proliferation of colorectal cancer cells through activation of the Hippo-Yap pathway (Zheng et al., 2019). Also in colorectal cancer, another microprotein, ASAP, promotes cancer progression modulating ATP synthase activity (Ge et al., 2021). Finally, microproteins have been related with metastasis, as it is the case of ZFAS1, which promotes cellular migration in hepatocellular carcinoma (HCC)(Guo et al., 2019).

Altogether, these evidences show the importance of the microproteome in the context of cancer. Although many microproteins have been directly linked to cancer, there are many others described in other contexts that, given their molecular function, we speculate over its possible a role regulating different aspects of cancer biology as well **(Fig 3).** For example, microproteins regulating DNA-damage repair such as MRI-2 or CYREN might as well be implicated in cancer progression as their function as their inhibition or overexpression respectively would result in genomic instability and genomic mutations. We have recently reviewed these evidences in (Merino-Valverde et al., 2020).

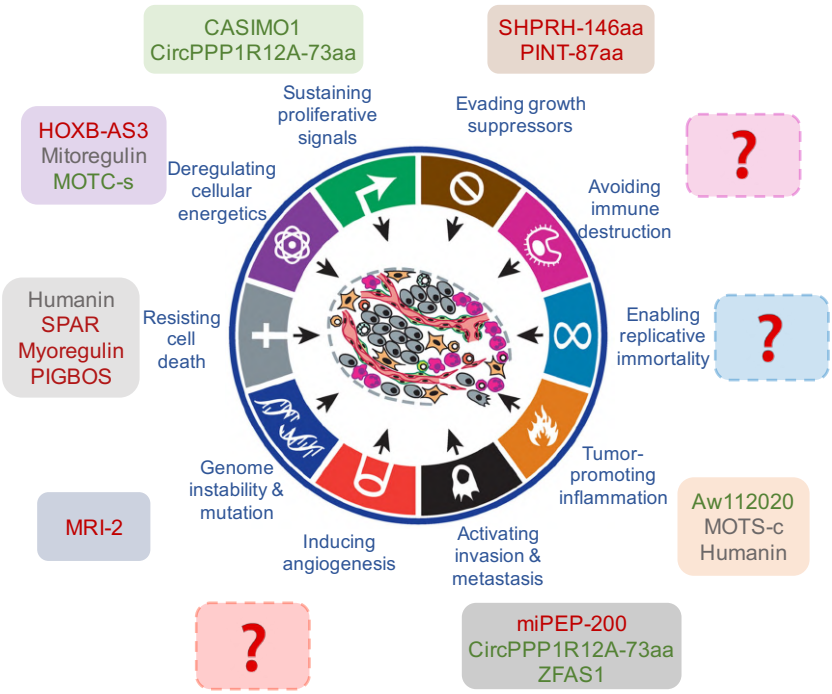


Figure 3. Microproteins as novel regulators of the hallmarks of cancer. A subset of microproteins have been functionally characterized and have been directly related with cancer; some others, based on their function, are likely to be related with cancer. The figure represents the hallmarks of cancer defined by Hanahan and Weinberg and their related microproteins. In green, the microproteins that promotes or activate the hallmark and, in red, the microproteins that function inhibiting or blocking the hallmark. The ones in grey need further investigation to be classified. Of interest, those depicted in red represent tumor suppressor microproteins with potential pharmacological activity, while those in green are pro-oncogenic proteins that could be targeted in the clinic.

1.2.3. Mitochondrial microproteins

Microproteins are distributed along different cellular compartments. The mitochondria, however, seems to be an organelle enriched in microproteins (Zhang et al., 2020), given that many of the microproteins that have been already characterized are located in the mitochondria.

There are two different types of mitochondrial microproteins according to their origin. Microproteins encoded by the mitochondrial genome are called mitochondrial-derived peptides (MDPs). To date, eight different MDPs have been already described, being the first one Humanin (Hashimoto et al., 2001). This MDP is encoded by the 16S ribosomal RNA in the mitochondrial genome and has been shown to be downregulated with age. Additionally, it has been demonstrated to display cytoprotective activity in prion-like diseases, such as Alzheimer's, and strokes (Kariya et al., 2005; Sponne et al., 2004; Xu et al., 2006). Also, humanin can ameliorate metabolic conditions such as obesity and diabetes by regulating insulin sensitivity (Hoang et al., 2010; Muzumdar et al., 2009).

Interestingly, in the same locus six more MDPs have been found and named small humanin-like peptides (SHLP1-6), and all of them display cytoprotective activities upon different metabolic stressors (Merry et al., 2020). The last MDP that has been characterized is MOTS-c, which is encoded by the 12S Ribosomal RNA. This protein has been linked to improve glucose metabolism, in part via AMPK. The expression of this MDP is also downregulated with age, it ameliorates insulin-sensitivity and enhances

muscle performance (Lee et al., 2015; Reynolds et al., 2021; Zarse and Ristow, 2015). Altogether, these evidences indicate that MDPs have a cytoprotective role, acting through metabolism and apoptosis regulation. Beside its cytoprotective activity, MDPs have been linked to aging and the process of cellular senescence, in which cells are metabolically very active. MDPs have been shown to be upregulated in senescent cells and to exacerbate the expression of certain Senescence-associated secretory phenotype (SASP) molecules (Kim et al., 2018; Mendelsohn and Larrick, 2018).

On the other hand, since 99% of the mitochondrial proteome comes from nuclear-encoded proteins, most of the mitochondrial microproteins have a nuclear origin. Most of them remain uncharacterized, but the few of them that have been characterized had been shown to be important regulators of mitochondrial biology. For example, some of them act on mitochondrial respiration, as Mitoregulin, which has been found in the mitochondrial inner membrane (IMM), displaying a role in the respirasome assembly, enhancing oxidative phosphorylation (OXPHOS) and promoting fatty acid oxidation as described before (Lin et al., 2019; Makarewich et al., 2018; Stein et al., 2018). Other examples are BRAWNIN microprotein, that has been identified in the IMM being a target of AMPK-PGC-1a axis and an essential component of the respiratory chain complex III, and the ASAP microprotein, that promotes the ATP synthase activity (Ge et al., 2021; Zhang et al., 2020). Another IMM microprotein is the already mentioned MP31, encoded by the 5'UTR region of the PTEN gene, which prevents the lactate-pyruvate conversion and bears tumor suppressor activity (Huang et al., 2021).

Regarding other aspects of mitochondrial biology, such as inter-organelle communication, PIGBOS has been described as an essential component of the mitochondrial-ER associated membranes. As mentioned above, it is located in the mitochondrial outer membrane and regulates the ER stress response (Chu et al., 2019). Another microprotein, MIEF1-MP has been identified as regulator mitochondrial translation via its binding to the mitoribosome (Rathore et al., 2018) and finally, another mitochondrial microprotein, Mm47, has been shown to regulate the activation of the inflammasome in macrophages, although its role in the mitochondria remains unknown (Bhatta et al., 2020).

In conclusion, microproteins have been revealed as a new class of fine-tune molecular regulators with crucial functions in many cellular processes. The discovery of the microproteome has expanded our view on the coding capacity of the genome, adding a new layer of complexity on biological processes with potential implications in both, basic and translational research.

2. Cellular Reprogramming and Cancer

2.1. Cellular plasticity

Life on Earth appeared approximately 3.5 billion years ago. Since then, different life forms have acquired more and more complexity, starting from unicellular organisms to multicellular ones. Over 600 million years ago animals evolved from single cell or colonial organisms to the complex systems that we see nowadays. This multicellularity arose from the specialization of cells in the organisms, that differentiated from one another to perform specific functions, acquiring their own cellular identity (Brunet and King, 2017).

Cellular plasticity is defined as the ability of a cell to change its original identity and transit between different cellular states. A paradigmatic example of cellular plasticity is represented by embryonic stem cells (ESCs). These cells are only present for a brief

period of time during embryo development, and they are pluripotent, meaning that they have the potential to differentiate into every cell type of our adult body. As embryonic development progresses, there is a decrease in differentiation potential and, as adults, we only possess multipotent adult stem cells, that are able to generate only some types of terminally differentiated cells. Terminally differentiated cells have specific functions and present unique properties that are usually maintained along its life, since they have limited proliferation capacity or do not proliferate at all. Under most circumstances, cells within the tissues maintain their cellular identities stable. However, upon specific environmental cues some cells can change their original identity and transit between different cellular states (Jopling et al., 2011; Raff, 2003). There are different types of conversions: on one hand, differentiated cells can be converted into another differentiated cell type through a process called transdifferentiation **(Fig 4).** On the other side, cells can go back to a more de-differentiated state, through a process called de-differentiation **(Fig 4)** (Merrell and Stanger, 2016).

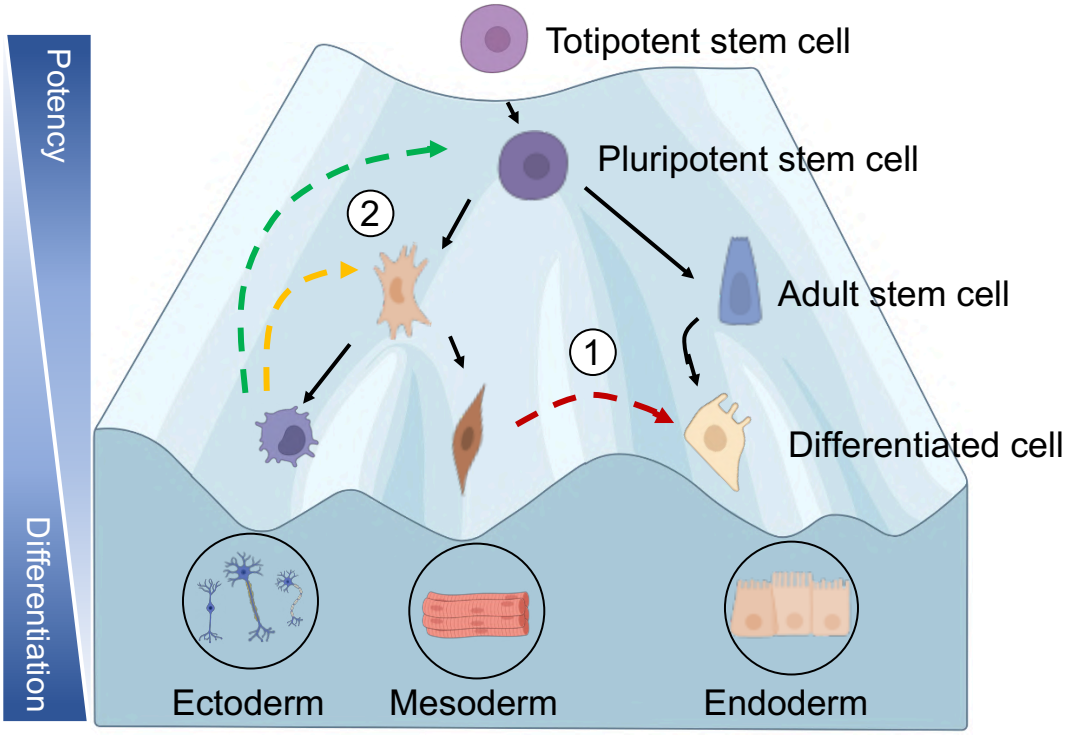


Figure 4. Cellular hierarchies and the conversions between them by cellular plasticity. Adaptation of the Waddington landscape in which there are depicted the differentiation stages and the possible conversions between them. **1,** the transdifferentiation process, by which a terminally differentiated cell loses its cellular identity and it is transformed into a different somatic cell type. **2**, cellular dedifferentiation process, in which cells gain differentiation potential. This process can be partial; generating a cell with limited potency (yellow arrow) or can generate a pluripotent stem cell, which can differentiate into the three germinal layers (in this case the process is known as cellular reprogramming, green arrow).

2.2. The process of cellular reprogramming

2.2.1. Discovery of cellular reprogramming

In 1958, John Gurdon and collaborators demonstrated that an adult differentiated cell could be experimentally transformed back to an embryonic state by transferring its nucleus into an enucleated oocyte, a technique called somatic cell nuclear transfer (SCNT), demonstrating for the first time that cellular identity and developmental processes are determined epigenetically, and thus, are reversible (Gurdon et al., 1958). This finding, together with the advances in genetics paved the way to the discovery of the process of cellular reprogramming by Shinya Yamanaka and Kazutoshi Takahasi in 2006. In this landmark study, these authors discovered that somatic cells can be converted into induced pluripotent stem cells (iPSCs) by the simple overexpression of four different embryonic transcription factors: *Oct4, Klf4, Sox2* and *c-Myc,* now widely known as the "Yamanaka factors" or OSKM. These reprogrammed cells are undistinguishable from ESCs: iPSCs can be maintained and propagated in a pluripotent state (they self-renew), and they could also be differentiated into the three germlines: endoderm, mesoderm and ectoderm (Takahashi and Yamanaka, 2006). This discovery revolutionized the stem cell field and has had important implications in the study of embryology, aging, regenerative medicine and multiple diseases including cancer (Pesaresi et al., 2019; Xiong et al., 2019).

Importantly, Yamanaka and Takahasi achieved the reprogramming to iPSCs *in vitro*, and this process was thought to don't have a biological counterpart. However, in 2013, María Abad in the lab of Manuel Serrano generated a mouse model that overexpresses the Yamanaka factors in an inducible manner (hereafter abbreviated as i4F), and demonstrated that *in vivo* conditions are in fact permissive to cellular de-differentiation and to reprogramming to iPSCs (Abad et al., 2013). Even more surprisingly, they observed that *in vivo* generated iPSCs could also contribute to trophectoderm lineage and that were able to generate embryo-like structures that express extra-embryonic markers, suggestive of totipotency features. In other words, when overexpressed *in vivo*, the four Yamanaka factors induced an even more de-differentiated and plastic phenotype (Abad et al., 2013).

2.2.2. Molecular mechanisms governing cellular reprogramming

To understand how cells lose their identity and are reprogrammed from a somatic differentiated state to pluripotency, biologists have been carefully dissecting the molecular pathways that govern the process of cellular reprogramming.

As first observed in Gurdon experiments, cell identity is regulated epigenetically. Accordingly, the initial induction of the OSKM generates a wide chromatin remodeling in the cells that targets developmental genes promoters to re-activate its transcription (Koche et al., 2011; Soufi et al., 2012). These early stages, however, are highly unstable and a continuous expression of the OSKM is required for them to progress, or cells differentiate back to the initial state. Changes in the transcriptional landscape makes cells to lose their cellular identity markers and start rapidly proliferating, promoting an epithelial-to-mesenchymal transition that its reversed during the late stages of reprogramming (Liu et al., 2013). After continuous OSKM expression, in the late phase of reprogramming endogenous expression of OSKM starts, together with other pluripotency factors such as *Nanog* and telomerase (Brambrink et al., 2008; Papp and Plath, 2013; Polo et al., 2012). This is accompanied by an open chromatin status that resembles that of ES cells (Papp and Plath, 2013). Additionally, iPSCs are characterized by their ability to proliferate indefinitely at high speed, thus, for a somatic cell to be reprogrammed they need to undergo a metabolic reprogramming that enables them to cope with the energetic

requirements (Folmes et al., 2011; Panopoulos et al., 2012; Teslaa and Teitell, 2015; Vander Heiden et al., 2009).

While cell-autonomous mechanisms are important for cellular reprogramming, evidences suggest that cellular context and the microenvironment are determinant for de-differentiation. The extracellular matrix and softer cellular substrates have been demonstrated to provide cells with the adequate mechanical signals that are transduced to promote reprogramming (Gerardo et al., 2019). Extracellular signals such as TNFα secreted by neighboring fibroblasts during reprogramming have been shown to promote dedifferentiation (Mahmoudi et al., 2019). Additionally, the overexpression of the OSKM factors *in vitro* and *in vivo* cause many of the cells to enter into an irreversible, non-proliferative, metabolically active state named cellular senescence. Senescent cells secrete a wide array of cytokines and soluble factors such as IL6 that provides neighboring cells with the signals necessary to reprogram (Mahmoudi et al., 2019; Mosteiro et al., 2016; Mosteiro et al., 2018).

2.2.3. Damage-induced dedifferentiation

Cellular plasticity is an essential cellular feature for living organisms to adapt and survive. In fact, acquisition of plasticity by de-differentiation is not restricted to laboratory conditions. In living organisms, we find many examples during the process of regeneration, when differentiated cells can acquire a more de-differentiated plastic state in order to repopulate the injured area. In amphibians, for example, following limb amputation of *Notophtalmus viridescens* salamander de-differentiation of cells have been observed to produce an activated pool of proliferating muscle cells, called blastema, that differentiates into a new completely functional limb (Iten and Bryant, 1973). In *Danio rerio*, de-differentiation plays a role in heart regeneration as well. In this case, upon amputation cardiomyocytes de-differentiate and start proliferating again to regrow cardiac muscle tissue (Jopling et al., 2010; Kikuchi et al., 2010). In mammals, however, regeneration is a process that has only been conserved to a limited extent. Nevertheless, several examples of de-differentiation have already been identified. In mouse, it has been observed that certain secretory progenitor cells of the intestine can revert to stem cells upon crypt damage in order to reconstruct the tissue (van Es et al., 2012). Additionally, if the stem cell compartment of the intestine is genetically ablated, it can be repopulated by their enterocyte-lineage daughter cells (Tetteh et al., 2016). De-differentiation upon damage also happens in other tissues such as the lung, where epithelial cells de-differentiate into stem cells upon stem cell niche ablation *in vivo* (Tata et al., 2013), and it has also been observed in the liver, pancreas and nervous system (Kopp et al., 2016; Lin et al., 2017; Painter et al., 2014). Consistently, Manuel Serrano's group has shown that tissue damage, through the induction of cellular senescence and the secretion of IL6 promotes *in vivo* cellular reprogramming (Mosteiro et al., 2016).

Altogether, these evidences reveal a link between damage and the induction of cellular plasticity in order to aim regeneration, although only a transient and limited potency can be achieved in nature. Of note, only a few cell types have been shown to be capable to de-differentiate in a tissue, and some cells are more prone to reprogramming than others (Abad et al., 2013; Mosteiro et al., 2016; Vidal et al., 2014), suggesting that cellular plasticity acquisition depends on both, cell and non-cell autonomous properties.

2.3. Connections between cellular reprogramming and cancer

During the past decade, it has become evident that cellular plasticity is a critical feature of cancer cells at every stage of carcinogenesis: tumor initiation, tumor maintenance, and metastasis. In addition, several evidences show that cancer cells acquire plasticity upon chemotherapy treatment, which promote therapy resistance and tumor relapse (Filipponi et al., 2019; Raghavan et al., 2021; Xiong et al., 2019). In agreement, it has been shown that typical indicators of de-differentiation are present in many cancer types and those cancers with high expression of the reprogramming factors correlate with a more aggressive, poor prognosis in the clinic (Boumahdi et al., 2014; Hepburn et al., 2019; Suva et al., 2013). For all these reasons, acquisition of plasticity has been recently proposed as an emergent hallmark of cancer (Hanahan, 2022)

In fact, since the discovery of cellular reprogramming, this process has been shown to share many characteristics with neoplastic transformation, and nowadays both processes can be seen as parallel processes governed by common molecular mechanisms. Firstly, although the identification of the tumor cell of origin has been challenging for many years, and still is in some cases, there is great number of evidences that suggest that the cell of origin is a somatic, differentiated cell that acquires unlimited proliferation (Friedmann-Morvinski and Verma, 2014), as it is in cellular reprogramming. Second, both processes require the expression of oncogenes; while this may be evident in the case of tumors, in the case of reprogramming the reprogramming factors are known to have oncogenic activity (Bass et al., 2009; Iglesias et al., 2017; Rodini et al., 2012; Saiki et al., 2009). Additionally, the acquisition of an unlimited proliferation during reprogramming requires a metabolic rewire in the cell that switches from an OXPHOS dependent status to a more glycolytic phenotype **(Fig 5).** In cancer, this has been widely studied as the Warburg effect, although this is highly dependent on tumor type and cell populations within the tumor, and recent evidences suggest that OXPHOS is also an important feature in certain tumor contexts (DeBerardinis and Chandel, 2020).

Finally, in order to change their cell identity, cells that are subjected to oncogenic transformation or cellular reprogramming need to overcome the induction of tumor suppressor responses, mainly apoptosis and cellular senescence. Consistently, tumor suppressor genes had been described as molecular barriers for both processes **(Fig 5).**

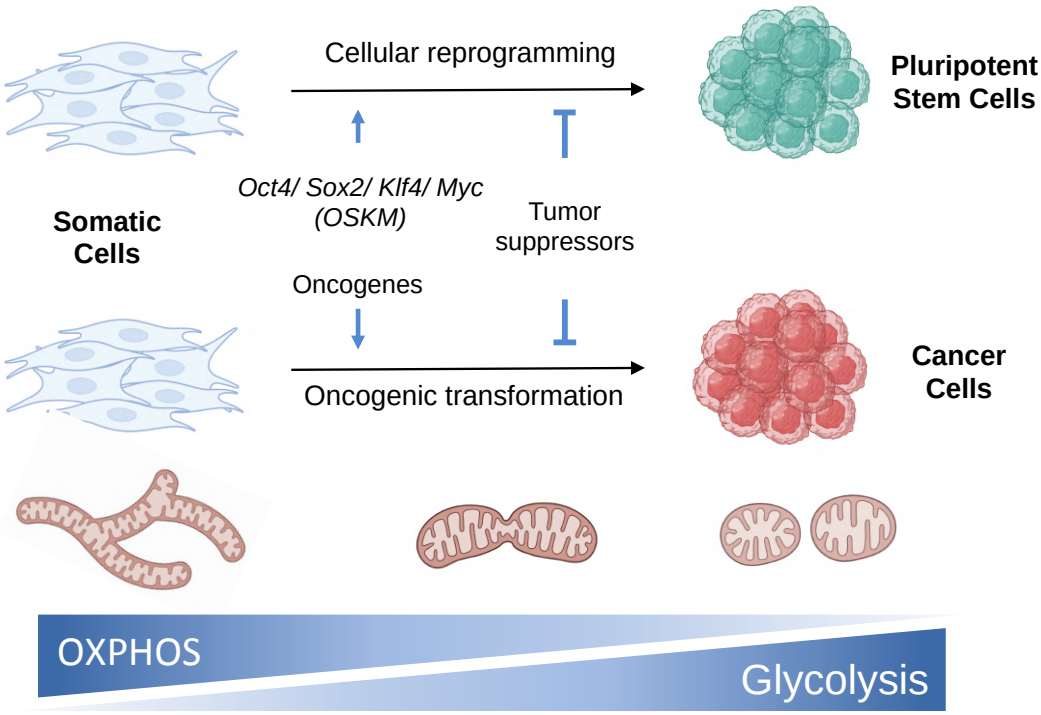


Figure 5. Parallelisms between cellular reprogramming and cancer. Cellular reprogramming as well as neoplastic transformation starts from a population of somatic cells with limited proliferation capacity and finish with an immortal population of cells that self-renew. Additionally, both processes are triggered by oncogenic signals and are blocked by the activation of tumor suppressor genes that protect cellular identity. Finally, both processes are subjected to a metabolic switch, from an OXPHOS-based energy source to a glycolytic one.

2.3.1. Tumor suppressors as guardians of cell identity

Upon damage, cells that de-differentiate switch to more proliferative states that resemble those acquired during neoplastic transformation, as discussed above. Hyperproliferative states need to be negatively regulated to warrant that this status is transiently maintained and do not disrupt tissue homeostasis. To avoid unlimited proliferation and maintain their original identity, cells have molecular barriers that tightly regulates cell cycle progression; those are tumor suppressor genes.

Besides, hyperproliferation generates replicative stress that makes cells prone to mutation and genome instability (Zeman and Cimprich, 2014). Moreover, replicative stress is accompanied by mitochondrial stress that generates reactive oxygen species (ROS) what further promotes DNA-damage. To ensure that replicative stress do not cause alterations in DNA replication or DNA-damage, activation of tumor suppressor genes generates a wide array of outcomes such as cell-cycle arrest, apoptosis or senescence, that limits the detrimental effects of rapid replication (Hills and Diffley, 2014). These tumor suppressor genes are highly upregulated at early phases of cellular reprogramming and need to be inactivated in cancer progression. Accordingly, several evidences indicated that inactivation of key tumor suppressor genes such as Ink4-Arf, Rb, p53 or p21 greatly increased the cellular reprogramming efficiency (Hong et al., 2009; Kareta et al., 2015; Li et al., 2009; Marion et al., 2009). Inactivation or loss of these tumor suppressor genes allow cells to avoid cell death, cell cycle arrest and senescence and successfully change their original identity. Moreover, it has been demonstrated that loss of tumor suppression genes promote the metabolic reprogramming needed for reprogramming and neoplastic transformation, switching from the slow but highly efficient OXPHOS to the less efficient

but fast glycolysis (Liu et al., 2019). For example, loss of tumor suppressor LKB1 promotes this process through HIF1-a activation (Faubert et al., 2014). Accordingly, activation of AMPK protein, a downstream effector of LKB1, has been shown to inhibit both cellular reprogramming and cancer impairing the metabolic switch (Vazquez-Martin et al., 2012b). Finally, the metabolic switch also needs mitochondrial remodeling from a more elongated to a more fragmented phenotype. For this reason, alteration of the pathways that negatively regulate mitochondrial dynamics, glycolysis or mitophagy are also barriers for cellular transformation (Prieto et al., 2016; Wang et al., 2020a).

In summary, in our view, tumor suppressor genes can be viewed as guardians of cell identity, that prevents cells to be transformed into cancer cells, but also to acquire more dedifferentiated stages.

3. The Mitochondria in Cellular Reprogramming and Cancer

3.1. Mitochondrial functions

Eukaryotic life emerged from prokaryotic life forms as early as 1.6 billion years ago. These eukaryotic life forms were the product of endosymbiotic events crucial to increase cellular complexity (Lane and Martin, 2010; Leander, 2020). One of these endosymbiotic events, specifically the engulfment of an α-proteobacterium 1.45 billion years ago, generated the organelles known as mitochondria. This independent origin of the mitochondria provides this organelle with unique features. First of all, it has two membranes, an outer membrane (OM) and an inner membrane (IM) separated by the intermembrane space. These two membranes contain the matrix, where the mitochondrial DNA (mtDNA) is contained. The mitochondrial genome is another unique characteristic of this organelle and a vestigial evidence of its prokaryotic past. Although mtDNA has been reduced along evolution due to nuclear transfer, it still provides mitochondria with most of the necessary proteins for one of its main roles, aerobic respiration. Aerobic respiration is what made mitochondria be named as “the powerhouses” of the cell. Indeed, mitochondria produce ATP through oxidative phosphorylation (OXPHOS). Mitochondria oxidize metabolic substrates obtained through glycolysis, fatty acid ß-oxidation or other catabolic pathways, producing electron carriers such as NADH and FADH through the tricarboxylic acid cycle (TCA-cycle). The electron carriers are the molecules that transfer the electrons to the electron transport chain (ETC) in the IM. The ETC is composed by five protein complexes named Complex I – V and transfer the electrons from Complex I to Complex IV pumping out protons to the intermembrane space in the process. The differential concentration of protons between the matrix and the intermembrane space generates an electrochemical membrane potential used by Complex V (ATP-synthase) to phosphorylate ADP into ATP. OXPHOS is a very efficient way to obtain energy in aerobic conditions. Importantly, electron leakage mainly through the ETC complexes I and III generate reactive oxygen species (ROS). ROS can generate damage to cell lipids, nucleic acids and proteins, but there are also multiple evidences that suggest that ROS are important signaling molecules for cellular homeostasis, regulating processes such as cellular differentiation, proliferation or the stress response upon hypoxia (Hamanaka and Chandel, 2010). Additionally, the metabolic activity of the mitochondria provides the cell with numerous metabolites that play a role in the regulation of multiple enzymes, including chromatin modifications such as histone acetylation (Martinez-Reyes et al., 2016).

Another important role of the mitochondria is the regulation of intracellular Ca^{2+}. Together with the ER, the mitochondrion is one of the organelles where Ca^{2+} can be

transported and accumulated. Indeed, mitochondria and ER are physically connected and its association allows the cell to buffer intracellular Ca^{2+} to prevent cell death, as Ca^{2+} overload is one of the main signals for apoptotic triggering (Lee et al., 2018).

Besides the metabolic activity, one of the major and most studied functions of mitochondria is the regulation of cell death through apoptosis. Mitochondria participate in the apoptotic process through the release of molecules, such as cytochrome c, from the intermembrane space to the cytoplasm. The release of these signaling molecules occurs when the integrity of the outer membrane is compromised through of a process called mitochondrial outer membrane permeabilization (MOMP). MOMP is regulated by a family of proteins called BCL-2, that comprises pro-apoptotic members like BAX and BAK, which interact and are inhibited by anti-apoptotic members, such as the BCL-2 like proteins (BCL-2, BCL-x_L, BCL-w and others) (Chipuk et al., 2006; Kalkavan and Green, 2018). The interaction of BCL-proteins is regulated by the integration of cellular signals. Upon cellular damage or prolonged stress, anti-apoptotic members are inhibited and its interaction with the pro-apoptotic member is impaired, allowing pro-apoptotic members to display their function, opening the pores of the mitochondrial outer membrane and triggering apoptosis **(Fig 6).**

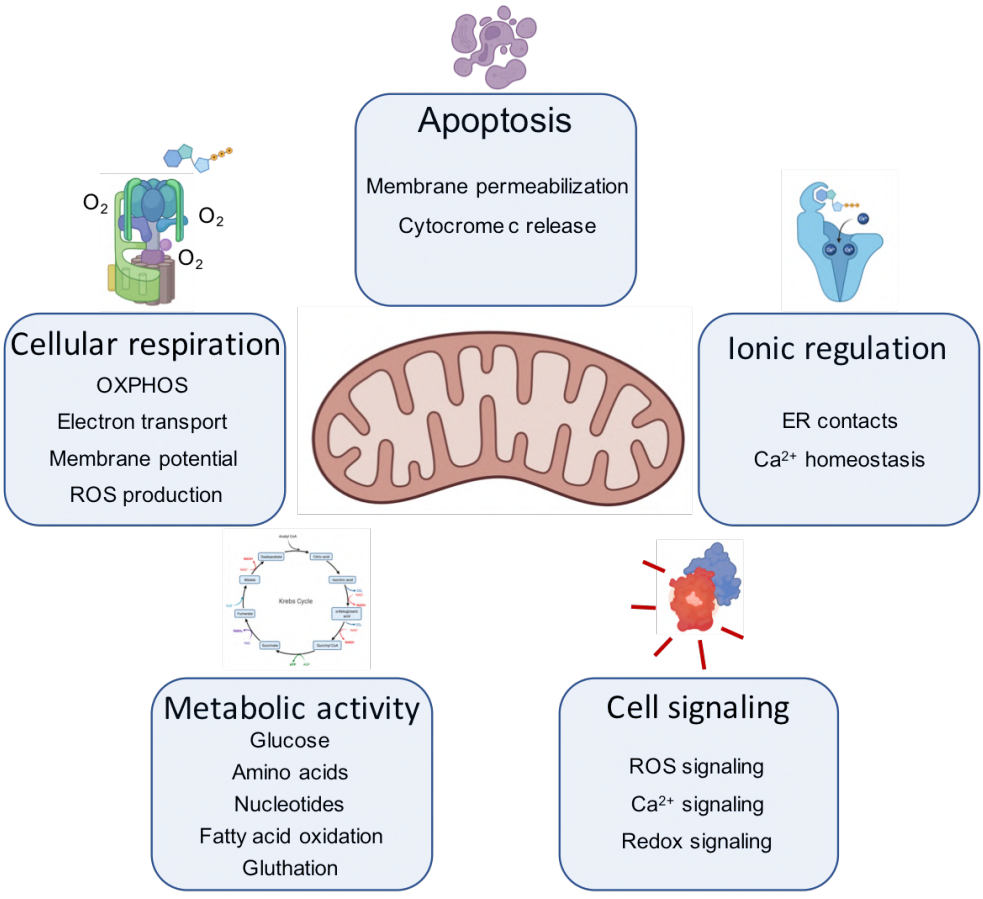


Figure 6. Summary of mitochondrial functions. Mitochondria have multiple functions that influence different aspects of cell biology. It regulates apoptosis through permeabilization of its membrane and releasing of Cytochrome c. The channels of the mitochondria and the interactions with the ER regulate ionic exchange. Mitochondria are also signaling hubs through export of molecules such as ROS or calcium. Additionally, mitochondria regulate different catabolic and anabolic processes of different cellular substrates. Finally, mitochondria produce energy through aerobic respiration.

As explained above, mitochondria are key regulators of multiple cellular processes. To adapt to cellular requirements and act as metabolic signaling centers, mitochondria form a complex network that changes dynamically and interacts with different organelles.

3.2. Mitochondrial dynamics

Mitochondria are highly dynamic organelles, and they adapt to cellular requirements changing their shape undergoing coordinated cycles of fusion and fission events. This is what has been named as mitochondrial dynamics.

3.2.1. Regulation of mitochondrial dynamics

The transitions that regulate the mitochondrial morphology network can be divided into two different and antagonistic events: mitochondrial fusion, when different mitochondria join together to form long tubular mitochondria and mitochondrial fission, when they fragment into smaller daughter mitochondria through a process similar to mitosis. Both processes are regulated by post-translational modifications of different large GTPases (Praefcke and McMahon, 2004).

Mitochondrial fusion takes place by fusion of the lipid membranes of the mitochondria, and different machineries are responsible for each of the two lipid membranes. The outer membrane fusion is the first to take place and begins with two approaching mitochondria tethering together. Mitochondrial outer membranes are negatively charged and spontaneous lipid fusion is not possible. Instead, a group of highly conserved GTPases named mitofusins, through their GTPase activity, regulate the binding of both lipid layers (Gao and Hu, 2021). In mammals, there are two different mitofusins, named MFN1 and MFN2, that have different roles but cooperate forming molecular complexes through homo and hetero-dymerization (Koshiba et al., 2004). These molecular complexes are further regulated by different post-translational modifications that favor or impair MFN function such as ubiquitylation, summoylation and phosphorylation, among others (Senft and Ronai, 2016). In normal conditions, fusion of both outer and inner membranes happens in a coordinated manner to complete the process. The main player for the inner membrane fusion is OPA-1. This protein undergoes proteolytic processing prior to gain its fusogenic activity, and then it bends the mitochondrial cristae of fusing mitochondria together, what ultimately fuse cristae (Ban et al., 2010; Rujiviphat et al., 2015).

Mitochondrial fusion and its machinery allows to distribute the mitochondria in the periphery of the cell, and coordinate the interaction with other organelles such as the ER. This is especially important for muscle and neuron performance, since they have unusually large cytoplasms and need to distribute mitochondria accordingly (Lewis et al., 2018; Mishra et al., 2015(Lewis et al., 2018; Mishra et al., 2015). Also, in muscle and neural tissues ER-mitochondria interaction is crucial for the maintenance of tissue function and, consequently, defects in mitochondrial fusion machinery are associated to severe pathologies in those tissues (Hernandez-Alvarez et al., 2019; Schneeberger et al., 2013; Sebastian et al., 2016). Finally, fusion is required to form a continuous network of elongated mitochondria that provides ATP in a very efficient way. Upon cellular stresses that induces highly demands ATP, such as fasting or low glucose conditions, mitochondrial fusion is critical (Gomes et al., 2011; Rambold et al., 2011).

The process of fusion is complemented by the antagonistic event: mitochondrial fission. As in the process of fusion, fission is also regulated by large GTPases of the Dynamin family that undergo post-translational modifications to perform their activity (Chang and Blackstone, 2010; Praefcke and McMahon, 2004). In mammals, mitochondrial

fission begins when Dynamin-related protein 1 (DRP-1), a cytoplasmatic protein, is recruited to mitochondrial outer membrane and interacts with mitochondrial outer membrane receptors, such us the mitochondrial fission factor (Mff), Fission protein 1 (Fis1), and mitochondrial dynamic proteins 49 and 51 KDa (MiD49 and MiD51) (Loson et al., 2013). Once DRP-1 is located in the outer membrane, it self-assembles with other DRP-1 monomers forming a curved ring that strangles and constrict mitochondria using its GTPase activity until both membranes are separated. Mitochondrial fission is an important mechanism for mitochondrial quality control, mitochondrial biogenesis and apoptosis.

Defects in the mitochondrial dynamics may lead to several pathologies and thus, these two processes are tightly regulated. Cellular regulation of mitochondrial dynamics is very context-specific, but nevertheless, major common regulators have been identified. The cAMP-dependent kinase A (PKA) is one of the main sensors of the second messenger cAMP, and integrates extracellular signals to phosphorylate different substrates. PKA has been shown to phosphorylate DRP1 in different residues depending on the cellular context (Cribbs and Strack, 2007). In addition, the MAPK that are activated to promote cell cycle and proliferation also promotes cellular fragmentation through DRP1 phosphorylation by extracellular regulated kinase (ERK) (Cook et al., 2017). Mitochondrial dynamics are also regulated to respond upon energy requirements by adenosine monophosphate- activated protein kinase (AMPK). This protein senses AMP/ATP ratio in the cell and it regulates mitochondrial dynamics, sometimes in opposing ways depending on cellular context, bioenergetics status or oxidative stress (Kang et al., 2016; Toyama et al., 2016; Wang et al., 2017). Many regulators of mitochondrial dynamics display their function in different contexts interconnecting with other pathways, such as sirtuins, mTOR pathway and HIF-1a (Lang et al., 2017; Li et al., 2019; Morita et al., 2017).

Altogether, mitochondrial dynamics regulation is a complex system of interconnected pathways that balances fusion and fission events in a highly-coordinated manner to maintain cellular homeostasis. However, this regulation is highly specific for cell type and cellular context, and this specificity should be taken into account when studying particular biological processes **(Fig 7)**.

3.2.2. Mitochondrial dynamics and mitophagy

As mitochondria are essential components of the cell, an adequate quality control is essential in order to maintain cellular homeostasis upon dysfunctional or defective mitochondria. This quality control is performed by a specific mitochondrial autophagy, named mitophagy, that eliminates and digest defective mitochondria in a coordinated manner. Additionally, mitophagy is used by the cells that need to reduce the mitochondrial mass in order to adapt to new metabolic requirements.

Mitochondrial dynamics is a key regulator of the mitophagy process, as mitochondria need to fragment to isolate dysfunctional regions of the network. At the molecular level, there are multiple mechanisms by which the cells can eliminate mitochondria through mitophagy and these differ depending on the stimuli and the cellular context (Ashrafi and Schwarz, 2013). These mechanisms can be classified into two different groups: ubiquitin-dependent and ubiquitin-independent mitophagy.

The ubiquitin-dependent mechanism is mainly regulated by PINK1-PARKIN axis. The dissipation of membrane potential stabilizes PINK1 protein in the OM, what leads to PARKIN translocation to the OM as well. PINK1 phosphorylates PARKIN, and the poly-Ubiquitin chains generated by PARKIN in the dysfunctional mitochondria acts as a signal for mitochondrial degradation (Khaminets et al., 2016; Pickles et al., 2018; Sekine and Youle, 2018). PINK1-PARKIN pathway indirectly activates DRP1, promoting the fission of defective mitochondria for degradation. In addition, this pathway triggers MFN proteasome

degradation to abolish mitochondrial fusion (Chen and Dorn, 2013). Interestingly, the degradation of mitofusins disrupts ER-mitochondria contact sites what further promotes autophagosomal formation for mitochondrial degradation (Gelmetti et al., 2017)

On the other hand, Ubiquitin-independent mitophagy requires the recruitment and stabilization of OM-proteins that directly interact with the autophagosome. These proteins, NIX, BNIP3 and FUNDC1 are regulated post-translationally and although they act through different mechanisms, most of them result in the recognition of the autophagosomal receptor LC3. These proteins, additionally, influence mitochondrial dynamics promoting mitochondrial fission. BNIP3, for example, promotes DRP1 and PINK1 stabilization in the OM, while FUNDC1 promotes DRP1 recruitment to the mitochondria (Palikaras et al., 2018; Quinsay et al., 2010; Wu et al., 2016).

In homeostatic conditions, mitophagy is a housekeeping activity of the cell used to recycle old and damaged mitochondria. However, under specific damage or stress-signals mitophagy is further promoted to cope with the different stresses inducing an acute mitochondrial clearance. Appropriate balance between mitochondrial biogenesis and mitophagy is essential for the cell to maintain the homeostasis, and mitochondrial dynamics play a critical role in the regulation of this balance

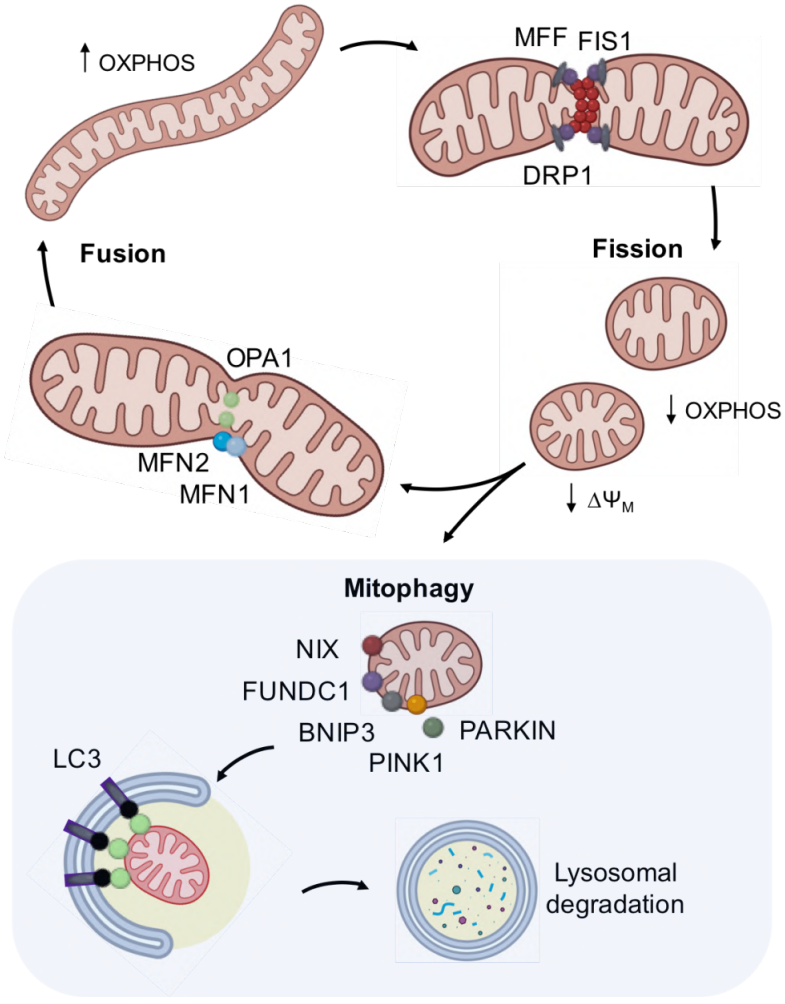


Figure 7. Mitochondrial dynamics regulation. Mitochondrial dynamics balances the number of mitochondria through cyclic fusion and fission events that are regulated through different proteins of the Dynamin GTPases family. Mitochondrial fission is also required for mitophagy elimination, where different proteins are recruited to the outer mitochondrial membrane to interact with the autophagosome that is later engulfed for lysosomal degradation.

3.2.3. Mitochondrial dynamics and cellular reprogramming

As discussed in sections 2.2 and 2.3 cellular reprogramming requires to modify the metabolic activity of cells undergoing de-differentiation. During the conversion to iPSC, mitochondria change their shape from a tubular interconnected network towards a more fragmented, individual-shaped mitochondrion. Accordingly, iPSC and ESCs cells have been shown to have low levels of Mfn1/2 (Prieto et al., 2016; Son et al., 2015).

Although the specific events of mitochondrial dynamics during cellular reprogramming are not completely understood, Prieto and collaborators have reported that early events of reprogramming activate proliferative signals like ERK1/2 that promotes the activating phosphorylation of DRP-1 and the inhibition of Mfn1/2 activity. These events combined ultimately lead to mitochondrial fragmentation through fission process (Prieto et al., 2016). Importantly, the inhibition of DRP-1 at early stages was sufficient to impair reprogramming, while Mfn1/2 deletion favors cellular reprogramming facilitating the transition from an OXPHOS phenotype to a glycolytic one.

Fission is also required for the elimination of mitochondria through mitophagy. It has been proposed that mitophagy is necessary during cellular reprogramming to reduce the mitochondrial mass, since less use of the OXPHOS is required (Xiang et al., 2017). In addition, some studies have demonstrated the importance of autophagy during cell reprogramming and stemness maintenance (Ma et al., 2015; Prigione and Adjaye, 2010; Wang et al., 2013b; Wu et al., 2015). Consistently, inhibition of autophagy during cellular reprogramming reduces the reprogramming efficiency due to the accumulation of abnormal mitochondria (Liu et al., 2016).

Altogether, mitochondrial dynamics is an important component of the metabolic switch necessary for cellular reprogramming to pluripotency and for stemness maintenance. Although much more research is needed to fully understand the process, the general idea is that mitochondrial fission takes place during pluripotency acquisition and mitophagy complements this fission to clear abnormal mitochondria and to reduce the mitochondrial network to single, globular-shaped mitochondria.

3.2.4. Mitochondrial dynamics and cancer

Similar to reprogramming, during carcinogenesis there is a metabolic switch from OXPHOS to a glycolytic phenotype. This is known in the cancer field as the Warburg effect. To accomplish this metabolic status, cancer cells follow a similar strategy to reprogramming cells regarding mitochondrial dynamics. It has been reported that cells from several cancer types display more fragmented mitochondria, together with an increase in DRP1 activation and/or MFN inhibition (Rehman et al., 2012; Serasinghe et al., 2015; Wan et al., 2014). Moreover, similarly to cellular reprogramming, several typical oncogenes are important regulators of metabolic activity. Oncogenic RAS activates the MAPK cascade, activating ERK1/2 which in turn induces the activation of DRP1 (Kashatus et al., 2015). On the contrary MYC oncogene, also a member of the 4 Yamanaka Factors, is a crucial regulator of mitochondrial biogenesis and fusion events in the cell that increase the OXPHOS activity. Although this seems to be counterintuitive, MYC expression favors mitochondrial biogenesis, and its function as a fusogenic factor is override by mitochondrial fission signals promoted by the MAPK (Vyas et al., 2016). These oncogenic events promoted by MYC generate a greater production of mitochondrial ROS, which in turn favors mitochondrial stress, the induction of mitochondrial fission and mitophagy (Graves et al., 2012; Vafa et al., 2002).

It is important to mention that, although the activation of DRP1 is required for cellular transformation, already transformed cancer cells can respond to nutrients availability or to different treatments by changing mitochondrial phenotype to a more interconnected mitochondrial network, indicating that cancer cells maintain their mitochondrial dynamics

capacity completely functional, at least to a certain extent (Caino et al., 2015; Rossignol et al., 2004).
Besides the oncogenic activation, cancer cells have deregulated their tumor suppression barriers to achieve neoplastic transformation. Accordingly, tumor suppressor genes also have an effect in the regulation of mitochondrial dynamics. P53 role in mitochondrial dynamics is currently being investigated, however, several studies suggest that p53 acts inhibiting the translocation of DRP-1 to the mitochondria, impairing the process of fission necessary during cellular reprograming and oncogenic transformation (Kim et al., 2021; Kim et al., 2020; Wang et al., 2013a). Finally, upregulation of AMPK, a downstream effector of LKB1 tumor suppressor, has been correlated with an impairment of neoplastic transformation and cellular reprogramming (as seen in 2.3.4), and it has also been described as a regulator of mitochondrial dynamics (as seen in 3.2.1), although its function as tumor suppressor highly depends on the cellular context (Liang and Mills, 2013)

Collectively, the regulation of mitochondrial dynamics is crucial for both, cellular reprogramming and neoplastic transformation, since both processes require changing their original identity and to adapt to new metabolic and proliferative needs.

Objectives

1. Identify novel microproteins related with celular plasticity and cancer.

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

3. Characterize the molecular functions of the microprotein, including its subcellular localization and the regulation of its expression.

4. Study the role of the microprotein in cellular plasticity assays, such us cellular reprogramming and oncogenic transformation.

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

Materials & Methods

1. Mouse experiments

1.1. Mouse housing

Mice were housed at the specific pathogen-free (SPF) barrier area of the Vall d'Hebron Institute of Oncology (VHIO) in Barcelona. Animals coming from other laboratories were re-derived to SPF conditions. All animal procedures were approved by the Animal Care and Use Ethical Committee of animal experimentation (CEEA) of the Vall d'Hebron Research Institute (VHIR) and the Catalan Government and performed according to European legal framework for Research animal use and bioethics. Observation of the animals was performed daily and euthanized upon signs for humane endpoint.

1.2. Animal models

1.2.1. Reprogrammable mouse (i4F) model.

We used the reprogrammable mouse line known as i4F-B generated by María Abad (Abad et al., 2013), which carries a ubiquitous doxycycline-inducible OSKM transgene, abbreviated as i4F, inserted into the Pparg gene, and the rtTA transactivator under the control of the Rosa26 promoter. This mouse strain was used to obtain mouse embryonic fibroblasts.

1.2.2. RCAS/GFAP-tva glioblastoma model.

Transgenic mice that express the Replication-Competent Avian leukosis virus Splice acceptor (RCAS) vectors receptor (*tva*) under the control of GFAP promoter (Ozawa et al., 2014). This model was kindly provided by Massimo Squatrito (Principal Investigator of Seve-Ballesteros Foundation Brain Tumor Group (CNIO, Madrid)).

1.2.3. MASALA-deficient mouse model

This animal model was generated by Sagrario Ortega (Head of the Transgenic Mice Unit at CNIO, Madrid). It is a CRISPR- knock-in, that was generated using a homologous recombination template that include the mutation of the MASALA start codon to a stop codon. Both, a plasmid coding the Cas9 and the gRNA and the homologous recombination template were delivered by microinjection in mouse embryos.
The homologous template was designed including an Age I restriction site, so the modified allele (KI allele) could be detected by enzymatic digestion. Briefly, DNA was extracted from the tail and *Smim10l2a* locus was amplified by PCR using *Smim10l2a_locus* pair of primers (Table X). Once amplified, the PCR product was incubated with AgeI restriction enzyme (New England Biolabs) overnight at 37ºC and run in agarose gel to test digestion. Separately, the PCR product was treated with ExoSAP-IT™ PCR Product Cleanup Reagent (ThermoFischer) following manufacturer instructions and sequenced by Sanger Sequencing to confirm the modification.
The successfully modified animals were used as founders to generate MASALA-deficient colony (MASALA_KO)

1.2.4. Mice genotyping

DNA was extracted from ear-punch tissue fragment and genotyped by PCR using specific primer pairs (Table 2).
For MASALA-KO colony, specific forward primers were designed to detect wild-type and knock-in alleles (MASALA_WT and MASALA_KO) together with a common reverse primer (MASALA_genotyping_Rvs).
Reprogrammable (i4F) mice were genotyped using OSKM and rtTA pair of primers.
GFAP/tva mice were genotyped using *Tv-a*_transgene pair of primers.

1.3. In vivo experiments

1.3.1. Xenograft of transformed NIH3T3 cells

NIH3T3 cells previously transduced with pLIX403-MASALA-HA or its empty vector were transduced with pLXSN-HRAS^{v12} retroviral vector and selected with geneticin at 500µg/ml. MASALA overexpression was induced using doxycycline at 1µg/ml from the first day of selection. Once selected, cells were trypsinized and counted and $1 \cdot 10^6$ cells were injected subcutaneously into athymic-nude mice in both flanks. Tumor growth was monitored every 2 days and scored using the formula height × width × width × (π/6) and animals were sacrificed when tumors reached 1.2 cm^3.

1.3.2. Glioblastoma formation with RCAS/ GFAP-tva model

DF-1 cells were transfected with RCAS viral vectors encoding the glioblastoma oncogenic drivers PDGFA, shRNA-p53, shRNA-Nf1 and an additional vector containing MASALA-HA or the empty vector as control. As RCAS vectors auto-transduce DF-1 cells in cell culture, cells were left for several passages until 100% transduction of the plate was achieved.
Newborn mice (P1-P3) were injected intracranially with 200,000 DF-1 cells/mouse (50,000 cells for each one of the vectors) or 20,000 (5,000 cells/vector) using a Hamilton syringe. Mice were monitored every 2 days for weight loss or neurological symptoms (seizures, aggressiveness or unbalance). Humane end-point was stablished upon neurological symptoms appearance or 20% weight loss

2. General Cell Culture and treatments

2.1. Cell Culture Conditions

HEK293T, NIH3T3, HCT116, DF-1 and MEFs cells were cultured in DMEM with GlutaMAX supplemented with 10% of fetal bovine serum (FBS) and 1% of Penicillin-Streptomycin (P/S) (Gibco). A549 cell line was cultured with RPMI (Gibco) supplemented with 10% FBS and 1% P/S. Mouse induced pluripotent stem cells (iPSCs) were cultured over feeder layers in DMEM GlutaMax supplemented with 15% FBS (or KO serum for cellular reprogramming experiments), 50 mM ß-mercaptoethanol, 1% NEAA (Invitrogen), 1% P/S and 1000 U/ml LIF (ESGRO, Chemicon). All cells were incubated at 37ºC with 5%CO_2.

2.2. Treatments

To induce expression of Tet-on inducible systems, cells were treated when indicated with Doxycycline (Sigma Aldrich) at 1µg/ml.
To induce genotoxic stress and/or p53 activation, cells were incubated with 1µM doxorubicin (Sigma Aldrich) or 10µM nutlin-3a (Sigma Aldrich) for 24 hours.
For apoptosis induction cells were incubated with 4µM staurosporine (Sigma Aldrich) for 4 hours or 50µM Navitoclax (MedChem Express) for 1.5 hours.
To induce mitochondrial stress, cells were incubated with 40µM Carbonyl cyanide-*p*-trifluoromethoxyphenylhydrazone (FCCP) for 4 hours or 100µM Hydrogen Peroxide (H_2O_2) for 30 minutes.
Protein stabilization experiments were performed incubating cells for the indicated timepoints with Cycloheximide (Merck) at 100µg/ml to inhibit protein translation.

2.3. Mouse embryonic fibroblasts (MEFs) isolation

Mouse embryos were extracted at 13.5 dpc and suspended in PBS supplemented with 1% P/S. The head and fetal liver were removed and then, embryos were chopped and incubated in 0.2% Trypsin for 20 minutes at 37ºC in a cell culture incubator, pipetted up and down 20 times, and incubated for 20 more minutes at 37ºC. Trypsin was inactivated adding DMEM GlutaMAX supplemented with 10%FBS and 1%P/S, and cells were cultured in normal conditions or frozen.

2.4. Isolation of mouse iPSC lines

Individual iPSC colonies derived from cellular reprogrammed MEFs were picked using a 25G syringe and transferred into a feeder-layered 96-well plate. Once confirmed that a single colony was transferred to the well, colonies were trypsinized and reseeded into a feeder-layered 96-well plate for amplification.

2.5. Generation of feeders

Exponentially growing MEFs were treated with 1µg/ml Mitomycin C (provided by Vall d'Hebron Hospital) during 2.5 hours at 37ºC. Then, cells were washed, trypsinized, counted and frozen in cryotubes for storage or seeded immediately in 0,1% gelatin-coated plates.

3. Cloning Procedures

MASALA mouse and human ORFs were synthesized (IDT technologies) fused with a flexible linker (3xGGGGS) and an HA tag epitope at the C-terminal part of the microprotein and flanked by *EcoRI* enzyme restrictions sites at both ends (Complete Sequences can be found in Table 4. After enzymatic digestion, constructs were ligated into the pENTR1A vector (Addgene).
For the lentiviral vectors, the MASALA-HA tag construct was obtained by recombining donor pENTR1A-MASALA-HA vector with the lentiviral inducible system pINDUCER20 (Invitrogen) or the pLIX-403 inducible system using the Gateway Cloning Technology LR clonase (Invitrogen), following manufacturer's instructions.
RCAS viral vectors containing the glioblastoma oncogenic drivers (PDGFA, shRNA-p53, shRNA-Nf1) and RCAS backbone were kindly provided by Massimo Squatrito. To produce RCAS-MASALA-HA viral vector pENTR1A- MASALA-HA donor vector was recombined

into RCAS empty backbone using Gateway Cloning Technology following manufacturer's instructions. A complete list of all the plasmid used can be found in Table X.

4. Retro- and lentiviral infections

HEK293T cells were used as packaging cells and transfected with the lenti- or retroviral plasmids and the packaging plasmids indicated in Table X using Fugene HD (Promega) or Polyethylenimine (PEI) (Polyscience Europe GmbH) reagent 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 400-600µg/ml (pINDUCER-20) or puromycin (VWR) selection 1-2µg/ml (pLIX403) in the case of inducible vectors.

5. Protein analysis by Western blot

Cells and tissues were homogenized in 2% SDS lysis buffer (50 mM Tris-HCl, pH 8.0, 1 mM EDTA, 2% SDS) supplemented with protease (Roche) and phophatase (SigmaAldrich) inhibitors cocktails except when indicated otherwise. Protein concentration was determined using the Pierce™ BCA Protein Assay Kit (Thermo Fisher). Lysates were loaded in 12% bis-tris acrylamide gels and transferred to nitrocellulose membranes. Primary antibodies were incubated overnight at 4° C. Secondary HRP-conjugated antibodies were incubated the following day for 1 h at room temperature, and ECL Prime Western Blotting Detection Reagent (Fisher Scientific) or SuperSignal™ West Dura Extended Duration Substrate were used as a chemiluminescent reagent for protein detection. Antibodies and dilutions are listed in Table X.

6. 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 Extended Data Table 5 and 6. Cycle threshold (Ct) values were normalized to GAPDH.

7. Ribosome profiling analysis

We retrieved a public ribosome profiling dataset from human and mouse brain (ArrayExpress accession number E-MTAB-7247) (Wang et al., 2020b) and we adapted a computational approach to identify translated sORFs (Ruiz-Orera et al., 2018) .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) and human genome (hg38). Next, mapped reads from experimental replicates were merged and we used the ribORF algorithm (Ji et al., 2015) to predict translated sORFs with significant read uniformity and frame periodicity (score ≥0.7), as this feature is indicative of active ORF translation.

8. MASALA interactome analysis by mass-spectrometry

NIH3T3 cells overexpressing MASALA-HA or the Empty vector for 24 hours were lysed in a buffer containing 50 mM Tris-Hcl pH 7.5–8, 150 mM NaCl, 1%Triton X-100 and protease and phosphatase inhibitors and homogenized for 30 min in a rotor wheel. 3 mg of lysates were immunoprecipitated with 5 µg of monoclonal HA-tag antibody (Sigma) overnight at 4° C. Immunocomplexes were collected using PureProteome™ Protein A Magnetic Beads (MERCK) and eluted by competition incubating with synthetic HA peptide (Sigma) 5 hours at room temperature. Eluate was digested with trypsin and analyzed by liquid chromatography-mass spectrometry 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 mouse proteins). 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 Immunoprecipitate and control groups were considered significantly differential.

9. Immunofluorescence

9.1. Immunofluorescence of cultured cells

Cells were seeded in coverslips (Sigma Aldrich) previously coated with fibronectin (Santa Cruz Biotechnologies). When desired, cells were fixed in 4% paraformaldehyde for fifteen minutes and permeabilized with 0.5% Triton X-100 for seven minutes at room temperature. Blocking step was made in 3% Bovine Serum Albumin (BSA) for one hour. Cells were incubated overnight at 4°C with the primary antibody diluted in blocking buffer. Next day, secondary antibodies were incubated for one hour at room temperature in the dark. Finally, cells were mounted in Prolong Mounting Medium with DAPI (Invitrogen). Images were taken in a Nikon Eclipse Ti-E inverted microscope system or Nikon C2 Plus Confocal Microscope. Antibodies and dilutions are listed in Table 2.

9.2. Immunofluorescence in tissues

Extracted organs were fixed in 4% Paraformaldehyde for 24 hour and then incubated in 30% Sucrose for 24-72h. Fixed organs were embebed in OCT Tissue-Tek® (Sakura Finetek) and frozen at -80ºC. 5µm cryosections were permeabilized incubating in 3%Triton X-100 for 10 minutes at room temperature. Blocking was made with 3%BSA and 10% goat serum incubation for 1 hour at room temperature. Primary antibodies were incubated in 3% BSA and 10% goat serum overnight at 4ºC. Next day, secondary antibodies were incubated for 1 hour at room temperature in the dark. Tissue sections were mounted in Prolong Mounting Medium with DAPI (Invitrogen). Images were taken in a Nikon C2 Plus Confocal Microscope.

10. Cellular Reprogramming

MEFs were transduced with Tet-O-FUW-OSKM inducible lentiviral vector or FUGW-GFP vector as transduction control together with pINDUCER20-MASALA-HA or its empty vector. Cells were selected with geneticin at 400µg/ml (pINDUCER20). Transduction efficiency was assessed by Flow Cytometry using the FUGW-GFP infected cells. After selection, MEFs were plated in a 6-well plate at $5 \cdot 10^5$ cells/well. Once attached, OSKM cassette and MASALA expression was induced by treating with Doxycycline at 1µg/ml in cellular reprogramming medium. Medium was changed every other day during 12-14 days, until iPSC colonies were evident.
For reprogramming experiments without MASALA overexpression, MEFs isolated from i4F mice were used. In this case, i4F MEFs were directly seeded at $5 \cdot 10^5$ cells/well and induced with Doxycycline at 1µg/ml in cellular reprogramming medium. Medium was changed every other day during 12-14 days, until iPSC colonies were evident
To determine cellular reprogramming efficiency in both cases, mIPSC colonies were quantified manually after alkaline phosphatase staining (Promega) following manufacturer instructions.

11. Oncogenic Transformation

MEFs previously transduced with pINDUCER20-MASALA-HA or empty vector were transduced with pLPC-E1a-IRES-HRAS^{v12} retroviral vector and selected with puromycin 1.5µg/ml. Once selected, $1.2 \cdot 10^5$ cells/plate were seeded in 60mm plates. Overexpression of MASALA was induced with 1µg/ml doxycycline treatment refreshed every other day for 15-20 days. Transformation efficiency was quantified manually by counting transformation foci using crystal violet staining.

12. ROS quantification

12.1. Total ROS

MEFs previously transduced with pINDUCER20-MASALA-HA or its empty vector were seeded in a black 96-well plate at 10^4cells/well density. Then, cells were induced for MASALA overexpression with doxycycline at 1µg/ml for 24h. ROS accumulation was quantified using 2′,7′-Dichlorofluorescin diacetate (DCFH-DA) fluorescent probe. Briefly, cells were incubated with DCFH-DA 10µM for 1h at 37ºC, then washed with PBS and fluorescence was quantified every 30 minutes up to 3 hours with a Spark 10M (TECAN) microplate reader.

12.2. Mitochondrial ROS

$5 \cdot 10^5$ cells were seeded in a 24-well plate and treated with 1µg/ml doxycycline for 24 hours (for MASALA overexpression). Cells were incubated with MitoSOX fluorescent probe (Thermo Fischer) at 2.5µM in Hank's balanced salt solution (HBSS) for 15 minutes at 37ºC, then washed with PBS, trypsinized and resuspended in FACS buffer (PBS + 0.1%BSA + EDTA 1mM). For fluorescence quantification $2 \cdot 10^5$ events were acquired using Navios Cytometer (Beckman Coulter).

13. Mitochondrial membrane potential quantification

MASALA overexpression was induced for 24 hours in MEFs previously transduced with pINDUCER20-MASALA-HA or the empty vector as control. Cells were incubated with Tetramethylrhodamine methyl ester (TMRM) fluorescent probe at a final concentration of 100nM for 30 minutes at 37ºC. We used cells treated with 20µM Oligomycin as negative control and cells treated with FCCP 30µM as positive control. After treatment, cells were washed twice with PBS and suspended in FACS buffer. Fluorescence quantification was performed in Navios Cytometer (Beckman Coulter).

14. Oxygen consumption rate (OCR) measurement

OCR was measured using Seahorse XFe24 Mito Stress Assay (Agilent). The day prior to the experiment cells were trypsinized, counted and seeded in Seahorse XFe24 cell culture plates at $4 \cdot 10^4$ cells/well (overexpression experiments) or $4.5 \cdot 10^4$ (for loss of function experiments). Next day, cells were analyzed in Seahorse XFe24 Analyzer following Mito Stress Assay manufacturer instructions using 1.5µM Oligomycin, 2µM FCCP and 0.5 µM Rotenone + Antimycin A. After the measurements, cells were fixed in methanol and counted manually for normalization.

15. Mitochondria-enriched fraction isolation

For mitochondrial enrichment, we adapted the protocol described previously in(Kojima et al., 2016). A549 cells previously transduced with pLIX-403-MASALA-HA or its empty vector and were induced for MASALA-HA overexpression for 24 hours. Then, cells were suspended on ice-cold HKSS buffer [70 mM sucrose, 220 mM D-Mannitol, 10 mM KAc, 20 mM HEPES-KOH pH 7.5, protease inhibitors] and disrupted with 25-30 passages of the pestle in a Douncer homogenizer on ice. Resulting lysate was centrifuged 10 minutes at 800xG and 4ºC to remove debris. Part of the supernatant was taken as total fraction and the rest of the lysate was further centrifuged at 10,000x G for 25 minutes at 4ºC to obtain the mitochondrial-enriched fraction in the precipitate and the cytosolic fraction in the supernatant. Cytosolic fraction was taken apart and mitochondrial-enriched fraction was washed by resuspension in HKSS buffer and centrifugation 10,000x G for 25 minutes at 4ºC. Then, supernatant was discarded and mitochondrial-enriched precipitate was resuspended in 2%SDS protein extraction buffer (described before). Fraction extracts were quantified for protein concentration and immunoblotted.

16. Protease K protection assay

A549 cells previously transduced with pLIX-403-MASALA-HA or its empty vector were induced for MASALA-HA overexpression for 24 hours. Mitochondrial-enriched fraction was extracted as described before but the final resuspension was done in two different buffers for each condition. For intact mitochondria, precipitate was resuspended in HKSS buffer without protease inhibitors and for mitoplast generation, precipitate was resuspended in hypotonic buffer HK without protease inhibitors [10mM KAc, 20mM HEPES-KOH pH 7.5]. Then, the two conditions (MASALA and control) were separated into three different tubes for Protease K treatment with its positive and negative controls. Corresponding tubes were incubated with Protease K at 12.5µg/ml for test condition or Protease K 12.5µg/ml + Triton X-100 1% final concentration in the case of the positive control during 12 minutes at 37ºC. Negative control tubes were also incubated at 37ºC for 12 minutes without any treatment. Protease K was inactivated by adding

phenylmethylsulfonyl fluoride (PMSF) to the tubes at a final concentration of 5μM. Finally, lysates were quantified for protein concentration and immunoblotted

17. *In silico* analyses

17.1. Codon conservation

SMIM10L2A coding potential was assessed using PhyloCSF (Lin et al., 2011), a comparative genomics algorithm that analyzes a multispecies nucleotide sequence alignment and score it according to phylogenetic codon conservation.

17.2. lncRNA structure prediction

Smim10l2a secondary RNA structure was predicted using MFold web-server software (Zuker, 1994, 2003) using the cDNA sequence.

17.3. Protein features prediction

The different characteristics of the protein were predicted using publicly available prediction software in different web-servers. Using amino acid sequence of MASALA, transmembrane domain was predicted by TMHMM (Krogh et al., 2001) and target peptide of the N-terminal region was determined by TargetP (Almagro Armenteros et al., 2019). Three dimensional protein-structure prediction was obtained with C-Quark software (Mortuza et al., 2021) from the amino acidic sequence. Subcellular protein localization prediction was determined using DeepLoc 1.0 software (Almagro Armenteros et al., 2017) for both mouse and human versions of MASALA. Probability of a mitochondrial targeting presequence and mitochondrial processing peptidase cleavage site were predicted using MitoFates web-server software (Fukasawa et al., 2015). Finally, protein membrane topology was predicted using Protter (Omasits et al., 2014).

17.4. Image analysis

17.4.1. MASALA colocalization analysis

NIH3T3 cells transduced with pINDUCER20-MASALA-HA or the empty vector were induced 24 hours for overexpression of the construct. Cells were processed for immunofluorescence as indicated above using anti-HA-tag antibody (Sigma) 1:150 dilution and anti-AIF antibody (Santa Cruz Biotechnologies) at 1:100 dilution. Images were acquired in Nikon C2 Plus Confocal Microscope at 60X magnification and analyzed using ImageJ Coloc-2 software for signal correlation analysis.

17.4.2. Mitochondrial morphology analysis

MEFs previously transduced with pINDUCER20-MASALA-HA or the empty vector were further transduced with pDsRed-Mito for mitochondria visualization. Then, cells were treated with doxycycline for the indicated times. 100 cells per timepoint were analyzed and classified according to their mitochondrial morphology. Cells were classified according to the mitochondrial phenotype in three different categories: Fragmented when the majority of the cell presented single globular mitochondria, tubulated when the majority of the mitochondria presented a tubular, networked shape and intermediate when cells presented a mixed phenotype between the two. Results **Fig 38** for visual reference.

18. Transmission electron microscopy

MEFs were collected in cold Phosphate buffer (0.1M pH7.4). Then were fixed in 2% Paraformaldehyde, 2,5% glutaraldehyde in phosphate buffer for 30 minutes on ice and resuspended in phosphate buffer (0.1M pH7.4) for storage at 4ºC.
Cell pellet was then treated with 1% Osmium tetraoxyde and 0.8% potassium fericyanide in fixing solution for 2h (4ºC). Next, cells were washed with MiliQ water and dehydrated with increasing concentrations of acetone at 4ºC.
Inclusion of the samples was perfomred in Epoxy resin (Spurr) and survey cuts were performed using a glass blade in an ultramicrotome (Leica). Once the area of interest was selected, final cuts were performed in an ultramicrotome (Leica) using a diamond blade (Diatome), collecting the cuts in a copper grid. Final cuts were contrasted with uranyl acetate and lead citrate.
Images were aquired by Dr. Miguel Angel Lafarga Coscojuela at SERMET (Universidad de Cantabria) using a Jeol Jem 2100 electron microscope.

19. Statistical analysis

Data is presented as mean ± SD or mean ± SEM unless indicated otherwise. In data obtained from biological replicates n correspond to individual animals or individual MEF preparations. Statistical significance was evaluated with GraphPad Software using Student's t-tests with Welch correction or two-way Anova, unless indicated otherwise. Significant differences were established based on P-value (ns= $p > 0.05$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$).

Table 1. Antibodies

Target	Reference	Dilution WB	Dilution IF
HA-Tag	Ab9110, Abcam	1:2500	
HA-Tag	H6908, Sigma		1:150
GAPDH	AM4300, Thermo Fisher Scientific	1:10000	
MASALA	Proteogenix	1:200	1:10
AIF	SC-13116, Santa Cruz		1:250
p53	SC-126, Santa Cruz	1:500	
p21	MS-891-P0, Thermo Scientific		
Tubulin	SC-32293, Santa Cruz	1:1000	
VDAC-1	SC-390996, Santa Cruz	1:500	
TOMM20	GTX133756, GeneTex	1:1000	
TIMM23	GTX66539, GeneTex	1:500	
SUCLG2	GTX107002, GeneTex	1:10.000	
SERCA2	NB300-581, Novus Biologicals		1:100
DRP1	8750, Cell Signaling	1:1000	
pDRP1(S616)	3455, Cell Signaling	1:1000	

MFN1+2	ab57602, Abcam	1:2000	
OPA1	612606, BD Biosciences	1:1000	

Table 2. Primers

Human primers

Primer Name	Forward	Reverse
GAPDH	GGACTCATGACCACAGTCCATGCC	TCAGGGATGACCTTGCCCACAG
SMIM10L2A	CGGCGACTTGACTTTCCAG	AGGCCCTTCTTCTCCTCTTG
CDKN1A	TGTCACTGTCTTGTACCCTTG	GGCGTTTGGAGTGGTAGAA
CCND1	ATGTTCGTGGCCTCTAAGATGA	CAGGTTCCACTTGAGCTTGTTC

Mouse primers

Primer Name	Forward	Reverse
mGAPDH	TGTGTCCGTCGTGGATCTGA	TTGCTGTTGAAGTCGCAGGAG
Smim10l2a	TACGAACAGAACAGGAGGAC	CTCTCCACATTGTGGCTTTG
Smim10l2a_locus	GGGTGGAGCATTGAAGGCT	CATCACTGAGGCCACCATGT
MASALA_WT	GCGGTGAGCCAGGCCAATGG	
MASALA_KO	GCGGTGAGCCAGGCTATAAC	
MASALA_genotyping_Rvs	GCCGCAGTCACCGGTTATAG	
OSKM	ACTGCCCCTGTCGCACAT	CATGTCAGACTCGCCAGGTG
rtTA	AAAGTCGCTCTGAGTTGTTAT	GGAGCGGGAGAAATGGATATG
	GCGAAGAGTTTGTCCTCAACC	
*Tv-a*_transgene	CAGATTTGAGAGCTCATGTCCA	ACGGACAACGGCACAGAG

Table 3. Plasmids

Retroviral packaging vectors

pCL-Ampho
pCL-Eco

Retroviral vectors

pLP-HRASV12-IRES-E1A

Lentiviral packaging vectors

pLP1
pLP2
pLP- VSVG

Lentiviral vectors

pFUW-TetO-OSKM
pFUGW
pLIX-403
pINDUCER20
pENTR1A
pDsRED-mito

RCAS vectors

RCAS-Y (empty vector)
RCAS-MASALA-HA
RCAS-PDGFA
RCAS-shp53
RCAS-shNf1

Table 4. Overexpression Constructs

Human MASALA-HA

*taagcaGAATTCCTCGAGCGCCACC*ATGGCGGCGTCGGCGGCTCTGTCTGCAGCGGCGGCTGCG
GCGGCCCTGTCTGGCCTGGCGGTGCGGCTGTCGCGCTCAGCTGCGGCCCGAGGCTCGTAC
GGCGCCTTCTGCAAGGGGCTCACGCGCACGCTGCTCACCTTCTTCGACCTGGCCTGGCGGC
TGCGCATGAACTTCCCCTACTTCTACATCGTGGCCTCGGTGATGCTTAACGTCCGCCTGCAGG
TGCGGATCGAGGGCGGTGGTGGCAGTGGTGGCGGAGGAAGCGGTGGGGGAGGCAGCTATC
CGTATGATGTGCCGGATTATGCGTGA*GAATTCtgctta*

Mouse MASALA-HA

*taagcaGAATTCCTCGAGCGCCACC*ATGGCGGTGACTGCGGCCCTCTCTGCTGCTGCTGCGGCGG
CGGCCCTGTCTGGCCTGGCAGTGCGACTGTCGCGTTGGGCGGCGACACGCAGCTCCTACGG
TGCGTTCTGCAAGGGGCTCACGCGCACGCTACTCACCTTCTTCGACCTAGCCTGGCGCCTGC
GCGTGAACTTCCCCTACTTCTACATGGTGGCCTCAGTGATGCTCAACGTCCGCCTGCAGGTG
CGGATCGAGGGCGGTGGTGGCAGTGGTGGCGGAGGAAGCGGTGGGGGAGGCAGCTATCCG
TATGATGTGCCGGATTATGCGTGA*GAATTCtgctta*

Results

1. Identification of MASALA microprotein

1.1. The lncRNA SMIM10L2A encodes for a 78-amino acid microprotein.

To identify novel microproteins involved in cancer cell plasticity we followed a computational approach. We focused our search on lncRNAs-encoded microproteins, given that many lncRNAs are annotated and we could find information about their expression and regulation. Additionally, many lncRNAs are expressed in a tissue specific manner (Cabili et al., 2011), what may help to understand their biological role.

Using previous literature and public databases, we selected lncRNAs that are deregulated in cancer and are regulated during cellular reprogramming/differentiation processes **(Fig 8A)**.

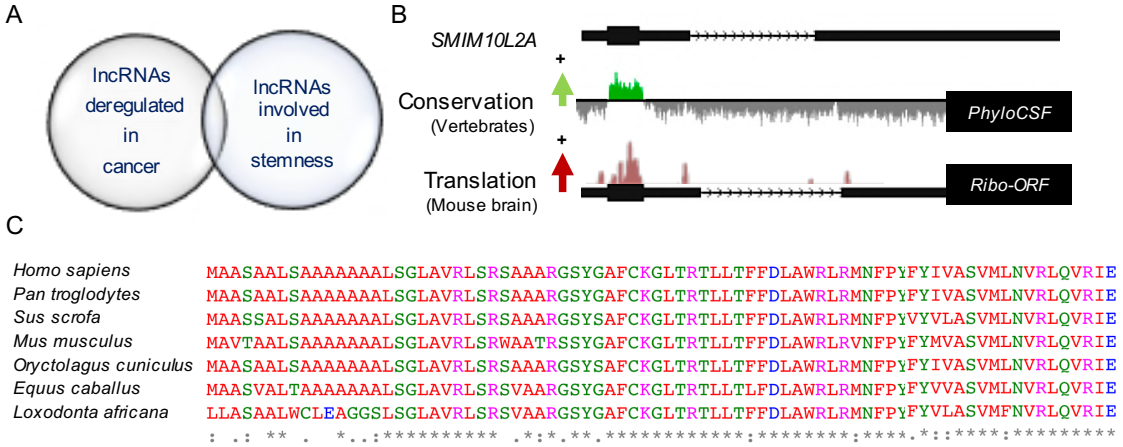


Figure 8. *SMIM10L2A* is a lncRNA with a highly conserved sORF and is actively translated. **(A)** Strategy for the discovery of novel microproteins involved in cancer cell plasticity. **(B)** Analysis of codon conservation across *SMIM10L2A* locus by PhyloCSF software, which shows a positive deflection (in green) corresponding to MASALA sORF. In red, representation of Ribo-Seq performed in mouse brain and analyzed with RibORF. **(C)** Amino acid conservation of *SMIM10L2A* sORF across placental mammals.

Then, we look for the presence of evolutionary conserved sORFs in those lncRNAs. For that purpose we used PhyloCSF, a comparative genomics algorithm that quantifies the codon conservation across 35 vertebrate species (Lin et al., 2011). Following this approach, we found that *SMIM10L2A (LINC00086)*, a gene annotated as a lncRNA, contains a sORF highly conserved across placental mammals **(Fig 8C)**. To validate the coding potential of our candidate, we analyzed already published ribosome profiling experiments performed in brain (Wang et al., 2020b) using RiboORF, a software that evaluates reads per frame and read uniformity (Ruiz-Orera et al., 2018). This analysis revealed that *SMIM10L2A* is translated in mouse and human brain, encoding a 78-amino acid microprotein **(Fig 8B)**. We have named this new microprotein MASALA.

We performed different *in silico* analyses to determine the main characteristics of MASALA microprotein. Using Target P (Almagro Armenteros et al., 2019), we observed that MASALA has a predicted targeting peptide at the N-terminal region from the start M1 to R22, where there is a Mitochondrial Processing Peptidase (MPP) cleavage site **(Fig 9A)**.

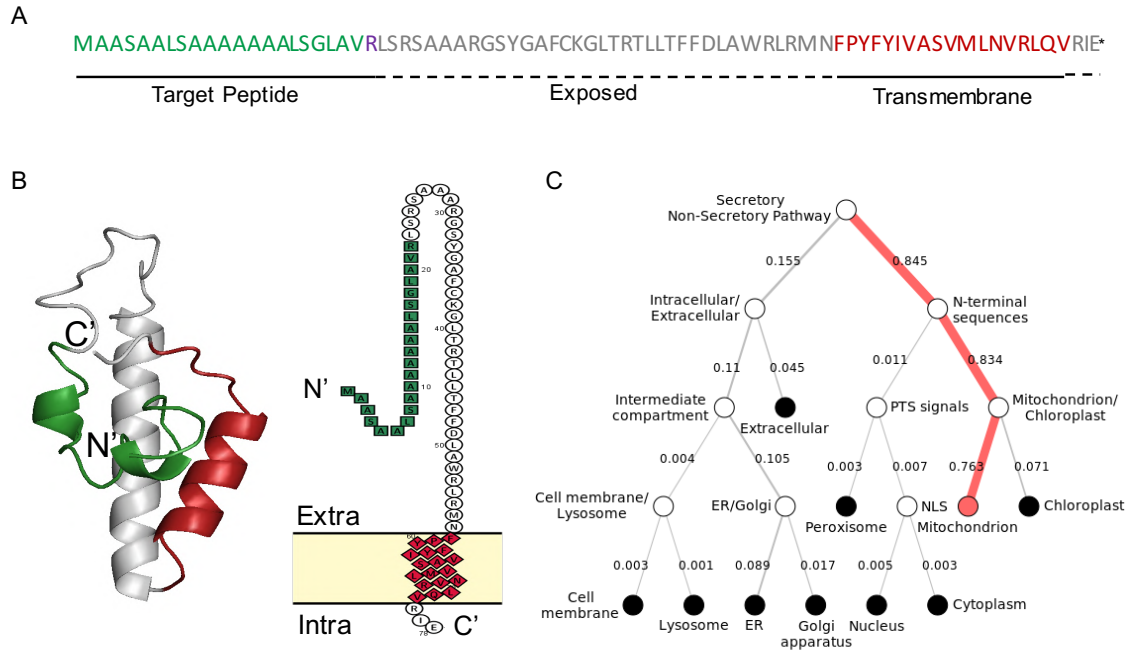


Figure 9. *In silico* characterization of MASALA microprotein. (A) Main features of the protein over its primary structure. Green color represents the targeting peptide in N-terminal domain and red color represents the transmembrane domain in the C-Terminal. Purple "R" amino acid corresponds to mitochondrial peptidase cleavage site at R22. **(B)** Predicted tertiary structure of MASALA by C-Quark and its topology within a membrane predicted by Protter. Green color represents N-terminal targeting peptide and red color represents the C-Terminal transmembrane domain. **(C)** Predicted mitochondrial localization of MASALA using DeepLoc 1.0. Numbers in the branches indicate the probability of the protein being at that compartment.

Additionally, MASALA has a predicted transmembrane domain at the C-terminal part (F58-V75). Protter software (Omasits et al., 2014) predicts that MASALA's N-terminal domain is oriented towards the outer part of the membrane **(Fig 9B).**

Finally, to predict the subcellular localization of the microprotein we used DeepLoc 1.0 predictor (Almagro Armenteros et al., 2017), which predicted MASALA to be located in the mitochondria with a 0.76 probability **(Fig 9C)**. Moreover, another mitochondrial presequence predictor (MitoFates, (Fukasawa et al., 2015)) identified a putative mitochondrial presequence with a probability of 0.796 before the mitochondrial peptidase cleavage site at R22.

1.2. Generation of different tools to study MASALA microprotein

To further characterize the function of the microprotein we generated gain- and loss-of function tools. To asses MASALA gain of function, we generated a MASALA-overexpressing vector by cloning the sORFs of the human or murine MASALA tagged with an HA **(Fig 10)** into the pINDUCER20 or pLIX-403 doxycycline-inducible lentiviral vector. To minimize the possible effect of the tag over MASALA, we separated MASALA and the HA by a linker of low interacting amino acids (3xGGGGS).

We transduced MEFs with the pINDUCER20-MASALA-HA (murine) and A549 with pLIX-403-MASALA-HA (human) lentiviral vectors. Successfully transduced cells were treated with 1µg/ml doxycycline treatment for 24 hours to induce microprotein overexpression **(Fig 10)**. Importantly, we detected the microprotein by immunoblotting against the HA-tag, meaning that MASALA microprotein can be expressed and is stable in cellular conditions.

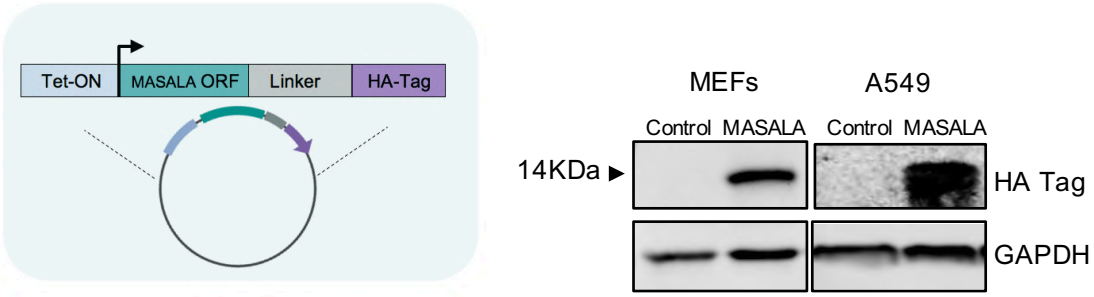


Figure 10. **Generation of a MASALA overexpression vector. (A)** Design of the doxycycline-inducible MASALA overexpressing construct. **(B)** Detection of mouse and human MASALA overexpression in the indicated cells lines after 24 hours of induction with 1 µg/ml doxycycline. Control cells were transduced with the empty backbone. Protein expression was detected by immunoblot using an anti-HA-tag antibody

To assess MASALA loss of function, we have generated a MASALA-deficient mouse model using CRISPR-Cas9 technology. It is known that lncRNAs functionality often depends on their secondary and tertiary structures (Zampetaki et al., 2018),(Johnsson et al., 2014). To disrupt MASALA translation without modifying the function of *SMIM10L2A* lncRNA, we introduced as few changes as possible into the genomic sequence. We used a homologous recombination template in which we substituted MASALA start codon by a stop codon and mutated the PAM sequence to avoid recognition of the gRNA-Cas9 complex once the genomic region has been modified. In addition, we introduced a G>C mutation to generate a restriction site to be able to detect the recombination by enzymatic digestion **(Fig 11).**

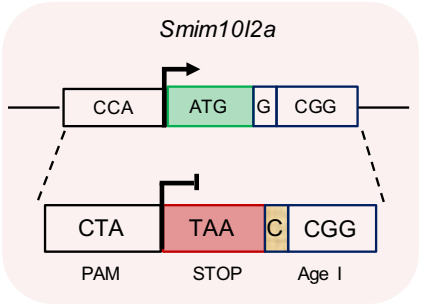


Figure 11. Strategy for the generation of a MASALA-deficient mouse strain using CRISPR-Cas9 technology. Image showing the MASALA WT allele and the MASALA-KO allele. We substituted the start codon (green) by a stop codon (red), mutated the PAM sequence, and introduced a G>C mutation to include a restriction site (in yellow).

We analyzed the wild type and knock-out *SMIM10L2A* transcript sequences using Mfold lncRNA structure predictor (Zuker, 2003), and we only observed a minimal change in the ATG loop that did not significantly affect the stability of the entire molecule **(Fig 12A).**

In collaboration with Sagrario Ortega (Head of the Transgenic Mice Unit at the CNIO), we delivered the Cas9, gRNA and the homologous recombination template by microinjection of mice embryos and we successfully obtained the desired modification in several embryos that we used as founders to establish a mouse colony of MASALA-deficient mice **(Fig 12B).**

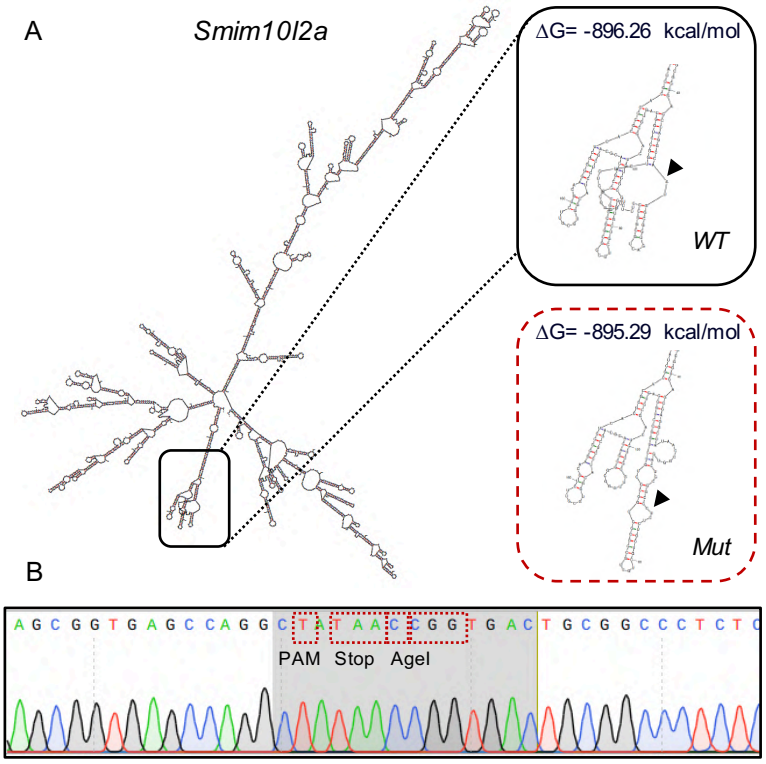


Figure 12. Modification of the start codon does not affect RNA stability or structure. (A) Left, predicted structure of *SMIM10L2A* using Mfold software. Right, detail of the modification in RNA structure with the minimum free energy of both WT and mutated molecules. The arrows indicate the start codon position. **(B)** result of the sequencing of *Smim10l2a* locus in one of the founders carrying the modification. Bordered in red are the changes that are introduced in the locus: mutation of the PAM sequence, STOP codon and a restriction site for genotyping.

Finally, in order to detect the expression of the endogenous MASALA microprotein, we have generated a custom rabbit polyclonal antibody. We subcontracted the generation of the antibody to the company Proteogenix, which synthesized the MASALA protein and use it to immunize the rabbits. We obtained a purified antibody that we have validated in mouse tissues (see 1.3.2)

1.3. Analysis of *SMIM10L2A*/MASALA expression and regulation

1.3.1. *SMIM10L2A* is expressed in brain, testis and adrenal gland

Many lncRNAs present a tissue-specific expression pattern. We checked the expression pattern of *SMIM10L2A* in mouse and human tissues. For human tissues, we used GTEX public data. *SMIM10L2A* is mainly expressed in brain, testis and adrenal gland **(Fig 13A).** We further corroborate the expression of *Smim10l2a* in these three organs in mouse by qRT-PCR, being the testis the one with the higher expression **(Fig 13B)**.

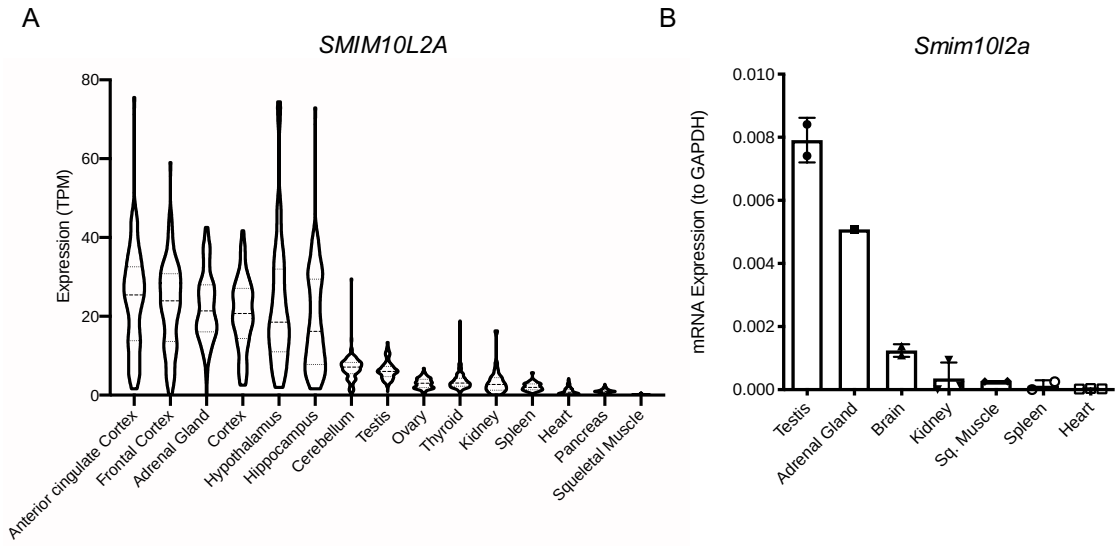


Figure 13. *SMIM10L2A* is mainly expressed in brain, testis and adrenal gland. (A) Violin plot of *SMIM10L2A* expression in human tissues. Data extracted from Gtex database. **(B)** mRNA expression levels of *Smim10l2a* in 8-12 weeks old mouse tissues analyzed by qRT-PCR. Each dot represents a single animal, bars represent mean expression ± SEM.

1.3.2. MASALA microprotein is detected in the adrenal gland.

We used our custom-made antibody against MASALA to test the expression of the endogenous MASALA protein in different organs. As a negative control, we used our MASALA knock-out mice. So far, we have not been able to detect a specific and reliable band by Western blot in mouse tissues (see 2.2.1 for the detection in MEFs during reprogramming). We tried by immunofluorescence, and although we did not detect MASALA in the testes or the brain, we found the expression of the protein in the *Zona glomerulosa* of the adrenal gland cortex in WT mice, whereas there was no signal in MASALA KO mice **(Fig 14A).** Importantly, we observed that MASALA is localized in the cytoplasm of the cell **(Fig 14B).**

This result further confirms that *SMIM10L2A* was misannotated as a lncRNA, and in fact is translated producing MASALA microprotein.

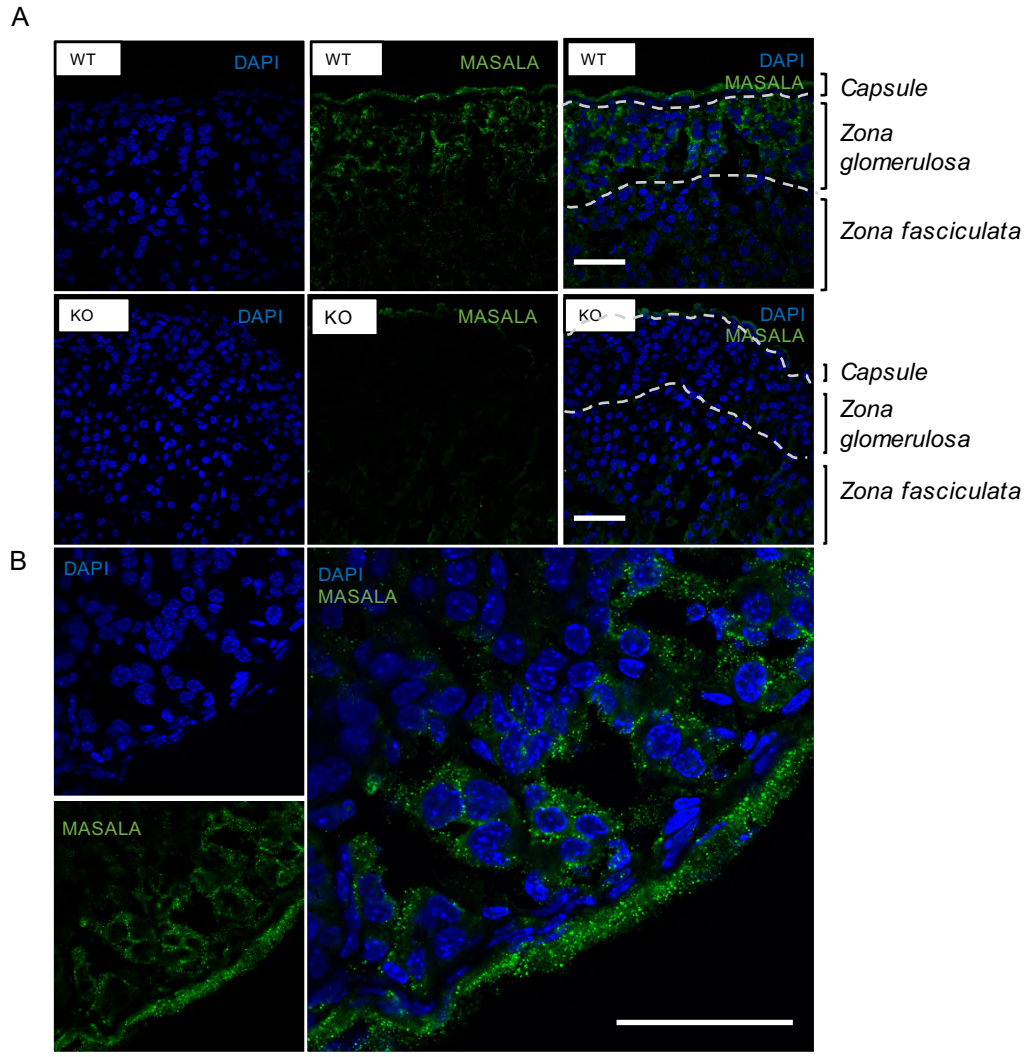


Figure 14. MASALA protein is detected in adrenal gland. (A) Representative immunofluorescence images of MASALA (green) in WT and MASALA KO adrenal glands obtained from 10-week-old mouse. MASALA was stained using an anti-MASALA polyclonal custom-made antibody. Nuclei are counterstained with DAPI. The different areas of the adrenal gland cortex are shown. Scale bar = 50µm. **(B)** Digital magnification of the *Zona glomerulosa* in a WT mouse. Scale bar = 50µm.

1.3.3. *SMIM10L2A* is upregulated upon damage in a p53 dependent manner

Previous studies have reported that *SMIM10L2A* is upregulated by p53 stabilization upon Nutlin-3a treatment (Leveille et al., 2015). Nutlin-3a is an inhibitor of MDM2, a E3 ubiquitine-ligase that promotes p53 degradation (Vassilev et al., 2004). We confirmed the regulation of *SMIM10L2A* by p53 using the isogenic cell lines HCT116 p53-WT and HCT116 p53-KO. We treated both cell lines with Nutlin-3a and measured the expression of the lncRNA by qRT-PCR. Indeed, we observed that *SMIM10L2A* is upregulated by Nutlin-3a only in HCT116 p53 WT cells **(Fig 15).**

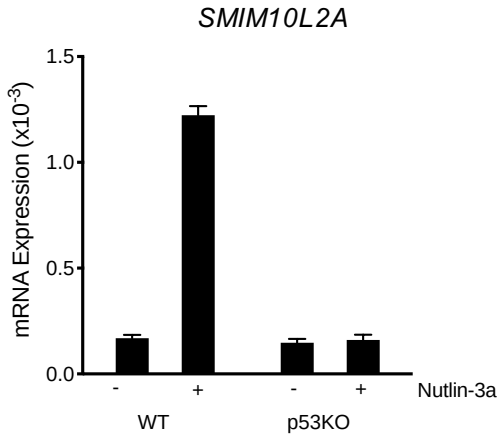


Figure 15. *SMIM10L2A* is upregulated by p53 stabilization. Expression analysis of *SMIM10L2A* by RT-qPCR in HCT116 p53-WT or p53-KO cells treated with Nutlin-3a 10μM for 24h. Bars represent mean expression ± SD.

In addition, we wondered if the upregulation of the lncRNA was still dependent of p53 in a context of general damage. We treated p53-WT and p53-KO HCT116 cells with doxorubicin, a chemotherapeutic drug that induces genotoxic stress, and we observed that *SMIM10L2A* was upregulated upon damage in a p53 dependent manner **(Fig 16).**

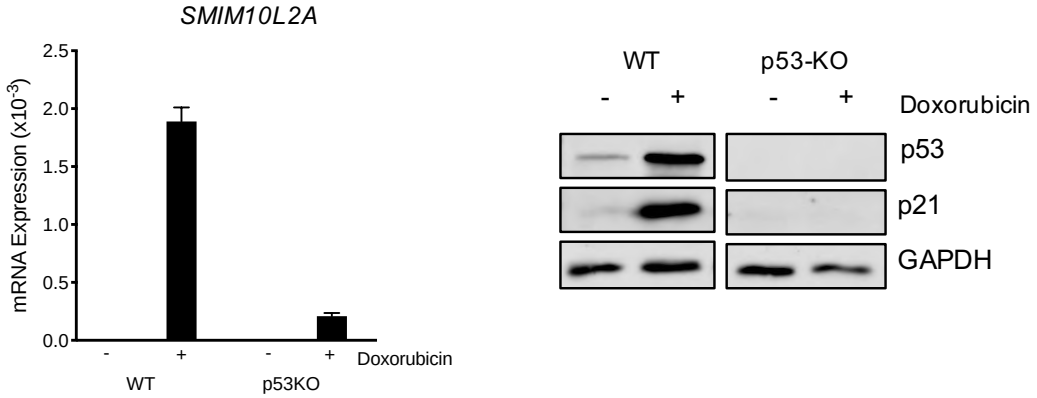


Figure 16. *SMIM10L2A* is upregulated upon doxorubicin treatment in a p53-dependent manner. On the left, expression of *SMIM10L2A* in HCT116 p53-WT or p53-KO cell lines treated with 1μM doxorubicin for 24h, measured by RT-qPCR. Bars represent mean expression ± SD. On the right, Western blot showing p53 stabilization and upregulation of its downstream effector p21.

1.3.4. *SMIM10L2A* is downregulated in some types of cancer

Given that MASALA is regulated by p53, the tumor suppressor most frequently mutated in human tumors, we wanted to study the regulation of *SMIM10L2A* in cancer. In fact, several studies have previously reported a positive correlation of *SMIM10L2A* expression with a better outcome in nasopharyngeal carcinoma and gastric cancer (Guo et al., 2017; Li et al., 2016).

We analyzed the expression of *SMIM10L2A* in tumors originated in the brain and in the adrenal gland, tissues in which *SMIM10L2A* is highly expressed. More specifically, using the Genotype-Tissue Expression (GTEx) Project database and the Cancer Genome Atlas (TCGA) database, we compared the expression of *SMIM10L2A* in the brain with its

expression in Low-grade glioma and Glioblastoma, and the expression in the adrenal gland with the expression in Pheocromocytoma & Paraganglioma. We observed that *SMIM10L2A* is downregulated in tumors compared to their healthy tissue in all of the cases **(Fig 17).**

Altogether, these observations suggest a potential role of MASALA microprotein as a tumor suppressor.

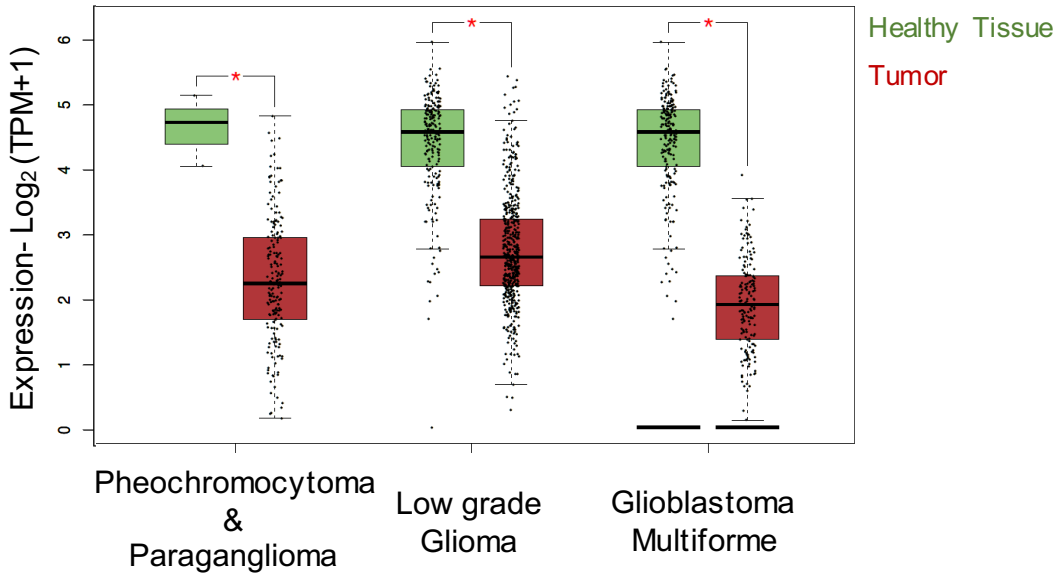


Figure 17. *SMIM10L2A* is downregulated in cancer. Comparison between *SMIM10L2A* expression in healthy brain and adrenal gland (retrieved from from GTEx, in green) and in tumors derived from those tissues (from TCGA, in red). Each point represents a single patient. * p< 0.01 One-way ANOVA.

2. Functional characterization of MASALA

2.1. MASALA in cancer

2.1.1. MASALA overexpression does not affect cell growth in cancer cell lines

As we have described above, MASALA is regulated by p53 protein and its expression is downregulated in several tumors. Thus, we considered of interest to test if MASALA overexpression could affect cancer cell growth. As a first approach to study this, we transduced A549 lung cancer cells and GL261 mouse glioblastoma cells with the MASALA-inducible lentiviral vector or control vector, and measured cell proliferation after inducing MASALA with doxycycline at different time points. MASALA overexpression over several days did not produce any significant effect in the growth rate of any of the cell lines **(Fig 18).**

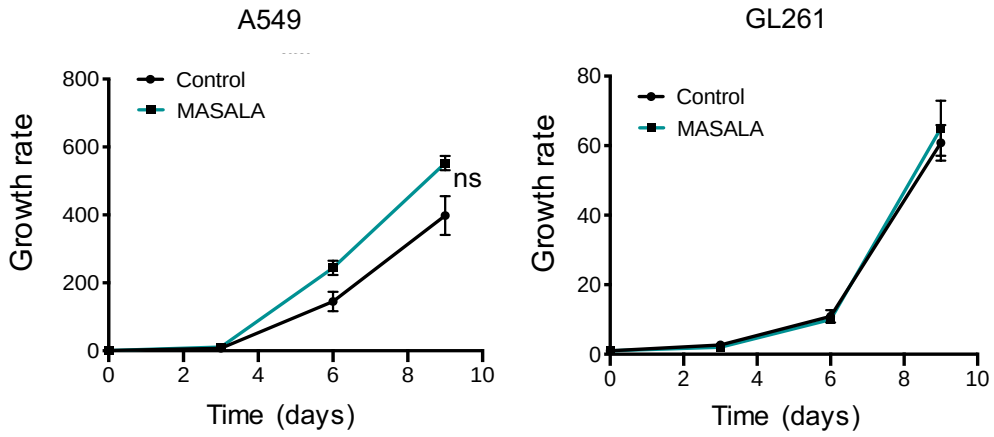


Figure 18. MASALA does not affect cell growth in cancer cell lines. Growth rate of A549 and GL261 cell lines overexpressing MASALA over 9 days. Growth rate was calculated by counting the number of cells at the indicated timepoints and normalized to their respective number of cells at day zero. Each point represents the mean ±SD of three technical replicates.

Next, we transduced A549 cells with pLIX-403-MASALA-HA inducible vector and overexpressed the microprotein for 24 hours and did not observe significant differences in a Cell titer-Glo assay, which measures ATP as an indicator of the number of viable cells in the culture **(Fig 19A).** In addition, since *SMIM10L2A* lncRNA has been shown to regulate *CDKN1A* expression (Leveille et al., 2015), we wanted to check if MASALA affects the expression of *CDKN1A* gene together with *CCND1,* another cell cycle gene. Consistently with previous results, we did not observe significant differences in the expression of these genes upon 24 hours of MASALA induction in A549 cells **(Fig 19B).**

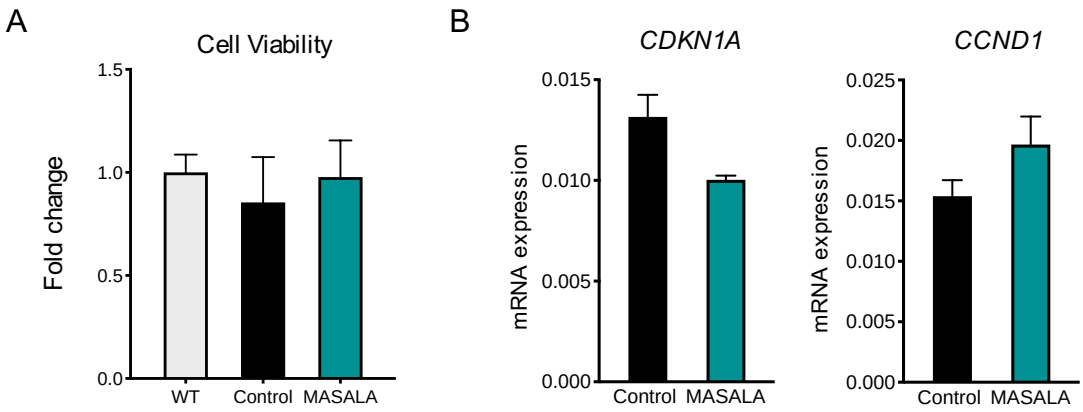


Figure 19. MASALA overexpression does not affect the number of viable cells or the expression of cell cycle genes. (A) Cell titer-Glo analysis in A549 cells overexpressing MASALA for 24h. Bars represent mean ± SD of five technical replicates. WT column represent A549 cells that have not been lentivirally transduced. Control column represent A549 cells transduced with pLIX-403 empty vector. **(B)** mRNA expression analysis by RT-qPCR of the indicated genes in A549 cells overexpressing MASALA or transduced with its empty vector for 24h. Bars represent mean ±SD of three technical replicates.

2.1.2. MASALA impairs oncogenic transformation *in vitro*

Although we could not observe any effect of MASALA on cell cycle in cancer cell lines, we know that *SMIM10L2A* is downregulated in several tumors. Since *SMIM10L2A* is regulated by p53, and the inactivation of this gene (or the inactivation of p53 pathway) is usually one of the first oncogenic hits during carcinogenesis, we wondered if the overexpression of MASALA could have a tumor suppressor effect in an oncogenic transformation experiment.

To test this, we transduced MEFs with the murine version of the MASALA-inducible vector or control vector, together with a construct that constitutively express oncogenic RAS ($HRAS^{G12V}$) and the E1a protein, which inactivates retinoblastoma (*Rb)* and induces p53 stabilization. The combination of the expression of oncogenic RAS with *Rb* inactivation and the stabilization of p53 promote surviving primary cells to transform into cancer cells. After the infection, we plated cells at low density and induced the expression of MASALA over 15-21 days. Of note, MASALA overexpression drastically reduced the efficiency of oncogenic transformation, quantified as the number of transformation foci stained with Crystal violet **(Fig 20)**, suggesting that MASALA may be a molecular barrier for oncogenic transformation.

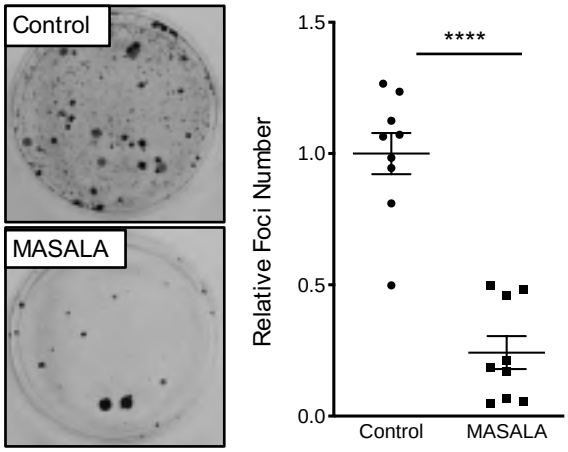


Figure 20. MASALA significantly reduces oncogenic transformation by E1a-Ras *in vitro*. On the left, representative images of MASALA and control transformation foci. MEFs were infected with E1a-HRAS^{G12V}, together with pINDUCER20-MASALA-HA or the empty vector. Successfully infected cells were plated at low density, and treated for MASALA overexpression with 1µg/ml doxycycline every other day and stained with crystal violet 15-21 days after plating. On the right, quantification of the number of transformation foci, represented as relative to the average number in their respective controls. Each dot represents a biological replicate (n=8) of three different experiments. Plot represents mean ±SEM. ****p< 0.0001 obtained with unpaired student t-test with Welch correction.

2.1.3. MASALA does not impair oncogenic transformation in the RCAS/GFAP-*tv-a* Glioblastoma model

Following the observation that MASALA reduces oncogenic transformation *in vitro,* we wanted to test this phenotype *in vivo*. Given that *SMIM10L2A* is highly expressed in brain and it is downregulated in glioblastoma, we chose to use a well-established mouse model for glioblastoma, the RCAS/ GFAP-*tv-a* (Hambardzumyan et al., 2009; Holland et al., 1998) . This model is a genetically engineered mouse that expresses the *tv-a* receptor under the promoter of *GFAP*. This cell surface receptor allows the targeting of specific cells (in our case GFAP-expressing cells) with replication-competent avian leukosis virus splice-acceptor (RCAS) viruses encoding for different drivers of glioblastoma (Ozawa et al., 2014). Thus, this system models the origin of glioblastoma from glial precursors and astrocytes, which express *GFAP.*

In our case, we followed a very aggressive combination of oncogenic drivers that includes the overexpression of *PDGFA* together with two shRNAs against p53 and Nf1. To study the effect of MASALA, we added another viral vector encoding our microprotein or an empty vector for the control condition **(Fig 21)**.

Figure 21. Schematic representation of the RCAS/ GFAP/ *tv-a* mouse model. Depiction of the different RCAS vectors carrying the glioblastoma drivers (shp53, shNf1 and PDGFA) and the RCAS vector for MASALA overexpression. Injection of DF-1 cells producing the RCAS viral vectors were injected intracranially into GFAP/*tv-a* newborn mice.

We injected intracranially newborn mice with $2{\cdot}10^5$ DF-1 cells producing the glioblastoma drivers and followed their evolution until mice reached humane end point due to weight loss or symptoms of neural damage, when the mice were sacrificed. We did not observe any effect of MASALA overexpression in the survival of the inoculated mice **(Fig 22A)**. We have also, repeated the same experiment inoculating $2{\cdot}10^4$ cells observing the same result **(Fig 22B)**. It is possible that the aggressiveness of this model (with a median survival of 52 and 55 days in the Control and MASALA conditions, respectively) is sufficient to overcome the effect of MASALA as a tumor suppressor. There are different combinations of oncogenic drivers that renders a milder phenotype, for example using only PDGFA + shNf1 (Ozawa et al., 2014), which could be used to further check the role of MASALA in this model.

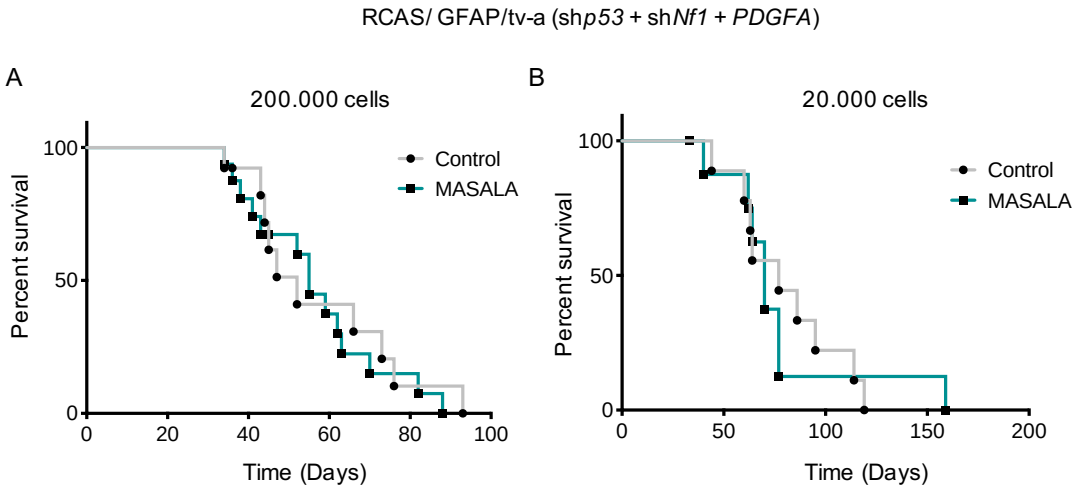


Figure 22. MASALA does not improve survival in RCAS/GFAP-*tv-a* glioblastoma mouse model. A) Kaplan-Meyer curve representing survival over time. $2{\cdot}10^5$ cells were injected. Each dot represents a single mouse. Control n= 12; MASALA n= 14. Median survival for control = 52 days; Median survival for MASALA = 56 days. **B)** Kaplan-Meyer curve representing survival over time. $2{\cdot}10^4$ cells were injected. Each dot represents a single mouse. Control n= 10; MASALA n= 16. Median survival for control = 77 days; Median survival for MASALA = 70 days

2.1.4. MASALA does not impair NIH3T3 fibroblast transformation *in vivo*

Continuing to address the possible effect of MASALA impairing oncogenic transformation *in vivo,* we decided to use the immortal mouse fibroblasts cell line NIH3T3. Since NIH3T3 are already immortalized by inactivation of p16, a simple oncogenic hit is sufficient to transform them into cancer cells. We transduced NIH3T3 with oncogenic $HRAS^{G12V}$ and with MASALA or control vector and right after the selection we injected them subcutaneously into both flanks of immunosuppressed mice. We measured the tumor growth over time until mice reached humane endpoint. Tumors grew at great speed and MASALA did not display any effect on tumor incidence or tumor growth **(Fig 23A)**. All tumors developed at similar rates and the incidence was comparable in both control and MASALA conditions **(Fig 23B).**

Altogether, MASALA microprotein acts as a barrier for oncogenic transformation by E1a-Ras *in vitro*, but at the moment we have not seen that effect in two different models *in vivo*. This suggest that the role of MASALA as a tumor suppressor could be context dependent, and the target cell type and oncogenic drivers are important for MASALA activity (see Discussion).

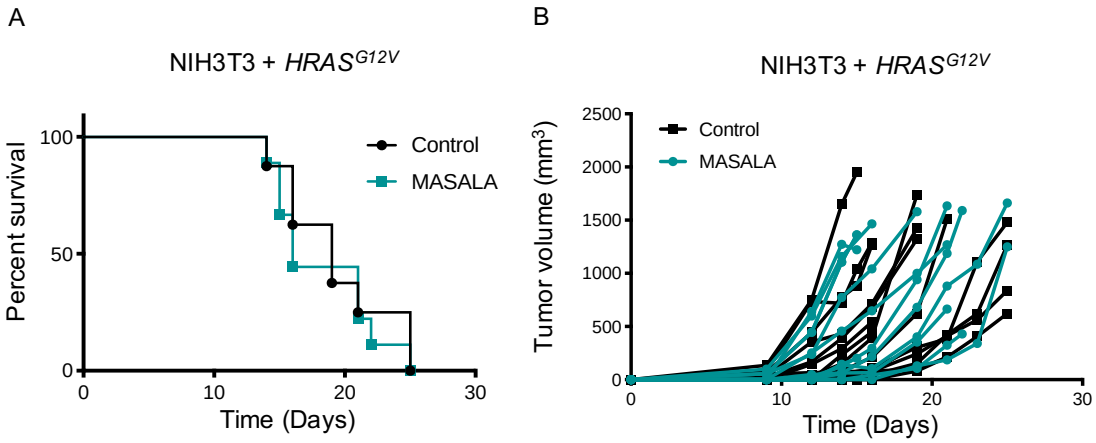


Figure 23. MASALA does not impair NIH3T3 oncogenic transformation *in vivo*. (A) Kaplan-Meyer representation of mice survival across time. Each point represents a single animal. Control n= 8; MASALA n=9. Median survival control = 19 days; Median survival MASALA= 16 days. **(B)** Representation of tumor growth over time. Each point is a measurement while each line represents a single tumor.

2.2. MASALA in cellular reprogramming

2.2.1. MASALA expression during cellular reprogramming

In our initial computational screening, we aimed to find novel microproteins with a potential role in cancer and in cellular plasticity. One of the best models to study cellular plasticity is the process of cellular reprogramming to iPSCs and, importantly, tumor suppressor genes are also molecular barriers for cellular reprogramming. Therefore, we wondered if MASALA could play a role in this process. First, we checked the expression of *SMIM10L2A* in mouse or human iPSCs compared to primary fibroblasts, and we observed that *SMIM10L2A* is upregulated in the iPSCs **(Fig 24).**

Additionally, we wanted to follow the dynamics of *SMIM10L2A* transcriptional activity and the expression of MASALA protein during cellular reprogramming by Oct4, Sox2, Klf4 and c-Myc (OSKM). For that, we extracted MEFs from i4F reprogrammable mice, which ubiquitously express OSKM in a doxycycline inducible manner (Abad et al., 2013). i4F-MEFs were treated with doxycycline to induce OSKM expression for 12 days, when iPSC colonies were clearly visible. We monitored *SMIM10L2A* mRNA expression and MASALA protein expression every two days. Regarding the transcriptional dynamics of the gene, we observed an initial upregulation with its maximum at day 4, after which the expression progressively drops until day 10 where it gets stabilized. This pattern is consistent with the expression pattern of a tumor suppressor gene, which gets upregulated upon cellular stress **(Fig 24B)**. Of note, MASALA protein level progressively increases during the first 6 days, and gets stabilized during the rest of the experiment **(Fig 24B)**. Therefore, MASALA is regulated during cellular reprogramming and its protein levels increases during the process.

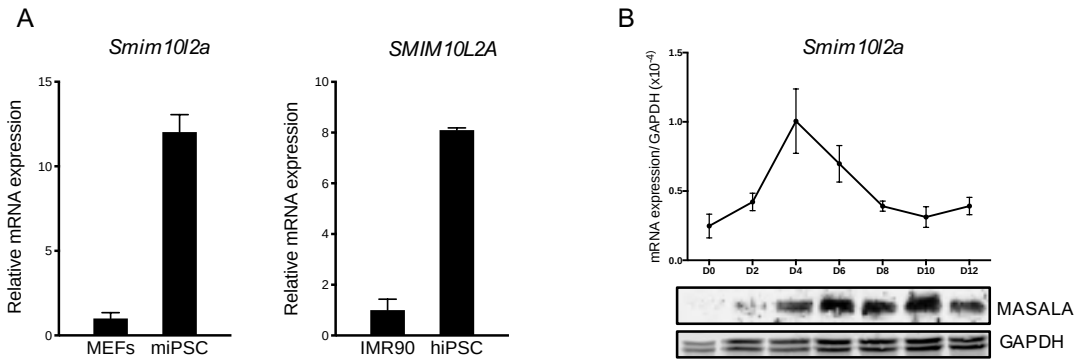


Figure 24. *SMIM10L2A* is upregulated during iPS reprogramming. (A) Expression analysis of *SMIM10L2A* in both mouse (left) and human (right) iPSCs compared to primary fibroblasts. Bars represent mean ± SD of three technical replicates. **(B)** Analysis of *SMIM10L2A* mRNA expression by RT-qPCR of i4F-MEFs n=3, and immunoblot of endogenous MASALA in one i4F-MEF line at the indicated times during cellular reprogramming.

2.2.2. MASALA overexpression impairs cellular reprogramming

As discussed previously, both oncogenic transformation and cellular reprogramming share many characteristics. One of them is that tumor suppressors are barriers for both processes. Since we hypothesized a potential role of MASALA as a tumor suppressor, we decided to study the effect of MASALA during cellular reprogramming.

WT MEFs were transduced with a doxycycline inducible lentiviral vector encoding the 4 Yamanaka (OSKM) factors together with the murine MASALA-inducible vector or the empty control. We induced the expression of both MASALA and OSKM for 12-14 days, when iPS colonies were observed. Consistently with our hypothesis, MASALA expression induced a ≈75% reduction in the reprogramming efficiency compared to control condition, as visualized by Alkaline Phosphatase positive colonies **(Fig 25).**

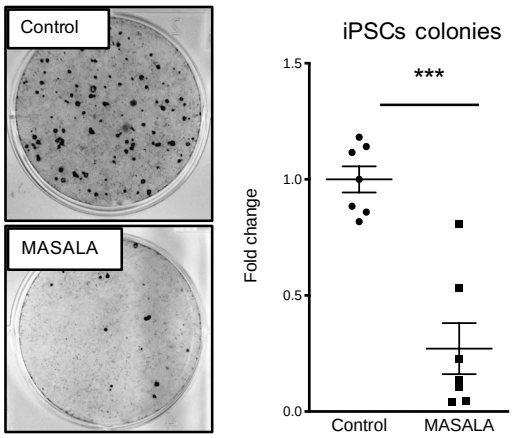


Figure 25. MASALA overexpression impairs cellular reprogramming. Left, representative images of alkaline phosphatase-positive iPSCs colonies derived from MEFs overexpressing the 4 Yamanaka factors together with MASALA or the empty vector as control cells. Right, quantification of iPSCs colonies, represented as the number of colonies relative to the control condition. Each dot represents a single biological replicate (n=7) of three different experiments. *** p= 0.0002 Student t-test with Welch correction.

Next, we wanted to test the effect of MASALA deficiency in cellular reprogramming. To test this, we isolated MEFs from MASALA KO or WT animals and transduce them with an inducible lentiviral vector to express OSKM. Surprisingly, we did not observe an increase nor a decrease of the reprogramming efficiency in MASALA-deficient MEFs compared to WT MEFs **(Fig 26)**.

Altogether, these experiments suggest that MASALA overexpression is a barrier for cellular reprogramming, but its deficiency does not alter the efficiency of the process, probably due to a compensation by other gene (see Discussion). Currently, we are analyzing the role of MASALA on *in vivo* reprogramming using i4F; MASALA-KO mice, and we are confident that these experiments will provide valuable information on the role of MASALA.

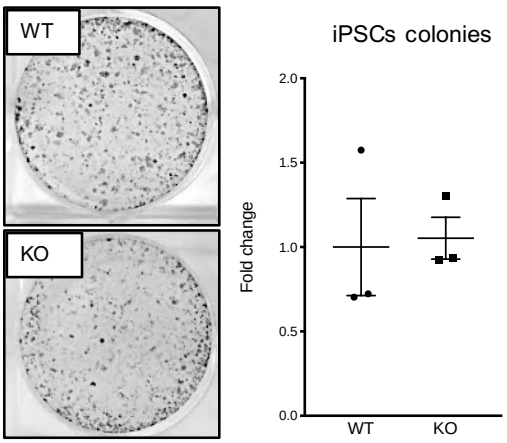


Figure 26. MASALA deficiency does not affect cellular reprogramming efficiency. Representative images of alkaline phosphatase-positive iPSCs colonies derived from MEFs overexpressing the 4 Yamanaka factors in both MASALA WT and KO MEFs (left). Quantification of iPSCs colonies (right). Each point represents a single biological replicate (n=3).

3. Molecular mechanisms behind MASALA's function

3.1. MASALA is located in the outer mitochondrial membrane

In order to understand how MASALA impairs cellular reprogramming and oncogenic transformation *in vitro*, we performed a set of experiments that helps us understand the molecular mechanisms behind this phenotype. As a first step, we focused on determining its subcellular location, as this could give important information about MASALA's function.

We previously obtained evidences for a possible mitochondrial localization of MASALA **(Fig 9C).** For this reason, we performed immunofluorescence experiments to test co-localization of MASALA with mitochondrial markers by confocal microscoscopy in NIH3T3 transduced with exogenous MASALA. We observed a clear co-localization of MASALA (detected with an anti-HA antibody) with the mitochondrial protein AIF **(Fig 27A)**. A co-localization image analysis showed that both signals' intensities have a correlation value of 0.8 **(Fig 27B)**, confirming the *in silico* prediction. We further confirmed the mitochondrial localization of MASALA by performing a subcellular fractionation for mitochondrial enrichment. We determined that MASALA microprotein mainly localizes in the mitochondria, however, a small amount of the microprotein seems to remain cytoplasmic **(Fig 27C)**.

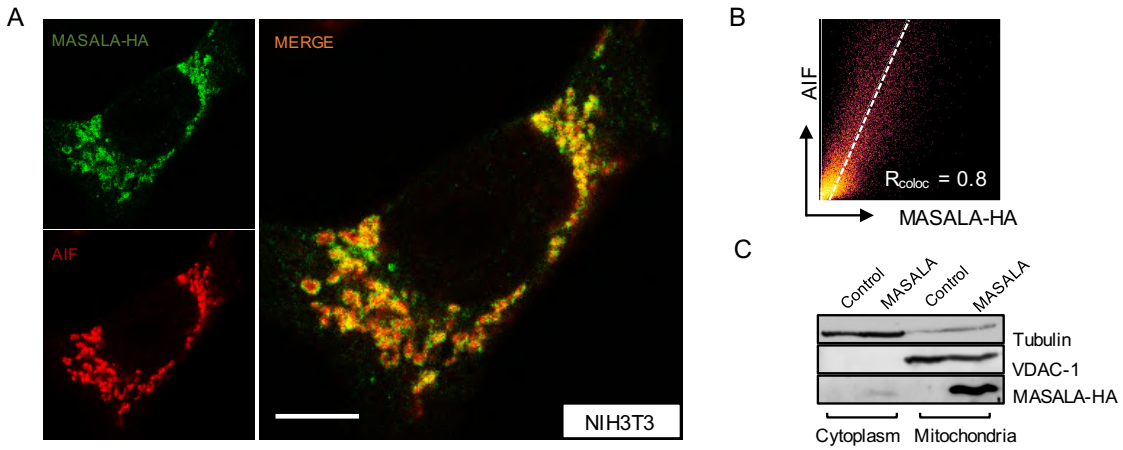


Figure 27. MASALA localizes in the mitochondria. (A) Representative immunofluorescence images of MASALA-HA (in green) co-localizing with mitochondrial protein AIF in NIH3T3 cells. Scale bar = 10µm. **(B)** Co-localization correlation analysis performed by Coloc2 software. **(C)** Subcellular fractionation of cytoplasm and mitochondrial fractions in A549 cells overexpressing MASALA. VDAC-1 is a marker of mitochondrial fraction. Tubulin is a marker for cytoplasmic fraction.

Another feature of the protein predicted *in silico* was the presence of a transmembrane domain at the C-terminal part. Since mitochondria have two membranes with very specific functions and protein composition, we aimed to determine in which of the membranes MASALA localizes.

To asses this, we performed a Protease K Protection Assay in A549 cells overexpressing MASALA. This assay allows to discriminate in which mitochondrial compartment a protein is, by treating mitochondria and mitoplast (mitochondria without the outer membrane) with protease K. If the protein is in the outer membrane, it will not be protected to degradation by protease K (see TOM20 protein as a control, **Fig 28**).

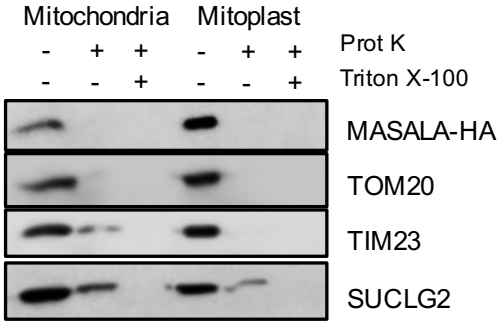


Figure 28. MASALA localizes in the mitochondrial outer membrane. Immunoblot analysis of mitochondrial proteins and MASALA after Protease K protection assay in A549 cells previously transduced with pLIX-403-MASALA-HA and induced with 1µg/ml doxycycline for 24 hours. TOM20 is a marker for mitochondrial outer membrane proteins; TIM23 is a marker of inner membrane proteins and SUCGL2 is a marker for mitochondrial matrix proteins. Triton X-100 was used to disaggregate the membranes, exposing all the proteins to Protease K degradation.

However, if it is in the inner membrane, it will be protected by the outer membrane in the untreated mitochondria, while exposed to protease K degradation in the mitoplast (see

TIM23 as a control in **Fig 28**). In the case of a mitochondrial matrix protein, it will be protected from degradation in both cases (see SUCLG2 as a control in **Fig 28**). With this approach, we observed that MASALA displays the pattern of an outer membrane protein **(Fig 28)**.

This assay, however, does not isolate pure mitochondria from their associated membranes, mainly ER-associated membranes. For this reason, we wanted to test the possible co-localization of MASALA with ER by immunofluorescence, using an antibody against the ER-membrane protein SERCA-2 in NIH3T3 cells overexpressing MASALA. In this experiment, we could only see some spots of co-localization in some of the areas, consistent with the regular ER-mitochondria interactions **(Fig 29B)**. Thus, although MASALA is mainly located in the mitochondria, we cannot discard a possible localization of MASALA in the ER membrane as well.

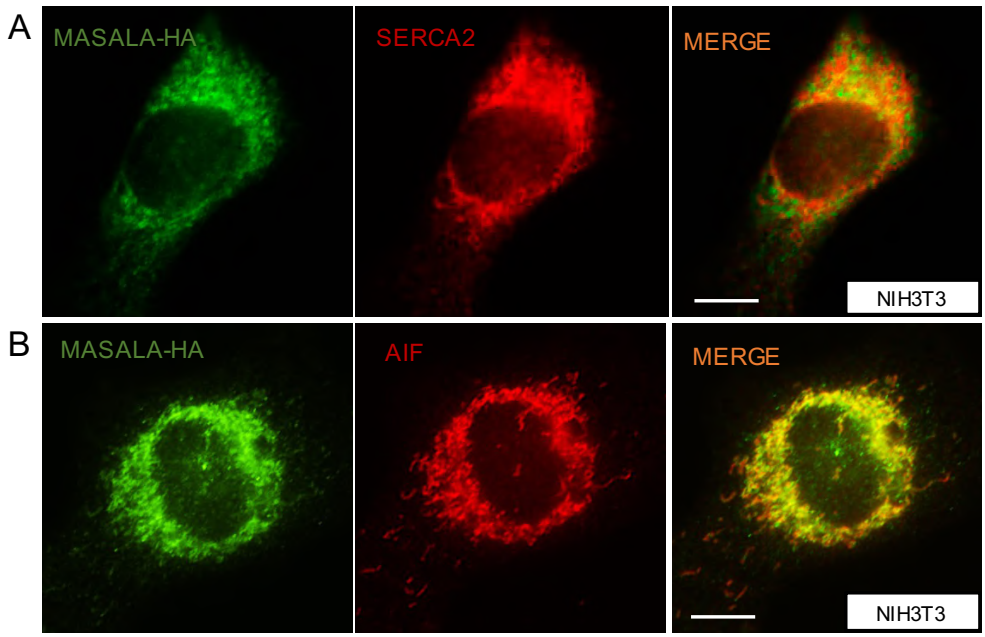


Figure 29. MASALA co-localizes only partially with endoplasmic reticulum (ER). (A) Representative immunofluorescence images of MASALA-HA and the ER marker SERCA2 in NIH3T3 cells transduced with pLIX-403-MASALA-HA and induced for 24 hours. **(B)** Representative immunofluorescence images of MASALA-HA and the mitochondrial marker AIF in NIH3T3 cells transduced with pLIX-403-MASALA-HA and induced with doxycycline for 24 hours. Scale bars = 10µm

3.2. MASALA does not induce or sensitize to apoptosis

Mitochondrial outer membrane and the ER are known to be important compartments for the apoptotic signaling by regulating intracellular Ca^{2+} concentration (Lee et al., 2018; Pinton et al., 2008). Additionally, mitochondrial outer membrane bears important pro and anti-apoptotic proteins, such as the BCL-2 protein family, that controls the opening of the mitochondrial pores required to trigger apoptosis (Singh et al., 2019).

The mitochondrial localization of MASALA, together with its upregulation by p53 and the function of *SMIM10L2A* as a p21 enhancer pointed out to a possible role in apoptosis. To assess this hypothesis, we first tested in WT MEFs if the overexpression of MASALA,

with or without the induction of cellular damage, could affect the cell number. We treated MASALA-overexpressing and control MEFs with doxorubicin or with DMSO, and we stained the plates with crystal violet after 24 hours. As expected, doxorubicin treatment reduced the number of cells in the plates; however, MASALA overexpression had no effect on either the doxorubicin-treated or the control MEFs **(Fig 30A)**. Then, we tested whether MASALA sensitizes to apoptosis upon other apoptotic triggers, such as the pro-apoptotic agent staurosporine and the BCL-2 inhibitor Navitoclax. In order to have a more sensitive readout, we measured the activation of Caspase 3/7 in MASALA-overexpressing or MASALA-deficient MEFs 24 h after the treatments, and we did not observe significant differences in the caspase activation between the different conditions **(Fig 30B)**. This indicates that MASALA overexpression or its deficiency does not trigger apoptosis by itself, and MASALA does not sensitize to apoptosis triggered by different agents **(Fig 30)**.

Together, MASALA, at least in these settings, does not seem to play a role in the induction of apoptosis.

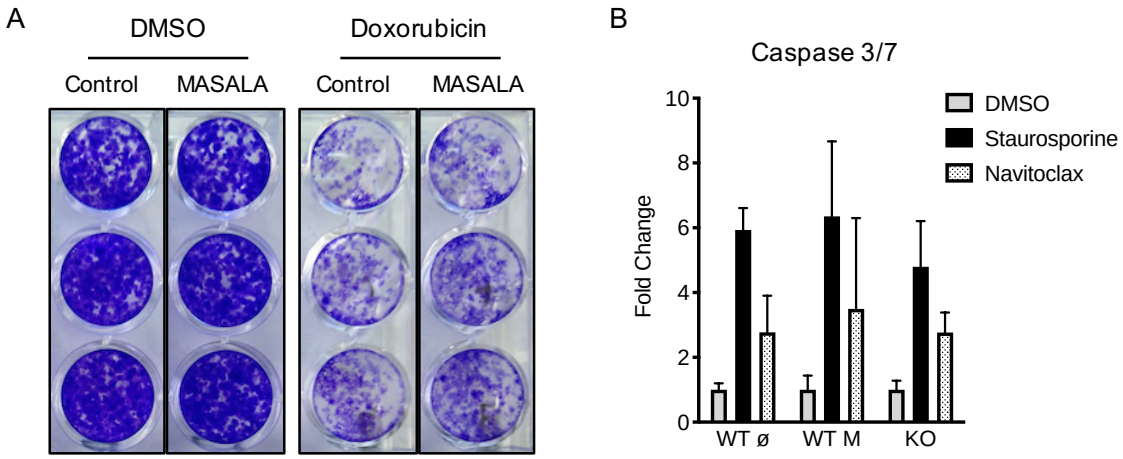


Figure 30 MASALA does not affect apoptotic response. (A) Crystal violet staining of MEFs overexpressing MASALA and treated with 1µM doxorubicin or DMSO, both for 24h. **(B)** Quantification of Caspase 3/7-Glo assay. Activation of Caspase 3/7 was measured in 3 biological replicates of WT MEFs transduced with the empty vector (WTø), MEFs overexpressing MASALA for 24 hours (WT M) or MASALA KO MEFs (KO) upon the 24h of treatment with the indicated pro-apoptotic agents.

3.3. MASALA does not affect mitochondrial or cellular ROS

Next, we wondered if MASALA could be affecting other processes related with mitochondrial biology. Mitochondria are the organelles responsible of aerobic respiration in the cells. This process can generate Reactive Oxygen Species (ROS), which are highly reactive metabolites derived from electron leakage in the electron transport chain. This leakage generates a partial reduction of the oxygen molecules that are transformed into superoxide. These ROS play important roles in physiological and pathological processes, including oncogenesis and cellular reprogramming. Indeed, an optimal amount of ROS is required during cellular reprogramming to be successful (Zhou et al., 2016). For these reasons, we wondered whether MASALA could have role regulating the production of ROS.

As a first approach, we wanted to test if MASALA affects total cellular ROS. For that, we used the DCFH-DA fluorescent probe in MEFs overexpressing MASALA for 24 hours. We observed that MASALA overexpression did not affect the amount of cellular ROS in MEFs **(Fig 31)**.

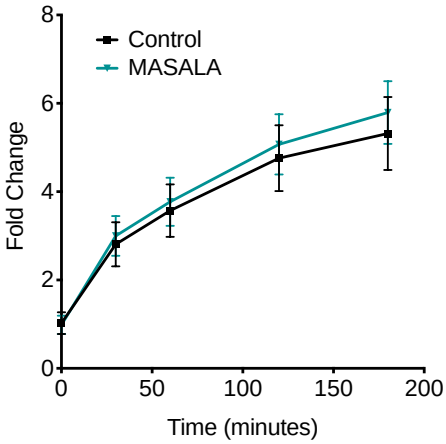


Figure 31. MASALA does not affect total ROS production. Quantification of total ROS measured by fluorescence intensity over time in MEFs overexpressing MASALA for 24 hours and treated with DCFH-DA fluorescent probe. Data is represented as mean fluorescence intensity (MFI) relative to MFI at time 0. Each line represents the mean intensity in 3 biological replicates.

Next, we tested if MASALA could affect specifically mitochondrial ROS (mtROS). To study this, we used MitoSOX fluorescent probe and quantified the signal intensity in WT, MASALA-overexpressing or MASALA-deficient MEFs by FACs analysis. MASALA overexpression did not change the amount of mtROS, by itself or after Antimycin A treatment **(Fig 32)**. We observed a slight reduction in mtROS in MASALA-KO MEFs, although not statistically significant **(Fig 32).**

These results indicate that MASALA does not affect ROS production or the antioxidant response in MEFs.

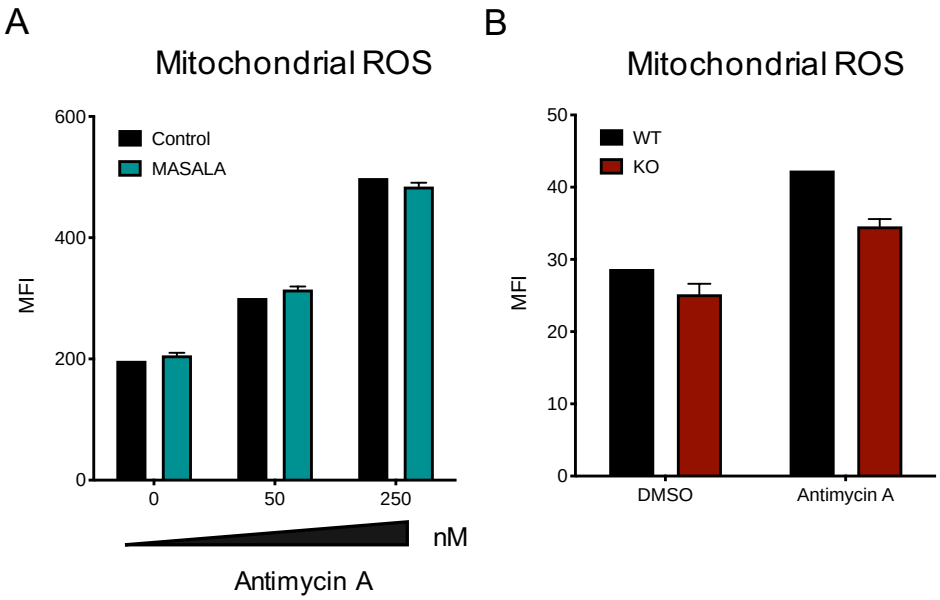


Figure 32. MASALA does not affect mitochondrial ROS. (A) Representation of median fluorescence intensity of MEFs overexpressing MASALA for 24 hours detected by FACs using MitoSOX probe. Cells were treated with the specified concentration of Antimycin A as a positive control for ROS production. **(B)** Median fluorescence intensity of WT or MASALA-KO MEFs treated with 250nM or DMSO.

3.4. MASALA does not affect mitochondrial membrane potential

Mitochondria are the main producers of ATP in the cell. To achieve this, the electron transport chain generates a proton gradient across the inner membrane by using NADH and oxidizable substrates from the Krebs cycle and fatty acid oxidation, and by pumping protons out of the matrix. This differential proton gradient creates a membrane potential used by the Complex V-ATP synthase to generate the ATP, allowing protons to go back into the matrix. This membrane potential represents a good readout of the energetic status of the cell and the fitness of the mitochondria.

We wondered if MASALA overexpressing cells could have an alteration in the membrane potential, due to mitochondrial stress or affectations in the mitochondrial homeostasis. To study this, we used the fluorescent probe TMRM in MEFs transduced with MASALA-overexpressing or control vector. As controls, we treated WT MEFs with FCCP and with Oligomycin A. As expected, FCCP treatment decreased the membrane potential, and Oligomycin A increased the membrane potential, but we couldn't observe any significant differences with the induction of MASALA **(Fig 33).**

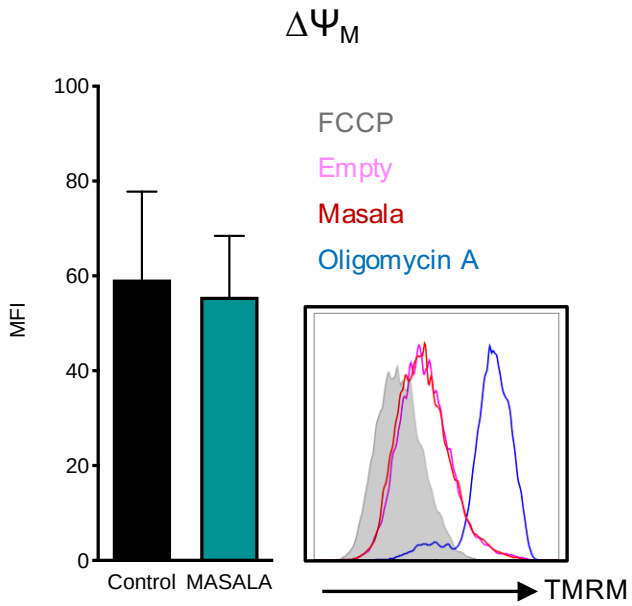


Figure 33. MASALA does not affect mitochondrial membrane potential. On the left, quantification of membrane potential in MEFs control or overexpressing MASALA for 24 hours, measured by FACs and represented as mean fluorescence intensity of 3 biological replicates ±SD. On the right, histogram representation of the fluorescence intensities. Treatment with 30μM FCCP and 20μM Oligomycin A for 30 minutes were used as negative and positive controls, respectively.

3.5. MASALA reduces 2-Hydroxyglutarate in MEFs

As discussed before, mitochondria produce energy oxidizing different substrates in the TCA cycle. In order to investigate a possible role of MASALA in the metabolic processing of glucose we performed a Glucose C^{13} tracing experiment in MEFs overexpressing MASALA. MEFs where induced with doxycycline in regular cell culture medium and, 24 hours later we changed to a Glucose- C^{13} labelled medium with doxycycline. We incubated the MEFs for 12 hours in these conditions to let the glucose get metabolized and we analyzed the metabolic flux, by looking at common cellular metabolites using mass-spectrometry.

This analysis revealed that metabolic flux was very similar in MASALA-overexpressing cells compared to the empty vector-transduced MEFs. We only observed a significant difference in the 2-Hydroxyglutarate (2-HG) metabolite isotopologue M+4 **(Fig 34A).**

This metabolite comes from the processing of α-ketoglutarate, an intermediate metabolite from the TCA cycle, that is converted into 2-HG by the isocitrate dehydrogenase enzyme (IDH1/2). 2-HG has been identified as an oncometabolite with an active oncogenic role in tumors harboring IDH1/2 mutations (Du and Hu, 2021).

However, it is difficult to interpret the effect of MASALA reducing this metabolite in primary fibroblasts, in which this metabolite is not very abundant. In fact, if we take into account the values of all the 2-HG carbons that can be labelled from the IDH1/2 activity (isotopologues) we found that no differences between the MASALA-overexpressing and the control cells **(Fig 34B).**

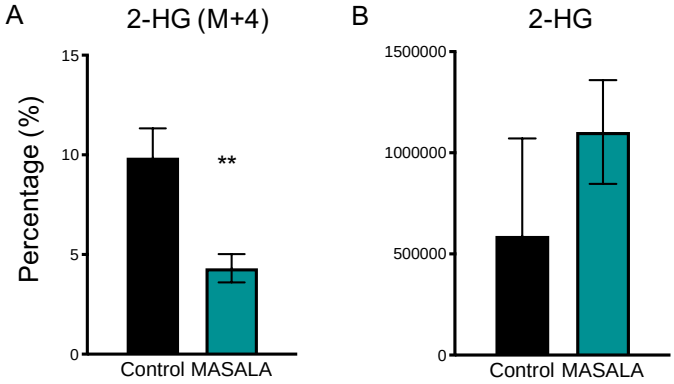


Figure 34. MASALA reduces the production of 2-Hydroxyglutarate (M+4) in MEFs. (A) Percentage of 2-Hydroxyglutarate M+4 isotopologue detected in MEFs overexpressing MASALA for 24 hours or the empty vector as control. Bars represent the mean value ±SD of n=4 biological replicates. **p≤ 0.01 Unpaired student's t-test with Welch correction. **(B)** Summed values of all the isotopologues of the 2-Hydroxyglutarate in MEFs overexpressing MASALA for 24 hours or the empty vector as control cells. Bars represent the mean value ±SD of n=4 biological replicates

3.6. MASALA reduces mitochondrial respiration and spare respiratory capacity

Although we could not observe differences in the membrane potential or in the amount of ROS, we wanted to further discard an effect on mitochondrial respiration and OXPHOS, by using the Seahorse Mito-Stress Assay. This technique allows you to determine, among other parameters, the level of basal respiration and the responding capacity of the mitochondria if forced to maximal respiration. It does so by measuring the Oxygen Consumption Rate (OCR) of the cells and the Extracellular Acidification Rate (ECAR).

We tested the effect of MASALA overexpression or its deficiency on these parameters in MEFs. We could not find any differences when compared the MASALA-KO MEFs to WT MEFs, indicating that MASALA deficiency does not affect mitochondrial respiration **(Fig 35)**. After 24h of MASALA induction we found no effect on cellular respiration or on respiratory capacity as well **(Fig 36 A).**

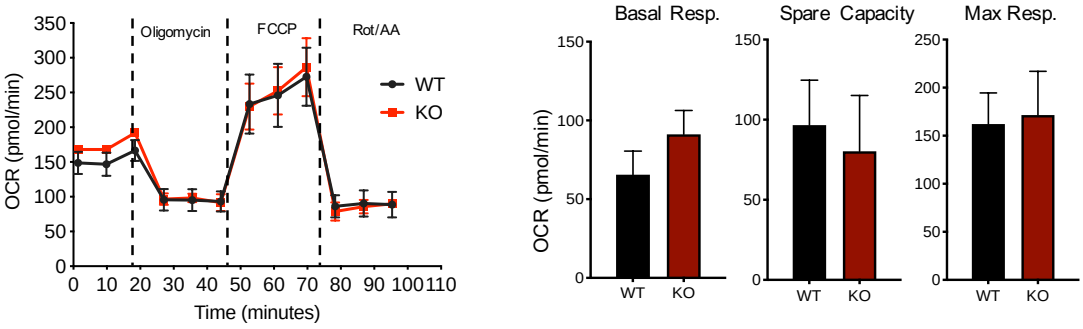


Figure 35. MASALA deficiency does not affect cellular respiration. Seahorse mito-stress assay in MEFs WT or MASALA-KO, n=3 (left). Quantification of respiratory parameters (right). Each bar represents the mean ±SEM of 3 biological replicates.

However, since cellular reprogramming and oncogenic transformation are long processes that involve several days, we wonder if the effect of MASALA could only be apparent after several days. Following this rationale, we decided to test cellular respiration after 5 days of MASALA induction. Importantly, we observed a slight but reliable reduction in the basal OCR of the cells overexpressing MASALA and, moreover, the spare respiratory capacity and the maximal respiration rates were drastically reduced after 5 days of MASALA overexpression **(Fig 36 B)**.

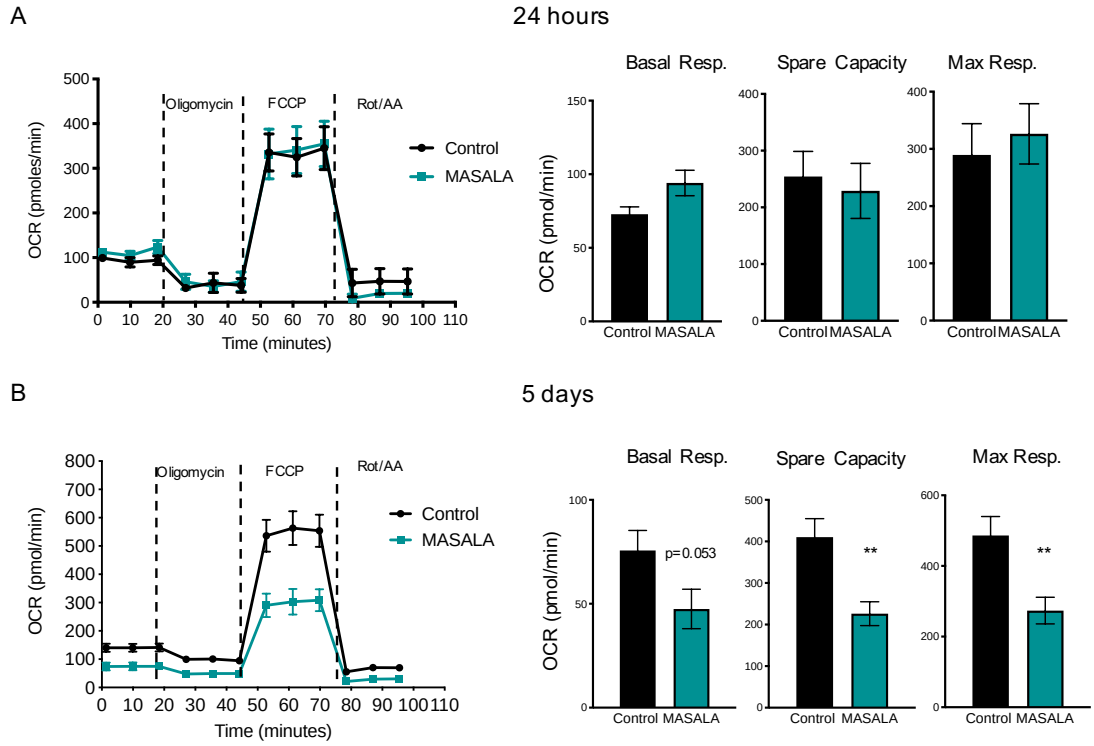


Figure 36. Overexpression of MASALA reduces cellular respiration and spare capacity in MEFs after 5 days of induction. (A) Representation of Seahorse Mito-stress assay of n=3 MEFs overexpressing MASALA for 24h (left). Quantification of the different mitochondrial parameters (right). Each bar represents the mean ±SEM of n=3 biological replicates. **(B)** Representation of Seahorse Mito-stress assay in MEFs overexpressing MASALA for 5 days (left). Quantification of the different parameters (right). Each bar represents the mean ±SEM of n=3 biological replicates. **p≤ 0.01 Unpaired Student t-test with Welch correction.

These data suggest that MASALA is not required for the normal functioning of the mitochondria, and its deficiency might be compensated by other means. Nevertheless, MASALA overexpression, after several days, disrupts mitochondrial physiology by reducing cellular respiration and spare capacity.

3.7. MASALA overexpression is stabilized upon mitochondrial damage

Once we observed a defect in the spare capacity of the cell, we investigated if MASALA is regulated by mitochondrial stressors. We observed that the detection of exogenous MASALA by Western blot was greatly improved upon inducing mitochondrial uncoupling by FCCP treatment, suggesting that MASALA protein is stabilized by FCCP **(Fig 37A)**. In order to discard an artefactual effect of the experiment, we tested MASALA stability under mitochondrial damaging agents using the translation inhibitor cycloheximide (CHX). We observed that exogenous MASALA was stabilized upon FCCP treatment, and even more when treated with H_2O_2 **(Fig 37B)**. Interestingly, both treatments result in mitochondrial damage that trigger events of mitophagy.

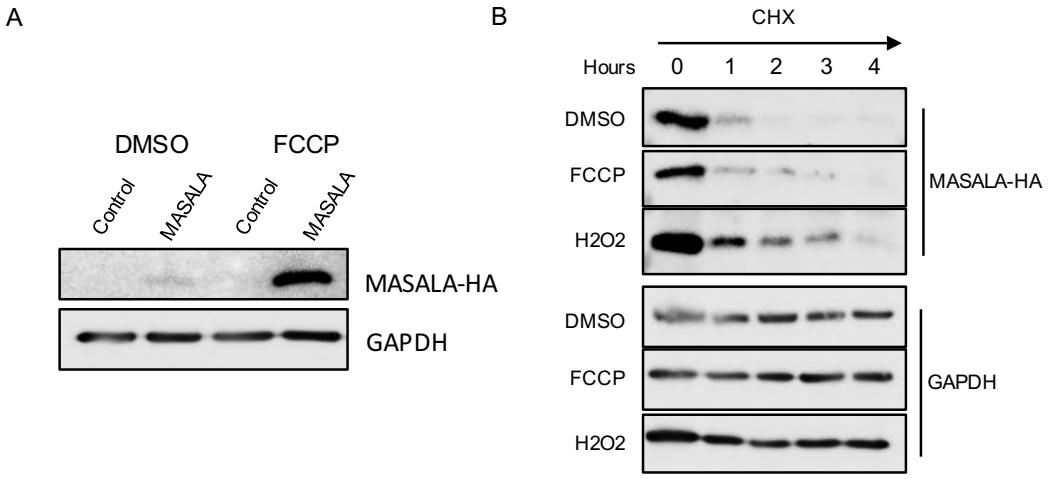


Figure 37. Exogenous MASALA is stabilized upon mitochondrial insults. (A) Western blot analysis of exogenous MASALA using an anti-HA antibody in NIH3T3 cells overexpressing MASALA and treated with 40μM FCCP for 4h. **(B)** Analysis of exogenous MASALA stability. NIH3T3 cells overexpressing MASALA were treated with FCCP (10μM for 4h) or H_2O_2 (100μM for 30min) and then, cells were treated with Cycloheximide 100μg/ml to stop translation. Cells were collected at the indicated time points and analyzed by western blot using and anti-HA antibody. GAPDH was used as loading control.

3.8. MASALA promotes mitochondrial fragmentation

Mitochondria are dynamic organelles that transit between fused and fragmented morphologies in order to adapt to cellular metabolic requirements.

The proper regulation of mitochondrial dynamics is critical to be proficient in many physiological processes. Typically, fusion events take place when mitochondria are more oxidative or upon nutrient scarce. Fission events, on the other hand, are needed to form smaller mitochondria generally associated with glycolytic phenotypes, but they are also required to eliminate unhealthy mitochondria through mitophagy.

To discriminate between the different mitochondrial phenotypes, we classified cells as fragmented, tubulated or intermediate **(Fig 38)**. We induced the expression of MASALA in MEFs and quantified the frequency of cells presenting the different phenotypes over time during 8 days.

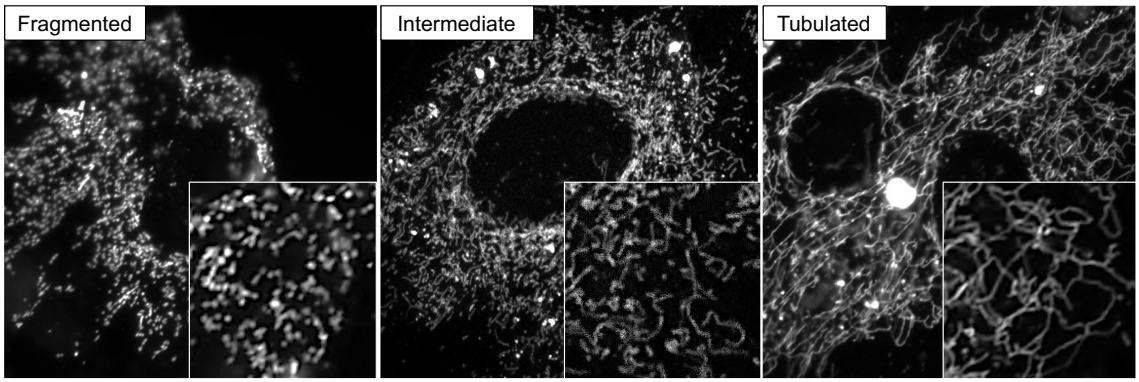


Figure 38. Representative examples of the different mitochondrial phenotypes. The images show MEFs transduced with the mitochondrial fluorescent tracker mito-dsRED and visualized by fluorescence microscopy.

Our analysis revealed that MASALA overexpression triggers a higher frequency of cells with fragmented phenotype **(Fig 39A)**, reaching the highest proportion at day 6 after induction **(Fig 39B)**.

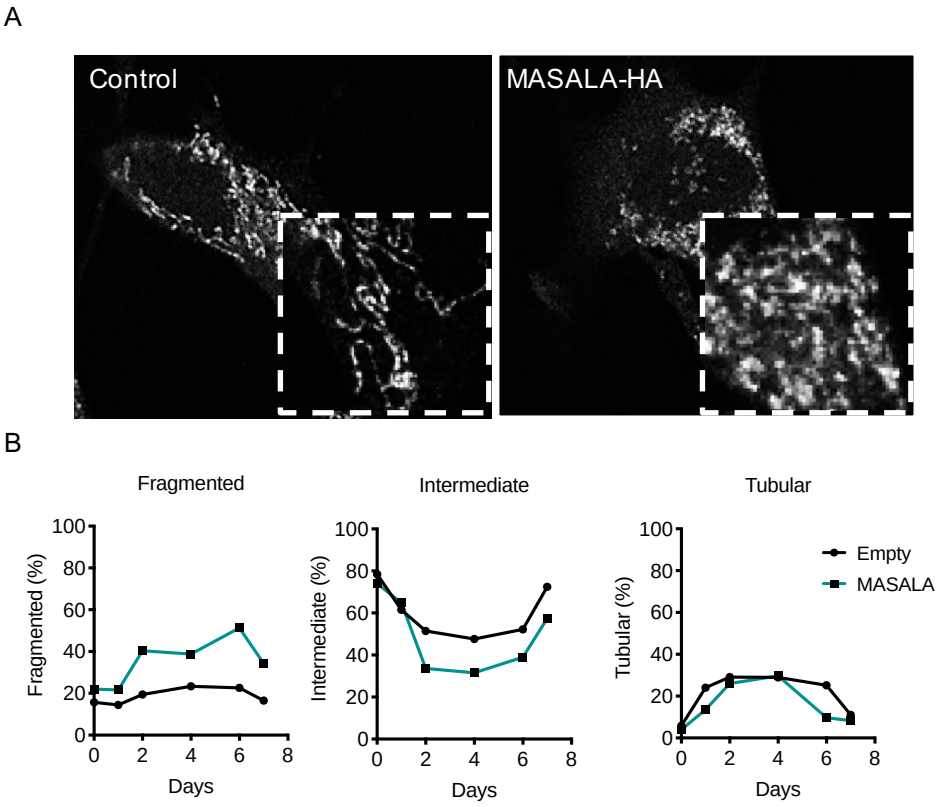


Figure 39. MASALA induces mitochondrial fragmentation. (A) Representative immunofluorescence images of MEFs control or overexpressing MASALA for 5 days and stained for AIF mitochondrial protein. **(B)** Quantification of the mitochondrial phenotypes frequencies in MEFs control or overexpressing MASALA at the indicated time points. At least 100 cells were scored at each time point.

Then, we decided to test different fission and fusion markers upon MASALA induction in three different MEF preparations (derived from three different embryos). At 24h of induction we observed a heterogeneous response. In one of the MEF preparations

we saw a clear phosphorylation of DRP1 at Ser579 (that corresponds to Ser616 in human), which indicates activation of the fission process, while no major changes were apparent in the other two biological replicates **(Fig 39A)**.

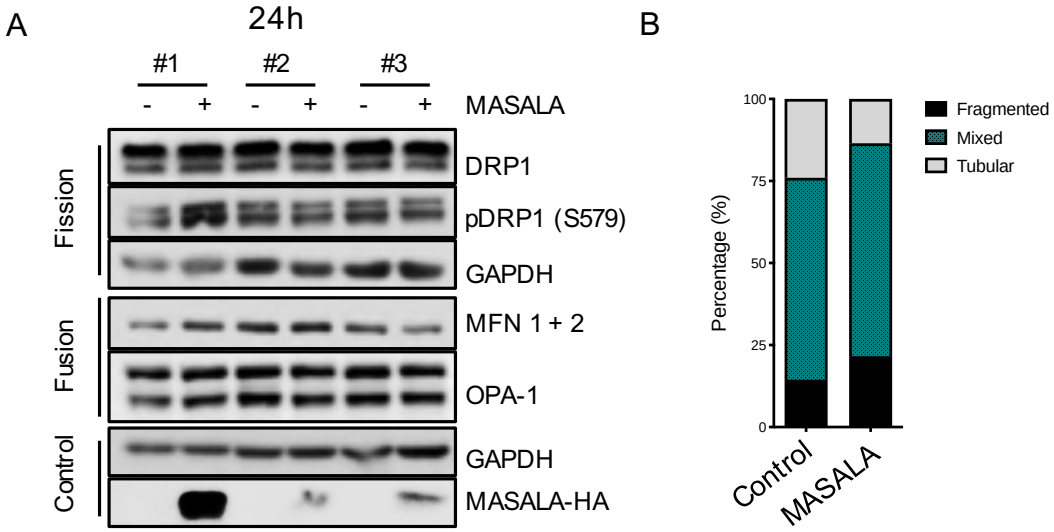


Figure 40. MASALA produces fission activation in a heterogeneous manner after 24h of overexpression. (A) Western blot analysis of exogenous MASALA (using and anti-HA antibody) and the indicated mitochondrial dynamics markers in MEFs control or overexpressing MASALA. n=3 biological replicates. **(B)** Quantification of the different mitochondrial phenotypes frequencies in MEFs overexpressing MASALA-HA for 24h.

We did not observe differences in the fusion markers MFN or OPA-1 in any of the MEF preparations **(Fig 40A)**. Accordingly, when we analyzed by immunofluorescence the frequency of mitochondrial phenotypes (as indicated above) after 24h of MASALA overexpression we saw only a small increase in the frequency of fragmented mitochondria compared to the control **(Fig 40B)**. Of note, when the same analysis was performed after 7 days of MASALA overexpression we observed a clear increase in the number of cells with fragmented mitochondria **(Fig 41B)**. This is consistent with previous results **(Fig 39B)** that indicates that the effect of MASALA overexpression is only apparent at longer time points. Surprisingly, after 7 days of MASALA expression we observed a decrease in the fission marker pDRP1 (S579) **(Fig 41A)**, possibly indicating that this is a dynamic process and at day 7 cells try to compensate the excess of mitochondrial fragmentation decreasing the pro-fission signals.

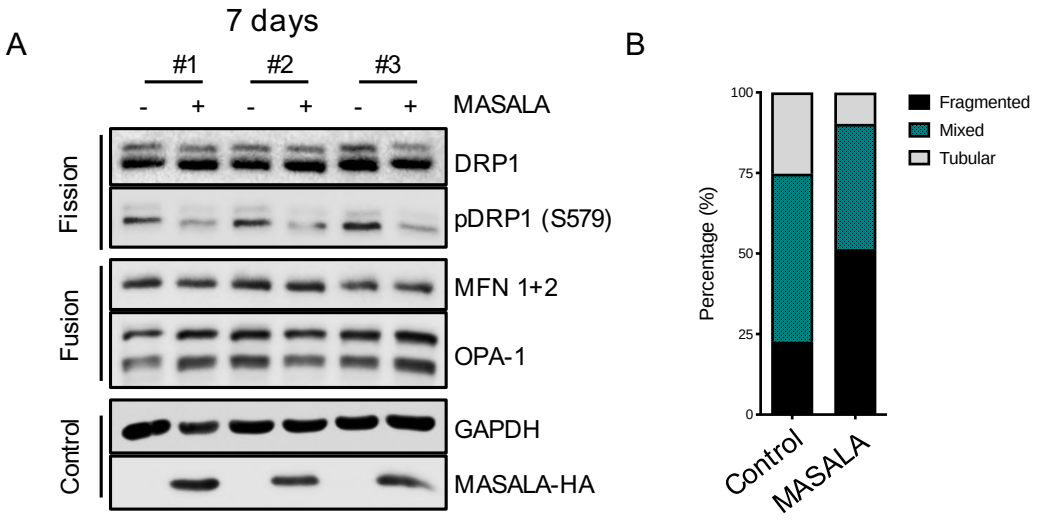


Figure 41. MASALA promotes mitochondrial fission after 7 days of overexpression. (A) Western blot analysis of exogenous MASALA (using and anti-HA antibody) and the indicated mitochondrial dynamics markers in MEFs control or overexpressing MASALA. n=3 biological replicates. **(B)** Quantification of the different mitochondrial phenotypes frequencies in MEFs overexpressing MASALA for 7 days.

As we observed at 24h, we did not observe changes in the fusion markers MFN1 and 2 or OPA-1, suggesting that MASALA is specifically affecting mitochondrial fission.

To further corroborate these observations, we decided to analyze by Transmission Electron Microscopy (TEM) MEFs overexpressing MASALA for 6 days. This analysis confirmed that cells overexpressing MASALA presented more rounded, small mitochondria compared to control cells, consistent with a role of MASALA inducing mitochondrial fission **(Fig 42)**.

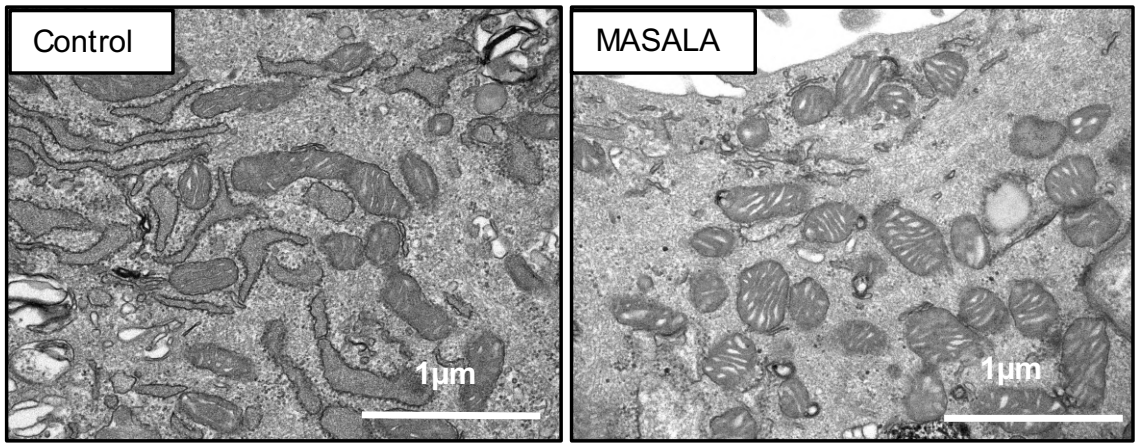


Figure 42. MASALA promotes mitochondrial fission after 6 days of overexpression. Representative TEM images of MEFs overexpressing MASALA-HA for 6 days compared to control MEFs.

3.9. MASALA produces mitochondrial alterations and glycogen accumulation

Besides the confirmation of the induction of mitochondrial fragmentation upon MASALA overexpression, TEM analyses also revealed several events associated to the overexpression of MASALA. Although further quantification is needed, differences were noticeable at sight. The first observation was that mitochondria of cells overexpressing MASALA have a significant amount of matrix swelling events **(Fig 43)**. These events normally appear associated to different mitochondrial dysregulations, such as oxidative stress or deficient calcium handling by ER/mitochondria that causes an ionic unbalance resulting in a matrix swelling (Webster, 2012).

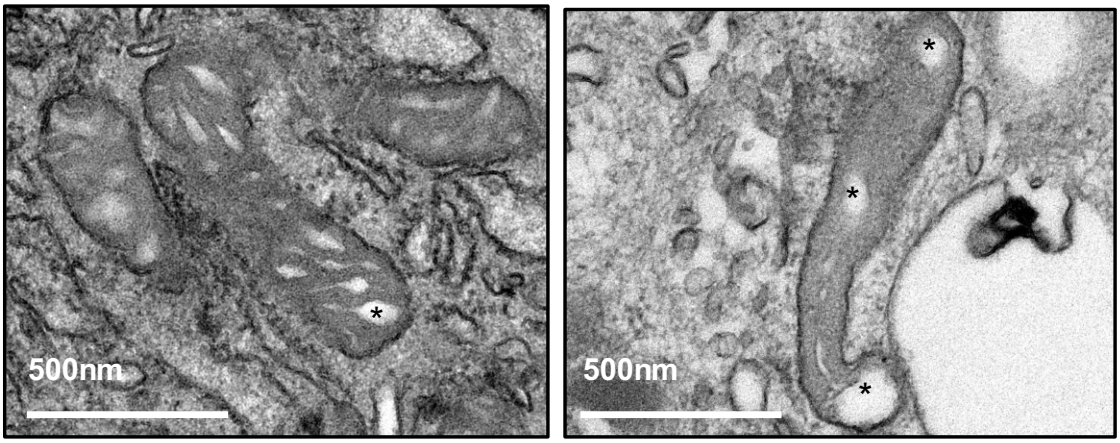


Figure 43. MASALA promotes mitochondrial matrix swelling after 6 days of overexpression. Two representative TEM images of mitochondrial in the mitochondria of MEFs overexpressing MASALA-HA for 5 days. Asterisks are pointing to swelling events.

Accordingly, consistent with an altered mitochondrial biology, another observation that caught our attention was the increased frequency of mitochondrial spheroids in MEFs overexpressing MASALA. These mitochondrial spheroids originate from the invagination of the mitochondria **(Fig 44A)**, that ultimately generates a cavity in the middle of the organelle that acquires a sphere vesicle-like shape **(Fig 44B).**

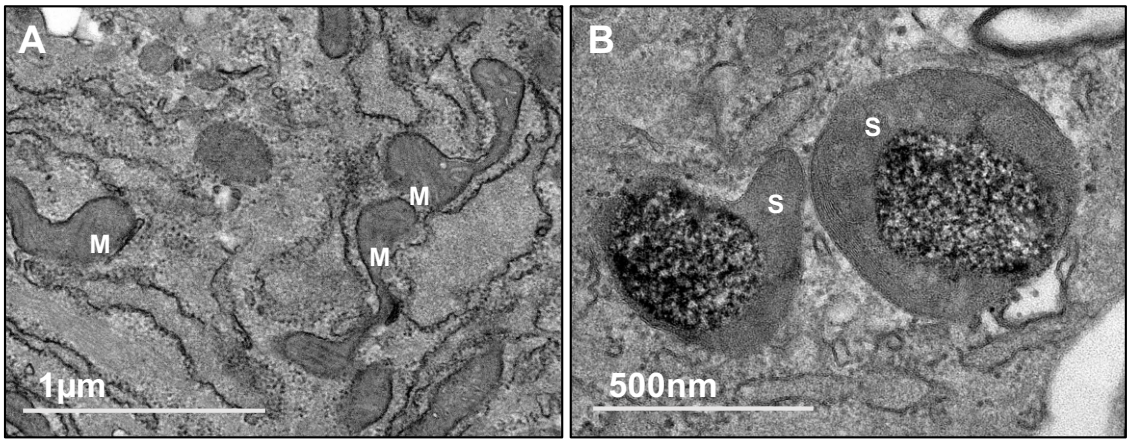


Figure 44. MASALA overexpression produces mitochondrial spheroid formation. (A) Representative TEM image of mitochondrial spheroid precursors coming from mitochondria invagination events (M) in MEFs overexpressing MASALA for 6 days. **(B)** Representative TEM image of mitochondrial spheroids forming a vesicle-like organelle that contains electro-dense particles inside (S) in MEFs overexpressing MASALA for 6 days.

Although the biological relevance of these mitochondrial spheroids has not been fully investigated, several evidences suggest that mitochondria acquire this shape upon uncoupling or oxidative stress (Ding et al., 2012; Miyazono et al., 2018).

Finally, in MEFs overexpressing MASALA we observed a higher accumulation of the isolated alpha-glycogen or aggregated beta-glycogen particles in the cytoplasm, observed as very electro-dense particles **(Fig 45).**

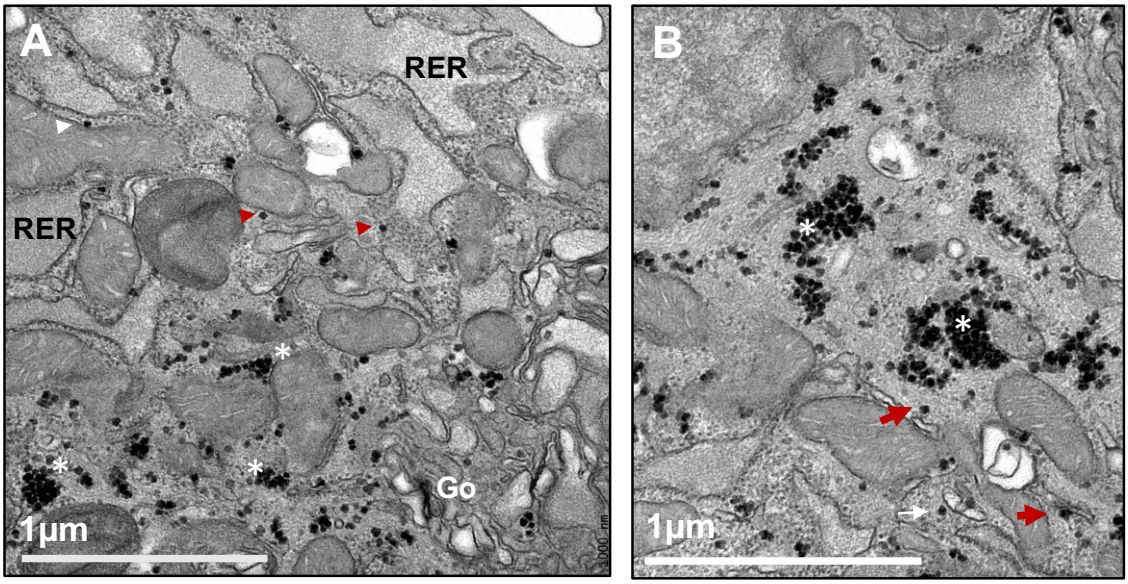


Figure 45. MASALA produces alpha and beta-glycogen accumulation. (A) Representative TEM images of MASALA overexpressing MEFs. The image shows the rough endoplasmic reticulum (RER), the Golgi apparatus (Go) and alpha- (arrows) and beta- (asterisks) glycogen particles. **(B)** Detailed magnification showing alpha and beta-glycogen particles (arrows and asterisks, respectively). Scale bars =1µm

This glycogen accumulation might be indicative of a dysregulation in the glycolytic pathway or a defective activity of the glycogen synthase kinase 3 (GSK-3) pathway. Further studies are needed to corroborate and interpret the meaning of this interesting results.

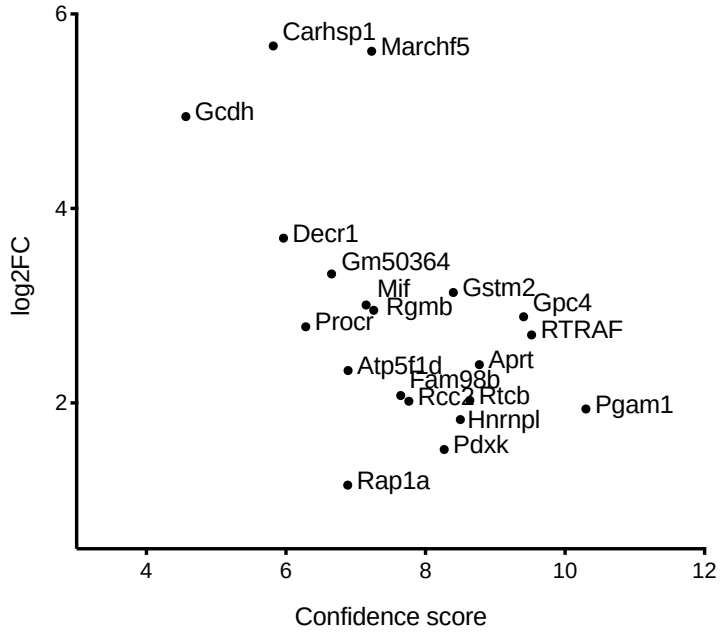
3.10. MASALA might interact with MARCH5 protein

In order to fully comprehend the molecular mechanism behind MASALA's phenotype we wanted to determine MASALA's interactome. We overexpressed MASALA in NIH3T3 cells and performed MASALA immunoprecipitation (using an anti-HA antibody) followed by mass-spectrometry. This analysis revealed several potential interactor candidates. We ranked the candidates according to their fold change in the MASALA condition (FC) and the confidence score obtained from the spectra **(Fig 46A)**. Of note, from the top 15 ranked proteins, 8 of them localize in the Mitochondria (4) or the Endoplasmic Reticulum (4) **(Fig 46B)**.

Interestingly, the interactor candidate with the best FC and good confidence score was the Membrane-associated RING finger protein 5 (Marchf5 or MITOL). This protein has been reported to be located in the outer mitochondrial membrane, and it is known to be an important regulator of the mitochondrial dynamics acting as a E3-ubiquitin ligase that regulates the ubiquitylation of several mitochondrial proteins like Drp1 or Mitofusins (Karbowski et al., 2007; Sugiura et al., 2013; Yonashiro et al., 2006). We are currently

trying to validate this interaction by co-immunoprecipitation. In case we demonstrate it, we will perform functional assays to decipher the nature and the outcomes of this interaction in connection with mitochondrial dynamics (see Discussion).

A



B

Mitochondria	4	Marchf5, Gcdh, Decr1, Atp5f1d
ER	4	Gstm2, Rgmb,Fam98b, Rtcb
Cytosol	3	Aprt, Carhsp1, RTRAF
Secreted	2	Mif, Gpc4
Plasma memb.	2	Gna13, Gpc4

Figure 46. MASALA interactome candidates. (A) Graphical representation of the interaction candidates identified by mass-spectrometry, represented according to the Fold change (Y axis) and its confidence score (X axis). **(B)** Table of the TOP15 candidate's subcellular localization. Localization of the proteins was obtained from UniProt DB.

Discussion

After the complete sequencing of model organisms and human genomes it was assumed that the coding genome represents only a small fraction of the whole genome (1% in humans) (Pennisi, 2012). At that time, the other 99% of the genome was considered useless or "junk DNA". This assumption has been changing over the years, as with large-scale transcriptomic analyses many different types of RNA molecules, such as lincRNAs and miRNAs emerged as important components of many cellular processes, despite they were considered to lack coding potential. More recently, advances in proteomics, ribosome profiling and computational technologies have revealed that many assumed non-coding regions actually code for small functional proteins, called microproteins, adding a new layer of complexity to biological processes. In this thesis, we have focused on finding novel microproteins coded by lncRNAs with a potential role in cellular plasticity and cancer.

1. Identification of MASALA microprotein

1.1. The discovery of MASALA.

In this work we have identified MASALA, a novel 78 amino acid microprotein encoded by *SMIM10L2A* (or *LINC0086),* a gene previously annotated as a lncRNA that is located in the X chromosome in mouse and human. We identified MASALA sORF in the first exon of the gene, and observed that it is highly conserved across placental mammals at codon level **(Fig 8C)**. Expression of this gene has been observed in brain, testis and adrenal gland **(Fig 13 & 14)**. Accordingly, using our custom-made antibody against MASALA we have detected the endogenous protein in the adrenal gland cortex. We further obtained evidence of MASALA translation in mouse and human brain by analyzing ribosome profiling data **(Fig 8B)** and, moreover, public mass-spectrometry data showed that MASALA peptides can be detected in human hippocampus and testis (He et al., 2018; Hwang et al., 2015; Sun et al., 2018; Wei et al., 2016). In testis, despite its high transcription level, MASALA protein is found in low-abundance by mass spectrometry, which could be caused by issues with sample processing or by technique sensitivity limitations due to its small size. Alternatively, it is possible that transcription and translation of this gene are uncoupled in this organ, as it happens with other mRNA transcripts that do not correlate with their protein levels (Koussounadis et al., 2015). Given that this gene also has a function as an RNA molecule in the nucleus (Leveille et al., 2015), it is possible that different cellular contexts might favor transcription over translation, impairing the nuclear export of the RNA molecule to the cytoplasm where translation occurs. This would explain the absence of MASALA signal in our immunofluorescence experiments in testis.

In addition, and in line with our first criteria of analyzing lncRNAs dysregulated in cancer and in stemness, we observed that *SMIM10L2A* is highly expressed in stem cells **(Fig 24A)** and it is downregulated in different cancer types **(Fig 17)**. Accordingly, we have observed an increase of the expression of MASALA microprotein during the process of cellular reprogramming to iPSCs **(Fig 24B)**, and using publicly available mass-spectrometry data we have seen that the protein has been detected in human embryonic stem cells (Weldemariam et al., 2018).

Altogether, we concluded that *SMIM10L2A* was miss-annotated as a lncRNA and codes for a novel 78 aa microprotein that we have named MASALA.

1.2. MASALA upregulation upon stress

In order to understand the biological relevance of this novel microprotein, we first investigated how the gene is regulated. Previous studies identified the function of *SMIM10L2A* as an enhancer RNA upregulated upon p53 stabilization that promotes the expression of *CDKN1A* (Leveille et al., 2015). For this reason, we tested the regulation of *SMIM10L2A* upon chemically induced DNA damage with doxorubicin and observed an upregulation of the transcript in a p53 dependent manner **(Fig 15 & 16).** p53 is a profoundly studied transcription factor known as the guardian of the genome, which senses cellular damage and regulates the expression of several genes implicated in stress responses such as cell cycle arrest, apoptosis and cellular senescence (Hafner et al., 2019; Kastenhuber and Lowe, 2017). Accordingly, p53 is the most frequently mutated tumor suppressor gene in many cancer types (Kandoth et al., 2013). The upregulation of MASALA transcript upon damage, and the fact that its upregulation is dependent on p53, suggests a potential role of the microprotein in response to stress and tumor suppression.

2. Functional characterization of MASALA

2.1. MASALA in cancer

The expression of *SMIM10L2A* is downregulated in different cancer types **(Fig 17)** and its expression has been correlated with better patient outcome in nasopharyngeal carcinoma and gastric cancer, (Guo et al., 2017; Li et al., 2016). This, together with its regulation by stress and by p53, suggest a possible role of the MASALA as a tumor suppressor. In order to investigate the function of MASALA microprotein independently of *SMIM10L2A* function as a lncRNA, we used a gain-of-function approach using a lentiviral vector in which we cloned MASALA sORF tagged with an HA. The overexpression of the sORF encoding MASALA and not the complete lncRNA allows us to separate the function of the lncRNA from the function of MASALA, because we only overexpress a small fragment of *SMIM10L2A*, and lncRNAs' activity is highly dependent on their molecular structure (Graf and Kretz, 2020; Zampetaki et al., 2018).

We could not observe any effect of MASALA overexpression on cell growth or in cell cycle in A549 and GL261 cancer cell lines **(Fig 18 & 19).** Of note, *CDKN1A* gene was not upregulated upon MASALA overexpression **(Fig 19)**, what further confirms that our overexpression construct does not mimic *SMIM10L2A* function. Although we could not observe any effect of MASALA in already established cancer cell lines, we wondered if MASALA overexpression does have an effect in the neoplastic transformation of primary cells. Importantly, our data indicates that MASALA overexpression drastically reduces the efficiency of oncogenic transformation of primary fibroblasts with mutant HRAS and E1A protein **(Fig 20)**. To further test the role of MASALA as tumor suppressor we also plan to repeat this oncogenic transformation experiments in MEFs deficient for MASALA where we expect to see an increase in the transformation efficiency.

Next, as the expression of *SMIM10L2A* is downregulated in glioblastoma and gliomas, we decided to test the tumor suppressor activity of MASALA *in vivo*, using the RCAS/GFAP-*tv-a* glioblastoma mouse model with PDGFA, shp53 and shNF1 as oncogenic drivers. Surprisingly, we could not observe differences in tumor incidence or in the survival of mice overexpressing MASALA over the control group **(Fig 22)**. This, however, is a very aggressive model with a very fast tumor onset and it is possible that MASALA overexpression is not enough to overcome the effect of the oncogenic drivers. Another possibility for further investigation would be to use a different and less aggressive combination of oncogenic drivers.

To further test the tumor suppressor effect of MASALA in a different model, we transformed the immortalized-fibroblast cell line NIH3T3 with $HRAS^{G12V}$ in combination or not with MASALA, and injected them into the flanks of immunodeficient mice. As observed in the RCAS/GFAP-*tv-a* model, MASALA did not impair or delay the tumor formation **(Fig 23)**. Although we have not seen differences in the tumor onset, we need to confirm the presence of MASALA in the tumors together with a careful examination of the histopathological differences between the tumors expressing MASALA and the control tumors. The differences observed between oncogenic transformation *in vitro* and *in vivo* can be due to several reasons including, for instance, the type of target cells and the oncogenic drivers used. The comparison between the transformation of MEFs *in vitro* by E1A/RasV12 and the transformation of NIH3T3 cells *in vivo* by Ras V12 suggest that the presence of E1A is needed for MASALA to inhibit oncogenic transformation. E1A oncoprotein is known to induce p53 stabilization (Nakajima et al., 1998), what makes cells to be more sensitive to proapoptotic cues. Thus, it is possible that MASALA impairs oncogenic transformation by priming cells to apoptosis in cooperation with E1A. Although we have not observed differences in the induction of apoptosis upon different stimuli **(Fig 30)**, it is possible that times longer than 24-48 hours are needed for MASALA to display an effect on apoptosis.

2.2. MASALA in cellular reprogramming

The acquisition of cellular plasticity is a feature of both cellular reprogramming and neoplastic transformation, being recently recognized as a new a new hallmark of cancer (Hanahan, 2022). It makes perfect sense that, when we were looking for novel regulators of cancer cell plasticity, we found a microprotein that is upregulated in pluripotent stem cells, the paradigm of cellular plasticity.

Cellular reprogramming and neoplastic transformation share many characteristics, one of them is the activation of tumor suppressor genes, which act as barriers for both processes (Kawamura et al., 2009; Li et al., 2009; Marion et al., 2009; Utikal et al., 2009). Consistent with its role as a tumor suppressor, our data indicate that the expression of MASALA increases during the process of cellular reprogramming, with a peak in the levels of mRNA expression at day 4 followed by a subsequent decrease, while the expression of the microprotein increases and gets stabilized along the process **(Fig 24B)**. In addition, as MASALA drastically reduces oncogenic transformation of primary fibroblasts, we wanted to test its ability to inhibit the process of cellular reprogramming to iPSCs. We observed that MASALA significantly impairs the reprogramming of MEFs to iPSCs by OSKM **(Fig 25)**, once again suggesting that MASALA acts as a tumor suppressor in this process. In order to gain more insight on MASALA effect during reprogramming, we used our MASALA-KO mouse model to test if MASALA deficiency could improve reprogramming efficiency. Surprisingly, MASALA-KO MEFs reprogrammed as efficiently as MASALA-WT MEFs (**Fig 26).** Microproteins generally act as fine-tune regulators of cellular processes, thus, it is likely that MASALA-deficient cells compensate MASALA deficiency by upregulating or activating other tumor suppressor proteins. We are currently studying the differences between WT and MASALA-KO iPSCs to know the role of MASALA in iPSC self-renewal and differentiation.

Additionally, we are performing *in vivo* reprogramming in MASALA-deficient mice (by crossing them with the reprogrammable i4F mice) to test a possible improvement in the reprogramming dynamics *in vivo*. Although we have not seen an effect of MASALA deficiency *in vitro*, it might be that different cellular contexts changes the tumor suppressor

activity of MASALA, as it has been observed with other tumor suppressor genes (Mosteiro Science 2016).

2.3. Role of MASALA in homeostasis

In this work, we have mainly focused on determining the role of MASALA in cancer and cellular reprogramming. Nevertheless, based on the expression pattern of *SMIM10L2A* in human and mouse organs, together with our evidences from the characterization of the protein, we can hypothesize over the role of MASALA in physiological conditions.

Expression of *SMIM10L2A* in brain, testis and adrenal gland suggests that MASALA could be related to the Hypothalamic-Pituitary-Adrenal (HPA) axis or the Hypothalamic-Pituitary-Gonadal axis (HPG). The HPA axis is the set of interactions between the hypothalamus, the pituitary gland and the adrenal gland. This axis and the organs implicated is one of the major components of the neuroendocrine system, with important roles in the regulation of different body responses to stress, such as metabolic activity, immune system and emotions (Smith and Vale, 2006). The HPG axis is also part of the neuroendocrine system and regulates the production of sexual hormones in both male and female organisms (Dagklis et al., 2015; O'Shaughnessy et al., 2003). Interestingly, a common feature of both axes is the biosynthesis of steroid hormones derived from cholesterol, such as testosterone, progesterone, mineralocorticoids or glucocorticoids. Our analyses indicate that MASALA is expressed in the cortical area of the adrenal gland rather than the medullar area. More specifically, the expression is restricted to the *zona glomerulosa* **(Fig 14)**. The hormone production in the adrenal gland cortex is very compartmentalized into the three different layers, being the *zona glomerulosa* the area where mineralocorticoids are produced (Ehrhart-Bornstein and Hilbers, 1998). The main mineralocorticoid produced in the body is the aldosterone. Aldosterone is a hormone of the renin-angiotensin axis that regulates blood tension by promoting Na^+ and water retention in the plasma, acting on distal tubules and collecting ducts of the nephron (Atlas, 2007; Danziger and Zeidel, 2015). MASALA localization in the *zona glomerulosa* suggests that it might be implicated in the biosynthesis of this hormone from cholesterol. Additionally, aldosterone biosynthesis is promoted by mitochondrial calcium concentrations (Bird et al., 1995; Rossier et al., 1996). Given that at the subcellular level MASALA localizes in the outer membrane of the mitochondria, it is possible that MASALA plays a role in the mitochondrial calcium channeling. This could be in accordance with TEM observations of mitochondrial matrix swelling events that have been correlated with osmotic imbalances **(Fig 43)** (Webster, 2012). Finally, it has been demonstrated that aldosterone treatment induces mitochondrial fission in a p53/DRP1 dependent manner (Yuan et al., 2018), what correlates with the upregulation of MASALA upon p53 stabilization and its role in mitochondrial dynamics. An important experiment that we will perform in this regard is to analyze the levels of aldosterone in the plasma of MASALA-KO mice, in order to know if MASALA deficiency in the adrenal gland can have an impact into mineralocorticoid production. We will also study the possible effects of aldosterone dysregulation in the blood pressure and the stress response. Similarly, although we have not yet detected endogenous expression of MASALA microprotein in testis, and we have not observed a defect of MASALA-KO mice in fertility, it would be interesting to test whether MASALA deficiency is affecting the production of steroid sexual hormones.

Finally, although we have not deeply characterized MASALA-KO phenotype in the mice, a preliminary study has shown no major macroscopic alterations in any of the organs. We are currently performing a histological analysis of the different tissues in which MASALA is normally expressed looking for abnormalities in MASALA deficient animals.

3. Molecular mechanism

3.1. MASALA in mitochondria

The discovery that cells can be reprogrammed *in vivo* opened the possibility of using this technology to promote regeneration of damaged tissues. However, the induction of dedifferentiation *in vivo* can also result in the formation of tumors, as it has been previously observed (Abad et al., 2013; Goding et al., 2014; Ohnishi et al., 2014). For this reason, a good understanding of the molecular processes behind the intersection between cellular reprogramming and neoplastic transformation is crucial to develop novel therapeutic strategies in regenerative medicine and cancer.

In order to determine how MASALA impairs cellular reprogramming and oncogenic transformation, we studied the molecular characteristics of the microprotein as well as the cellular processes in which it might be implicated. Our *in silico* analyses revealed that MASALA is a 78-amino acid microprotein that has a mitochondrial targeting sequence in the N-terminal region and a transmembrane domain in the C-terminal part **(Fig 9)**. Accordingly, we have observed that MASALA is mainly located in the mitochondrial outer membrane **(Fig 27C and 28)**. The mitochondrial outer membrane has a very specific proteome, implicated in a plethora of mitochondria-related functions.

To characterize how MASALA affects mitochondrial biology we performed a set of experiments testing different mitochondria-related processes, concluding that MASALA is involved in the regulation of mitochondrial dynamics **(Fig 38, 39 and 40).**

3.1.1. MASALA in apoptosis

The mitochondrial outer membrane is an essential compartment in the apoptosis signaling cascade. Regulation of the membrane permeability by anti and pro-apoptotic proteins is crucial for the release of cytochrome c that ultimately triggers Caspase activation (Singh et al., 2019). As MASALA is located in the outer mitochondrial membrane and it is regulated upon stress in a p53 dependent manner **(Fig 15 & 16)**, we wondered if MASALA microprotein could play a role in the apoptotic process. Our results concluded that MASALA microprotein overexpression, by itself, could not reduce cell number or reduce cell viability **(Fig 30)**. Moreover, MASALA overexpression or MASALA deficiency did not have an effect in the induction of apoptosis by treatment with staurosporine or the BCL-2 inhibitor Navitoclax, measured by Caspase 3/7 activation **(Fig 30B)**. We concluded that MASALA is not directly implicated in the apoptotic process. This data, however, was obtained 24 hours after MASALA induction. Although the expression of MASALA for longer periods did not alter the cell growth **(Fig 18)**, it is possible that those cells are primed to the induction of apoptosis after several days of treatment, as discussed above. Further analyses are needed to completely discard an effect of MASALA in the apoptotic process. Indeed, longer treatment experiments have revealed that MASALA-overexpressing cells have several mitochondrial alterations compatible with affectations in apoptotic sensitivity. Analysis by immunofluorescence and TEM revealed that cells that overexpressed MASALA for 5 days or more have smaller mitochondria and a stressed mitochondrial phenotype **(Fig 38, 41, 42 and 43)**. As others have described, mitochondrial fragmentation is an early and necessary event of the apoptotic process (Cereghetti et al., 2010; Karbowski and Youle, 2003). In addition, the observed mitochondrial swelling **(Fig 42)** might indicate a defect in calcium concentrations that cause osmotic imbalances in the mitochondrial matrix (Webster, 2012). As mitochondrial intracellular calcium concentration is crucial for the apoptosis-signaling cascade (Orrenius et al., 2003), these results are compatible with a role of MASALA in apoptosis through the

regulation of calcium dynamics. This could explain how MASALA reduces the efficiency of cellular reprogramming or oncogenic transformation. Moreover, Ca^{2+} concentrations have been directly linked to mitochondrial dynamics regulation. It has been reported that a continuous Ca^{2+} influx in hippocampal neurons and astrocytes promotes mitochondrial fragmentation in a DRP-1 dependent manner (Li et al., 2004; Tan et al., 2011; Yi et al., 2004). Consistently, induction of mitochondrial fission drastically reduces mitochondrial Ca^{2+} uptake independently of the calcium uniporter expression, while mitochondrial fusion increase mitochondrial Ca^{2+} concentrations (Favaro et al., 2019; Kowaltowski et al., 2019). In any case, we think that the role of MASALA in apoptosis would be indirect, and not as an effector in the signaling cascade.

3.1.2. MASALA and mitochondrial oxidative stress

Mitochondrial activity, even in physiological conditions, produces ROS as a byproduct of the oxidative phosphorylation and the ETC activity. These ROS are further increased upon mitochondrial dysfunction, with important consequences in DNA damage, protein instability and lipid-membranes oxidation (Schieber and Chandel, 2014). Although ROS are generally viewed as damaging byproducts of cellular respiration, they can also act as signaling molecules in biological processes such as cell differentiation. Stem cells are intrinsically low ROS producing cells and the increase in ROS concentrations is a necessary event of cellular differentiation (Sinenko et al., 2021). Additionally, ROS play a dual role in cellular reprogramming and in cancer initiation processes. A certain level of ROS is crucial for an efficient reprogramming to iPSC, as both depletion or excessive ROS can ablate de process (Zhou et al., 2016). In cancer, oncogenic mutations contribute to increase ROS, what generates a positive feedback loop of DNA-damage, genome instability and inactivation of crucial proteins for tumor suppression (Sabharwal and Schumacker, 2014). We hypothesized a potential implication of MASALA in the cellular stress response and envisioned a role of MASALA in ROS regulation. Our data indicates that the stability of MASALA is increased upon oxidizing agents **(Fig 37)**. Importantly, the stabilization of exogenous MASALA upon the uncoupling agent FCCP and more evidently with H_2O_2 suggested a role of MASALA in the antioxidant response. The overexpression of MASALA or its deficiency, however, did not affect the amount of ROS in the cell or the amount of mitochondrial ROS **(Fig 31 and 32)**, indicating that MASALA is not affecting ROS production or the antioxidant response.

Nevertheless, the stabilization of MASALA upon oxidative agents is suggestive of a role in the mitochondrial damage response. FCCP and H_2O_2 produce a high mitochondrial stress that ultimately leads to the clearance of damaged mitochondria through mitophagy (Frank et al., 2012; Kondapalli et al., 2012; Matsuda et al., 2010). This process eliminates unhealthy mitochondria through engulfment of damaged mitochondria in doubled-membrane vesicles (autophagosomes) that ultimately digest them by fusing with lysosomes (Ma et al., 2020). Our observations indicate that cells overexpressing MASALA for 6 days present mitochondrial alterations suggestive of defective mitophagy. Firstly, mitochondria of MASALA-overexpressing cells more frequently present a stressed phenotype with matrix swelling events and secondly, we observe an increase in the frequency of mitochondrial spheroids **(Fig 43 and 44)**.

These observations could indicate an accumulation of damaged mitochondria in the cytoplasm due to a defective elimination process. Whether MASALA is related with the process of mitophagy is not yet known and further experiments are currently being performed to address this question.

3.1.3. MASALA in mitochondrial metabolism

Mitochondria generate ATP from different metabolic substrates using a series of chemical transformations that oxidize those substrates in a cyclic manner. This is known as the TCA cycle or Krebs cycle and it is crucial to provide an electron flux to the ETC for an effective OXPHOS. Our metabolomic analysis of MEFs overexpressing MASALA for 24 hours did not show relevant changes in the metabolic flux or the amount of the different TCA cycle metabolites. Our only significant observation was a reduction in the levels of one specific isotopologue of the 2-HG oncometabolite, suggesting a possible effect of MASALA in IDH1/2 enzymatic activity. 2-HG metabolite has been implicated in different cancers with IDH1/2 mutations such as gliomas and acute myeloid leukemia (Mardis et al., 2009; Parsons et al., 2008), promoting glutamine metabolism and regulating histone methylation through the modulation of the TET enzyme (Du and Hu, 2021; Hollinshead et al., 2018; Shim et al., 2014). This reduction in 2-HG upon MASALA expression is not relevant for primary fibroblasts in homesotasis, where this metabolite is produced at very low levels, but could be of importance in oncogenic transformation and during cellular reprogramming, where TET enzyme is known to be needed for DNA demethylation (Bagci and Fisher, 2013; Rasmussen and Helin, 2016; Singh et al., 2020). Further experiments are needed to test the effect of MASALA in 2-HG production during these processes.

Even if MASALA is located in the mitochondrial outer membrane, and the mitochondrial respiration occur in the matrix and the inner membrane, the measurement of cellular oxygen consumption rate can provide useful information to determine the metabolic status of the cell and infer possible mitochondrial alterations. Our data shows that overexpression of MASALA for 24 hours is not sufficient to change the oxygen consumption rate or the membrane potential generated through the ETC **(Fig 36A)**. Moreover, MEFs deficient for MASALA do not present any differences in mitochondrial respiration compared to WT MEFs **(Fig 35)**. Five days of MASALA overexpression in MEFs, however, were sufficient to observe a reduction in mitochondrial oxygen consumption rate and a reduced spare respiratory capacity **(Fig 36B)**. Spare respiratory capacity is the parameter that measures the ability of the mitochondria to respond and cope with rapid increasing needs in energy demands. This indicates that after 5 days of MASALA expression cells are not able to meet the energy requirements, reducing the metabolic capacity of the mitochondria and suggesting a less dependency of the cells on OXPHOS. Several studies have addressed the respiration dynamics in the context of cellular reprogramming. Some of them observed that cells with reduced spare respiratory capacity have a better reprogramming efficiency (Zhou et al., 2017), but it seems that reprogramming is a dynamic process that requires both, an increase of respiratory capacity at early phases and a reduction later in the process (Kida et al., 2015). We observed an upregulation of MASALA protein during the process of reprogramming, indicating that it may participate in the reduction of the mitochondrial respiration during reprogramming, and this is also consistent with its high expression in iPSCs, which have very low respiration rates and more glycolysis dependency (Bahat and Gross, 2019). We hypothesize that, when MASALA is overexpressed from the very beginning of the process it might be impairing the oxygen consumption increase necessary for cellular reprogramming, what ultimately have a detrimental effect on the iPSC formation efficiency.

Interestingly, a reduction of mitochondrial respiration is characteristic of cells with a more fragmented mitochondrial phenotype (Vorobjev and Zorov, 1983; Zorov et al., 2019), consistent with our microscopy analysis that show that MASALA induces mitochondrial fragmentation **(Fig 38 & 41)**. Thus, the reduction in mitochondrial respiration can be an indirect effect of MASALA regulating the mitochondrial dynamics.

3.2. MASALA and mitochondrial dynamics in cellular reprogramming and cancer

Mitochondria are dynamic organelles that undergo fusion and fission cycles in order to correctly display their function. A correct regulation of the mitochondrial dynamics is crucial for the cell to cope with metabolic requirements and maintain the homeostasis. Dysregulation of mitochondrial dynamics has been shown to be highly detrimental for cell viability, with implications in apoptosis and mitophagy that ultimately can lead to metabolic, muscular and neurological pathologies, which at the end also impact on cancer and aging (Dai and Jiang, 2019; Scaini et al., 2021; Sebastian et al., 2017).

Importantly, mitochondrial fission is a key step in the process of mitophagy that ensures proper elimination of defective or unhealthy mitochondria (Kobayashi et al., 2020; Ma et al., 2020). Our data indicates that MASALA promotes mitochondrial fragmentation, as observed by mitochondria immunostaining and electron microscopy experiments in cells overexpressing MASALA **(Fig 39 & 42)**. Moreover, we have observed that MASALA modifies the activating phosphorylation of DRP1 in Ser579. Although not very consistently, MEFs overexpressing MASALA for 24 hours presented an increase in pDRP1(S579), indicating that MASALA might promote mitochondrial fission regulating the post-translational modification of DRP1 **(Fig 39).** This activating effect, however, turned the opposite when MASALA expression was maintained for more than 5 days **(Fig 41)**. Reduction in DRP1 phosphorylation after long MASALA expression might be indicative of an accumulative effect of MASALA in the mitochondria, that results in a compensatory effect of the cell.

Mitochondrial fission produced by DRP1 activation is a necessary early event in the process of cellular reprogramming and in tumor initiation (Prieto and Torres, 2017). In this sense, we would expect that the pro-fission effect of MASALA would result in an increase of both cellular reprogramming and oncogenic transformation, and not the opposite. However, as we have discussed earlier, these processes are very dynamic and different energetic requirements need to be addressed at different times of the process. As we have observed, prolonged expression of MASALA microprotein results in less phosphorylation of DRP1 (Ser579) what could explain the impairment in cellular reprogramming and oncogenic transformation processes, in line with some reports indicating that DRP1 inactivation drastically reduces cellular reprogramming (Vazquez-Martin et al., 2012a). Although not mutually exclusive, another explanation is that MASALA produces mitochondrial fragmentation but at the same time it negatively affects the process of mitophagy. This would produce an excess of fragmented mitochondria that are not completely functional. In this line, evidences suggest that mitophagy and autophagy are necessary for the mitochondrial clearance during cellular reprogramming (Wang et al., 2013b; Xiang et al., 2017)

3.2.1. MASALA and MARCH5

To further investigate the molecular mechanisms behind MASALA's effect on cellular reprogramming, we studied MASALA's interactome. For that, we performed an immunoprecipitation of MASALA microprotein followed by mass-spectrometry. This analysis revealed several MASALA interactor candidates, that we ranked based on their fold change and confidence score **(Fig 45).** The candidates ranked with the highest score localize in the mitochondria and the ER, as we expected based on MASALA localization. The most interesting candidate was the membrane-associated ring protein-CH 5 (MARCH5), an E3-ubiquitin ligase located in the mitochondrial outer membrane.

Importantly, this protein has been implicated in the process of mitochondrial dynamics, regulating DRP1, FIS1 and MFN2 ubiquitination (Nakamura et al., 2006; Yonashiro et al., 2006). Moreover, Karbowski et al. demonstrated that mutant versions of MARCH5 or its silencing produces mitochondrial elongation and aberrant mitochondrial localization, indicating that DRP1 activation through ubiquitination by MARCH5 is necessary for the fission process (Karbowski et al., 2007). Even more, MARCH5 has been demonstrated to be essential for ES pluripotency maintenance and for cellular reprogramming to iPSCs (Gu et al., 2015). Altogether, these evidences reinforce the idea that a potential interaction of MASALA with MARCH5 could alter its function as an E3-ubiquitin ligase. It is possible that overexpression of MASALA improves MARCH5 activity resulting in a more fragmented mitochondrial phenotype that, maintained over long periods of time could induce a compensatory mechanism that negatively regulates DRP1 phosphorylation at Ser579. MARCH5 has been shown to activate DRP1, and DRP1 activation has been shown to promote iPSC reprogramming. This does not fit with our results that show that MASALA impairs reprogramming. However, the post-translational modifications carried out by MARCH5 are not completely understood and other evidences suggest that ubiquitination of DRP1 by MARCH5 produced a more elongated mitochondrial phenotype that would reduce the reprogramming efficiency, as others have observed (Das et al., 2022; Nakamura et al., 2006; Son et al., 2015; Yonashiro et al., 2006). Further investigation is needed in order to clarify the effects of DRP1 ubiquitylation. The abovementioned contradictions may be explained by the possibility that DRP1 can be ubiquitinated in different lysines by different proteins, with different effects. It is possible that the increased activity of MARCH5 induced by MASALA induces a ubiquitination on DRP1 that makes it unstable or marks it for degradation. If we finally confirm the interaction of MASALA with MARCH5, we will characterize the ubiquitination pattern of DRP1 and the other MARCH5 targets upon MASALA overexpression and deficiency.

4. Final considerations and proposed working model

Our studies have demonstrated that the lncRNA *SMIM10L2A* encodes a 78-amino acid microprotein that is upregulated upon damage and is located in the outer mitochondrial membrane. We have observed that MASALA overexpression impairs the process of cellular reprogramming to iPSC as well as oncogenic transformation *in vitro* (**Fig 46**).

The overexpression of MASALA over long periods of time (5 or more days) reduces cellular respiration and spare respiratory capacity. In agreement, we have observed that MASALA overexpression induces a more fragmented mitochondrial phenotype **(Fig 39, 41 & 42)** indicating a role of MASALA in the regulation of mitochondrial dynamics regulation. We have observed that overexpression of MASALA affects DRP1 phosphorylation at Ser579, possibly by its interaction with MARCH5 E3-ubiquitin ligase **(Fig 41 & 46)**. Although we have not completely characterized the molecular mechanism by which MASALA impairs cellular reprogramming, our data suggest that it could be due to a dysregulation of mitochondrial dynamics, probably through its interaction with MARCH5.

This, however, does not exclude other mechanisms. For instance, although we have detected MASALA in the mitochondria, due to the close contact of this organelle with the ER it is possible that MASALA is located also in some regions of ER, as we have observed by immunofluorescence analysis **(Fig 29)**. ER-mitochondrial contacts are crucial for many cellular processes, including Ca^{2+} transport and mitochondrial dynamics. Additionally, MASALA could have a role in mitophagy, which fits with the stabilization of the microprotein upon mitochondrial stress **(Fig 37)** and could potentially explain the impairment of cellular reprogramming and tumor initiation. Additional mechanisms can

involve, as we previously mentioned, a priming to apoptosis, or the impairment of DNA demethylation by regulation of TET enzyme **(Fig 49).**

Based on the evidences that we have, we propose a working model in which MASALA is upregulated upon p53 activation and increases MARCH5 activity, modifying the ubiquitination of DRP1 which, in turn, induces mitochondrial fission **(Fig 48)**. It is known that changes in cell identity need the regulation of mitochondrial dynamics and, therefore, we think that MASALA acts as a barrier for reprogramming and for neoplastic transformation regulating mitochondrial dynamics.

In conclusion, microproteins have emerged as novel class of molecular regulators with important roles in different biological processes, including cancer. In this study, we have identified MASALA, a novel microprotein encoded by a gene annotated as non-coding, that inhibits cellular plasticity through the regulation of mitochondrial dynamics. Our results increase the body of knowledge of the molecular mechanisms behind cellular plasticity and add a new molecular player in this process, MASALA. From a more general point of view, our results have uncovered that the microproteome regulates mitochondrial dynamics, and could hide important players involved in homeostasis and disease that remain to be identified

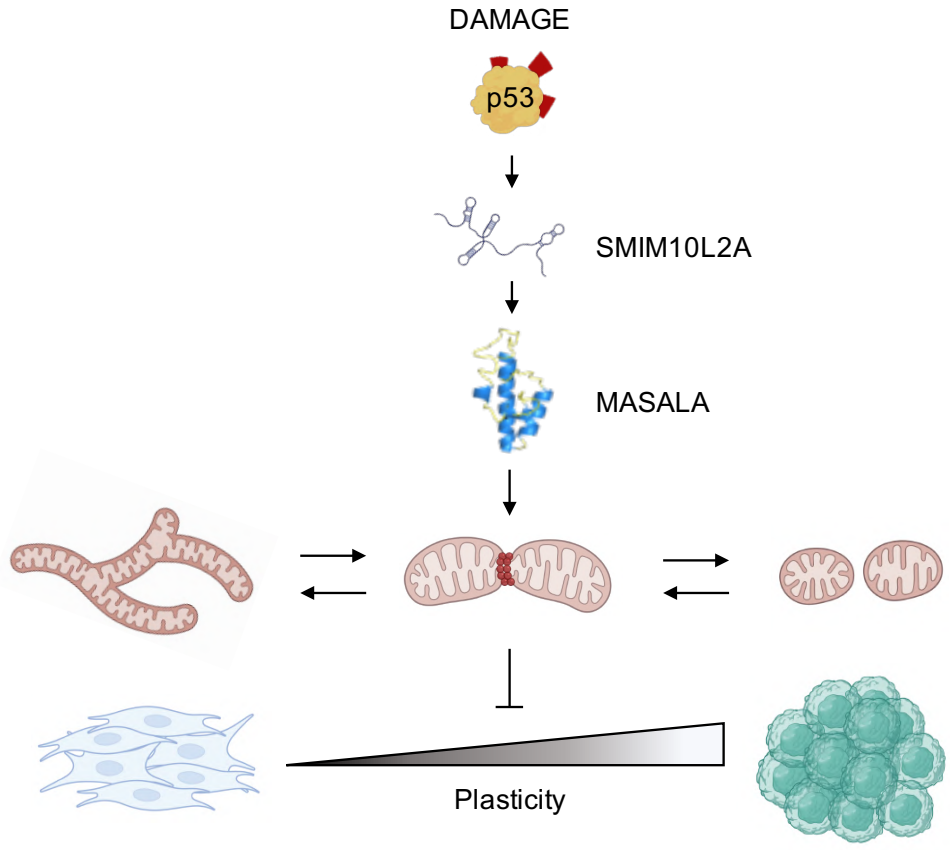


Figure 47. Role of MASALA in damage-induced plasticity. Activation of p53 induces de the expression of SMIM10L2A, which in turn encodes for MASALA microprotein. MASALA localizes to the mitochondrial outer membrane and alters mitochondrial dynamics impairing changes in cell identity.

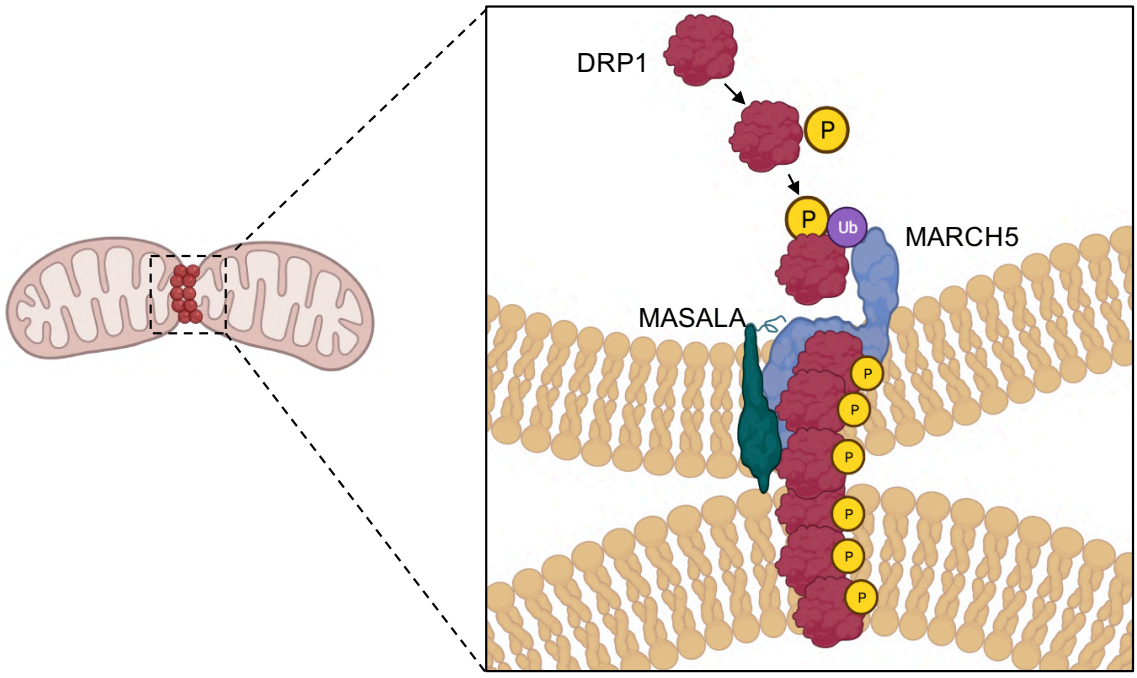


Figure 48. Proposed model for MASALA-MARCH5 interaction effect. Interaction of MASALA with MARCH5 protein would alter the normal ubiquitylation pattern of DRP-1 protein towards a more active phenotype, promoting mitochondrial fission.

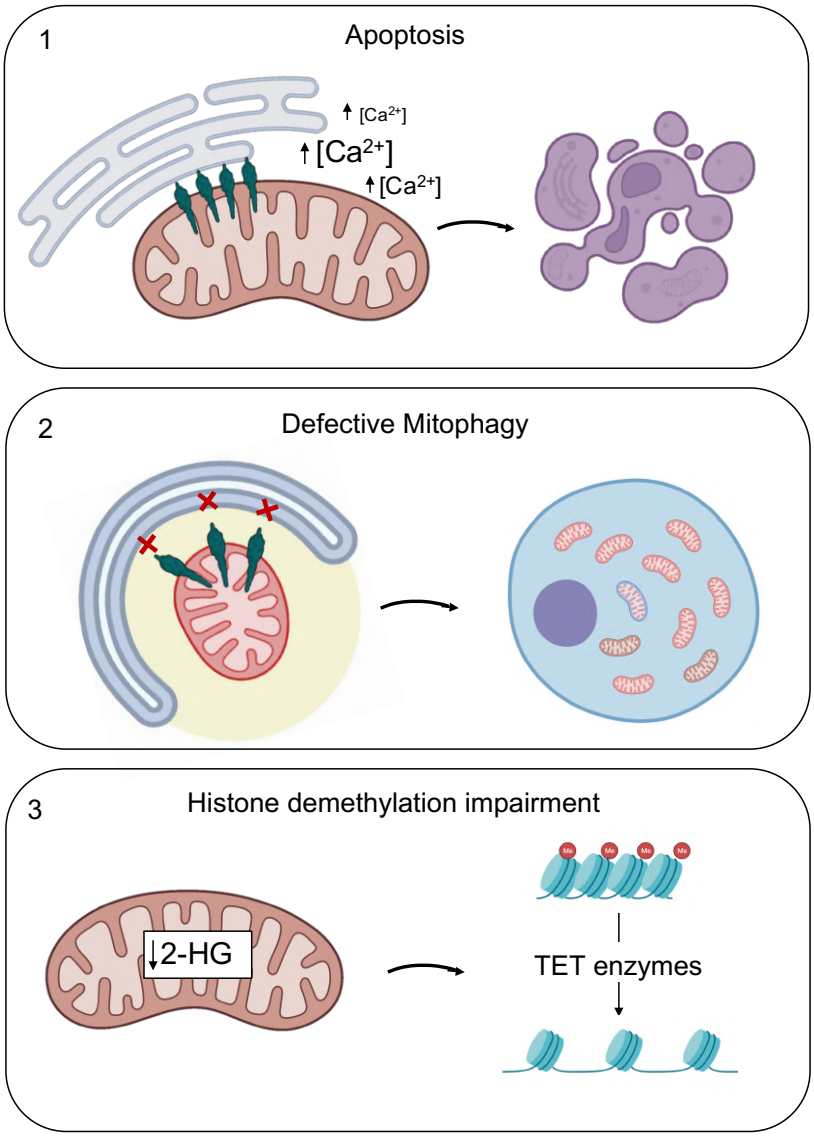


Figure 49. Alternative models for MASALA's way of action. 1. Dysregulation of Ca^{2+} homeostasis that sensitizes cells to apoptosis. **2.** MASALA could interfere with mitophagy-mediated elimination of defective mitochondria, leading to the accumulation of non-functional mitochondria. **3.** Reduction of 2-HG metabolite that impairs histone demethylation by TET enzymes during cellular reprogramming and oncogenic transformation.

Conclusions

1. *SMIM10L2A* contains a highly conserved sORF encoding a 78-amino acid microprotein that we have named MASALA.

2. *SMIM10L2A* is naturally expressed in brain, testis, adrenal gland and in embryonic stem cells. Additionally, it is upregulated upon genotoxic damage in a p53 dependent manner, and downregulated in glioblastomas, gliomas, pheocromocytomas and paragangliomas.

3. MASALA is located in the outer mitochondrial membrane and is stabilized upon mitochondrial stress.

4. MASALA overexpression impairs cellular reprogramming from MEFs to iPSCs.

5. MASALA overexpression in MEFs impairs oncogenic transformation by E1a/HRAS^{G12V}.

6. Continued overexpression of MASALA increases mitochondrial fragmentation, reducing the oxygen consumption rate and the spare respiratory capacity.

7. MASALA interacts with several mitochondrial and ER proteins, being particularly interesting MARCH5, an E3 ligase involved in mitochondrial dynamics.

Bibliography

Abad M (2020). The hidden world of non-canonical ORFs. Exp Cell Res *396*, 112267.

Abad, M., Mosteiro, L., Pantoja, C., Canamero, M., Rayon, T., Ors, I., Grana, O., Megias, D., Dominguez, O., Martinez, D*., et al.* (2013). Reprogramming in vivo produces teratomas and iPS cells with totipotency features. Nature *502*, 340-345.

Almagro Armenteros, J.J., Salvatore, M., Emanuelsson, O., Winther, O., von Heijne, G., Elofsson, A., and Nielsen, H. (2019). Detecting sequence signals in targeting peptides using deep learning. Life Sci Alliance *2*.

Almagro Armenteros, J.J., Sonderby, C.K., Sonderby, S.K., Nielsen, H., and Winther, O. (2017). DeepLoc: prediction of protein subcellular localization using deep learning. Bioinformatics *33*, 3387-3395.

Anderson, D.M., Anderson, K.M., Chang, C.L., Makarewich, C.A., Nelson, B.R., McAnally, J.R., Kasaragod, P., Shelton, J.M., Liou, J., Bassel-Duby, R*., et al.* (2015). A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. Cell *160*, 595-606.

Anderson, D.M., Makarewich, C.A., Anderson, K.M., Shelton, J.M., Bezprozvannaya, S., Bassel-Duby, R., and Olson, E.N. (2016). Widespread control of calcium signaling by a family of SERCA-inhibiting micropeptides. Sci Signal *9*, ra119.

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

Arnoult, N., Correia, A., Ma, J., Merlo, A., Garcia-Gomez, S., Maric, M., Tognetti, M., Benner, C.W., Boulton, S.J., Saghatelian, A*., et al.* (2017). Regulation of DNA repair pathway choice in S and G2 phases by the NHEJ inhibitor CYREN. Nature *549*, 548-552.

Ashrafi, G., and Schwarz, T.L. (2013). The pathways of mitophagy for quality control and clearance of mitochondria. Cell Death Differ *20*, 31-42.

Atlas, S.A. (2007). The renin-angiotensin aldosterone system: pathophysiological role and pharmacologic inhibition. J Manag Care Pharm *13*, 9-20.

Bagci, H., and Fisher, A.G. (2013). DNA demethylation in pluripotency and reprogramming: the role of tet proteins and cell division. Cell Stem Cell *13*, 265-269.

Bahat, A., and Gross, A. (2019). Mitochondrial plasticity in cell fate regulation. J Biol Chem *294*, 13852-13863.

Ban, T., Heymann, J.A., Song, Z., Hinshaw, J.E., and Chan, D.C. (2010). OPA1 disease alleles causing dominant optic atrophy have defects in cardiolipin-stimulated GTP hydrolysis and membrane tubulation. Hum Mol Genet *19*, 2113-2122.

Basrai, M.A., Hieter, P., and Boeke, J.D. (1997). Small open reading frames: beautiful needles in the haystack. Genome Res *7*, 768-771.

Bass, A.J., Watanabe, H., Mermel, C.H., Yu, S., Perner, S., Verhaak, R.G., Kim, S.Y., Wardwell, L., Tamayo, P., Gat-Viks, I*., et al.* (2009). SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. Nat Genet *41*, 1238-1242.

Bhatta, A., Atianand, M., Jiang, Z., Crabtree, J., Blin, J., and Fitzgerald, K.A. (2020). A Mitochondrial Micropeptide Is Required for Activation of the Nlrp3 Inflammasome. J Immunol *204*, 428-437.

Bird, I.M., Mathis, J.M., Mason, J.I., and Rainey, W.E. (1995). Ca(2+)-regulated expression of steroid hydroxylases in H295R human adrenocortical cells. Endocrinology *136*, 5677-5684.

Boumahdi, S., Driessens, G., Lapouge, G., Rorive, S., Nassar, D., Le Mercier, M., Delatte, B., Caauwe, A., Lenglez, S., Nkusi, E.*, et al.* (2014). SOX2 controls tumour initiation and cancer stem-cell functions in squamous-cell carcinoma. Nature *511*, 246-250.

Brambrink, T., Foreman, R., Welstead, G.G., Lengner, C.J., Wernig, M., Suh, H., and Jaenisch, R. (2008). Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. Cell Stem Cell *2*, 151-159.

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

Brunet, T., and King, N. (2017). The Origin of Animal Multicellularity and Cell Differentiation. Dev Cell *43*, 124-140.

Cabili, M.N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., and Rinn, J.L. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev *25*, 1915-1927.

Caino, M.C., Ghosh, J.C., Chae, Y.C., Vaira, V., Rivadeneira, D.B., Faversani, A., Rampini, P., Kossenkov, A.V., Aird, K.M., Zhang, R.*, et al.* (2015). PI3K therapy reprograms mitochondrial trafficking to fuel tumor cell invasion. Proc Natl Acad Sci U S A *112*, 8638-8643.

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

Cereghetti, G.M., Costa, V., and Scorrano, L. (2010). Inhibition of Drp1-dependent mitochondrial fragmentation and apoptosis by a polypeptide antagonist of calcineurin. Cell Death Differ *17*, 1785-1794.

Chai, G.S., Duan, D.X., Ma, R.H., Shen, J.Y., Li, H.L., Ma, Z.W., Luo, Y., Wang, L., Qi, X.H., Wang, Q.*, et al.* (2014). Humanin attenuates Alzheimer-like cognitive deficits and pathological changes induced by amyloid beta-peptide in rats. Neurosci Bull *30*, 923-935.

Chang, C.R., and Blackstone, C. (2010). Dynamic regulation of mitochondrial fission through modification of the dynamin-related protein Drp1. Ann N Y Acad Sci *1201*, 34-39.

Chen, Y., and Dorn, G.W., 2nd (2013). PINK1-phosphorylated mitofusin 2 is a Parkin receptor for culling damaged mitochondria. Science *340*, 471-475.

Chipuk, J.E., Bouchier-Hayes, L., and Green, D.R. (2006). Mitochondrial outer membrane permeabilization during apoptosis: the innocent bystander scenario. Cell Death Differ *13*, 1396-1402.

Chu, Q., Martinez, T.F., Novak, S.W., Donaldson, C.J., Tan, D., Vaughan, J.M., Chang, T., Diedrich, J.K., Andrade, L., Kim, A.*, et al.* (2019). Regulation of the ER stress response by a mitochondrial microprotein. Nat Commun *10*, 4883.

Chu, Q., and Saghatelian, A. (2019). A Hidden ORF Reveals an Immune Protector. Biochemistry *58*, 1022-1023.

Cook, S.J., Stuart, K., Gilley, R., and Sale, M.J. (2017). Control of cell death and mitochondrial fission by ERK1/2 MAP kinase signalling. FEBS J *284*, 4177-4195.

Cribbs, J.T., and Strack, S. (2007). Reversible phosphorylation of Drp1 by cyclic AMP-dependent protein kinase and calcineurin regulates mitochondrial fission and cell death. EMBO Rep *8*, 939-944.

D'Lima, N.G., Ma, J., Winkler, L., Chu, Q., Loh, K.H., Corpuz, E.O., Budnik, B.A., Lykke-Andersen, J., Saghatelian, A., and Slavoff, S.A. (2017). A human microprotein that interacts with the mRNA decapping complex. Nat Chem Biol *13*, 174-180.

Dagklis, T., Ravanos, K., Makedou, K., Kourtis, A., and Rousso, D. (2015). Common features and differences of the hypothalamic-pituitary-gonadal axis in male and female. Gynecol Endocrinol *31*, 14-17.

Dai, W., and Jiang, L. (2019). Dysregulated Mitochondrial Dynamics and Metabolism in Obesity, Diabetes, and Cancer. Front Endocrinol (Lausanne) *10*, 570.

Danziger, J., and Zeidel, M.L. (2015). Osmotic homeostasis. Clin J Am Soc Nephrol *10*, 852-862.

Das, R., Kamal, I.M., Das, S., Chakrabarti, S., and Chakrabarti, O. (2022). MITOL-mediated DRP1 ubiquitylation and degradation promotes mitochondrial hyperfusion in a CMT2A-linked MFN2 mutant. J Cell Sci *135*.

DeBerardinis, R.J., and Chandel, N.S. (2020). We need to talk about the Warburg effect. Nat Metab *2*, 127-129.

Ding, W.X., Guo, F., Ni, H.M., Bockus, A., Manley, S., Stolz, D.B., Eskelinen, E.L., Jaeschke, H., and Yin, X.M. (2012). Parkin and mitofusins reciprocally regulate mitophagy and mitochondrial spheroid formation. J Biol Chem *287*, 42379-42388.

Du, X., and Hu, H. (2021). The Roles of 2-Hydroxyglutarate. Front Cell Dev Biol *9*, 651317.

Ehrhart-Bornstein, M., and Hilbers, U. (1998). Neuroendocrine properties of adrenocortical cells. Horm Metab Res *30*, 436-439.

Faubert, B., Vincent, E.E., Griss, T., Samborska, B., Izreig, S., Svensson, R.U., Mamer, O.A., Avizonis, D., Shackelford, D.B., Shaw, R.J.*, et al.* (2014). Loss of the tumor suppressor LKB1 promotes metabolic reprogramming of cancer cells via HIF-1alpha. Proc Natl Acad Sci U S A *111*, 2554-2559.

Favaro, G., Romanello, V., Varanita, T., Andrea Desbats, M., Morbidoni, V., Tezze, C., Albiero, M., Canato, M., Gherardi, G., De Stefani, D.*, et al.* (2019). DRP1-mediated mitochondrial shape controls calcium homeostasis and muscle mass. Nat Commun *10*, 2576.

Fickett, J.W. (1995). ORFs and genes: how strong a connection? J Comput Biol *2*, 117-123.

Filipponi, D., Emelyanov, A., Muller, J., Molina, C., Nichols, J., and Bulavin, D.V. (2019). DNA Damage Signaling-Induced Cancer Cell Reprogramming as a Driver of Tumor Relapse. Mol Cell *74*, 651-663 e658.

Folmes, C.D., Nelson, T.J., Martinez-Fernandez, A., Arrell, D.K., Lindor, J.Z., Dzeja, P.P., Ikeda, Y., Perez-Terzic, C., and Terzic, A. (2011). Somatic oxidative bioenergetics transitions into pluripotency-dependent glycolysis to facilitate nuclear reprogramming. Cell Metab *14*, 264-271.

Frank, M., Duvezin-Caubet, S., Koob, S., Occhipinti, A., Jagasia, R., Petcherski, A., Ruonala, M.O., Priault, M., Salin, B., and Reichert, A.S. (2012). Mitophagy is triggered by mild oxidative stress in a mitochondrial fission dependent manner. Biochim Biophys Acta *1823*, 2297-2310.

Friedmann-Morvinski, D., and Verma, I.M. (2014). Dedifferentiation and reprogramming: origins of cancer stem cells. EMBO Rep *15*, 244-253.

Fukasawa, Y., Tsuji, J., Fu, S.C., Tomii, K., Horton, P., and Imai, K. (2015). MitoFates: improved prediction of mitochondrial targeting sequences and their cleavage sites. Mol Cell Proteomics *14*, 1113-1126.

Gao, S., and Hu, J. (2021). Mitochondrial Fusion: The Machineries In and Out. Trends Cell Biol *31*, 62-74.

Ge, Q., Jia, D., Cen, D., Qi, Y., Shi, C., Li, J., Sang, L., Yang, L.J., He, J., Lin, A.*, et al.* (2021). Micropeptide ASAP encoded by LINC00467 promotes colorectal cancer progression by directly modulating ATP synthase activity. J Clin Invest *131*.

Gelmetti, V., De Rosa, P., Torosantucci, L., Marini, E.S., Romagnoli, A., Di Rienzo, M., Arena, G., Vignone, D., Fimia, G.M., and Valente, E.M. (2017). PINK1 and BECN1 relocalize at mitochondria-associated membranes during mitophagy and promote ER-mitochondria tethering and autophagosome formation. Autophagy *13*, 654-669.

Gerardo, H., Lima, A., Carvalho, J., Ramos, J.R.D., Couceiro, S., Travasso, R.D.M., Pires das Neves, R., and Graos, M. (2019). Soft culture substrates favor stem-like cellular phenotype and facilitate reprogramming of human mesenchymal stem/stromal cells (hMSCs) through mechanotransduction. Sci Rep *9*, 9086.

Goding, C.R., Pei, D., and Lu, X. (2014). Cancer: pathological nuclear reprogramming? Nat Rev Cancer *14*, 568-573.

Gomes, L.C., Di Benedetto, G., and Scorrano, L. (2011). Essential amino acids and glutamine regulate induction of mitochondrial elongation during autophagy. Cell Cycle *10*, 2635-2639.

Graf, J., and Kretz, M. (2020). From structure to function: Route to understanding lncRNA mechanism. Bioessays *42*, e2000027.

Graves, J.A., Wang, Y., Sims-Lucas, S., Cherok, E., Rothermund, K., Branca, M.F., Elster, J., Beer-Stolz, D., Van Houten, B., Vockley, J.*, et al.* (2012). Mitochondrial

structure, function and dynamics are temporally controlled by c-Myc. PLoS One *7*, e37699.

Gu, H., Li, Q., Huang, S., Lu, W., Cheng, F., Gao, P., Wang, C., Miao, L., Mei, Y., and Wu, M. (2015). Mitochondrial E3 ligase March5 maintains stemness of mouse ES cells via suppression of ERK signalling. Nat Commun *6*, 7112.

Guo, B., Wu, S., Zhu, X., Zhang, L., Deng, J., Li, F., Wang, Y., Zhang, S., Wu, R., Lu, J.*, et al.* (2020). Micropeptide CIP2A-BP encoded by LINC00665 inhibits triple-negative breast cancer progression. EMBO J *39*, e102190.

Guo, J., Ma, J., Zhao, G., Li, G., Fu, Y., Luo, Y., and Gui, R. (2017). Long Noncoding RNA LINC0086 Functions as a Tumor Suppressor in Nasopharyngeal Carcinoma by Targeting miR-214. Oncol Res *25*, 1189-1197.

Guo, Z.W., Meng, Y., Zhai, X.M., Xie, C., Zhao, N., Li, M., Zhou, C.L., Li, K., Liu, T.C., Yang, X.X.*, et al.* (2019). Translated Long Non-Coding Ribonucleic Acid ZFAS1 Promotes Cancer Cell Migration by Elevating Reactive Oxygen Species Production in Hepatocellular Carcinoma. Front Genet *10*, 1111.

Gurdon, J.B., Elsdale, T.R., and Fischberg, M. (1958). Sexually mature individuals of Xenopus laevis from the transplantation of single somatic nuclei. Nature *182*, 64-65.

Hafner, A., Bulyk, M.L., Jambhekar, A., and Lahav, G. (2019). The multiple mechanisms that regulate p53 activity and cell fate. Nat Rev Mol Cell Biol *20*, 199-210.

Hamanaka, R.B., and Chandel, N.S. (2010). Mitochondrial reactive oxygen species regulate cellular signaling and dictate biological outcomes. Trends Biochem Sci *35*, 505-513.

Hambardzumyan, D., Amankulor, N.M., Helmy, K.Y., Becher, O.J., and Holland, E.C. (2009). Modeling Adult Gliomas Using RCAS/t-va Technology. Transl Oncol *2*, 89-95.

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

Hashimoto, Y., Niikura, T., Tajima, H., Yasukawa, T., Sudo, H., Ito, Y., Kita, Y., Kawasumi, M., Kouyama, K., Doyu, M.*, et al.* (2001). A rescue factor abolishing neuronal cell death by a wide spectrum of familial Alzheimer's disease genes and Abeta. Proc Natl Acad Sci U S A *98*, 6336-6341.

He, C., Sun, J., Shi, J., Wang, Y., Zhao, J., Wu, S., Chang, L., Gao, H., Liu, F., Lv, Z.*, et al.* (2018). Digging for Missing Proteins Using Low-Molecular-Weight Protein Enrichment and a "Mirror Protease" Strategy. J Proteome Res *17*, 4178-4185.

Hepburn, A.C., Steele, R.E., Veeratterapillay, R., Wilson, L., Kounatidou, E.E., Barnard, A., Berry, P., Cassidy, J.R., Moad, M., El-Sherif, A.*, et al.* (2019). The induction of core pluripotency master regulators in cancers defines poor clinical outcomes and treatment resistance. Oncogene *38*, 4412-4424.

Hernandez-Alvarez, M.I., Sebastian, D., Vives, S., Ivanova, S., Bartoccioni, P., Kakimoto, P., Plana, N., Veiga, S.R., Hernandez, V., Vasconcelos, N.*, et al.* (2019). Deficient Endoplasmic Reticulum-Mitochondrial Phosphatidylserine Transfer Causes Liver Disease. Cell *177*, 881-895 e817.

Hills, S.A., and Diffley, J.F. (2014). DNA replication and oncogene-induced replicative stress. Curr Biol *24*, R435-444.

Hoang, P.T., Park, P., Cobb, L.J., Paharkova-Vatchkova, V., Hakimi, M., Cohen, P., and Lee, K.W. (2010). The neurosurvival factor Humanin inhibits beta-cell apoptosis via signal transducer and activator of transcription 3 activation and delays and ameliorates diabetes in nonobese diabetic mice. Metabolism *59*, 343-349.

Holland, E.C., Hively, W.P., DePinho, R.A., and Varmus, H.E. (1998). A constitutively active epidermal growth factor receptor cooperates with disruption of G1 cell-cycle arrest pathways to induce glioma-like lesions in mice. Genes Dev *12*, 3675-3685.

Hollinshead, K.E.R., Munford, H., Eales, K.L., Bardella, C., Li, C., Escribano-Gonzalez, C., Thakker, A., Nonnenmacher, Y., Kluckova, K., Jeeves, M.*, et al.* (2018). Oncogenic IDH1 Mutations Promote Enhanced Proline Synthesis through PYCR1 to Support the Maintenance of Mitochondrial Redox Homeostasis. Cell Rep *22*, 3107-3114.

Hong, H., Takahashi, K., Ichisaka, T., Aoi, T., Kanagawa, O., Nakagawa, M., Okita, K., and Yamanaka, S. (2009). Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. Nature *460*, 1132-1135.

Huang, J.Z., Chen, M., Chen, Gao, X.C., Zhu, S., Huang, H., Hu, M., Zhu, H., and Yan, G.R. (2017). A Peptide Encoded by a Putative lncRNA HOXB-AS3 Suppresses Colon Cancer Growth. Mol Cell *68*, 171-184 e176.

Huang, N., Li, F., Zhang, M., Zhou, H., Chen, Z., Ma, X., Yang, L., Wu, X., Zhong, J., Xiao, F.*, et al.* (2021). An Upstream Open Reading Frame in Phosphatase and Tensin Homolog Encodes a Circuit Breaker of Lactate Metabolism. Cell Metab *33*, 128-144 e129.

Hwang, H., Park, G.W., Kim, K.H., Lee, J.Y., Lee, H.K., Ji, E.S., Park, S.K., Xu, T., Yates, J.R., 3rd, Kwon, K.H.*, et al.* (2015). Chromosome-Based Proteomic Study for Identifying Novel Protein Variants from Human Hippocampal Tissue Using Customized neXtProt and GENCODE Databases. J Proteome Res *14*, 5028-5037.

Iglesias, J.M., Gumuzio, J., and Martin, A.G. (2017). Linking Pluripotency Reprogramming and Cancer. Stem Cells Transl Med *6*, 335-339.

Iten, L.E., and Bryant, S.V. (1973). Forelimb regeneration from different levels of amputation in the newt,Notophthalmus viridescens: Length, rate, and stages. Wilhelm Roux Arch Entwickl Mech Org *173*, 263-282.

Ji, Z., Song, R., Regev, A., and Struhl, K. (2015). Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. Elife *4*, e08890.

Johnsson, P., Lipovich, L., Grander, D., and Morris, K.V. (2014). Evolutionary conservation of long non-coding RNAs; sequence, structure, function. Biochim Biophys Acta *1840*, 1063-1071.

Jopling, C., Boue, S., and Izpisua Belmonte, J.C. (2011). Dedifferentiation, transdifferentiation and reprogramming: three routes to regeneration. Nat Rev Mol Cell Biol *12*, 79-89.

Jopling, C., Sleep, E., Raya, M., Marti, M., Raya, A., and Izpisua Belmonte, J.C. (2010). Zebrafish heart regeneration occurs by cardiomyocyte dedifferentiation and proliferation. Nature *464*, 606-609.

Kalkavan, H., and Green, D.R. (2018). MOMP, cell suicide as a BCL-2 family business. Cell Death Differ *25*, 46-55.

Kandoth, C., McLellan, M.D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J.F., Wyczalkowski, M.A*., et al.* (2013). Mutational landscape and significance across 12 major cancer types. Nature *502*, 333-339.

Kang, S.W., Haydar, G., Taniane, C., Farrell, G., Arias, I.M., Lippincott-Schwartz, J., and Fu, D. (2016). AMPK Activation Prevents and Reverses Drug-Induced Mitochondrial and Hepatocyte Injury by Promoting Mitochondrial Fusion and Function. PLoS One *11*, e0165638.

Karbowski, M., Neutzner, A., and Youle, R.J. (2007). The mitochondrial E3 ubiquitin ligase MARCH5 is required for Drp1 dependent mitochondrial division. J Cell Biol *178*, 71-84.

Karbowski, M., and Youle, R.J. (2003). Dynamics of mitochondrial morphology in healthy cells and during apoptosis. Cell Death Differ *10*, 870-880.

Kareta, M.S., Gorges, L.L., Hafeez, S., Benayoun, B.A., Marro, S., Zmoos, A.F., Cecchini, M.J., Spacek, D., Batista, L.F., O'Brien, M*., et al.* (2015). Inhibition of pluripotency networks by the Rb tumor suppressor restricts reprogramming and tumorigenesis. Cell Stem Cell *16*, 39-50.

Kariya, S., Hirano, M., Nagai, Y., Furiya, Y., Fujikake, N., Toda, T., and Ueno, S. (2005). Humanin attenuates apoptosis induced by DRPLA proteins with expanded polyglutamine stretches. J Mol Neurosci *25*, 165-169.

Kashatus, J.A., Nascimento, A., Myers, L.J., Sher, A., Byrne, F.L., Hoehn, K.L., Counter, C.M., and Kashatus, D.F. (2015). Erk2 phosphorylation of Drp1 promotes mitochondrial fission and MAPK-driven tumor growth. Mol Cell *57*, 537-551.

Kastenhuber, E.R., and Lowe, S.W. (2017). Putting p53 in Context. Cell *170*, 1062-1078.

Kawamura, T., Suzuki, J., Wang, Y.V., Menendez, S., Morera, L.B., Raya, A., Wahl, G.M., and Izpisua Belmonte, J.C. (2009). Linking the p53 tumour suppressor pathway to somatic cell reprogramming. Nature *460*, 1140-1144.

Khaminets, A., Behl, C., and Dikic, I. (2016). Ubiquitin-Dependent And Independent Signals In Selective Autophagy. Trends Cell Biol *26*, 6-16.

Kida, Y.S., Kawamura, T., Wei, Z., Sogo, T., Jacinto, S., Shigeno, A., Kushige, H., Yoshihara, E., Liddle, C., Ecker, J.R*., et al.* (2015). ERRs Mediate a Metabolic Switch Required for Somatic Cell Reprogramming to Pluripotency. Cell Stem Cell *16*, 547-555.

Kikuchi, K., Holdway, J.E., Werdich, A.A., Anderson, R.M., Fang, Y., Egnaczyk, G.F., Evans, T., Macrae, C.A., Stainier, D.Y., and Poss, K.D. (2010). Primary contribution to zebrafish heart regeneration by gata4(+) cardiomyocytes. Nature *464*, 601-605.

Kim, S.J., Mehta, H.H., Wan, J., Kuehnemann, C., Chen, J., Hu, J.F., Hoffman, A.R., and Cohen, P. (2018). Mitochondrial peptides modulate mitochondrial function during cellular senescence. Aging (Albany NY) *10*, 1239-1256.

Kim, Y.Y., Um, J.H., Shin, D.J., Jeong, D.J., Hong, Y.B., and Yun, J. (2021). p53-mediated regulation of mitochondrial dynamics plays a pivotal role in the senescence of various normal cells as well as cancer cells. FASEB J *35*, e21319.

Kim, Y.Y., Um, J.H., Yoon, J.H., Lee, D.Y., Lee, Y.J., Kim, D.H., Park, J.I., and Yun, J. (2020). p53 regulates mitochondrial dynamics by inhibiting Drp1 translocation into mitochondria during cellular senescence. FASEB J *34*, 2451-2464.

Kobayashi, S., Zhao, F., Zhang, Z., Kobayashi, T., Huang, Y., Shi, B., Wu, W., and Liang, Q. (2020). Mitochondrial Fission and Mitophagy Coordinately Restrict High Glucose Toxicity in Cardiomyocytes. Front Physiol *11*, 604069.

Koche, R.P., Smith, Z.D., Adli, M., Gu, H., Ku, M., Gnirke, A., Bernstein, B.E., and Meissner, A. (2011). Reprogramming factor expression initiates widespread targeted chromatin remodeling. Cell Stem Cell *8*, 96-105.

Kojima, W., Kujuro, Y., Okatsu, K., Bruno, Q., Koyano, F., Kimura, M., Yamano, K., Tanaka, K., and Matsuda, N. (2016). Unexpected mitochondrial matrix localization of Parkinson's disease-related DJ-1 mutants but not wild-type DJ-1. Genes Cells *21*, 772-788.

Kondapalli, C., Kazlauskaite, A., Zhang, N., Woodroof, H.I., Campbell, D.G., Gourlay, R., Burchell, L., Walden, H., Macartney, T.J., Deak, M.*, et al.* (2012). PINK1 is activated by mitochondrial membrane potential depolarization and stimulates Parkin E3 ligase activity by phosphorylating Serine 65. Open Biol *2*, 120080.

Kopp, J.L., Grompe, M., and Sander, M. (2016). Stem cells versus plasticity in liver and pancreas regeneration. Nat Cell Biol *18*, 238-245.

Koshiba, T., Detmer, S.A., Kaiser, J.T., Chen, H., McCaffery, J.M., and Chan, D.C. (2004). Structural basis of mitochondrial tethering by mitofusin complexes. Science *305*, 858-862.

Koussounadis, A., Langdon, S.P., Um, I.H., Harrison, D.J., and Smith, V.A. (2015). Relationship between differentially expressed mRNA and mRNA-protein correlations in a xenograft model system. Sci Rep *5*, 10775.

Kowaltowski, A.J., Menezes-Filho, S.L., Assali, E.A., Goncalves, I.G., Cabral-Costa, J.V., Abreu, P., Miller, N., Nolasco, P., Laurindo, F.R.M., Bruni-Cardoso, A.*, et al.* (2019). Mitochondrial morphology regulates organellar Ca(2+) uptake and changes cellular Ca(2+) homeostasis. FASEB J *33*, 13176-13188.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol *305*, 567-580.

Lane, N., and Martin, W. (2010). The energetics of genome complexity. Nature *467*, 929-934.

Lang, A., Anand, R., Altinoluk-Hambuchen, S., Ezzahoini, H., Stefanski, A., Iram, A., Bergmann, L., Urbach, J., Bohler, P., Hansel, J.*, et al.* (2017). SIRT4 interacts with OPA1 and regulates mitochondrial quality control and mitophagy. Aging (Albany NY) *9*, 2163-2189.

Leander, B.S. (2020). Predatory protists. Curr Biol *30*, R510-R516.

Lee, C., Zeng, J., Drew, B.G., Sallam, T., Martin-Montalvo, A., Wan, J., Kim, S.J., Mehta, H., Hevener, A.L., de Cabo, R.*, et al.* (2015). The mitochondrial-derived peptide MOTS-c promotes metabolic homeostasis and reduces obesity and insulin resistance. Cell Metab *21*, 443-454.

Lee, K.S., Huh, S., Lee, S., Wu, Z., Kim, A.K., Kang, H.Y., and Lu, B. (2018). Altered ER-mitochondria contact impacts mitochondria calcium homeostasis and contributes to neurodegeneration in vivo in disease models. Proc Natl Acad Sci U S A *115*, E8844-E8853.

Leveille, N., Melo, C.A., Rooijers, K., Diaz-Lagares, A., Melo, S.A., Korkmaz, G., Lopes, R., Akbari Moqadam, F., Maia, A.R., Wijchers, P.J.*, et al.* (2015). Genome-wide profiling of p53-regulated enhancer RNAs uncovers a subset of enhancers controlled by a lncRNA. Nat Commun *6*, 6520.

Lewis, T.L., Jr., Kwon, S.K., Lee, A., Shaw, R., and Polleux, F. (2018). MFF-dependent mitochondrial fission regulates presynaptic release and axon branching by limiting axonal mitochondria size. Nat Commun *9*, 5008.

Li, C.Y., Liang, G.Y., Yao, W.Z., Sui, J., Shen, X., Zhang, Y.Q., Peng, H., Hong, W.W., Ye, Y.C., Zhang, Z.Y.*, et al.* (2016). Integrated analysis of long non-coding RNA competing interactions reveals the potential role in progression of human gastric cancer. Int J Oncol *48*, 1965-1976.

Li, H., Collado, M., Villasante, A., Strati, K., Ortega, S., Canamero, M., Blasco, M.A., and Serrano, M. (2009). The Ink4/Arf locus is a barrier for iPS cell reprogramming. Nature *460*, 1136-1139.

Li, H.S., Zhou, Y.N., Li, L., Li, S.F., Long, D., Chen, X.L., Zhang, J.B., Feng, L., and Li, Y.P. (2019). HIF-1alpha protects against oxidative stress by directly targeting mitochondria. Redox Biol *25*, 101109.

Li, Z., Okamoto, K., Hayashi, Y., and Sheng, M. (2004). The importance of dendritic mitochondria in the morphogenesis and plasticity of spines and synapses. Cell *119*, 873-887.

Liang, J., and Mills, G.B. (2013). AMPK: a contextual oncogene or tumor suppressor? Cancer Res *73*, 2929-2935.

Lin, B., Coleman, J.H., Peterson, J.N., Zunitch, M.J., Jang, W., Herrick, D.B., and Schwob, J.E. (2017). Injury Induces Endogenous Reprogramming and Dedifferentiation of Neuronal Progenitors to Multipotency. Cell Stem Cell *21*, 761-774 e765.

Lin, M.F., Jungreis, I., and Kellis, M. (2011). PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. Bioinformatics *27*, i275-282.

Lin, Y.F., Xiao, M.H., Chen, H.X., Meng, Y., Zhao, N., Yang, L., Tang, H., Wang, J.L., Liu, X., Zhu, Y., *et al.* (2019). A novel mitochondrial micropeptide MPM enhances mitochondrial respiratory activity and promotes myogenic differentiation. Cell Death Dis *10*, 528.

Liu, J., Zhang, C., Hu, W., and Feng, Z. (2019). Tumor suppressor p53 and metabolism. J Mol Cell Biol *11*, 284-292.

Liu, K., Zhao, Q., Liu, P., Cao, J., Gong, J., Wang, C., Wang, W., Li, X., Sun, H., Zhang, C., *et al.* (2016). ATG3-dependent autophagy mediates mitochondrial homeostasis in pluripotency acquirement and maintenance. Autophagy *12*, 2000-2008.

Liu, X., Sun, H., Qi, J., Wang, L., He, S., Liu, J., Feng, C., Chen, C., Li, W., Guo, Y., *et al.* (2013). Sequential introduction of reprogramming factors reveals a time-sensitive requirement for individual factors and a sequential EMT-MET mechanism for optimal reprogramming. Nat Cell Biol *15*, 829-838.

Loson, O.C., Song, Z., Chen, H., and Chan, D.C. (2013). Fis1, Mff, MiD49, and MiD51 mediate Drp1 recruitment in mitochondrial fission. Mol Biol Cell *24*, 659-667.

Ma, K., Chen, G., Li, W., Kepp, O., Zhu, Y., and Chen, Q. (2020). Mitophagy, Mitochondrial Homeostasis, and Cell Fate. Front Cell Dev Biol *8*, 467.

Ma, T., Li, J., Xu, Y., Yu, C., Xu, T., Wang, H., Liu, K., Cao, N., Nie, B.M., Zhu, S.Y., *et al.* (2015). Atg5-independent autophagy regulates mitochondrial clearance and is essential for iPSC reprogramming. Nat Cell Biol *17*, 1379-1387.

Mahmoudi, S., Mancini, E., Xu, L., Moore, A., Jahanbani, F., Hebestreit, K., Srinivasan, R., Li, X., Devarajan, K., Prelot, L., *et al.* (2019). Heterogeneity in old fibroblasts is linked to variability in reprogramming and wound healing. Nature *574*, 553-558.

Makarewich, C.A. (2020). The hidden world of membrane microproteins. Exp Cell Res *388*, 111853.

Makarewich, C.A., Baskin, K.K., Munir, A.Z., Bezprozvannaya, S., Sharma, G., Khemtong, C., Shah, A.M., McAnally, J.R., Malloy, C.R., Szweda, L.I., *et al.* (2018). MOXI Is a Mitochondrial Micropeptide That Enhances Fatty Acid beta-Oxidation. Cell Rep *23*, 3701-3709.

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

Mardis, E.R., Ding, L., Dooling, D.J., Larson, D.E., McLellan, M.D., Chen, K., Koboldt, D.C., Fulton, R.S., Delehaunty, K.D., McGrath, S.D., *et al.* (2009). Recurring mutations found by sequencing an acute myeloid leukemia genome. N Engl J Med *361*, 1058-1066.

Marion, R.M., Strati, K., Li, H., Murga, M., Blanco, R., Ortega, S., Fernandez-Capetillo, O., Serrano, M., and Blasco, M.A. (2009). A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. Nature *460*, 1149-1153.

Martinez-Reyes, I., Diebold, L.P., Kong, H., Schieber, M., Huang, H., Hensley, C.T., Mehta, M.M., Wang, T., Santos, J.H., Woychik, R., *et al.* (2016). TCA Cycle and

Mitochondrial Membrane Potential Are Necessary for Diverse Biological Functions. Mol Cell *61*, 199-209.

Matsuda, N., Sato, S., Shiba, K., Okatsu, K., Saisho, K., Gautier, C.A., Sou, Y.S., Saiki, S., Kawajiri, S., Sato, F.*, et al.* (2010). PINK1 stabilized by mitochondrial depolarization recruits Parkin to damaged mitochondria and activates latent Parkin for mitophagy. J Cell Biol *189*, 211-221.

Matsumoto, A., Pasut, A., Matsumoto, M., Yamashita, R., Fung, J., Monteleone, E., Saghatelian, A., Nakayama, K.I., Clohessy, J.G., and Pandolfi, P.P. (2017). mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide. Nature *541*, 228-232.

Mendelsohn, A.R., and Larrick, J.W. (2018). Mitochondrial-Derived Peptides Exacerbate Senescence. Rejuvenation Res *21*, 369-373.

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

Merrell, A.J., and Stanger, B.Z. (2016). Adult cell plasticity in vivo: de-differentiation and transdifferentiation are back in style. Nat Rev Mol Cell Biol *17*, 413-425.

Merry, T.L., Chan, A., Woodhead, J.S.T., Reynolds, J.C., Kumagai, H., Kim, S.J., and Lee, C. (2020). Mitochondrial-derived peptides in energy metabolism. Am J Physiol Endocrinol Metab *319*, E659-E666.

Mishra, P., Varuzhanyan, G., Pham, A.H., and Chan, D.C. (2015). Mitochondrial Dynamics is a Distinguishing Feature of Skeletal Muscle Fiber Types and Regulates Organellar Compartmentalization. Cell Metab *22*, 1033-1044.

Miyazono, Y., Hirashima, S., Ishihara, N., Kusukawa, J., Nakamura, K.I., and Ohta, K. (2018). Uncoupled mitochondria quickly shorten along their long axis to form indented spheroids, instead of rings, in a fission-independent manner. Sci Rep *8*, 350.

Morita, M., Prudent, J., Basu, K., Goyon, V., Katsumura, S., Hulea, L., Pearl, D., Siddiqui, N., Strack, S., McGuirk, S.*, et al.* (2017). mTOR Controls Mitochondrial Dynamics and Cell Survival via MTFP1. Mol Cell *67*, 922-935 e925.

Mortuza, S.M., Zheng, W., Zhang, C., Li, Y., Pearce, R., and Zhang, Y. (2021). Improving fragment-based ab initio protein structure assembly using low-accuracy contact-map predictions. Nat Commun *12*, 5011.

Mosteiro, L., Pantoja, C., Alcazar, N., Marion, R.M., Chondronasiou, D., Rovira, M., Fernandez-Marcos, P.J., Munoz-Martin, M., Blanco-Aparicio, C., Pastor, J.*, et al.* (2016). Tissue damage and senescence provide critical signals for cellular reprogramming in vivo. Science *354*.

Mosteiro, L., Pantoja, C., de Martino, A., and Serrano, M. (2018). Senescence promotes in vivo reprogramming through p16(INK)(4a) and IL-6. Aging Cell *17*.

Muzumdar, R.H., Huffman, D.M., Atzmon, G., Buettner, C., Cobb, L.J., Fishman, S., Budagov, T., Cui, L., Einstein, F.H., Poduval, A.*, et al.* (2009). Humanin: a novel central regulator of peripheral insulin action. PLoS One *4*, e6334.

Nakajima, T., Morita, K., Tsunoda, H., Imajoh-Ohmi, S., Tanaka, H., Yasuda, H., and Oda, K. (1998). Stabilization of p53 by adenovirus E1A occurs through its amino-terminal region by modification of the ubiquitin-proteasome pathway. J Biol Chem *273*, 20036-20045.

Nakamura, N., Kimura, Y., Tokuda, M., Honda, S., and Hirose, S. (2006). MARCH-V is a novel mitofusin 2- and Drp1-binding protein able to change mitochondrial morphology. EMBO Rep *7*, 1019-1022.

Nelson, B.R., Makarewich, C.A., Anderson, D.M., Winders, B.R., Troupes, C.D., Wu, F., Reese, A.L., McAnally, J.R., Chen, X., Kavalali, E.T.*, et al.* (2016). A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. Science *351*, 271-275.

O'Shaughnessy, P.J., Fleming, L.M., Jackson, G., Hochgeschwender, U., Reed, P., and Baker, P.J. (2003). Adrenocorticotropic hormone directly stimulates testosterone production by the fetal and neonatal mouse testis. Endocrinology *144*, 3279-3284.

Ohnishi, K., Semi, K., Yamamoto, T., Shimizu, M., Tanaka, A., Mitsunaga, K., Okita, K., Osafune, K., Arioka, Y., Maeda, T.*, et al.* (2014). Premature termination of reprogramming in vivo leads to cancer development through altered epigenetic regulation. Cell *156*, 663-677.

Omasits, U., Ahrens, C.H., Muller, S., and Wollscheid, B. (2014). Protter: interactive protein feature visualization and integration with experimental proteomic data. Bioinformatics *30*, 884-886.

Orrenius, S., Zhivotovsky, B., and Nicotera, P. (2003). Regulation of cell death: the calcium-apoptosis link. Nat Rev Mol Cell Biol *4*, 552-565.

Ozawa, T., Riester, M., Cheng, Y.K., Huse, J.T., Squatrito, M., Helmy, K., Charles, N., Michor, F., and Holland, E.C. (2014). Most human non-GCIMP glioblastoma subtypes evolve from a common proneural-like precursor glioma. Cancer Cell *26*, 288-300.

Painter, M.W., Brosius Lutz, A., Cheng, Y.C., Latremoliere, A., Duong, K., Miller, C.M., Posada, S., Cobos, E.J., Zhang, A.X., Wagers, A.J.*, et al.* (2014). Diminished Schwann cell repair responses underlie age-associated impaired axonal regeneration. Neuron *83*, 331-343.

Palikaras, K., Lionaki, E., and Tavernarakis, N. (2018). Mechanisms of mitophagy in cellular homeostasis, physiology and pathology. Nat Cell Biol *20*, 1013-1022.

Panopoulos, A.D., Yanes, O., Ruiz, S., Kida, Y.S., Diep, D., Tautenhahn, R., Herrerias, A., Batchelder, E.M., Plongthongkum, N., Lutz, M.*, et al.* (2012). The metabolome of induced pluripotent stem cells reveals metabolic changes occurring in somatic cell reprogramming. Cell Res *22*, 168-177.

Papp, B., and Plath, K. (2013). Epigenetics of reprogramming to induced pluripotency. Cell *152*, 1324-1343.

Parsons, D.W., Jones, S., Zhang, X., Lin, J.C., Leary, R.J., Angenendt, P., Mankoo, P., Carter, H., Siu, I.M., Gallia, G.L.*, et al.* (2008). An integrated genomic analysis of human glioblastoma multiforme. Science *321*, 1807-1812.

Patraquim, P., Mumtaz, M.A.S., Pueyo, J.I., Aspden, J.L., and Couso, J.P. (2020). Developmental regulation of canonical and small ORF translation from mRNAs. Genome Biol *21*, 128.

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

Pennisi, E. (2012). Genomics. ENCODE project writes eulogy for junk DNA. Science *337*, 1159, 1161.

Pesaresi, M., Sebastian-Perez, R., and Cosma, M.P. (2019). Dedifferentiation, transdifferentiation and cell fusion: in vivo reprogramming strategies for regenerative medicine. FEBS J *286*, 1074-1093.

Pickles, S., Vigie, P., and Youle, R.J. (2018). Mitophagy and Quality Control Mechanisms in Mitochondrial Maintenance. Curr Biol *28*, R170-R185.

Pinton, P., Giorgi, C., Siviero, R., Zecchini, E., and Rizzuto, R. (2008). Calcium and apoptosis: ER-mitochondria Ca2+ transfer in the control of apoptosis. Oncogene *27*, 6407-6418.

Polo, J.M., Anderssen, E., Walsh, R.M., Schwarz, B.A., Nefzger, C.M., Lim, S.M., Borkent, M., Apostolou, E., Alaei, S., Cloutier, J.*, et al.* (2012). A molecular roadmap of reprogramming somatic cells into iPS cells. Cell *151*, 1617-1632.

Polycarpou-Schwarz, M., Gross, M., Mestdagh, P., Schott, J., Grund, S.E., Hildenbrand, C., Rom, J., Aulmann, S., Sinn, H.P., Vandesompele, J.*, et al.* (2018). The cancer-associated microprotein CASIMO1 controls cell proliferation and interacts with squalene epoxidase modulating lipid droplet formation. Oncogene *37*, 4750-4768.

Praefcke, G.J., and McMahon, H.T. (2004). The dynamin superfamily: universal membrane tubulation and fission molecules? Nat Rev Mol Cell Biol *5*, 133-147.

Prieto, J., Leon, M., Ponsoda, X., Sendra, R., Bort, R., Ferrer-Lorente, R., Raya, A., Lopez-Garcia, C., and Torres, J. (2016). Early ERK1/2 activation promotes DRP1-dependent mitochondrial fission necessary for cell reprogramming. Nat Commun *7*, 11124.

Prieto, J., and Torres, J. (2017). Mitochondrial Dynamics: In Cell Reprogramming as It Is in Cancer. Stem Cells Int *2017*, 8073721.

Prigione, A., and Adjaye, J. (2010). Modulation of mitochondrial biogenesis and bioenergetic metabolism upon in vitro and in vivo differentiation of human ES and iPS cells. Int J Dev Biol *54*, 1729-1741.

Quinsay, M.N., Thomas, R.L., Lee, Y., and Gustafsson, A.B. (2010). Bnip3-mediated mitochondrial autophagy is independent of the mitochondrial permeability transition pore. Autophagy *6*, 855-862.

Raff, M. (2003). Adult stem cell plasticity: fact or artifact? Annu Rev Cell Dev Biol *19*, 1-22.

Raghavan, S., Winter, P.S., Navia, A.W., Williams, H.L., DenAdel, A., Lowder, K.E., Galvez-Reyes, J., Kalekar, R.L., Mulugeta, N., Kapner, K.S.*, et al.* (2021).

Microenvironment drives cell state, plasticity, and drug response in pancreatic cancer. Cell *184*, 6119-6137 e6126.

Rambold, A.S., Kostelecky, B., Elia, N., and Lippincott-Schwartz, J. (2011). Tubular network formation protects mitochondria from autophagosomal degradation during nutrient starvation. Proc Natl Acad Sci U S A *108*, 10190-10195.

Rasmussen, K.D., and Helin, K. (2016). Role of TET enzymes in DNA methylation, development, and cancer. Genes Dev *30*, 733-750.

Rathore, A., Chu, Q., Tan, D., Martinez, T.F., Donaldson, C.J., Diedrich, J.K., Yates, J.R., 3rd, and Saghatelian, A. (2018). MIEF1 Microprotein Regulates Mitochondrial Translation. Biochemistry *57*, 5564-5575.

Rehman, J., Zhang, H.J., Toth, P.T., Zhang, Y., Marsboom, G., Hong, Z., Salgia, R., Husain, A.N., Wietholt, C., and Archer, S.L. (2012). Inhibition of mitochondrial fission prevents cell cycle progression in lung cancer. FASEB J *26*, 2175-2186.

Reynolds, J.C., Lai, R.W., Woodhead, J.S.T., Joly, J.H., Mitchell, C.J., Cameron-Smith, D., Lu, R., Cohen, P., Graham, N.A., Benayoun, B.A.*, et al.* (2021). MOTS-c is an exercise-induced mitochondrial-encoded regulator of age-dependent physical decline and muscle homeostasis. Nat Commun *12*, 470.

Rodini, C.O., Suzuki, D.E., Saba-Silva, N., Cappellano, A., de Souza, J.E., Cavalheiro, S., Toledo, S.R., and Okamoto, O.K. (2012). Expression analysis of stem cell-related genes reveal OCT4 as a predictor of poor clinical outcome in medulloblastoma. J Neurooncol *106*, 71-79.

Rossier, M.F., Burnay, M.M., Brandenburger, Y., Cherradi, N., Vallotton, M.B., and Capponi, A.M. (1996). Sources and sites of action of calcium in the regulation of aldosterone biosynthesis. Endocr Res *22*, 579-588.

Rossignol, R., Gilkerson, R., Aggeler, R., Yamagata, K., Remington, S.J., and Capaldi, R.A. (2004). Energy substrate modulates mitochondrial structure and oxidative capacity in cancer cells. Cancer Res *64*, 985-993.

Ruiz-Orera, J., Verdaguer-Grau, P., Villanueva-Canas, J.L., Messeguer, X., and Alba, M.M. (2018). Translation of neutrally evolving peptides provides a basis for de novo gene evolution. Nat Ecol Evol *2*, 890-896.

Rujiviphat, J., Wong, M.K., Won, A., Shih, Y.L., Yip, C.M., and McQuibban, G.A. (2015). Mitochondrial Genome Maintenance 1 (Mgm1) Protein Alters Membrane Topology and Promotes Local Membrane Bending. J Mol Biol *427*, 2599-2609.

Sabharwal, S.S., and Schumacker, P.T. (2014). Mitochondrial ROS in cancer: initiators, amplifiers or an Achilles' heel? Nat Rev Cancer *14*, 709-721.

Saiki, Y., Ishimaru, S., Mimori, K., Takatsuno, Y., Nagahara, M., Ishii, H., Yamada, K., and Mori, M. (2009). Comprehensive analysis of the clinical significance of inducing pluripotent stemness-related gene expression in colorectal cancer cells. Ann Surg Oncol *16*, 2638-2644.

Scaini, G., Mason, B.L., Diaz, A.P., Jha, M.K., Soares, J.C., Trivedi, M.H., and Quevedo, J. (2021). Dysregulation of mitochondrial dynamics, mitophagy and apoptosis in major depressive disorder: Does inflammation play a role? Mol Psychiatry.

Schieber, M., and Chandel, N.S. (2014). ROS function in redox signaling and oxidative stress. Curr Biol *24*, R453-462.

Schneeberger, M., Dietrich, M.O., Sebastian, D., Imbernon, M., Castano, C., Garcia, A., Esteban, Y., Gonzalez-Franquesa, A., Rodriguez, I.C., Bortolozzi, A.*, et al.* (2013). Mitofusin 2 in POMC neurons connects ER stress with leptin resistance and energy imbalance. Cell *155*, 172-187.

Sebastian, D., Palacin, M., and Zorzano, A. (2017). Mitochondrial Dynamics: Coupling Mitochondrial Fitness with Healthy Aging. Trends Mol Med *23*, 201-215.

Sebastian, D., Sorianello, E., Segales, J., Irazoki, A., Ruiz-Bonilla, V., Sala, D., Planet, E., Berenguer-Llergo, A., Munoz, J.P., Sanchez-Feutrie, M.*, et al.* (2016). Mfn2 deficiency links age-related sarcopenia and impaired autophagy to activation of an adaptive mitophagy pathway. EMBO J *35*, 1677-1693.

Sekine, S., and Youle, R.J. (2018). PINK1 import regulation; a fine system to convey mitochondrial stress to the cytosol. BMC Biol *16*, 2.

Senft, D., and Ronai, Z.A. (2016). Regulators of mitochondrial dynamics in cancer. Curr Opin Cell Biol *39*, 43-52.

Serasinghe, M.N., Wieder, S.Y., Renault, T.T., Elkholi, R., Asciolla, J.J., Yao, J.L., Jabado, O., Hoehn, K., Kageyama, Y., Sesaki, H.*, et al.* (2015). Mitochondrial division is requisite to RAS-induced transformation and targeted by oncogenic MAPK pathway inhibitors. Mol Cell *57*, 521-536.

Shim, E.H., Livi, C.B., Rakheja, D., Tan, J., Benson, D., Parekh, V., Kho, E.Y., Ghosh, A.P., Kirkman, R., Velu, S.*, et al.* (2014). L-2-Hydroxyglutarate: an epigenetic modifier and putative oncometabolite in renal cancer. Cancer Discov *4*, 1290-1298.

Sinenko, S.A., Starkova, T.Y., Kuzmin, A.A., and Tomilin, A.N. (2021). Physiological Signaling Functions of Reactive Oxygen Species in Stem Cells: From Flies to Man. Front Cell Dev Biol *9*, 714370.

Singh, A.K., Zhao, B., Liu, X., Wang, X., Li, H., Qin, H., Wu, X., Ma, Y., Horne, D., and Yu, X. (2020). Selective targeting of TET catalytic domain promotes somatic cell reprogramming. Proc Natl Acad Sci U S A *117*, 3621-3626.

Singh, R., Letai, A., and Sarosiek, K. (2019). Regulation of apoptosis in health and disease: the balancing act of BCL-2 family proteins. Nat Rev Mol Cell Biol *20*, 175-193.

Slavoff, S.A., Heo, J., Budnik, B.A., Hanakahi, L.A., and Saghatelian, A. (2014). A human short open reading frame (sORF)-encoded polypeptide that stimulates DNA end joining. J Biol Chem *289*, 10950-10957.

Slavoff, S.A., Mitchell, A.J., Schwaid, A.G., Cabili, M.N., Ma, J., Levin, J.Z., Karger, A.D., Budnik, B.A., Rinn, J.L., and Saghatelian, A. (2013). Peptidomic discovery of short open reading frame-encoded peptides in human cells. Nat Chem Biol *9*, 59-64.

Smith, S.M., and Vale, W.W. (2006). The role of the hypothalamic-pituitary-adrenal axis in neuroendocrine responses to stress. Dialogues Clin Neurosci *8*, 383-395.

Son, M.J., Kwon, Y., Son, M.Y., Seol, B., Choi, H.S., Ryu, S.W., Choi, C., and Cho, Y.S. (2015). Mitofusins deficiency elicits mitochondrial metabolic reprogramming to pluripotency. Cell Death Differ *22*, 1957-1969.

Soufi, A., Donahue, G., and Zaret, K.S. (2012). Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. Cell *151*, 994-1004.

Sponne, I., Fifre, A., Koziel, V., Kriem, B., Oster, T., and Pillot, T. (2004). Humanin rescues cortical neurons from prion-peptide-induced apoptosis. Mol Cell Neurosci *25*, 95-102.

Stein, C.S., Jadiya, P., Zhang, X., McLendon, J.M., Abouassaly, G.M., Witmer, N.H., Anderson, E.J., Elrod, J.W., and Boudreau, R.L. (2018). Mitoregulin: A lncRNA-Encoded Microprotein that Supports Mitochondrial Supercomplexes and Respiratory Efficiency. Cell Rep *23*, 3710-3720 e3718.

Sugiura, A., Nagashima, S., Tokuyama, T., Amo, T., Matsuki, Y., Ishido, S., Kudo, Y., McBride, H.M., Fukuda, T., Matsushita, N.*, et al.* (2013). MITOL regulates endoplasmic reticulum-mitochondria contacts via Mitofusin2. Mol Cell *51*, 20-34.

Sun, J., Shi, J., Wang, Y., Chen, Y., Li, Y., Kong, D., Chang, L., Liu, F., Lv, Z., Zhou, Y.*, et al.* (2018). Multiproteases Combined with High-pH Reverse-Phase Separation Strategy Verified Fourteen Missing Proteins in Human Testis Tissue. J Proteome Res *17*, 4171-4177.

Suva, M.L., Riggi, N., and Bernstein, B.E. (2013). Epigenetic reprogramming in cancer. Science *339*, 1567-1570.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Tan, A.R., Cai, A.Y., Deheshi, S., and Rintoul, G.L. (2011). Elevated intracellular calcium causes distinct mitochondrial remodelling and calcineurin-dependent fission in astrocytes. Cell Calcium *49*, 108-114.

Tata, P.R., Mou, H., Pardo-Saganta, A., Zhao, R., Prabhu, M., Law, B.M., Vinarsky, V., Cho, J.L., Breton, S., Sahay, A.*, et al.* (2013). Dedifferentiation of committed epithelial cells into stem cells in vivo. Nature *503*, 218-223.

Teslaa, T., and Teitell, M.A. (2015). Pluripotent stem cell energy metabolism: an update. EMBO J *34*, 138-153.

Tetteh, P.W., Basak, O., Farin, H.F., Wiebrands, K., Kretzschmar, K., Begthel, H., van den Born, M., Korving, J., de Sauvage, F., van Es, J.H.*, et al.* (2016). Replacement of Lost Lgr5-Positive Stem Cells through Plasticity of Their Enterocyte-Lineage Daughters. Cell Stem Cell *18*, 203-213.

Toyama, E.Q., Herzig, S., Courchet, J., Lewis, T.L., Jr., Loson, O.C., Hellberg, K., Young, N.P., Chen, H., Polleux, F., Chan, D.C.*, et al.* (2016). Metabolism. AMP-activated protein

kinase mediates mitochondrial fission in response to energy stress. Science *351*, 275-281.

Utikal, J., Polo, J.M., Stadtfeld, M., Maherali, N., Kulalert, W., Walsh, R.M., Khalil, A., Rheinwald, J.G., and Hochedlinger, K. (2009). Immortalization eliminates a roadblock during cellular reprogramming into iPS cells. Nature *460*, 1145-1148.

Vafa, O., Wade, M., Kern, S., Beeche, M., Pandita, T.K., Hampton, G.M., and Wahl, G.M. (2002). c-Myc can induce DNA damage, increase reactive oxygen species, and mitigate p53 function: a mechanism for oncogene-induced genetic instability. Mol Cell *9*, 1031-1044.

van Es, J.H., Sato, T., van de Wetering, M., Lyubimova, A., Yee Nee, A.N., Gregorieff, A., Sasaki, N., Zeinstra, L., van den Born, M., Korving, J.*, et al.* (2012). Dll1+ secretory progenitor cells revert to stem cells upon crypt damage. Nat Cell Biol *14*, 1099-1104.

Vander Heiden, M.G., Cantley, L.C., and Thompson, C.B. (2009). Understanding the Warburg effect: the metabolic requirements of cell proliferation. Science *324*, 1029-1033.

Vassilev, L.T., Vu, B.T., Graves, B., Carvajal, D., Podlaski, F., Filipovic, Z., Kong, N., Kammlott, U., Lukacs, C., Klein, C.*, et al.* (2004). In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. Science *303*, 844-848.

Vazquez-Martin, A., Cufi, S., Corominas-Faja, B., Oliveras-Ferraros, C., Vellon, L., and Menendez, J.A. (2012a). Mitochondrial fusion by pharmacological manipulation impedes somatic cell reprogramming to pluripotency: new insight into the role of mitophagy in cell stemness. Aging (Albany NY) *4*, 393-401.

Vazquez-Martin, A., Vellon, L., Quiros, P.M., Cufi, S., Ruiz de Galarreta, E., Oliveras-Ferraros, C., Martin, A.G., Martin-Castillo, B., Lopez-Otin, C., and Menendez, J.A. (2012b). Activation of AMP-activated protein kinase (AMPK) provides a metabolic barrier to reprogramming somatic cells into stem cells. Cell Cycle *11*, 974-989.

Vidal, S.E., Amlani, B., Chen, T., Tsirigos, A., and Stadtfeld, M. (2014). Combinatorial modulation of signaling pathways reveals cell-type-specific requirements for highly efficient and synchronous iPSC reprogramming. Stem Cell Reports *3*, 574-584.

Vorobjev, I.A., and Zorov, D.B. (1983). Diazepam inhibits cell respiration and induces fragmentation of mitochondrial reticulum. FEBS Lett *163*, 311-314.

Vyas, S., Zaganjor, E., and Haigis, M.C. (2016). Mitochondria and Cancer. Cell *166*, 555-566.

Wan, Y.Y., Zhang, J.F., Yang, Z.J., Jiang, L.P., Wei, Y.F., Lai, Q.N., Wang, J.B., Xin, H.B., and Han, X.J. (2014). Involvement of Drp1 in hypoxia-induced migration of human glioblastoma U251 cells. Oncol Rep *32*, 619-626.

Wang, D.B., Garden, G.A., Kinoshita, C., Wyles, C., Babazadeh, N., Sopher, B., Kinoshita, Y., and Morrison, R.S. (2013a). Declines in Drp1 and parkin expression underlie DNA damage-induced changes in mitochondrial length and neuronal death. J Neurosci *33*, 1357-1365.

Wang, L., Xu, X., Jiang, C., Ma, G., Huang, Y., Zhang, H., Lai, Y., Wang, M., Ahmed, T., Lin, R.*, et al.* (2020a). mTORC1-PGC1 axis regulates mitochondrial remodeling during reprogramming. FEBS J *287*, 108-121.

Wang, Q., Zhang, M., Torres, G., Wu, S., Ouyang, C., Xie, Z., and Zou, M.H. (2017). Metformin Suppresses Diabetes-Accelerated Atherosclerosis via the Inhibition of Drp1-Mediated Mitochondrial Fission. Diabetes *66*, 193-205.

Wang, S., Xia, P., Ye, B., Huang, G., Liu, J., and Fan, Z. (2013b). Transient activation of autophagy via Sox2-mediated suppression of mTOR is an important early step in reprogramming to pluripotency. Cell Stem Cell *13*, 617-625.

Wang, Z.Y., Leushkin, E., Liechti, A., Ovchinnikova, S., Mossinger, K., Bruning, T., Rummel, C., Grutzner, F., Cardoso-Moreira, M., Janich, P.*, et al.* (2020b). Transcriptome and translatome co-evolution in mammals. Nature *588*, 642-647.

Webster, K.A. (2012). Mitochondrial membrane permeabilization and cell death during myocardial infarction: roles of calcium and reactive oxygen species. Future Cardiol *8*, 863-884.

Wei, W., Luo, W., Wu, F., Peng, X., Zhang, Y., Zhang, M., Zhao, Y., Su, N., Qi, Y., Chen, L.*, et al.* (2016). Deep Coverage Proteomics Identifies More Low-Abundance Missing Proteins in Human Testis Tissue with Q-Exactive HF Mass Spectrometer. J Proteome Res *15*, 3988-3997.

Weldemariam, M.M., Han, C.L., Shekari, F., Kitata, R.B., Chuang, C.Y., Hsu, W.T., Kuo, H.C., Choong, W.K., Sung, T.Y., He, F.C.*, et al.* (2018). Subcellular Proteome Landscape of Human Embryonic Stem Cells Revealed Missing Membrane Proteins. J Proteome Res *17*, 4138-4151.

Wu, S., Zhang, L., Deng, J., Guo, B., Li, F., Wang, Y., Wu, R., Zhang, S., Lu, J., and Zhou, Y. (2020). A Novel Micropeptide Encoded by Y-Linked LINC00278 Links Cigarette Smoking and AR Signaling in Male Esophageal Squamous Cell Carcinoma. Cancer Res.

Wu, W., Lin, C., Wu, K., Jiang, L., Wang, X., Li, W., Zhuang, H., Zhang, X., Chen, H., Li, S.*, et al.* (2016). FUNDC1 regulates mitochondrial dynamics at the ER-mitochondrial contact site under hypoxic conditions. EMBO J *35*, 1368-1384.

Wu, Y., Li, Y., Zhang, H., Huang, Y., Zhao, P., Tang, Y., Qiu, X., Ying, Y., Li, W., Ni, S.*, et al.* (2015). Autophagy and mTORC1 regulate the stochastic phase of somatic cell reprogramming. Nat Cell Biol *17*, 715-725.

Xiang, G., Yang, L., Long, Q., Chen, K., Tang, H., Wu, Y., Liu, Z., Zhou, Y., Qi, J., Zheng, L.*, et al.* (2017). BNIP3L-dependent mitophagy accounts for mitochondrial clearance during 3 factors-induced somatic cell reprogramming. Autophagy *13*, 1543-1555.

Xiong, S., Feng, Y., and Cheng, L. (2019). Cellular Reprogramming as a Therapeutic Target in Cancer. Trends Cell Biol *29*, 623-634.

Xu, X., Chua, C.C., Gao, J., Hamdy, R.C., and Chua, B.H. (2006). Humanin is a novel neuroprotective agent against stroke. Stroke *37*, 2613-2619.

Yi, M., Weaver, D., and Hajnoczky, G. (2004). Control of mitochondrial motility and distribution by the calcium signal: a homeostatic circuit. J Cell Biol *167*, 661-672.

Yonashiro, R., Ishido, S., Kyo, S., Fukuda, T., Goto, E., Matsuki, Y., Ohmura-Hoshino, M., Sada, K., Hotta, H., Yamamura, H.*, et al.* (2006). A novel mitochondrial ubiquitin ligase plays a critical role in mitochondrial dynamics. EMBO J *25*, 3618-3626.

Yuan, Y., Zhang, A., Qi, J., Wang, H., Liu, X., Zhao, M., Duan, S., Huang, Z., Zhang, C., Wu, L.*, et al.* (2018). p53/Drp1-dependent mitochondrial fission mediates aldosterone-induced podocyte injury and mitochondrial dysfunction. Am J Physiol Renal Physiol *314*, F798-F808.

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

Zarse, K., and Ristow, M. (2015). A mitochondrially encoded hormone ameliorates obesity and insulin resistance. Cell Metab *21*, 355-356.

Zeman, M.K., and Cimprich, K.A. (2014). Causes and consequences of replication stress. Nat Cell Biol *16*, 2-9.

Zhang, M., Zhao, K., Xu, X., Yang, Y., Yan, S., Wei, P., Liu, H., Xu, J., Xiao, F., Zhou, H.*, et al.* (2018). A peptide encoded by circular form of LINC-PINT suppresses oncogenic transcriptional elongation in glioblastoma. Nat Commun *9*, 4475.

Zhang, S., Reljic, B., Liang, C., Kerouanton, B., Francisco, J.C., Peh, J.H., Mary, C., Jagannathan, N.S., Olexiouk, V., Tang, C.*, et al.* (2020). Mitochondrial peptide BRAWNIN is essential for vertebrate respiratory complex III assembly. Nat Commun *11*, 1312.

Zheng, X., Chen, L., Zhou, Y., Wang, Q., Zheng, Z., Xu, B., Wu, C., Zhou, Q., Hu, W., Wu, C.*, et al.* (2019). A novel protein encoded by a circular RNA circPPP1R12A promotes tumor pathogenesis and metastasis of colon cancer via Hippo-YAP signaling. Mol Cancer *18*, 47.

Zhou, G., Meng, S., Li, Y., Ghebre, Y.T., and Cooke, J.P. (2016). Optimal ROS Signaling Is Critical for Nuclear Reprogramming. Cell Rep *15*, 919-925.

Zhou, Y., Al-Saaidi, R.A., Fernandez-Guerra, P., Freude, K.K., Olsen, R.K., Jensen, U.B., Gregersen, N., Hyttel, P., Bolund, L., Aagaard, L.*, et al.* (2017). Mitochondrial Spare Respiratory Capacity Is Negatively Correlated with Nuclear Reprogramming Efficiency. Stem Cells Dev *26*, 166-176.

Zhu, Y., Orre, L.M., Johansson, H.J., Huss, M., Boekel, J., Vesterlund, M., Fernandez-Woodbridge, A., Branca, R.M.M., and Lehtio, J. (2018). Discovery of coding regions in the human genome by integrated proteogenomics analysis workflow. Nat Commun *9*, 903.

Zorov, D.B., Vorobjev, I.A., Popkov, V.A., Babenko, V.A., Zorova, L.D., Pevzner, I.B., Silachev, D.N., Zorov, S.D., Andrianova, N.V., and Plotnikov, E.Y. (2019). Lessons from the Discovery of Mitochondrial Fragmentation (Fission): A Review and Update. Cells *8*.

Zuker, M. (1994). Prediction of RNA secondary structure by energy minimization. Methods Mol Biol *25*, 267-294.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res *31*, 3406-3415.

Annex

Additionally to the work presented in this thsis, the following publication included as an annex has been published

Experimental Cell Research 392 (2020) 111997

Contents lists available at ScienceDirect

Experimental Cell Research

journal homepage: www.elsevier.com/locate/yexcr

The microproteome of cancer: From invisibility to relevance

Iñaki Merino-Valverde[1], Emanuela Greco[1], María Abad[**]

Vall d'Hebron Institute of Oncology (VHIO), Barcelona, 08035, Spain

ARTICLE INFO

Keywords:
Microprotein
Micropeptide
SEPs
sORF
smORF
alt-ORF
Cancer
Tumor suppressors

ABSTRACT

Recent findings have revealed that many genomic regions previously annotated as non-protein coding actually contain small open reading frames, smaller that 300 bp, that are transcribed and translated into evolutionary conserved microproteins. To date, only a small subset of them have been functionally characterized, but they play key functions in fundamental processes such as DNA repair, RNA processing and metabolism regulation. This emergent field seems to hide a new category of molecular regulators with clinical potential. In this review, we focus on its relevance for cancer. Following Hanahan and Weinberg's classification of the hallmarks of cancer, we provide an overview of those microproteins known to be implicated in cancer or those that, based on their function, are likely to play a role in cancer. The resulting picture is that while we are at the very early times of this field, it holds the promise to provide crucial information to understand cancer biology.

1. Introduction

Recent advances coming from computational analyses, peptidomics and ribosome profiling have revealed that our proteome includes a new class of small proteins produced by the translation of small open reading frames shorter than 300 bp in length, generating proteins that have been called microproteins (also known as micropeptides or SEPs, from small ORF-encoded peptides) [1,2]. The main reason why microproteins have been overlooked till recently is that most ORF-prediction algorithms -including the one used by FANTOM annotation consortium- placed an arbitrary cut off of 300 bp, missing the proteins below 100 amino acids [3,4]. Although nomenclature has been inconsistent in the field, based on their location, the ORFs encoding microproteins can been classified as 1) small ORFs, “sORFs”, when they are located in assumed non-coding transcripts such as lncRNAs, miRNAs and circRNAs, or 2) “alt-ORFs”, when they are inside annotated coding genes starting from alternative start codons. Alt-ORFs can be located in the UTRs (typically called "uORFs" when they are in the 5′UTR) or overlapping the reference coding sequence. In this review, we focus on the microproteins derived from sORFs. Although only a subset of them have been functionally characterized, growing evidences demonstrate that microproteins are indeed active proteins playing important functions in a plethora of processes including RNA processing, DNA repair, metabolism regulation and regeneration [5–9]. Microproteins might be particularly well suited to fine-tune complex processes as regulation of enzyme activity, intracellular signal transduction and cell surface signaling, but extensive research in the field is needed to decipher additional functions of the microproteome [1]. These findings open a new category of molecular regulators with implications from basic research to the clinical setting.

Cancer is a complex and multistep disease in which normal cells, through the succession of several genetic and epigenetic events acquire the capacity to grow, invade adjacent tissues, disseminate and ultimately colonize distant organs. Although the specific mechanisms that allow neoplastic transformation and metastasis may vary between different cancer types, there are common regulatory circuits that collectively govern carcinogenesis. In 2000, Hanahan and Weinberg published a seminal paper in which they postulated six capabilities shared by most human tumors [10], which was revisited in 2011 to finally include eight hallmarks of cancer: sustained proliferative signaling, evasion of growth suppressors, resistance to cell death, replicative immortality, induction of angiogenesis, activation of invasion and metastasis, reprogramming of energy metabolism and the evasion of immune destruction (Fig. 1) [10]. Moreover, they proposed two enabling characteristics that represent the mechanisms by which the hallmarks of cancer are acquired: genome instability and the inflammatory state of premalignant lesions (Fig. 1). Although in the last decades we have witnessed a great advance in molecular oncology, we are still far from fully understanding how the hallmarks of cancer are acquired and maintained to sustain tumors. The new information emerging from the study of microproteins suggests that they constitute an important source of cancer regulators implicated in multiple hallmarks of cancer.

* Corresponding author.
E-mail address: mabad@vhio.net (M. Abad).
[1] Authors contributed equally to this work.

https://doi.org/10.1016/j.yexcr.2020.111997
Received 8 January 2020; Received in revised form 2 April 2020; Accepted 3 April 2020
Available online 14 April 2020

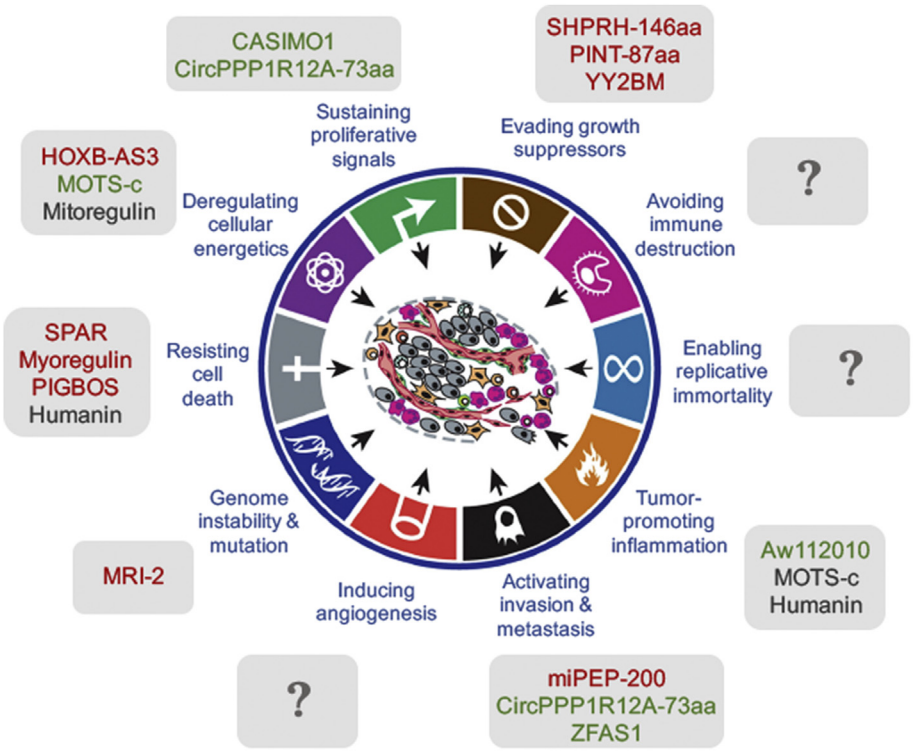


Fig. 1. Microproteins as novel regulators of the hallmarks of cancer. A subset of microproteins have been functionally characterized and have been directly related with cancer; some others, based on their function, are likely to be related with cancer. The figure represents the hallmarks of cancer defined by Hanahan and Weinberg and their related microproteins. In green, the microproteins that promotes or activate the hallmark and, in red, the microproteins that function inhibiting or blocking the hallmark. The ones in grey need further investigation to be classified. Of interest, those depicted in red represent tumor suppressor microproteins with potential pharmacological activity, while those in green are pro-oncogenic peptides that could be targeted in the clinic. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

1.1. Sustaining proliferation signals

In physiologic conditions, mitogenic signals are strictly controlled to ensure the maintenance of normal tissue architecture and homeostasis. By contrast, cancer cells are mitogenically overstimulated. Such mitogenic hyperactivity can be achieved in several ways: First, neoplastic cells can produce their own growth-promoting signals in an autocrine manner or, alternatively, produce paracrine factors that stimulate the release of mitogens by neighbor stromal cells. Second, specific somatic mutations can trigger constitutive activation of growth factor receptors or its downstream components, converting those elements of the signaling cascades in "*bona fide* oncogenes". Recent studies have shown the importance of some microproteins regulating mitogenic signaling. The first sORF described with oncogenic activity has been Cancer-Associated Small Integral Membrane Open reading frame 1 (CASIMO1). CASIMO1 is expressed in hormone-dependent breast cancer during all stages of malignancy, and its deficiency reduces cell proliferation in several breast cancer cell lines [11]. CASIMO1 interacts with squalene epoxidase (SQLE), an oncogene that promotes ERK phosphorylation and MAPK pathway activation [12]. Remarkably, CASIMO1 deficiency reduces ERK phosphorylation while exogenous overexpression of SQLE is sufficient to rescue the loss of CASIMO1, suggesting that CASIMO1 might be modulating ERK activation via SQLE interaction [11]. These observations are in line with a role of CASIMO1 as an oncogene, acting as a positive regulator of MAPK cascade in breast cancer cells. More recently, another microprotein has been proposed to act as a positive regulator of the Hippo-Yap pathway in colon cancer. CircPPP1R12A-73aa is a microprotein encoded by CircPPP1R12A, the most abundant circular RNA (circRNA) in colon cancer, and its overexpression leads to increased cell proliferation [13]. CircPPP1R12A-73aa induces the transcriptional upregulation of Hippo-Yap pathway components and increases YAP1 protein levels, suggesting a possible role of CircPPP1R12A-73aa as an activator of the Hippo pathway [13]. Although CircPPP1R12A-73aa's mechanism of action needs to be further studied, together with CASIMO1 exemplify the relevance of sORFs regulating mitogenic signals, and how they can be exploited by cancer cells to sustain their proliferation needs.

1.2. Evading growth suppressors

In homeostasis, powerful signaling programs block the proliferation of damaged or potentially malignant cells. Signals that activate growth suppression are integrated by the cell in a highly complex manner to decide whether to halt cell-cycle progression, activate apoptotic programs or enter senescence [14,15]. These signaling programs are mainly governed by tumor suppressor proteins and, for neoplastic transformation to occur, these tumor-suppressing mechanisms need to be inactivated [16–18]. Thus, a good understanding of tumor suppression pathways is crucial and might reveal novel therapeutic options in cancer. In this regard, the sORF-encoded proteome can provide new insights on the biology of tumor suppression mechanisms and could represent a novel source of therapeutic agents. Several microproteins have already been shown to play a role in regulating tumor suppression mechanisms through different ways. The *LINC-PINT* gene was already reported to produce a lncRNA regulated by p53 [19], but in its circular form, it contains a sORF that encodes an 87-amino acid microprotein. This microprotein, PINT-87aa, suppresses tumorigenic capabilities of glioblastoma cell lines *in vitro* including cell proliferation, self-renewal and anchorage independent growth, and its deficiency results in increased tumor burden *in vivo*. The authors showed that this microprotein interacts with the polymerase associated factor (PAF1) complex, essential for RNA II polymerase binding and transcription elongation [20]. The interaction of PINT-87aa with PAF1 pauses RNA II polymerase at specific oncogene promoters -such as Cyclin D1, CPEB1, c-MYC and SOX2-impairing their transcription [21].

Although it is bigger than 100 amino acids, and therefore strictly not generated from a sORF, another interesting example of the coding potential of circ-RNAs is SHPRH-146aa. SHPRH is an E3-ubiquitin ligase that promotes the degradation of PCNA (Proliferating Cell Nuclear Antigen), impeding cell cycle progression through S-phase. Recently, Zhang and colleagues have revealed a circular RNA derived from the *SHPRH* primary transcript that codes for SHPRH-146aa. SHPRH-146aa stabilizes SHRPH by preventing its degradation, and thereby promoting the ubiquitination of PCNA. Accordingly, the overexpression of SHPRH-146aa reduces the proliferation of glioma cell lines *in vitro* and *in vivo*

[22]. Importantly, both microproteins, SHPRH-146aa and PINT-87aa are silenced or downregulated in glioblastoma [19][22], further supporting their tumor suppressor potential and opening the possibility of using these small proteins as therapeutic agents.

More recently, the Y-chromosome-linked lncRNA LINC00278 has been shown to encode YY1BM, a microprotein with tumor suppressor activity in esophageal squamous cell carcinoma (ESCC). This microprotein induces apoptosis through the androgen receptor pathway under nutrient deprivation. Interestingly, YY1BM is downregulated by cigarette smoking in human males with ESCC, increasing the survival of cancer cells under nutrient deprivation. Moreover, intratumoral injection of the purified microprotein showed a therapeutic effect in xenograft models, suggesting its potential as a tumor suppressor agent [23].

Last, it is worth to mention that some identified microproteins, like NoBody, have not been directly linked with cancer but they regulate fundamental processes that can impact on cancer cells, such as mRNA decay [5]. Nonsense Mediated Decay is a complex process that can be exploited by cancer cells to degrade the mRNA of tumor suppressor genes. On the other hand, it can also be used as a therapeutic intervention to target oncogene-encoding mRNAs [24]. Thus, the role of this microprotein (and many others) as an oncogene or as a tumor suppressor might be highly tumor specific and dependent on the cellular context.

1.3. Resisting cell death

Regulation of the balance between cell death and survival is critical for maintaining tissue homeostasis. The induction of programmed cell death by apoptosis is a natural barrier for neoplastic transformation [25]. Signals that activate apoptosis are sensed and integrated, among others, by the pro- and anti-apoptotic proteins of the Bcl-2 family. Anti-apoptotic members of the family (Bcl-2, Bcl-x_L, Bcl-w, Mcl1, A1) interact with pro-apoptotic proteins Bax and Bak, suppressing their function. Upon certain pro-apoptotic stimuli, this interaction is broken allowing Bax and Bak to disrupt mitochondrial membrane, which releases cytochrome c to the cytosol, activating in turn the caspases cascade that ultimately disassembles the cell [26]. Interestingly, certain sORFs appear to be important fine-tuning regulators of this process. The microprotein Humanin (HN) is encoded by a sORF in the mitochondrial 16s ribosomal RNA gene, although a nuclear origin of this microprotein has not been ruled out yet due to the presence of similar ORFs in the nuclear genome [27]. HN was first described as a neuroprotective agent in Alzheimer disease [28]. Further studies have shown that its cytoprotective activity is, at least in part, due to its ability to block apoptosis through direct interaction with Bax [29] or by binding and inhibiting the Bax activator BimEL [30]. Regarding cancer, HN is expressed in gastric and bladder cancer and induces chemotherapy resistance [31]. On the other hand, the cytoprotective activity of HN could be beneficial for normal tissues, given that HN also reduces the side-effects of chemotherapy in non-cancer cells [31]. Further studies are needed to clarify the potential role of HN in the clinical setting.

Together with apoptotic suppression, cancer cells activate cell survival mechanisms to avoid cell death. mTOR is a serine/threonine kinase that integrates multiple environmental cues, and it is activated to promote cell growth and survival in favorable conditions [32]. Hyperactivation of mTOR has been reported in more than 70% of cancers [33] and, therefore, mTOR inhibition is an approach currently used in anti-cancer therapies. SPAR is a lncRNA-encoded microprotein that inhibits mTORC1 by its interaction with the lysosomal v-ATPase. Depletion of SPAR *in vivo* has been shown to improve muscle regeneration through higher mTORC1 activity [34]. Among the multiple mechanisms that regulate mTORC1, SPAR appears to reduce the amino-acid sensing route of mTORC1 activation. Although it has not been tested, SPAR expression could have anti-tumor properties in specific cancer types where mTORC1 activation occurs through amino-acid sensing pathways, as it is the case of some lymphomas [35].

On the other hand, the activation of the cellular stress response is an essential mechanism that helps healthy cells to cope with stress and damage, and it is co-opted by malignant cells to thrive in highly adverse conditions and avoid cell death. Accordingly, Unfolded Protein Response (UPR) and endoplasmic reticulum (ER)-stress response have been reported to be upregulated in many cancers, helping them to deal with high protein synthesis demands while protecting them from stress-induced cell death [36,37]. These pro-survival responses are highly dependent on cytosolic Ca^{2+} concentration and the ER is the main responsible of intracellular Ca^{2+} storage [38]. MYOREGULIN (MLN), encoded by a previously annotated long non-coding RNA works as an inhibitor of the Ca^{2+} ATPase pump SERCA [39]. Inhibition of SERCA activity increases cytosolic Ca^{2+} concentration, sensitizing cells to cell death. For this reason, SERCA inhibition is being tested as a potential therapy in tumors with hyperactivation of Ca^{2+} channeling activity [40]. Even though MLN has not been related with cancer so far and in homeostasis its expression is restricted to skeletal muscle, several SERCA-inhibitory microproteins have been reported to control Ca^{2+} signaling in a tissue specific manner [41] and their expression might be useful to tackle cancers with high SERCA activity.

Finally, PIGBOS is a novel microprotein encoded by an antisense of the *PIGB* gene. This microprotein localizes in the mitochondrial outer membrane interacting with the ER through the CLCC1 protein. Downregulation of PIGBOS induces apoptosis by increasing sensitivity to chemically-induced UPR [42]. Whereas the molecular mechanism of PIGBOS function has not been described yet, it is reasonable to think that its inhibition in cancer cells may have therapeutic potential by sensitizing cells to UPR and forcing them to enter apoptosis.

1.4. Activation of invasion and metastasis

Carcinomas are tumors that arise from epithelial tissues. The progression from localized tumor to invasive carcinoma and distal metastasis requires changes in cell morphology and in cell-cell and cell-extracellular matrix (ECM) attachment. The “epithelial-to-mesenchymal transition” (EMT) is a cellular program involved in embryonic morphogenesis and wound healing. Cancer cells can also activate the EMT program to acquire the invasive phenotype needed for metastatic spread [43]. Biological traits acquired during EMT include the loss of adherent junctions, the acquisition of spindle/fibroblastic morphology, the expression of matrix-degradation enzymes, increased motility and resistance to apoptosis. Importantly, while EMT is needed for metastasis, it must be reversed to colonize a new organ through a process known as “mesenchymal-to-epithelial transition” (MET) [43]. Of note, it has been recently proposed that cancer cells may acquire a “plastic-hybrid state” or a “partial EMT” state. According to this vision, cancer cells acquire mesenchymal features while continuing to express epithelial traits, resulting in a selective advantage during the metastatic process [44].

Many non-coding RNAs have been shown to regulate EMT, highlighting the importance of miR-200a and miR-200b as key negative EMT regulators, which are usually epigenetically repressed in cancer cells [45]. Of note, a recent study has identified two potential microproteins encoded by the miR-200a and miR-200b pri-miRNAs, the precursor transcripts of the miRNAs, which have been named miPEP-200a and miPEP-200b. The expression of these sORFs seems to be associated with a decreased migration in wound healing assay and with diminished expression of the mesenchymal marker Vimentin [46]. More recently, the lncRNA ZFAS1 has been shown to translate a microprotein that is proposed to promote cell migration by elevating intracellular reactive oxygen species (ROS) production [47]. Additionally, the microprotein CircPPP1R12A-73aa (described above) increases cancer cell migration and invasion *in vitro* and metastasis development *in vivo* [13], possibly reflecting the activation of EMT by the Hippo-Yap pathway [48,49].

Finally, recent evidences suggest that cell-to-cell fusion events,

especially between cancer cells and immune cells, could contribute to the acquisition of metastatic behaviors [50]. Importantly, microproteins can regulate cell fusion, as it has been shown for MYOMIXER, an 84-amino acid peptide necessary for heterotypic fibroblast-myoblast fusion [51]. Although speculative, it is possible that there are non-identified microproteins playing important functions in metastasis by regulating cell fusion. Although more investigations are needed to have a complete picture, all together these discoveries point to the microproteome as a source of regulators of cancer invasion and metastasis.

1.5. Deregulating cellular energetics

Another important feature of cancer cells is that they reprogram their metabolism in different ways to comply with their highly demanding energetic needs. One of these strategies is to rely on glycolysis rather than on oxidative phosphorylation (OXPHOS) as their primary energy production mechanism, even in the presence of oxygen [52]. The process of metabolic reprogramming results from a complex regulation between mitochondrial genes, tumor suppressors and oncogenic pathways and its targeting is currently being tested as a therapeutic strategy [53]. Importantly, microproteins have emerged as important regulators of this process with potential clinical implications. The lncRNA HOXB-AS3 has been shown to produce a 53-amino acid microprotein involved in RNA splicing. Huang and collaborators showed that HOXB-AS3 is downregulated in colorectal cancer cells, which changes the splicing of Pyruvate Kinase M (PKM) pre-mRNA to re-express the embryonic isoform PKM2, that favors glycolytic activity. By contrast, expression of HOXB-AS3 peptide favors the expression of the adult isoform (PKM1), that promotes oxidative phosphorylation. Collectively, they demonstrate that overexpression of HOXB-AS3 peptide in colorectal cancer cell lines attenuates their oncogenic capacity by altering their use of glucose metabolism [6]. Other microproteins have also been described to play a role in metabolic regulation, although their role in cancer is yet to be investigated. Specifically, MITOREGULIN (also called MOXI or MPM) is a 56-amino acid microprotein encoded by LINC00116, a muscle-enriched lncRNA. The function of this protein has been proposed to rely on its interaction with different inner mitochondrial proteins, increasing respiratory efficiency through the stabilization of supercomplexes in the electron transport chain [54,55] and promoting long-chain fatty acid ß-oxidation [56]. It would be of interest to study the role of this microprotein in cancer cell metabolism. Finally, the mitochondrial genome also plays a role in metabolic activity and, despite its small size, it has been described to contain several sORFs, like the one encoding for MOTS-c (mitochondrial open reading frame of the 12*s* rRNA-c). *In vitro*, MOTS-c increases glucose uptake, glycolytic activity and AMPK activation, while reducing oxygen consumption rate. Accordingly, MOTS-c improves metabolic parameters associated with obesity *in vivo* [57]. Given that MOTS-c favors a glycolytic program and activates AMPK, it might be beneficial for cancer cells, which would be interesting to be addressed in the future.

1.6. Enabling characteristic: genomic instability and mutations

As proposed by Hanahan and Weinberg [10], accumulation of genomic alterations during carcinogenesis is an event that enables the acquisition of the core hallmarks discussed above. Tumor progression can be seen as a succession of clonal expansions: mutations are randomly accumulated in the pool of cancer cells and, eventually, advantaged genotypes enabling cell survival and growth are selected under the pressure of environmental stimuli. Due to the adaptive advantage that a high mutational rate confers to cancer cells, the components of the genomic maintenance machinery are often affected in neoplasia. Typically, the accumulation of mutations can be accelerated by defects in the sensors of DNA damage, in components of DNA repair machinery and/or in effectors that force damaged cells to enter senescence or apoptosis [58]. Thus, microproteins involved in the maintenance of genome stability may behave as tumor suppressor genes and are likely to be affected in cancer cells. Slavoff and collaborators identified MRI-2 as a 69-amino acid microprotein coded by *C7orf49* gene, also known as “modulator of retrovirus infection homolog” (MRI) [8]. MRI-2 interactome analysis revealed that this microprotein interacts with Ku70 and Ku80 proteins, the two subunits of the heterodimeric protein Ku, key effector of the non-homologous-end-joining (NHEJ) pathway for DNA double-strand break (DSB) repair. When a DNA DSB occurs, the first protein that binds to the break is the Ku heterodimer, which allows the recruitment of the DNA-dependent protein kinase (DNA-PK) and additional factors which in turn repair the DSB [59]. MRI-2 is accumulated in the nuclei of DBS-induced cells, and recombinant MRI-2 increases NHEJ *in vitro* [8]. The mechanism by which MRI-2 enhances NHEJ has not been fully addressed, but it is possible that the interaction of MRI-2 with Ku proteins may improve their DNA-binding affinity or facilitate the recruitment of other repair complex components. Although in this review we focus in sORF-derived microproteins, it is worth to mention that many alt-ORFs and uORFs can produce functional microproteins with potential roles in genome instability. For example, the AltMRVI1 microprotein is coded by an alt-ORF inside the MRVI gene, and it directly interacts with BRCA1, one of the key effectors of homologous recombination DNA repair machinery [60,61]. The discovery of microproteins involved in genomic stability maintenance suggests that there could be numerous “genomic regulator-microproteins” which help preventing accumulation of genomic alterations in cancer cells.

1.7. Enabling characteristic: tumor-promoting inflammation

Another characteristic that allows the acquisition of the core cancer hallmarks is inflammation [10]. Several microproteins already described in this review might be of particular interest regarding tumor inflammation. In particular, the MDPs MOTS-c and HN have demonstrated their capacity to exacerbate the pro-inflammatory effect of senescence-associated secretory phenotype (SASP) of senescent cells by modulating their mitochondrial activity [62]. Cellular senescence is activated by multiple cellular stressors and it is characterized by a stable cell-cycle arrest and a pro-inflammatory secretome. Senescence acts as a main tumor suppressor barrier that impedes neoplastic transformation of damaged cells and promotes tissue repair [17]. However, the accumulation of senescent cells in tissues could have detrimental effects, mainly because of the inflammatory SASP and, in tumors, the presence of senescent cells promotes cancer cell growth and metastasis [63,64]. In this regard, it would be interesting to study whether the pro-inflammatory cytokines upregulated by MDPs facilitate the immunoclearance of senescent tumor cells or, on the contrary, favor a pro-tumorigenic microenvironment. In addition, ribosome profiling of bone marrow-derived macrophages revealed Aw112010, a non-ATG-initiated microprotein that promotes a pro-inflammatory response increasing canonical inflammatory cytokines like IL-6 and IL-12p40 upon bacterial infection [65]. These findings suggest a role of microproteins in cancer development through regulating inflammation.

2. Concluding remarks

sORF-encoded proteins have expanded our view about the coding potential of the genome, adding a new layer of complexity in the regulation of biological processes. The emergent picture suggests that microproteins allow the fine-tuning of many of these processes to adapt to specific needs and cellular contexts. Here, we have summarized what might be their implication in cancer. Even if only a small subset of microproteins have been functionally characterized so far, there is evidence of many of them as regulators of most of the hallmarks of cancer and its enabling characteristics (Fig. 1). While, so far, microproteins have not been directly related to angiogenesis, replicative immortality and immune evasion, we should take into account that the

proposed hallmarks are interconnected and some of the already identified microproteins, upon further analysis, could be classified in several hallmarks at the same time.

Finally, it is worth mentioning that most sequencing efforts in cancer restrict their analysis to the annotated protein-coding genome, unintentionally ignoring the microproteome. Therefore, it remains unclear if mutations in microproteins are selected during cancer evolution. If this were the case, mutated microproteins could also be a source of cancer neoantigens that can be used to improve the development of personalized immunotherapy [66]. We envision that this is an area that is going to be intensively studied and expanded in the coming years, and will bring crucial information for the clinic.

Here, we have discussed a set of microproteins encoded by lncRNAs, miRNAs, rRNAs, and cirRNAs but many more are yet to be explored, including the ones coded by the so-called alt-ORFs, that we have not addressed in this review. We are at the beginning of a new set of discoveries in which the identification and characterization of the cellular microproteins repertoire -the microproteome- will help us to better understand how physiological and pathological processes are regulated at its finest level. We anticipate that advances in this field will bring new therapeutic opportunities for oncology.

CRediT authorship contribution statement

Iñaki Merino-Valverde: Conceptualization, Writing - original draft, Writing - review & editing. **Emanuela Greco:** Conceptualization, Writing - original draft, Writing - review & editing. **María Abad:** Conceptualization, Writing - original draft, Writing - review & editing, Funding acquisition.

Acknowledgements

M.A. is a recipient of a Ramón y Cajal contract from the Spanish Ministry of Economy (MINECO). I.M. is a recipient of an FPI contract from the Spanish Ministry of Economy (MINECO). Work in the laboratory of M.A. is funded by Fero Foundation and by grants from the Spanish Ministry of Economy (RTI2018-102046-B-I00), from *"la Caixa"* Foundation (HR18-00256), and from Mutua Madrileña Foundation.

References

[1] C.A. Makarewich, E.N. Olson, Mining for micropeptides, Trends Cell Biol. 27 (2017) 685–696.
[2] M.W. Orr, Y. Mao, G. Storz, S.B. Qian, Alternative ORFs and small ORFs: shedding light on the dark proteome, Nucleic Acids Res. 48 (3) (2019) 1029–1042, https://doi.org/10.1093/nar/gkz734.
[3] M.E. Dinger, K.C. Pang, T.R. Mercer, J.S. Mattick, Differentiating protein-coding and noncoding RNA: challenges and ambiguities, PLoS Comput. Biol. 4 (2008) e1000176.
[4] J.W. Fickett, ORFs and genes: how strong a connection? J. Comput. Biol. 2 (1995) 117–123.
[5] N.G. D'Lima, J. Ma, L. Winkler, Q. Chu, K.H. Loh, E.O. Corpuz, B.A. Budnik, J. Lykke-Andersen, A. Saghatelian, S.A. Slavoff, A human microprotein that interacts with the mRNA decapping complex, Nat. Chem. Biol. 13 (2017) 174–180.
[6] J.Z. Huang, M. Chen, X.C. Gao, S. Zhu, H. Huang, M. Hu, H. Zhu, G.R. Yan, A peptide encoded by a putative lncRNA HOXB-AS3 suppresses colon cancer growth, Mol Cell 68 (2017) 171–184 e176.
[7] B.R. Nelson, C.A. Makarewich, D.M. Anderson, B.R. Winders, C.D. Troupes, F. Wu, A.L. Reese, J.R. McAnally, X. Chen, E.T. Kavalali, et al., A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle, Science 351 (2016) 271–275.
[8] S.A. Slavoff, J. Heo, B.A. Budnik, L.A. Hanakahi, A. Saghatelian, A human short open reading frame (sORF)-encoded polypeptide that stimulates DNA end joining, J. Biol. Chem. 289 (2014) 10950–10957.
[9] F. Yeasmin, T. Yada, N. Akimitsu, Micropeptides encoded in transcripts previously identified as long noncoding RNAs: a new chapter in transcriptomics and proteomics, Front. Genet. 9 (2018) 144.
[10] D. Hanahan, R.A. Weinberg, Hallmarks of cancer: the next generation, Cell 144 (2011) 646–674.
[11] M. Polycarpou-Schwarz, M. Gross, P. Mestdagh, J. Schott, S.E. Grund, C. Hildenbrand, J. Rom, S. Aulmann, H.P. Sinn, J. Vandesompele, et al., The cancer-associated microprotein CASIMO1 controls cell proliferation and interacts with squalene epoxidase modulating lipid droplet formation, Oncogene 37 (2018) 4750–4768.
[12] Z. Sui, J. Zhou, Z. Cheng, P. Lu, Squalene epoxidase (SQLE) promotes the growth and migration of the hepatocellular carcinoma cells, Tumour Biol 36 (2015) 6173–6179.
[13] X. Zheng, L. Chen, Y. Zhou, Q. Wang, Z. Zheng, B. Xu, C. Wu, Q. Zhou, W. Hu, C. Wu, et al., A novel protein encoded by a circular RNA circPPP1R12A promotes tumor pathogenesis and metastasis of colon cancer via Hippo-YAP signaling, Mol. Canc. 18 (2019) 47.
[14] B.J. Aubrey, G.L. Kelly, A. Janic, M.J. Herold, A. Strasser, How does p53 induce apoptosis and how does this relate to p53-mediated tumour suppression? Cell Death Differ. 25 (2018) 104–113.
[15] B.G. Childs, D.J. Baker, J.L. Kirkland, J. Campisi, J.M. van Deursen, Senescence and apoptosis: dueling or complementary cell fates? EMBO Rep. 15 (2014) 1139–1153.
[16] A.R. Delbridge, L.J. Valente, A. Strasser, The role of the apoptotic machinery in tumor suppression, Cold Spring Harb Perspect Biol 4 (2012).
[17] V. Gorgoulis, P.D. Adams, A. Alimonti, D.C. Bennett, O. Bischof, C. Bishop, J. Campisi, M. Collado, K. Evangelou, G. Ferbeyre, et al., Cellular senescence: defining a path forward, Cell 179 (2019) 813–827.
[18] C.J. Sherr, F. McCormick, The RB and p53 pathways in cancer, Canc. Cell 2 (2002) 103–112.
[19] O. Marin-Bejar, F.P. Marchese, A. Athie, Y. Sanchez, J. Gonzalez, V. Segura, L. Huang, I. Moreno, A. Navarro, M. Monzo, et al., Pint lincRNA connects the p53 pathway with epigenetic silencing by the Polycomb repressive complex 2, Genome Biol. 14 (2013) R104.
[20] F.X. Chen, A.R. Woodfin, A. Gardini, R.A. Rickels, S.A. Marshall, E.R. Smith, R. Shiekhattar, A. Shilatifard, PAF1, a molecular regulator of promoter-proximal pausing by RNA polymerase II, Cell 162 (2015) 1003–1015.
[21] M. Zhang, K. Zhao, X. Xu, Y. Yang, S. Yan, P. Wei, H. Liu, J. Xu, F. Xiao, H. Zhou, et al., A peptide encoded by circular form of LINC-PINT suppresses oncogenic transcriptional elongation in glioblastoma, Nat. Commun. 9 (2018) 4475.
[22] M. Zhang, N. Huang, X. Yang, J. Luo, S. Yan, F. Xiao, W. Chen, X. Gao, K. Zhao, H. Zhou, et al., A novel protein encoded by the circular form of the SHPRH gene suppresses glioma tumorigenesis, Oncogene 37 (2018) 1805–1814.
[23] S. Wu, L. Zhang, J. Deng, B. Guo, F. Li, Y. Wang, R. Wu, S. Zhang, J. Lu, Y. Zhou, A novel micropeptide encoded by Y-linked LINC00278 links cigarette smoking and AR signaling in male esophageal squamous cell carcinoma, Cancer Res Epub ahead of print (2020), https://doi.org/10.1158/0008-5472.CAN-19-3440.
[24] M.W. Popp, L.E. Maquat, Nonsense-mediated mRNA decay and cancer, Curr. Opin. Genet. Dev. 48 (2018) 44–50.
[25] S.W. Lowe, E. Cepero, G. Evan, Intrinsic tumour suppression, Nature 432 (2004) 307–315.
[26] J.M. Adams, S. Cory, The Bcl-2 apoptotic switch in cancer development and therapy, Oncogene 26 (2007) 1324–1337.
[27] C. Lee, K. Yen, P. Cohen, Humanin: a harbinger of mitochondrial-derived peptides? Trends Endocrinol Metab 24 (2013) 222–228.
[28] Y. Hashimoto, T. Niikura, Y. Ito, H. Sudo, M. Hata, E. Arakawa, Y. Abe, Y. Kita, I. Nishimoto, Detailed characterization of neuroprotection by a rescue factor humanin against various Alzheimer's disease-relevant insults, J. Neurosci. 21 (2001) 9235–9245.
[29] B. Guo, D. Zhai, E. Cabezas, K. Welsh, S. Nouraini, A.C. Satterthwait, J.C. Reed, Humanin peptide suppresses apoptosis by interfering with Bax activation, Nature 423 (2003) 456–461.
[30] F. Luciano, D. Zhai, X. Zhu, B. Bailly-Maitre, J.E. Ricci, A.C. Satterthwait, J.C. Reed, Cytoprotective peptide humanin binds and inhibits proapoptotic Bcl-2/Bax family protein BimEL, J. Biol. Chem. 280 (2005) 15825–15835.
[31] C.F. Zuccato, A.S. Asad, A.J. Nicola Candia, M.F. Gottardo, M.A. Moreno Ayala, M.S. Theas, A. Seilicovich, M. Candolfi, Mitochondrial-derived peptide humanin as therapeutic target in cancer and degenerative diseases, Expert Opin. Ther. Targets 23 (2019) 117–126.
[32] R.A. Saxton, D.M. Sabatini, mTOR signaling in growth, metabolism, and disease, Cell 169 (2017) 361–371.
[33] S.A. Forbes, N. Bindal, S. Bamford, C. Cole, C.Y. Kok, D. Beare, M. Jia, R. Shepherd, K. Leung, A. Menzies, et al., COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer, Nucleic Acids Res. 39 (2011) D945–D950.
[34] A. Matsumoto, J.G. Clohessy, P.P. Pandolfi, SPAR, a lncRNA encoded mTORC1 inhibitor, Cell Cycle 16 (2017) 815–816.
[35] A. Ortega-Molina, N. Deleyto-Seldas, J. Carreras, A. Sanz, C. Lebrero-Fernandez, C. Menendez, A. Vandenberg, B. Fernandez-Ruiz, L. Marin-Arraiza, C. de la Calle Arregui, et al., Oncogenic Rag GTPase signaling enhances B cell activation and drives follicular lymphoma sensitive to pharmacological inhibition of mTOR, Nat Metab 1 (2019) 775–789.
[36] J.R. Cubillos-Ruiz, S.E. Bettigole, L.H. Glimcher, Tumorigenic and immunosuppressive effects of endoplasmic reticulum stress in cancer, Cell 168 (2017) 692–706.
[37] M. Wang, R.J. Kaufman, The impact of the endoplasmic reticulum protein-folding environment on cancer development, Nat. Rev. Canc. 14 (2014) 581–597.
[38] R. Bagur, G. Hajnoczky, Intracellular Ca(2+) sensing: its role in calcium homeostasis and signaling, Mol Cell 66 (2017) 780–788.
[39] D.M. Anderson, K.M. Anderson, C.L. Chang, C.A. Makarewich, B.R. Nelson, J.R. McAnally, P. Kasaragod, J.M. Shelton, J. Liou, R. Bassel-Duby, et al., A micropeptide encoded by a putative long noncoding RNA regulates muscle performance, Cell 160 (2015) 595–606.
[40] C. Cui, R. Merritt, L. Fu, Z. Pan, Targeting calcium signaling in cancer therapy, Acta Pharm. Sin. B 7 (2017) 3–17.
[41] D.M. Anderson, C.A. Makarewich, K.M. Anderson, J.M. Shelton, S. Bezprozvannaya,

R. Bassel-Duby, E.N. Olson, Widespread control of calcium signaling by a family of SERCA-inhibiting micropeptides, Sci. Signal. 9 (2016) ra119.
[42] Q. Chu, T.F. Martinez, S.W. Novak, C.J. Donaldson, D. Tan, J.M. Vaughan, T. Chang, J.K. Diedrich, L. Andrade, A. Kim, et al., Regulation of the ER stress response by a mitochondrial microprotein, Nat. Commun. 10 (2019) 4883.
[43] K. Polyak, R.A. Weinberg, Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits, Nat. Rev. Canc. 9 (2009) 265–273.
[44] H.C. Lo, X.H. Zhang, EMT in metastasis: finding the right balance, Dev. Cell 45 (2018) 663–665.
[45] L. Vrba, J.C. Garbe, M.R. Stampfer, B.W. Futscher, Epigenetic regulation of normal human mammary cell type-specific miRNAs, Genome Res. 21 (2011) 2026–2037.
[46] S.M. Jinbo Fang, Veena N. Rao, E. Shyam P. Reddy, Decoding of non-coding DNA and non-coding RNA: pri-micro RNA-encoded novel peptides regulate migration of cancer cells, Journal of Pharmaceutical Sciences and Pharmacology 3 (2017) 23–27.
[47] Z.W. Guo, Y. Meng, X.M. Zhai, C. Xie, N. Zhao, M. Li, C.L. Zhou, K. Li, T.C. Liu, X.X. Yang, et al., Translated long non-coding ribonucleic acid ZFAS1 promotes cancer cell migration by elevating reactive oxygen species production in hepatocellular carcinoma, Front. Genet. 10 (2019) 1111.
[48] M. Cordenonsi, F. Zanconato, L. Azzolin, M. Forcato, A. Rosato, C. Frasson, M. Inui, M. Montagner, A.R. Parenti, A. Poletti, et al., The Hippo transducer TAZ confers cancer stem cell-related traits on breast cancer cells, Cell 147 (2011) 759–772.
[49] W. Lehmann, D. Mossmann, J. Kleemann, K. Mock, C. Meisinger, T. Brummer, R. Herr, S. Brabletz, M.P. Stemmler, T. Brabletz, ZEB1 turns into a transcriptional activator by interacting with YAP1 in aggressive cancer types, Nat. Commun. 7 (2016) 10498.
[50] C.E. Gast, A.D. Silk, L. Zarour, L. Riegler, J.G. Burkhart, K.T. Gustafson, M.S. Parappilly, M. Roh-Johnson, J.R. Goodman, B. Olson, et al., Cell fusion potentiates tumor heterogeneity and reveals circulating hybrid cells that correlate with stage and survival, Sci Adv 4 (2018) eaat7828.
[51] P. Bi, A. Ramirez-Martinez, H. Li, J. Cannavino, J.R. McAnally, J.M. Shelton, E. Sanchez-Ortiz, R. Bassel-Duby, E.N. Olson, Control of muscle formation by the fusogenic micropeptide myomixer, Science 356 (2017) 323–327.
[52] N.N. Pavlova, C.B. Thompson, The emerging hallmarks of cancer metabolism, Cell Metabol. 23 (2016) 27–47.
[53] U.E. Martinez-Outschoorn, M. Peiris-Pages, R.G. Pestell, F. Sotgia, M.P. Lisanti, Cancer metabolism: a therapeutic perspective, Nat. Rev. Clin. Oncol. 14 (2017) 11–31.
[54] Y.F. Lin, M.H. Xiao, H.X. Chen, Y. Meng, N. Zhao, L. Yang, H. Tang, J.L. Wang, X. Liu, Y. Zhu, et al., A novel mitochondrial micropeptide MPM enhances mitochondrial respiratory activity and promotes myogenic differentiation, Cell Death Dis. 10 (2019) 528.
[55] C.S. Stein, P. Jadiya, X. Zhang, J.M. McLendon, G.M. Abouassaly, N.H. Witmer, E.J. Anderson, J.W. Elrod, R.L. Boudreau, Mitoregulin: a lncRNA-encoded microprotein that supports mitochondrial supercomplexes and respiratory efficiency, Cell Rep. 23 (2018) 3710–3720 e3718.
[56] C.A. Makarewich, K.K. Baskin, A.Z. Munir, S. Bezprozvannaya, G. Sharma, C. Khemtong, A.M. Shah, J.R. McAnally, C.R. Malloy, L.I. Szweda, et al., MOXI is a mitochondrial micropeptide that enhances fatty acid beta-oxidation, Cell Rep. 23 (2018) 3701–3709.
[57] C. Lee, J. Zeng, B.G. Drew, T. Sallam, A. Martin-Montalvo, J. Wan, S.J. Kim, H. Mehta, A.L. Hevener, R. de Cabo, et al., The mitochondrial-derived peptide MOTS-c promotes metabolic homeostasis and reduces obesity and insulin resistance, Cell Metabol. 21 (2015) 443–454.
[58] S.P. Jackson, J. Bartek, The DNA-damage response in human biology and disease, Nature 461 (2009) 1071–1078.
[59] N.R. Pannunzio, G. Watanabe, M.R. Lieber, Nonhomologous DNA end-joining for repair of DNA double-strand breaks, J. Biol. Chem. 293 (2018) 10512–10523.
[60] B. Vanderperre, J.F. Lucier, C. Bissonnette, J. Motard, G. Tremblay, S. Vanderperre, M. Wisztorski, M. Salzet, F.M. Boisvert, X. Roucou, Direct detection of alternative open reading frames translation products in human significantly expands the proteome, PloS One 8 (2013) e70698.
[61] W. Zhao, J.B. Steinfeld, F. Liang, X. Chen, D.G. Maranon, C. Jian Ma, Y. Kwon, T. Rao, W. Wang, C. Sheng, et al., BRCA1-BARD1 promotes RAD51-mediated homologous DNA pairing, Nature 550 (2017) 360–365.
[62] S.J. Kim, H.H. Mehta, J. Wan, C. Kuehnemann, J. Chen, J.F. Hu, A.R. Hoffman, P. Cohen, Mitochondrial peptides modulate mitochondrial function during cellular senescence, Aging (Albany NY) 10 (2018) 1239–1256.
[63] A. Calcinotto, J. Kohli, E. Zagato, L. Pellegrini, M. Demaria, A. Alimonti, Cellular senescence: aging, cancer, and injury, Physiol. Rev. 99 (2019) 1047–1078.
[64] D.V. Faget, Q. Ren, S.A. Stewart, Unmasking senescence: context-dependent effects of SASP in cancer, Nat. Rev. Canc. 19 (2019) 439–453.
[65] R. Jackson, L. Kroehling, A. Khitun, W. Bailis, A. Jarret, A.G. York, O.M. Khan, J.R. Brewer, M.H. Skadow, C. Duizer, et al., The translation of non-canonical open reading frames controls mucosal immunity, Nature 564 (2018) 434–438.
[66] A. Garcia-Garijo, C.A. Fajardo, A. Gros, Determinants for neoantigen identification, Front. Immunol. 10 (2019) 1392.