DEPARTAMENTO DE BIOLOGÍA CELULAR,
FISIOLOGÍA E INMUNOLOGÍA

UNIVERSIDAD DE CÓRDOBA

# Role and regulation of the novel splicing variants sst5TMD4 and In1-ghrelin in breast cancer.

## Papel y regulación de las nuevas variantes de splicing sst5TMD4 e In1-ghrelina en cáncer de mama.

David Rincón Fernández-Pacheco

Córdoba, 2016

TITULO: *Role and regulation of the novel splicing variants sst5TMD4 and In1-ghrelin in breast cancer*

AUTOR: *David Rincón Fernández-Pacheco*

---

© Edita: UCOPress. 2016
Campus de Rabanales
Ctra. Nacional IV, Km. 396 A
14071 Córdoba

www.uco.es/publicaciones
publicaciones@uco.es

---

# DEPARTAMENTO DE BIOLOGÍA CELULAR, FISIOLOGÍA E INMUNOLOGÍA

UNIVERSIDAD DE CÓRDOBA

# Role and regulation of the novel splicing variants sst5TMD4 and In1-ghrelin in breast cancer.

## Papel y regulación de las nuevas variantes de splicing sst5TMD4 e In1-ghrelina en cáncer de mama.

Memoria de Tesis Doctoral presentada por **David Rincón Fernández-Pacheco**, Licenciado en Biología, para optar al grado de **Doctor en Ciencias**.

Los Directores


**Dr. Justo P. Castaño Fuentes**
Catedrático de Biología Celular
de la Universidad de Córdoba
IMIBIC

**Dr. Manuel D. Gahete Ortiz**
Investigador Postdoctoral
Universidad de Córdoba
IMIBIC


En Córdoba, a 15 de Septiembre de 2016

D. Justo P. Castaño Fuentes, Catedrático del Departamento de Biología Celular, Fisiología e Inmunología de la Universidad de Córdoba y D. Manuel D. Gahete Ortiz, Investigador Post-doctoral del Instituto Maimónides de Investigación Biomédica de Córdoba (IMIBIC) y asociado al Departamento de Biología Celular, Fisiología e Inmunología de la Universidad de Córdoba,

INFORMAN

Que D. David Rincón Fernández-Pacheco, Licenciado en Biología, ha realizado bajo nuestra dirección el trabajo titulado "**Role and regulation of the novel splicing variants sst5TMD4 and In1-ghrelin in breast cancer"** y que bajo nuestro juicio reúne los méritos suficientes para optar al Grado de Doctor en Biomedicina.

Y para que conste, firmamos la presente en Córdoba, a 15 de Septiembre de 2016

Fdo.: Dr. Justo Pastor Castaño Fuentes

Fdo.: Dr. Manuel D. Gahete Ortiz

**TITULO DE LA TESIS:** Role and regulation of the novel splicing variants sst5TMD4 and In1-ghrelin in breast cancer.

**DOCTORANDO:** David Rincón Fernández-Pacheco

**INFORME RAZONADO DE LOS DIRECTORES DE LA TESIS**

(Se hará mención a la evolución y desarrollo de la tesis, así como a trabajos y publicaciones derivados de la misma).

Duránte el desarrollo de la presente Tesis Doctoral, el doctorando David Rincón Fernández-Pacheco ha alcanzado e incluso superado los objetivos planteados al comienzo de la misma, al tiempo que ha desarrollado técnicas experimentales de gran utilidad para el grupo de investigación, que le han permitido obtener resultados muy relevantes en el estudio de los procesos de splicing alternativo asociados a la aparición de las variantes de splicing sst5TMD4 e In1-ghrelina en cáncer de mama. Concretamente, como fruto de su trabajo Duránte este periodo, ha publicado dos trabajos directamente relacionados con su tesis doctoral, en las revistas ***Oncotarget*** y ***Cancer Research***(sometido), de gran relevancia dentro de nuestra área de investigación. Además, el doctorando ha realizado una estancia en la universidad de California, San Diego, donde amplió y perfeccionó los métodos de trabajo. Por último, el doctorando ha presentado sus resultados en diferentes congresos de ámbito nacional e internacional, de los que han derivado varios capítulos de libro.

Por todo ello, se autoriza la presentación de la tesis doctoral.

Córdoba, 15 de Septiembre de 2016

Fdo: D. Justo Pastor Castaño Fuentes

Fdo.: D. Manuel David Gahete Ortiz

Esta Tesis Doctoral ha sido realizada en el Departamento de Biología Celular, Fisiología e Inmunología de la Universidad de Córdoba y en el Instituto Maimónides de Investigación Biomédica de Córdoba (IMIBIC), bajo la dirección de los Dres. Justo P. Castaño Fuentes y Manuel D. Gahete Ortiz. Dicho proyecto ha sido subvencionado mediante los proyectos BFU2010-19300 (Ministerio de Educación y Ciencia) y BFU2013-43282-R (Ministerio de Economía y Competitividad).

## List of abbreviations

**Ang1/Ang2** - Angiopoietins 1 and 2

**BIRADS** - Breast Imaging Reporting and Data System

**bp** - Number of nucleotides (base pairs)

**CDS** – Coding sequence of the DNA

**CORT** - Cortistatin

**CSCs** - Cancer stem cells

**CTCs** - Circulating tumor cells

**DMSO** - Dimethyl-sulfoxide

**DPBS** - Dulbecco's phosphate-buffered saline

**EGF** - Epithelial growth factor

**EMT** - Epithelial to mesenchymal transition

**ESE/ISE** – Exonic/Intronic splicing enhancers

**ESS/ISS** - Exonic/Intronic splicing suppressors

**gDNA** - Genomic DNA

**GH** - Growth hormone

**GHRL** - Ghrelin gene

**GHRLOS** - Ghrelin Opposite Strand/antisense gene

**GHSR-1a** - Growth hormone secretagogue receptor 1a

**GHSR-1b** - Growth hormone secretagogue receptor 1b

**GOAT** - Ghrelin-O-acyl transferase

**GPCRs** - G-Protein coupled receptors

**HER2** - Human epidermal growth factor receptor type 2

**HIF-1** - Hypoxia-inducible factor 1

**hnRNPs** - Heterogeneous ribonucleoproteins

**HURS** - Hospital Universitario Reina Sofía

**IDCs** - Infiltrating ductal breast carcinomas

**In1** - Intron 1

**Jag1** - Jagged protein 1

**lncRNAs** - Long-non-coding RNAs

**miRNAs** - MicroRNAs

**MTT** - 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide

**NATs** - Natural antisense transcripts

**ncRNAs** - Non-coding RNAs

**pHEMA** - Poly(2-hydroxyethyl methacrylate)

**qPCR** - Quantitative real-time PCR

**SDS** - Sodium dodecyl sulfate

**SNPs** - Single-nucleotide polymorphisms

**snRNPs** - Small nucleolar ribonucleoproteins

**SR-proteins** - Serine-arginine proteins

**SST** - Somatostatin

**SSAs** - Somatostatin analogs

**ssts** - Somatostatin receptors

**TGF-β1** - Transforming growth factor β1

**TICs** - Tumor-initiating cells

**TMA** - Tissue microarray

**TMDs** - Transmembrane domains

**VEGF** - Vascular endothelial growth factor

**Index**


**Resumen** 1

**Summary** 5

**1. Introduction** 9

1.1. Cancer 11
1.1.1. Endocrine-related cancers 11
1.2. Splicing 12
1.2.1. Splicing process 13
1.2.2. Alternative splicing 15
1.2.3. Splicing regulation 16
1.2.4. Splicing in cancer 18
1.2.5. Genomic alteration and aberrant splicing in cancer 19
1.2.6. Altered regulation of splicing in cancer 19
1.2.6.1. Dysregulation of splicing factors 19
1.2.6.1.1. SR proteins in cancer 20
1.2.6.1.2. hnRNPs in cancer 21
1.2.6.1.3. Other RNA-binding regulatory proteins 21
1.2.6.2. miRNA and lncRNA dysregulation 22
1.3. Somatostatin/Cortistatin/Ghrelin Axis 23
1.3.1. SST and CORT system 23
1.3.1.1. SST/CORT/ssts system in cancer 24
1.3.2. Ghrelin system 25
1.3.2.1. Ghrelin system in cancer 28

**2. Objectives** 30

**3. Materials & Methods** 32

3.1. Human samples 33
3.2. Molecular biology: Nucleic acids 34
3.2.1. Nucleic acid extraction 34
3.2.1.1. Genomic DNA and total RNA from human samples 34
3.2.1.2. miRNAs from human samples 34
3.2.1.3. Total RNA from cell lines 35
3.2.2. Total RNA and miRNA retrotranscription 35
3.2.3. PCR and qPCR 36
3.2.4. Plasmid vectors 38
3.2.4.1. Cloning sequences 38
3.2.4.2. Restriction enzymes digestion 38
3.2.4.3. Plasmids and ghrelin sequences ligation 38
3.2.4.4. Competent bacteria transformation 38

3.2.4.5. Plasmid DNA amplification and purification 39
3.3. Molecular biology: Proteins 39
3.3.1. Peptides 39
3.3.2. Protein extraction 40
3.3.2.1. Tissue 40
3.3.2.2. Cell lines 40
3.3.3. Western Blot 40
3.3.4. Immunohistochemistry 41
3.4. Cellular biology 41
3.4.1. Cell lines 41
3.4.1.1. MDA-MB-231 41
3.4.1.2. MCF-7 42
3.4.2. Defreeze/Freeze and maintenance of cells in culture 42
3.4.3. Plasmid, miRNA and siRNA transfection 42
3.4.4. Generation of monoclonal stably transfected cells lines 43
3.4.5. Peptide treatments 43
3.4.6. Gene Expression microarray 43
3.4.7. ELISA measurements 44
3.5. Functional assays 44
3.5.1. Proliferation studies 44
3.5.2. Migration assays 45
3.5.3. Cell plasticity studies 45
3.5.4. Mammospheres generation assays 46
3.6. Xenografts studies 46
3.7. *In silico* studies 47
3.8. Statistical analysis 47
**4. Results** 49
4.1. sst5TMD4 50
4.1.1. Genesis and regulation of sst5TMD4 50
4.1.1.1. Single nucleotic polymorphisms (SNPs) or new mutation affecting splicing process 50
4.1.1.2. Splicing factors involved in splicing process 54
4.1.1.3. Regulation of sst5TMD4 transcription and splicing by miRNA 57
4.1.1.3.1. Effect of hsa-miR-346 on sst5TMD4 expression 58
4.1.2. Functional role of sst5TMD4 in breast cancer 60
4.1.2.1. Changes in gene expression patterns induced by sst5TMD4 overexpression 60
4.1.2.1.1. sst5TMD4 increased proangiogenic factors production in MCF-7 cells 62
4.1.2.1.2. sst5TMD4 increased *in vivo* VEGF production and angiogenic features 63
4.1.2.1.3. sst5TMD4 correlated with poor clinical outcome in breast cancer patients 64
4.1.2.2. Effects of sst5TMD4 on cellular dedifferentiation in breast cancer cells 67
4.2. In1-Ghrelin 68

4.2.1. Genesis and regulation of In1-ghrelin splicing variant 68
4.2.1.1. Genomic changes associated to the regulation of In1-ghrelin splicing process 69
4.2.1.2. Splicing factors involved in In1-ghrelin splicing regulation 69
4.2.1.3. Regulation of In1-ghrelin transcription by GHRLOS 72
4.2.2. Role of In1-ghrelin functional in breast cancer 74
4.2.2.1. Effects of In1-ghrelin in proliferation rate 76
4.2.2.2. Effects of In1-ghrelin in migration capacity 77
4.2.2.3. Effects of In1-ghrelin downregulation on proliferation rate and migration capacity 79
4.2.2.4. Effects of In1-ghrelin on cell dedifferentiation 80
4.2.2.5. Signaling pathways altered by In1-ghrelin overexpression 82
4.2.2.6. In1-ghrelin correlated with poor clinical outcome in breast cancer patients 83

**5. Discussion 85**

5.1. sst5TMD4 87
5.1.1. SST5 splicing regulation 87
5.1.2. sst5TMD4 functional role in cancer 91
5.2. In1-ghrelin 95
5.2.1. Ghrelin gene splicing regulation 95
5.2.2. Role of In1-ghrelin in breast cancer malignancy 98

**6. Conclusions 103**

**7. Bibliography 107**

# Resumen

# Resumen

El cáncer, en sus múltiples formas, es uno de los mayores retos a los que se enfrenta la sociedad actual. Una de las mayores limitaciones en la investigación del cáncer es la gran heterogeneidad existente entre las diferentes patologías tumorales e incluso entre pacientes con un mismo tipo de cáncer. A pesar de esto, estudios recientes han definido un conjunto de características comunes compartidas por los diferentes tipos de cáncer entre los que se encuentran la alteración del splicing alternativo. En concreto, el splicing alternativo es un mecanismo molecular por el que los organismos eucariotas pueden aumentar exponencialmente la cantidad de transcritos diferentes partiendo de un mismo genoma, a través de la reorganización de los diferentes elementos (exones e intrones) que componen los genes.

Entre las diferentes patologías tumorales, llama especialmente la atención el cáncer de mama, ya que constituye uno de los tipos de cáncer más importante en términos de incidencia tumoral pero también en términos de mortalidad. Este tipo de cáncer también se caracteriza por una desregulación de los procesos de splicing alternativo y, por lo tanto, por un perfil alterado de ciertas variantes de splicing. Así, nuestro grupo de investigación ha identificado la presencia de determinadas variantes de splicing de los ejes neuroendocrinos constituidos por las hormonas somatostatina (SST), cortistatina (CORT) y ghrelina y sus receptores (ssts y GHSRs), especialmente el receptor truncado sst5TMD4 y la variantes de splicing In1-ghrelina, en este tipo de patologías.

En concreto, la variante de splicing del receptor 5 de SST denominada sst5TMD4 codifica un receptor truncado de 4 dominios transmembrana (TMDs) que está sobreexpresado en diversos tipos tumorales (tumores hipofisarios y neuroendocrinos, así como en cáncer de tiroides y de mama), mientras que su expresión en tejidos sanos es muy reducida o nula. Además, la expresión del sst5TMD4 en estas patologías se asocia con una mayor malignidad tumoral, con una menor respuesta al tratamiento con análogos de SST y con un peor pronóstico clínico. Por otro lado, la In1-ghrelina, descubierta recientemente por nuestro grupo y que se genera gracias a un proceso de retención intrónica, presenta importantes implicaciones patológicas en tumores hipofisarios, neuroendocrinos y de mama donde se ha encontrado sobreexpresada y asociada a procesos de malignización tumoral.

Sin embargo, los mecanismos moleculares implicados en la regulación de la expresión de sst5TMD4 e In1-ghrelina, así como el papel preciso y las implicaciones clínicas de estas variantes de splicing en cáncer de mama no se han explorado aún con suficiente detalle. Por este motivo, el **objetivo principal** de esta Tesis Doctoral era profundizar en el conocimiento de los sistemas y/o mecanismos de regulación que determinan la expresión diferencial de las variantes de splicing sst5TMD4 e In1-ghrelina, así como el papel y las implicaciones clínicas que ambas variantes juegan en la fisiopatología del cáncer de mama. Para alcanzar este objetivo general se planteó el estudio de estas variantes de splicing desde una perspectiva multidisciplinar que incluyera estudios

sobre muestras humanas, modelos preclínicos de ratón, líneas celulares y aproximaciones moleculares *in vitro* e *in silico*.

De esta manera, los estudios realizados sobre el receptor truncado sst5TMD4 demostraron que la modulación de su expresión está bajo el control preciso de un complejo e intrincado sistema de elementos reguladores entre los que se incluyen la presencia de polimorfismos de nucleótido único (SNPs), la acción de determinados factores de splicing o la interacción con ciertos miRNAs. En concreto, a través del estudio de la variabilidad poblacional en la secuencia genómica del gen SST5, observamos la presencia de dos SNPs en el intrón críptico eliminado Duránte la generación del sst5TMD4, los cuales mostraron claras diferencias en cuanto a las frecuencias alélicas entre muestras tumorales y no tumorales y entre aquellas con alta y baja expresión del receptor truncado sst5TMD4 entre las muestras tumorales. Además, estudios *in silico* sobre la presencia de dianas para factores de splicing dentro de la secuencia del gen SST5 demostraron la existencia de una alta densidad de secuencias para factores inhibidores del splicing en el intrón críptico en comparación con el resto de la secuencia génica, lo que sugiere un papel relevante de estos factores en la modulación del splicing de este intrón. Por último, estudios *in silico* sugirieron el posible papel de ciertos miRNAs en la regulación de la expresión del sst5TMD4. Entre ellos, el hsa-miR-346 mostró una correlación negativa con los niveles de expresión del sst5TMD4 en una batería de muestras de mama. Además, estudios *in vitro* mostraron que el hsa-miR-346 es capaz de modular la expresión de sst5TMD4, sugiriendo un papel central de este miRNA en la regulación de la expresión del sst5TMD4.

Desde el punto de vista funcional, realizamos un array de expresión con el objetivo de identificar los mecanismos moleculares subyacentes al aumento de la malignidad tumoral asociada a la expresión del sst5TMD4 que demostró que la sobreexpresión de este receptor truncado está asociada a una fuerte desregulación de multitud de genes relacionados con el proceso de angiogénesis. Estudios más profundos demostraron que el receptor sst5TMD4 es capaz de aumentar la expresión de factores pro-angiogénicos como VEGF, EGF o Angiopoietina-1 en modelos *in vitro* e *in vivo* (modelos preclínicos de ratón), en los que además se correlacionó con una mayor vascularización tumoral. Además, la presencia del sst5TMD4 incrementó la desdiferenciación celular y la formación de mamosferas *in vitro*. Más aún, la expresión de sst5TMD4 se correlacionó con marcadores de angiogénesis en muestras humanas de cáncer de mama y con mayor incidencia de metástasis. Por último, la alta expresión del receptor sst5TMD4 en muestras de cáncer de mama se asoció a menor supervivencia libre de enfermedad, lo que refuerza la idea de una importante asociación entre el sst5TMD4 y un peor pronóstico en cáncer de mama.

Por otro lado, con la finalidad de profundizar en el conocimiento de los mecanismos moleculares que regulan la expresión de la In1-ghrelina, analizamos la variabilidad genómica del gen de la ghrelina, centrándonos en la región del intrón retenido para la generación de la In1-ghrelina. Sorprendentemente, estos estudios no detectaron variaciones entre las diferentes muestras analizadas. Además, estudiamos *in silico* la

presencia de dianas para factores de splicing en la secuencia del gen de la ghrelina, encontrando una proporción equilibrada de secuencias diana para factores de splicing inhibidores y estimuladores, entre los que destacan SRSF5 y hnRNP H1, por situarse en regiones del intrón conservadas entre diferentes especies. Por último, también se exploró el posible papel de los lncRNAs codificados por el gen antisentido de la ghrelina, GHRLOS. A través de este estudio se determinó que tres de las variantes de splicing del gen GHRLOS presentan patrones de expresión comparables a los de In1-ghrelina, pero no a los de ghrelina, lo que sugiere un posible efecto regulador específico de estos lncRNAs sobre la expresión de In1-ghrelina.

Con el motivo de estudiar el papel de la In1-ghrelina en la malignidad tumoral en cáncer de mama y los mecanismos celulares asociados, se estudiaron *in vitro* diferentes características tumorales en relación con la presencia de In1-ghrelina, demostrando que esta variante de splicing, pero no la ghrelina nativa, aumenta la capacidad de proliferación y migración, probablemente a través de un aumento en la señalización a través de MEK/ERK. Estos estudios fueron confirmados por medio de ensayos de reducción de la expresión endógena de In1-ghrelina que produjeron una disminución en ambas capacidades. Adicionalmente, comprobamos que la sobreexpresión de In1-ghrelina, pero no de ghrelina, promueve un estado de mayor desdiferenciación representado por un aumento de la plasticidad celular y de la capacidad de formar mamosferas, y por tanto, del porcentaje de células madre tumorales (CSCs). Estos cambios parecen deberse, en parte, a la inducción de las rutas de señalización Jag1/Notch y Wnt/β-catenina. Además, estas capacidades funcionales observadas *in vitro* son probablemente la base de las correlaciones clínicas determinadas en muestras de cáncer de mama. Específicamente, hemos comprobado que una mayor expresión de In1-ghrelina correlaciona con un aumento de la aparición de metástasis en nódulos linfáticos así como con la disminución de la esperanza de vida libre de enfermedad de pacientes de cáncer de mama.

Por todo ello, los resultados presentados en esta Tesis Doctoral sobre la regulación de la expresión y el papel patológico de las variantes de splicing sst5TMD4 e In1-ghrelina en el contexto del cáncer de mama, refuerzan la relevancia de la desregulación del proceso de splicing alternativo en el desarrollo y progresión tumoral. Específicamente, nuestros hallazgos demuestran que ambas variantes de splicing inducen importantes alteraciones en esta patología tumoral y podrían representar potenciales dianas terapéuticas relevantes para el cáncer de mama.

# Summary

# Summary

Cancer and tumoral pathologies represent one of the main and more complex public health problems for the human population worldwide. One of the most relevant limitations in cancer research is the extraordinary heterogeneity among the different tumor types, and even among patients with the same type of cancer. However, recent studies have defined a group of hallmarks shared by all cancer types, which includes alteration of alternative splicing. Specifically, alternative splicing is a molecular mechanism that allows eukaryotic organisms to exponentially increase the number of different transcripts generated from the same genome, through reassembly of the different elements (exons and introns) comprising the genes.

Among the different tumor pathologies, breast cancers are of special relevance, for they represent one of the most important groups in terms of tumor incidence, but also in terms of mortality rate. These cancers are also characterized by dysregulated alternative splicing processes and, therefore, exhibit aberrant expression of certain splice variants. Indeed, our research group has identified the presence of certain splice variants from the neuroendocrine axes comprised by the hormones somatostatin (SST), cortistatin (CORT) and ghrelin and their receptors (ssts and GHSR), specially the truncated receptor sst5TMD4 and the splicing variant In1-ghrelin, in these tumoral pathologies.

Specifically, the splicing variants of the SST receptor type 5 named sst5TMD4 encodes a truncated receptor with four transmembrane domains (TMDs) that is overexpressed in various tumoral types (pituitary and neuroendocrine tumors, as well as thyroid and breast cancers); while its expression in normal tissues is virtually absent or negligible. Furthermore, the expression of sst5TMD4 in these pathologies correlates with exacerbated tumor malignancy, with lower response to treatment with SST analogues and with poor prognosis of the patients. In the case of the ghrelin system, several splicing variants have been described, including In1-ghrelin, a splicing variant recently discovered by our group that arises from a process of intron retention and exhibits important pathological implications in pituitary and neuroendocrine tumors, as well as in breast cancer, where it has been found to be overexpressed and associated with tumor malignancy.

Nevertheless, the molecular mechanisms involved in the regulation of the expression of both, sst5TMD4 and In1-ghrelin, as well as the precise role and clinical implications of these splice variants in breast cancer have not yet been completely unveiled. Therefore, the **main objective** of this Thesis was to expand our knowledge in the regulatory systems and/or mechanisms that determine the expression of the sst5TMD4 and In1-ghrelin splicing variants, as well as the functional role and clinical implications of both variants in the pathophysiology of breast cancer. To achieve this main objective we deployed a multidisciplinary strategy that includes studies on human samples, preclinical mouse models, cell lines and molecular *in vitro* and *in silico* approaches.

The studies performed on the truncated receptor sst5TMD4 demonstrated that the regulation of its expression is under the precise control of a complex and intricate system of regulatory elements, including the presence of single nucleotide polymorphisms (SNPs), the activity of various splicing factors or the interaction with certain miRNAs. Specifically, through the study of population variability in the genomic sequence of SST5 gene, we detected the presence of two SNPs within the cryptic intron removed during the generation of sst5TMD4, which showed clear differences in allele frequencies among tumoral and non tumoral samples and within high and low expression of the truncated receptor sst5TMD4 in tumoral samples. Furthermore, *in silico* studies on the presence of targets for splicing factors within the sequence of the SST5 gene demonstrated the existence of a high density of splicing silencers sequences within the cryptic intron compared to other gene sequence, suggesting a role of these silencing splicing factors in the modulation of the splicing of this intron. Finally, *in silico* studies suggested the putative role of certain miRNAs in regulating the expression of sst5TMD4. Among them, the hsa-miR-346 showed a negative correlation with the expression levels of sst5TMD4 in a battery of breast samples. Furthermore, *in vitro* studies showed that hsa-miR-346 could modulate sst5TMD4 expression, suggesting a key role of this miRNA in the regulation of sst5TMD4 expression.

From a functional point of view, a gene expression array was implemented in order to identify the molecular mechanisms underlying the increased tumor malignancy associated with the expression of sst5TMD4, which revealed that overexpression of this truncated receptor is associated with a strong dysregulation of several genes involved in the angiogenic process. Further studies showed that sst5TMD4 was able to increase the expression of pro-angiogenic factors such as VEGF, EGF or Angiopoietin-1 in *in vitro* and *in vivo* models (preclinical mouse models), wherein it also correlated with increased tumor vascularization. Additionally, the presence of sst5TMD4 increased cell dedifferentiation and the proportion of cancer stem cells *in vitro*. Furthermore, sst5TMD4 expression correlated with angiogenesis markers in human breast cancer samples and was associated to increased incidence of lymph node and distant metastases. Finally, the high expression of sst5TMD4 receptor in samples of breast cancer was associated with lower disease-free survival, which reinforces the idea of a significant association between sst5TMD4 and poor prognosis in breast cancer.

Similarly, in order to further understand the molecular mechanisms that regulate the expression of In1-ghrelin, we analyzed the genomic variability of the ghrelin gene, focusing on the retained intron. Surprisingly, these studies did not show variations among the different samples analyzed. Furthermore, *in silico* studies on the presence of target sites for splicing factors in the ghrelin gene sequence revealed a balanced ratio of enhancer and silencer splicing sequences, including binding sites for SRSF5 and hnRNP H1, which could have special relevance, since these sites are located in regions of the intron 1 conserved among species. Finally, the putative role of the lncRNAs encoded by the antisense gene of ghrelin, GHRLOS was explored. This study demonstrated that three of the splicing variants of the GHRLOS gene presented similar

expression patterns than In1-ghrelin, but not ghrelin, suggesting a possible specific regulatory role of these lncRNAs on the expression of In1-ghrelin.

With the purpose of studying the role of In1-ghrelin in tumor malignancy and the associated cellular mechanisms, the relationship of In1-ghrelin expression with different tumoral features was explored *in vitro.* These studies showed that this splicing variant, but not native ghrelin, increased proliferation and migration capacity, probably through the increase of the MEK/ERK signaling. This was further confirmed by silencing the endogenous expression of In1-ghrelin, which caused a reduction in both capacities. Additionally, we found that overexpression of In1-ghrelin, but not ghrelin, promoted a greater dedifferentiated cellular state represented by an increase in cellular plasticity and in the ability to form mammospheres, and therefore the proportion of cancer stem cells (CSCs). These changes could be, at least in part, due to the induction of the Jag1/Notch and Wnt/β-catenin signaling pathways. Interestingly, these functional capabilities observed *in vitro* are probably the basis of specific clinical correlations observed in breast cancer samples. Specifically, we found that increased expression of In1-ghrelin correlated with an increased occurrence of lymph node metastases and the decrease in disease-free survival of breast cancer patients.

Altogether, the results presented in this Thesis on the regulation of the expression and the pathological role of the splicing variants sst5TMD4 and In1-ghrelin in the context of breast cancer reinforce the relevance of the dysregulation of alternative splicing process in the tumor development and progression. Specifically, our findings demonstrate that both splicing variants determine important changes in tumor pathology and could provide relevant potential therapeutic targets in breast cancer.

# Introduction

# 1. Introduction

Cancer represents one of the most severe and complex health threats for the human population to date, in spite of the great research and clinical efforts deployed over the last decades to fight this pathology, and the advances achieved thereby [1, 2]. The development and progression of cancer is a highly heterogeneous and variable process, strongly influenced by genetics, but also by metabolic, nutritional, ambient and life style factors [3]. Despite this complexity, most cancers share a group of common **hallmarks**, such as sustained proliferative signaling, evasion of growth suppressors, resistance to cell death, angiogenesis, activation of invasion and metastasis [4, 5], or altered alternative splicing processes [6]. In this scenario, aberrant splicing is gaining an unexpected relevance, in that recent studies point out that tumoral heterogeneity in outcome and cancer survival can be explained, at least in part, by genetic variations (such as splicing variants) present in the primary tumor [7].

Among the different types of tumor pathologies, this Thesis will be focused on breast cancer, a major cancer type in terms of incidence (i.e. the most common cancer type in women) but also in terms of mortality rate, as it represents the second leading cause of cancer-associated deaths in women [1]. In addition, breast cancer represents a classical paradigm as it displays frequent intra- and inter-tumor heterogeneity as the result of genetic and non-genetic alterations [8, 9].

In this context, during the last years, our group has been interested in exploring the role of several endocrine systems [particularly somatostatin (SST) and ghrelin systems] in the development and progression of different endocrine-related tumors, as well as in determining the suitability of certain members of these systems as novel biomarkers for the diagnosis, prognosis and/or putative therapeutic treatment of those endocrine-related tumors [10-19]. Indeed, our group has identified novel splicing variants of SST and ghrelin systems (specially, the truncated receptor sst5TMD4 and the splicing variant In1-ghrelin), which are overexpressed in different tumoral pathologies (pituitary, thyroid and neuroendocrine tumors), wherein they are associated with malignant phenotypes [10-19]. However, the molecular mechanisms implicated in their generation and their precise role and clinical implications in breast cancer are still to be fully elucidated.

## 1.1. Cancer

Current estimates indicate that tumoral pathologies and cancer still represent one of the leading and more serious public health problems for the human population worldwide [1, 2]. Indeed, 25% of deaths in developed countries are associated to tumoral pathologies [1] and, particularly, in Spain, there are more than 100.000 cancer-related deaths every year (source: AECC). Even more worrying is the fact that cancer is predicted to overtake heart disease as the leading cause of death across all age groups by 2030, translating to a 45% increase in the number of cancer diagnoses in the next 15 years (source: American Society of Clinical Oncology).

Cancer is a multifactorial, multistep, and complex disease that arises as a result of perturbed cellular homeostasis. In fact, cancer can affect almost every cell type in the body, irrespective of its origin, localization or metabolic status. Consequently, cancer (or tumoral pathologies) encompasses a wide variety of malignancies with a variable etiology and pathology. Consequently, the extraordinary variability, heterogeneity and complexity of cancer hamper the finding of common molecular elements, which could facilitate the development of more general and effective diagnostic and therapeutic strategies [9]. This tumoral heterogeneity can be related to the diverse etiology of these tumors (mutations, genetic alterations, splicing alterations) but can also be associated to the specific milieu in which the tumor develops and progresses [8]. In this regard, the notion that tumor development and progression is profoundly conditioned by metabolic-endocrine dysregulations, is a re-emerging concept especially relevant in the so-called endocrine-related cancers [20].

### 1.1.1. Endocrine-related cancers

The terms endocrine-related and hormone-dependent cancers classically include a group of sex steroid responsive cancers, such as cancers of the breast, endometrium, prostate, and testis, but also other cancers such as thyroid and ovary cancers that are responsive to pituitary hormones [20]. However, evidence gathered during the last years regarding the tumorigenic potential of additional endocrine systems has broadened this concept and, nowadays, most cancers that exhibit certain “hormone sensitivity”, at least at some stages of their development and/or progression, are considered as endocrine-related or hormone-dependent cancers [21]. Actually, the overt, often ectopic presence of components of several endocrine systems (neuropeptides, peptide hormones and/or their receptors) in tumoral pathologies is not rare, although their precise role in cancer is still poorly defined, and their therapeutic potential has been poorly explored hitherto, as compared to that of growth factors and chemokines.

Of special interest for this Thesis is **breast cancer,** the most frequent malignant tumor and the second leading cause of cancer-related death in female population worldwide

[22]. This cancer is an extremely complex and heterogeneous type of tumor, in which several clinical-pathological features are used for its diagnosis and prognosis, as well as for selecting the most appropriate therapy, including histological grade, lymph node status, hormone receptor status, and human epidermal growth factor receptor type 2 (HER2) status. Some of these factors have been associated with the survival rate of patients and their clinical outcome after treatment and are therefore considered predictive of prognosis and response to treatment. However, it has also been reported that some patients, despite bearing a similar combination of breast cancer features, may display quite distinct clinical outcomes. Thus, the role of these factors in determining diagnosis and prognosis and in predicting therapeutic outcomes in breast cancer remains limited.

In an effort to systematically and conceptually apprehend the extraordinary complexity and diversity of cancers and of their accompanying pathological alterations, many analysis have been proposed and implemented during the last years, aimed at establishing a common conceptual framework for their study. In this context, Hanahan and Weinberg proposed, in two landmark articles [4, 5], that most cancers share a group of common **"cancer hallmarks",** such as sustained proliferative signaling, evasion of growth suppressors, resistance to cell death, replicative immortality, angiogenesis, activation of invasion and metastasis, genome instability, inflammation, altered energy metabolism and evasion to immune destruction [4, 5]. This conceptual advance in the study of cancer has implied a great progress in that it is extraordinarily useful – from an experimental and therapeutic point of view – to consider the existence of common processes shared by all cancer types.

In this same scenario, recent studies have proposed the existence of additional common cancer hallmarks, shared by all tumor types, as it is the case of altered alternative splicing processes [6], which could significantly compromise the function of a high number of genes associated to these pathologies [6]. Indeed, a growing body of evidence indicates the existence of an association between the presence of aberrant alternative mRNA isoforms and the development and/or progression of different cancer types [7], and recent studies point out that intratumoral heterogeneity in outcome and cancer survival can be explained, at least in part, by genetic variations (such as splicing variants) present in the primary tumor [7].

## 1.2. Splicing

The human genome sequence has been estimated to be composed of approximately 27.000 protein-coding genes [23], an estimation that it is not far from those of simpler organisms such as *C. elegans*, which seems to be composed of 20.000 protein-coding genes [24]. Certainly, it is hard to conceive that the huge complexity of an organism like a human being can be structured with a similar number of genes than a simple worm, which is composed by only 1.031 cells [25]. However, there are a number of

additional concurrent processes capable to increase extraordinarily the complexity of gene-derived products such as the post-translational processing [26], the genomic rearrangement [27] or the **splicing process**, by which one gene can generate several different mRNAs through transcripts reassembly [28]. All these processes drastically increase protein diversity and could help to explain how this relatively small set of genes (as compared to simpler organisms) can support the complex development and daily maintaining of the entire organism observed in mammals and, specifically, in humans.

The vast majority of eukaryotic genes are composed by two distinct elements named exons and introns. The **exons** represent protein-coding sequences, scattered throughout the gene and flanked by noncoding-protein sequences, called **introns**. During the transcription of a given gene, both, exons and introns, are transcribed in the nascent mRNA transcript and, co-transcriptionally, during the process of messenger RNA (mRNA) maturation, the introns are selectively removed from the mature mRNA.

Actually, mRNA maturation is comprised by three basic steps: capping, addition of a poly-A tail, and splicing, processes that are strongly interrelated and occur concomitantly [29]. Among them, splicing is the process responsible for removing the introns from the nascent transcripts and joining the remaining exons together, ensuring the correct cutting and assembling. Splicing process occurs in the vast majority of mammalian genes, for it is calculated that only a 3% of human genes do not present introns in their sequence [30].

### 1.2.1. Splicing process

The splicing process of a pre-mRNA is a complex mechanism in which many different elements are involved. In order to fulfill the splicing of a nascent transcript, a ribonucleoproteic complex named ***spliceosome*** must be organized [31]. Specifically, there are two types of spliceosome complexes, major and minor, which share similar mechanisms of action but act on different type of introns [32]. The **major spliceosome** is the molecular machinery that catalyzes the splicing process of almost 99% of the nascent mRNA. This complex is primarily composed by five small nucleolar ribonucleoproteins (snRNP): U1, U2, U4, U5 and U6, which comprise the functional core. This spliceosome core can be accompanied by more than 300 different, additional proteins (***splicing factors and regulatory proteins***), which are involved in the fine regulation of the process, in accordance with the cellular environment [33].

The splicing process is encompassed by two consecutive steps of trans-esterification in each terminal region of the intron sequence, wherein the precise identification of the splicing sites is directed by canonical target sequences (**Figure I1**). Specifically, the 5’ splice site located at the 5’ end of the intron sequence is commonly comprised, in

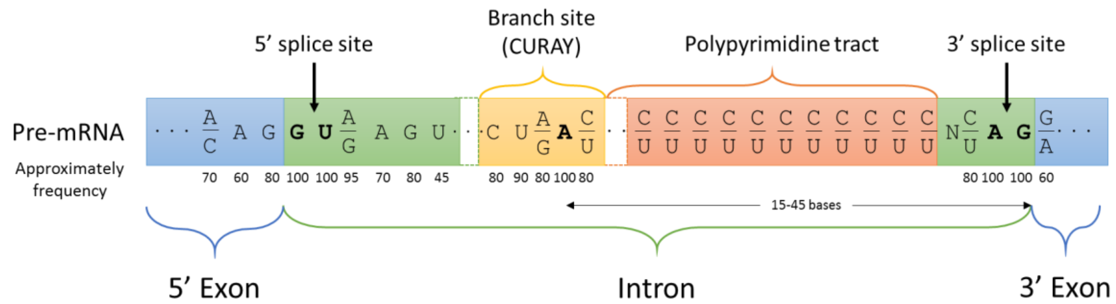


***Figure I1**: Representative scheme of canonical splicing consensus sequences within the intron.*

mammals, by two bases (GU). In the opposite side of the intron, the 3' splice site dinucleotide AG marks the end of the intron and the second site of trans-esterification. In addition, within the intron sequence, there are two more canonical splicing sequences - the branch site (CURAY sequence) and the polypyrimidine tract -, which are crucial for the appropriate folding of the intron during the splicing process [34]. Although these sequences, and especially those located in the extremes of the intron, can present some variation among different genes [35], nearly all introns belong to the so-called U2-type, which are spliced by the major spliceosome and are flanked by GT–AG splice site dinucleotides. The most frequent exception to this rule are the U2-type GC–AG splice sites, comprising ~0.9% of human splice sites [36].

On the other hand, the **minor spliceosome** is responsible for the processing of the remaining 1% of the introns (U12 introns), which are processed by a similar mechanism but recognizing and binding to different target sequences. The minor spliceosome shares the U5 RNP with the major spliceosome, but presents functional analogs to the other 4 snRNP, called U11, U12, U4atac and U6atac (functional analogs of U1, U2, U4 and U6, respectively) [37]. Specifically, although U12-introns were first described to have AT–AC dinucleotides at the intron/exon boundaries, the vast majority of them contain GT–AG sites [37]. Indeed, the AT–AC sites comprise only approximately 0.09% of the splice sites [36].

In addition to the well-known U2 and U12 recognition sites, introns with non-canonical splice sites (that is, with GC-AG, GG-AG, GT-TG, GT-CG or CT-AG dinucleotides at the intron/exon boundaries) have also been reported to be efficiently removed [36]. These reported non-canonical splice sites have U2/U12-like splice site consensus sequences (U2/U12-like non-canonical splice sites) and generate unexpected, and even as yet unpredictable, splicing variants with pathophysiological relevance [38-42].

Irrespective of the target sequences and the spliceosomal machinery involved, the splicing process initiates with the simultaneous binding of certain proteins to the splicing sequences (**Figure I2**). In the case of major spliceosome-regulated introns, the splicing process implies U1 binding to the nascent transcript at the 5' splice site, the binding of accessory proteins SF1 to the branch point, the binding of U2AF2 to the

polypyrimidine tract and the binding of U2AF1 to the 3' splice site, constituting a structure called **Complex E**. Then, U2 attaches the branch point displacing SF1 and modifies the RNA conformation generating the **Complex A** (pre-spliceosome). This new structure allows the binding of the trimeric element conformed by U4, U5 and U6 (U4/U6.U5 tri-snRNP). Subsequently, U5 binds to the upstream exon and to U6, which is formerly attached to U2. This structure is called **complex B1** (precatalytic spliceosome), and will induce the release of U1, the attachment of U6 to the 5' splice site and the translocation of U5 from the exon to the intron. This new organization is named **Complex B2**. At this point, U4 is released from the spliceosome, forming the **Complex C** (catalytic spliceosome), which induces a transesterification in the 5' splice site cleaving the intron from the upstream exon and promoting a ligation between the intron 5' end and an adenine that is located at the start of the polypyrimidine tract. As a result, a structure known as ***lariat*** is formed. U2, U5 and U6 remain bound to the *lariat* while 3' splice site is excised and both exons are assembled. Finally, the intron is degraded and the spliceosome components are recycled [32].

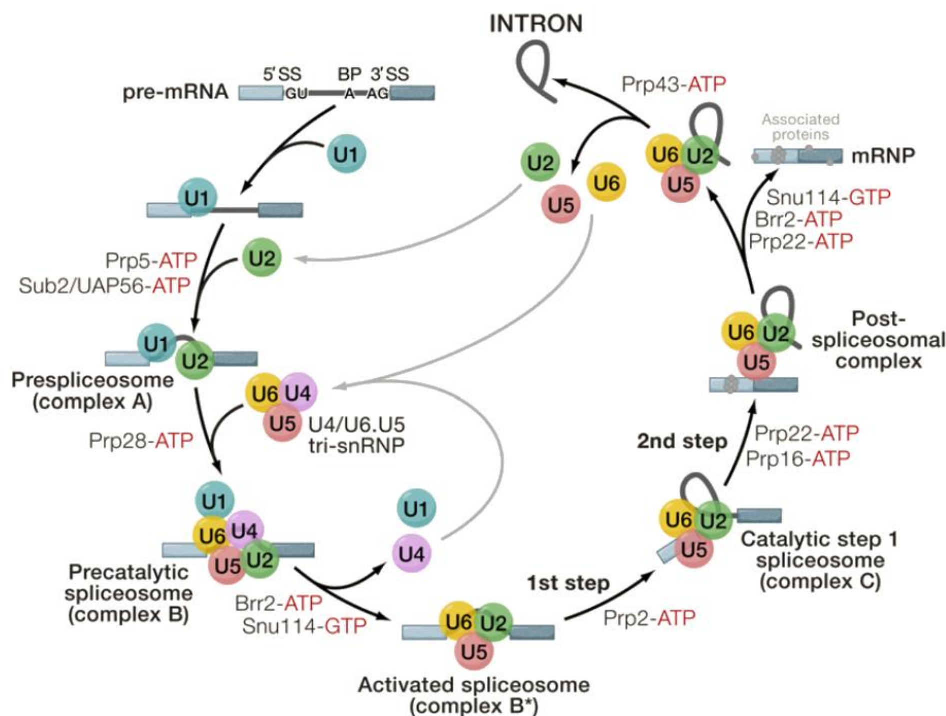


***Figure I2:*** *Representative scheme of the splicing process, adapted from [43]*

### 1.2.2. Alternative Splicing

**Alternative splicing** is the process by which several, different mRNAs can be generated from one gene through a series of rearrangements of its exons and introns. Thence, the components of a gene can be assembled in different combinations to generate several mRNA variants from the same gene. The number of splicing variants that can be generated by this process can vary remarkably, from genes that do not codify for more than one mRNA to genes like Dscam, a *D. melanogaster* gene, that underpin the

record of splicing variants codified by the same gene with 38.016 isoforms [44], which, somewhat curiously, represent more mRNA variants than genes exist in *D. melanogaster*. More than 90% of the human genes are reported to undergo alternative splicing processes, generating an estimated amount of 100.000 alternative splicing events [45, 46].

Alternative splicing processes can be classified in four main groups, depending on the sequences involved and the results it yields. Specifically, **exon skipping** is the process by which an exon is spliced out from the final transcript together with the flanking introns. This is a common alternative splicing event in higher eukaryotes, but extremely rare in lower eukaryotes [33]. Secondly, **alternative 3' and 5' splice site** represent two additional types of alternative splicing. They occur when two or more splice sites are recognized at one end of an exon, generating splicing variants with altered, incomplete exons [47]. An additional type of alternative splicing is the **intron retention**, a relatively common event in plants but less frequent in vertebrates, by which small introns are not spliced out from the transcript and, therefore, are retained in the mature mRNA. Additionally, there are other less frequent complex events of alternative splicing that give rise to alternative transcript variants including mutually exclusive exons, alternative promoter usage and alternative polyadenylation [33]. Another rare form of alternative splicing involves reactions between two primary transcripts in trans [33].

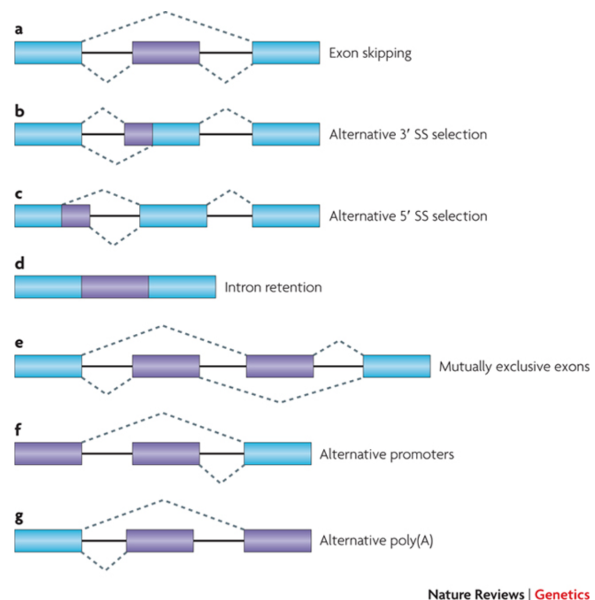


***Figure I3**: Alternative splicing processes (adapted from [33])*

### 1.2.3. Splicing regulation

The regulation of alternative splicing is a tightly controlled process in which an elevated number of **splicing factors** and regulatory proteins are involved. These splicing factors are RNA-binding proteins that modulate the splicing process interacting with specific RNA sequences or motifs [48]. Indeed, splicing factors are versatile modulator of splicing process that can bind to degenerated sequence motifs in the nascent transcript [49]. To date, over 71 different human splicing factors have been described [50], classified in **enhancers** and **silencers** of the splicing process, wherein some of them are able to induce both actions depending the specific sequence that bind. The target motifs of splicing factors are classified as:

- Exonic or intronic splicing enhancers (ESE/ISE): sequences localized at the introns (ISE) or exons (ESE) that are used to promote the splicing of a given intron or exon [51, 52].
- Exonic or intronic splicing suppressors (ESS/ISS): sequences involved in the inhibition of the splicing of introns or exons by blocking not only spliceosome components, but also enhancer proteins [53].

Splicing factors can be classified in two families, serine-arginine proteins (SR-proteins) and heterogeneous nuclear ribonucleoproteins (hnRNPs) [54]. **SR proteins** are usually enhancers of intron and exon splicing, acting through the recruitment of the spliceosome components [48, 55]. Many SR proteins are involved in other processes of gene regulation including RNA maturation, decay, transport and translation [56]. SR proteins and components of the spliceosome core can be recruited onto the nascent transcript by RNA Polymerase II [57]. On the other hand, heterogeneous nuclear ribonucleoproteins (**hnRNP**) are a family of splicing factors also involved in mRNA trafficking, stability and translation [58]. hnRNPs present one or more RNA-binding domains and a domain for protein-protein interaction. hnRNPs usually bind splicing silencer sequences blocking the splicing process [57]. Indeed, splicing inhibition mediated by hnRNPs can be induced by competing with the SR proteins for binding sites or by interactions with each other altering the structure of the pre-mRNA, and making some regions inaccessible for the spliceosome [48].

In conjunction, **the mature mRNA variant(s) generated from a given gene in a particular cellular environment represents the ultimate consequence from the dynamic interaction among splicing enhancers and silencers capable to bind the regulatory sequences of its introns and exons in order to define their precise assembling**.

However, the expression and stability of alternative spliced mRNAs can be also regulated by other trans-elements, such as microRNAs (**miRNAs**) and long-non-coding RNAs (**lncRNAs**). miRNAs and lncRNAs comprise a family of non-coding RNAs, which have been shown to be able to regulate several cellular processes by modulating the transcription and availability of target coding mRNAs [59-61].

**miRNAs** are short non-coding RNAs (ncRNAs) with 20-24 nucleotides that are usually encoded within the introns of other genes [62-64]. miRNAs are initially generated as long RNAs with a hairpin structure that are subsequently processed by a series of RNase-III enzymes (DROSHA and DICER) to form the mature, functional miRNA duplex [65]. These small RNA sequences recognize target sequences in coding mRNAs regulating its translation in a variety of manners, including translational repression, mRNA cleavage, and deadenylation [66]. However, in some cases, miRNAs have been shown to be able to enhance or promote the translation of certain mRNAs [67]. In this context, miRNAs can selectively target and, hence, regulate the expression of certain

alternative splicing variants [68]. Therefore, the expression pattern of cellular miRNAs represents another level of gene regulation, wherein the translation of the splicing variants could be selectively regulated by the miRNAs landscape.

**lncRNAs** are a large and greatly diverse class of transcribed RNA molecules with a length of more than 200 nucleotides that do not encode proteins. It is estimated that human genome encode for 23.000 lncRNA [69], although only 3.300 have been already identified [70]. A large and relevant group of lncRNAs are those transcribed from the antisense strand of known coding genes, which are also referred as **natural antisense transcripts** (NATs) [71], and can exhibit whole or partial overlapping with the sequence of the coding gene. Antisense-overlapping lncRNAs have a tendency to undergo fewer splicing events and typically show lower abundance than sense transcripts [72]. The basal expression levels of antisense-overlapping lncRNAs and sense mRNAs in different tissues can be either positively or negatively regulated [73, 74]. In general, lncRNAs are involved in epigenetic, transcription and post-transcriptional regulation acting as master regulators of mRNA expression. lncRNAs can alter the splicing process by masking splice sites [75], arresting splicing factors or modulating its phosphorylation [76].

Therefore, the splicing process is an extremely sophisticated mechanism, tightly regulated at multiples levels [expression and functionality of the different splicing factors, alteration of canonical splicing sites, presence and activity of regulatory ncRNAs (miRNAs or lncRNAs)], whose adequate balance is essential to maintain the appropriate cellular homeostasis. Indeed, alterations in the proper splicing process have been extensively associated to the development and/or progression of several types of cancer [6, 77].

#### 1.2.4. Splicing in cancer

As mentioned above, processes of alternative splicing are essential to maintain the appropriate cell physiology and, hence, they have to be exquisitely regulated. For this reason, it is reasonable to predict that the alteration of normal, physiological alternative splicing processes can lead to the development of a number of diverse pathologies [78-83]. In particular, alterations in the splicing process can induce the aberrant expression of certain splicing variants that result in altered proteins, which can act as true oncogenes [84], involved in the development and/or progression of certain tumoral pathologies. In this regard, there is mounting evidence supporting the intriguing relationship between many types of cancer, including breast cancer, and splicing associated dysfunctionalities [54, 83, 85, 86]. In fact, splicing dysregulation is now being proposed as a common hallmark shared by the vast majority of cancers [87]. Interestingly, although the splicing process as a whole is, by and large, downregulated in cancer [88], aberrant alternative splicing variants are common events in tumoral pathologies. Thus, data collected in EST databases have revealed

that a large fraction of the alternative splicing events is associated to the generation of tumor specific variants [88-90]. These splicing variants found to be expressed (or overexpressed) in tumoral pathologies provide, most frequently, an advantage to tumoral cells in order to improve their growth and survival. Indeed, tumor-derived splicing variants have been associated to the promotion of every malignant process from cancer initiation to progression [6, 91], affecting several cancer-associated processes such as dedifferentiation, apoptosis, glucose homeostasis, proliferation, angiogenesis, motility and invasion [92].

Due to the close relationship between the presence of aberrantly altered splicing variants and the development and/or progression of tumoral pathologies, many studies have been conducted to unveil the regulatory processes underlying the generation of tumor-related splicing variants. Results gathered hitherto indicate the existence of several mechanisms implicated in the generation of aberrantly expressed splice variants, which includes from mutations in genomic sequence to alterations in the regulatory mechanisms controlling the splicing process, as described below.

### 1.2.5. Genomic alteration and aberrant splicing in cancer

Large-scale analyses of splicing variants in several cancer pathologies have revealed massive alterations of splicing processes during tumor development or progression [93-95]. Many of these alterations are related to mutations in the genomic sequence, which can represent single-nucleotide polymorphisms (demographically described allelic variability characterized by a substitution of one nucleotide in the genomic sequence) or *de novo* cancer-associated mutations. In any case, the alteration of the genomic sequence can change the appropriate splicing process by modifying the splicing site sequence (named **splice site mutation** - a genetic mutation that inserts, deletes or alters a number of nucleotides in the specific site at which splicing takes place) or altering the sequence of relevant splicing factor target sites [96]. Therefore, any genomic alteration associated to splicing regulatory sequences can result in the aberrant expression of alternative splicing variants [97, 98].

### 1.2.6. Altered regulation of splicing in cancer

Alternatively, the presence of aberrant alternative splicing variants in tumoral cells can be related to the tumor-associated alteration of certain regulatory systems that maintain the appropriate splicing processes, as it is outlined below.

#### 1.2.6.1. Dysregulation of splicing factors

The expression and functionality (e.g. regulation of their activity and subcellular localization) of splicing factors are exquisitely controlled in the cells in order to maintain the appropriate gene expression patterns [99]. Thus, changes in the cellular machinery involved in the regulation of the splicing process can contribute to cancer

development and/or progression, favoring the expression of splicing variants that could contribute to malignancy-related processes such as tumor cell proliferation, migration, or metastasis [92]. In line with this, the expression of several splicing factors has been found to be altered in a number of tumoral pathologies [100]. However, the precise cellular and molecular changes induced by their dysregulation are still to be fully elucidated in the majority of the cases.

1.2.6.1.1. SR proteins in cancer

To date, several SR proteins have been found to be overexpressed in cancer [54], wherein SRSF1, SRSF2, SRSF3, SRSF5, SRSF6 and SRSF10 seem to have particular relevance.

- **SRSF1 (SF2/ASF):** This splicing factor is upregulated in different tumor types, where it can alter the normal alternative splicing pattern by modifying different aspects of cellular behavior. Indeed, SRSF1 has been shown to promote the appearance of splicing variants, showing loss of tumor-suppressor activities or gaining of oncogenic properties [101, 102].
- **SRSF2 (SC35):** Similarly to SRSF1, the appropriate activity of SRSF2 is associated to the maintenance of cell cycle regulation and genomic stability [103]. SRSF2 has been related with the inhibition of tumor suppressor KLF6 [104].
- **SRSF3 (SRp20):** SRSF3 is involved in alternative splicing, but also in different steps of mRNA maturation and export to the cytoplasm [105]. This splicing factor has been correlated with breast cancer tumorigenesis [106] and with the promotion of proliferation and dysregulation of cell cycle [105] through the dysregulation of the alternative splicing of several oncogenic and tumor suppressor genes [107, 108].
- **SRSF5 (SRp40):** It has been found to be overexpressed in breast cancer, wherein it correlates with alternative splicing of oncogenic genes [87, 109]. This SR protein induces the appearance of oncogenic splicing variants of CD44, a receptor implicated in proliferation, cell cycle and cytoskeleton regulation [109], but also in the formation of anti-apoptotic Mcl-s splice variants [110].
- **SRSF6 (SRp55):** Downregulation of this SR protein has been associated to estrogen receptor-mediated progression of breast cancer [111]. SRSF6 has been related with tumoral-induced angiogenic processes and with changes in alternative splicing that lead to accelerated tumor progression [112, 113].
- **SRSF10 (Tra2β):** It has been related to impaired DNA repair and cell cycle dysfunctionalities [114]. It has been found to be overexpressed in invasive breast cancer [85, 115].

#### 1.2.6.1.2. hnRNPs in cancer

As with SR proteins, hnRNPs have been associated to several mechanisms involved in cancer development, including processes of dedifferentiation and cell survival [116].

- **hnRNPA/B family:** This family of splicing silencers has been analyzed in several cancers [58, 87, 117, 118]. Interestingly, their functions are usually opposed to those exerted by SR proteins. In fact, some studies have shown that components of this splicing factor family can act as onco-repressors by preventing dedifferentiation processes, such as epithelial to mesenchymal transition (EMT) [119], or avoiding apoptosis scape of tumor cells [120]. However, these splicing factors can also exhibit oncogenic properties, by inducing the generation of splicing variants with anti-apoptotic activities [87], stabilizing the telomeres [117] or promoting cell proliferation and cell cycle progression [121].
- **PTB (hnRNP I):** PTB (polypyrimidine tract binding protein) is an hnRNP protein involved in numerous stages of RNA processing and translation, which has been found overexpressed in tumoral samples and breast cancer cell lines [122]. PTB usually acts as a splicing inhibitor by blocking spliceosome interactions with the pre-mRNA [122-124]. Similarly to proteins of hnRNPA/B family, PTB can exert both oncogenic [113, 122, 125, 126] and tumor-suppressor activities [127, 128].
- **hnRNP K:** hnRNP K is a splicing factor whose expression is regulated by EGF [129, 130]. In contrast with the other hnRNP splicing factors mentioned above, hnRNP K has only shown oncogenic properties stimulating cell cycle and viability [130- 132].

#### 1.2.6.1.3. Other RNA-binding regulatory proteins

In addition to the classical splicing factors described earlier, other RNA-binding proteins involved in the regulation of alternative splicing processes have been also related with a number of tumoral pathologies.

- **Sam68 (KHDRBS1, Src-associated in mitosis)**: It has been found to be overexpressed in breast cancer, wherein it shows cell growth stimulatory activities by modifying alternative splicing of different proteins related with cell cycle [133] and apoptosis [134]. Moreover, Sam68 promotes the expression of SRSF1 [135].
- **YB-1 (DNA-binding protein B1):** this splicing factor is a member of the Y-box family that has been found overexpressed in breast cancer [136, 137] wherein it increases proliferation [87] through the control of cell cycle [138] and genomic instability in breast cancer [139]. Moreover, it has been associated with resistance to breast cancer therapies [140, 141].

- **FOX2 (RBM9, Fxh):** Fox2 is overexpressed in basal breast cancer cells [142] and its expression has been associated to chemotherapy resistance [143, 144]. Intriguingly, Fox2 can influence a large number of splicing processes, including those involved in EMT [145, 146]. In addition, Fox2 can regulate alternative splicing of several factors including hnRNPs, SR proteins, and itself [147].

- **RBM5 (RNA-binding motif protein 5, LUCA15, H37):** It is another splicing factor with controversial functions. On one hand, RBM5 promotes apoptosis and inhibits cell cycle [148]; but, on the other hand, RBM5 stimulates the expression of anti-apoptotic splice variants [124].

- **RBM10 (RNA-binding motif protein 10, pS1-1, TARPS)**: It is a splicing factor that presents a close homology with RBM5. Indeed, both share the capacity to regulate apoptosis by modifying Fas and BCl-x genes alternative splicing [149]. In addition, RBM10 has been found overexpressed in breast cancer cell lines modulating TNF-a expression levels [150].

- **SPF45 (RBM17):** SPF45 acts as splicing factor by interacting with the spliceosome members SF1, U2AF65, and SF3b155 [54]. It has been found overexpressed in breast cancer tissues, wherein it correlates with tumoral severity and induces multidrug resistance [151, 152].

#### 1.2.6.2. miRNA and lncRNA dysregulation

Tumor-associated splicing alterations can also occur through dysregulations of splicing-regulatory **trans elements**, such as **ncRNAs**, including **miRNAs** and **lncRNAs**.

More than 3000 miRNAs have been described in human, which play important roles in virtually all biological pathways [153]. Therefore, is not surprising that many miRNAs have been reported to be involved in tumoral pathologies, where they can be related, for instance, to proliferation, cell cycle control, apoptosis, differentiation, migration and metabolism [153]. As mentioned above, mature miRNAs, together with the protein DICER can form a regulatory complex that binds to target mRNA sequences regulating the synthesis and/or stability of those mRNAs [68]. In this context, alteration of the expression of different miRNAs, which is usually observed in tumoral pathologies, could lead to the dysregulated expression of splicing variants from specific genes, which, in turn, would increase cell malignancy [154]. Moreover, dysregulations in miRNAs expression can globally modify the splicing expression pattern in tumoral cells by altering the generation of different splicing factors [154].

Additionally, splicing regulation by lncRNA has been also observed in tumor pathologies, wherein NATs can play an especially relevant role by regulating the alternative processing of their antisense genes. Indeed, a number of examples of NATs involved in cancer progression and malignancy have already been described in the bibliography [155].

Therefore, when all the available evidence is considered together, it can be surmised that the ultimate consequence of the dysregulation of the splicing process in cancer pathologies is the aberrant expression of alterative splicing variants, which can be involved in almost every known step encompassing cancer development and progression [156], and, consequently, can affect a number of cancer-associated processes such as dedifferentiation, apoptosis, glucose homeostasis, proliferation, angiogenesis, motility and invasion [156].

## 1.3. **Somatostatin/Cortistatin/Ghrelin Axes**

Hormone-related cancers comprise a heterogeneous and complex group of tumoral malignancies, whose regulation involves components of the endocrine system. During the last years, our group has been particularly interested in exploring the role of the components of two closely related endocrine systems—those comprised by SST, cortistatin (CORT), ghrelin and their receptors and associated proteins [10-19]— in the regulation of several types of hormone/endocrine-related tumors [157-159]. These studies have unveiled the existence of novel, previously unidentified splicing-derived variants of these systems, which exhibit a clear potential to be used in the development of new biomarkers for the diagnosis, prognosis and medical treatment of certain endocrine-related cancers [160-162]. In particular, in the present study, we were interested in exploring the regulation and functional role of two splicing variants of these systems (the **truncated receptor sst5TMD4** and the **splicing variant In1-ghrelin**) using breast cancer models.

### 1.3.1. SST and CORT system

SST and CORT are two neuropeptides that display similar sequences, structures and functionalities, likely due to their presumed shared evolutionary origin from a common antecessor gene [163]. SST was discovered in 1973, and was isolated from the ovine hypothalamus by its ability to inhibit GH secretion [164]; whereas, CORT was discovered more than 20 years later as a SST-related peptide in nervous system of amphibian, rodent and the human [165-167]. Generation of mature SST and CORT peptides comprises similar mechanisms and subsequent processing steps. Indeed, transcription of SST and CORT genes generates a pre-pro-peptide that, by post-translational maturation, yields the final active peptides (SST-14 and SST-28 from pre-pro-SST, and CORT-17 and CORT-29 from pre-pro-CORT; for review, see [168]).

SST is a truly pleiotropic neuropeptide, extensively distributed throughout the organism, with the capacity to modulate a plethora of physiological functions, from inhibition of basal and stimulated secretion from endocrine and exocrine cells, to inhibition of gastrointestinal motility, and modulation of neurotransmission,

metabolism and immune function, as well as inhibition of cell proliferation and differentiation of normal and tumoral cells [160, 164, 169-171]. On the other hand, CORT [165] is mainly produced in the cerebral cortex, where it was originally identified for its involvement in the modulation of sleep cycles, neuronal activity and immune system [165, 172-175]. CORT was initially suggested as an endocrine/metabolic sibling of SST; however, recent evidence has clearly established that CORT is able to trigger unique, and even opposite, endocrine and non-endocrine actions from those exerted by SST, including the regulation of endocrine secretions, the control of immune response or the modulation of neuronal activity [173, 176-184].

SST and CORT receptors

SST and CORT exert most of their actions through binding and activation a family of SST/CORT receptor named **ssts**, which are widely expressed throughout the organisms [170]. To date, five different intronless genes, which encode for distinct 5 receptors (sst1-5), as well as a carboxy-terminal spliced variant of the sst2 in mouse, named sst2B, have been identified and exhibit a comparable subnanomolar binding affinity for SST and CORT [185]. These receptors are currently classified as Class A G-protein coupled receptors (**GPCRs**), and display the typical molecular architecture shared by GPCRs, comprising seven transmembrane domains (**TMDs**), the conserved DRY motif, at the cytoplasmic region of the TMD3 and N-linked glycosilation sites in the N-terminal domain [185]. ssts can be subdivided into two groups, according to their sequence identity and pharmacological properties. Specifically, SST1 group is comprised by sst2, sst3 and sst5; while SST2 group includes sst1 and sst4. Sequences of ssts are highly conserved among species as well as among the sst-subtypes, despite their branched evolutive process, being more divergent in their N- and C-terminal domains [186].

The complexity and versatility of the sst family is considerably increased by the fact that several of these receptors can be simultaneously present in the same cells. In addition, ssts are able to functionally interact with each other or with other GPCR family members forming homo- and/or hetero-dimers complexes that can couple to different signaling cascades to mediate multiple actions [187].

#### 1.3.1.1. SST/CORT/ssts system in cancer

Synthetic SST analogs (SSAs) have been extensively used in clinical practice for the treatment of various hormone-related tumor pathologies, such as pituitary and neuroendocrine tumors, owing to the ubiquitous expression of ssts in normal and tumoral tissues, ad the well-known capacity of SST to inhibit hormone secretion and cell proliferation from a wide variety of cell types, including those from different types of tumors [170, 171, 188, 189]. Indeed, ssts are broadly expressed in several

endocrine-related tumors, such as those from the breast [190]. In general, sst2 is the most common sst subtype expressed in human tumors, followed by sst1, with sst3 and sst4 being less common. The expression of sst5 appears to be rather tumor-specific, with strong expression in some tumors (i.e. breast) and very low expression in others (i.e. pancreatic) [170, 191]. Unfortunately, clinical studies exploring the utility of SSAs in other ssts-positive, endocrine-related tumors, such as breast cancers, are lacking or unsatisfactory [192].

_Splicing variants of SST/CORT/ssts system_

In the process of further characterizing the sst family, our laboratory identified novel, functional truncated variants of the sst5, with less than 7TMDs (**Figure I4**), in various mammalian species (human, pig, mouse and rat) [11, 193-195]. These truncated receptors are originated by the elimination of a cryptic intron in the sst5 sequence during the mRNA maturation through a non-canonical splicing event. Remarkably, these truncated variants of the sst5 gene have unique ligand-selective signaling properties, distinct distribution in normal tissues and different subcellular localization to that shown by the originally identified, long sst5 isoform [195].

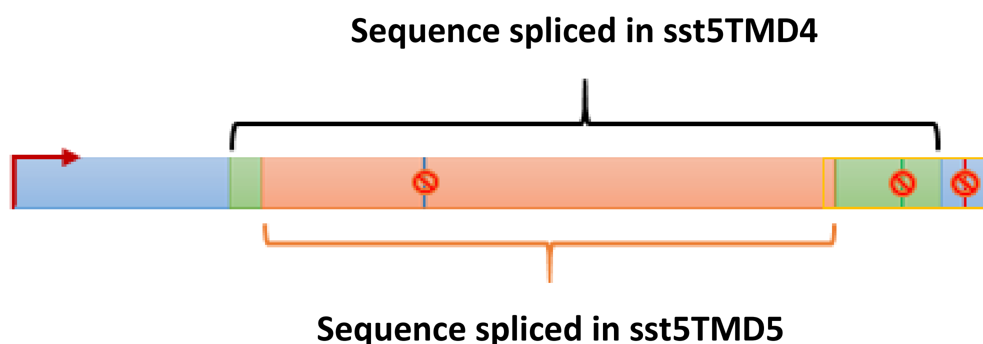


***Figure I4***: *Schematic image of SST5 gene indicating the cryptic intron eliminated during the generation of sst5TMD4 and sst5TMD5 splicing variants. Red circles indicate the presence of a stop codon.*

Interestingly, human sst5 truncated receptors, and specially the truncated receptor with 4TMDs (sst5TMD4), are barely expressed in normal tissues [11], but have been found to be highly expressed in a subset of endocrine-related tumoral pathologies such as pituitary tumors [11, 12, 14], NETs [19], thyroid cancer [10] or breast cancer [18], wherein its expression has been correlated to poorer prognosis. Thus, expression of sst5TMD4 has been correlated with impaired response to SSA treatment in pituitary adenomas [12] and, likely, in thyroid carcinoma [10]. In addition, sst5TMD4 expression was associated with increased aggressiveness features in thyroid cancer [10], NETs [19] and breast cancer [18]. The data gathered hitherto suggest that sst5TMD4 would act,

at least in part, through the blockade of the normal activity of full-length canonical receptors, particularly sst2, thus behaving as a dominant-negative receptor. Although sst5TMD4 has been linked to increased malignant phenotype in *in vitro* models of NETs [19], thyroid [10] and breast cancer [18] through increased proliferation, migration and invasion abilities [18], the regulation of sst5TMD4 expression and the molecular determinants underlying these and other actions remain to be determined.

### 1.3.2. Ghrelin system

Ghrelin is a 28-amino acid peptide hormone, originally isolated from stomach by its ability to induce the release of growth hormone (GH) [196] through the activation of the, until then, orphan receptor for synthetic GH-secretagogues 1a (GHSR-1a) [197]. Thereafter, ghrelin has arisen as a pleiotropic hormone, involved in the regulation of many bodily functions and capable to interact with a number of related endocrine systems in a wide variety of tissues [198]. Indeed, ghrelin functions are widely distributed through the organism, exhibiting endocrine and not endocrine actions. The endocrine functions of ghrelin are mainly triggered through the regulation of the hypothalamus-pituitary axis, where ghrelin is involved in the modulation of the secretion of several hormones including GH, prolactin and adrenocorticotropin (ACTH) [199]. However, ghrelin also exerts other relevant endocrine actions, regulating pancreatic and gastrointestinal tract functions [200]. Moreover, it is also able to modulate a variety of non-endocrine functions through the regulation of immune, digestive and nervous systems to maintain whole body homeostasis [201] (**Figure I5**).

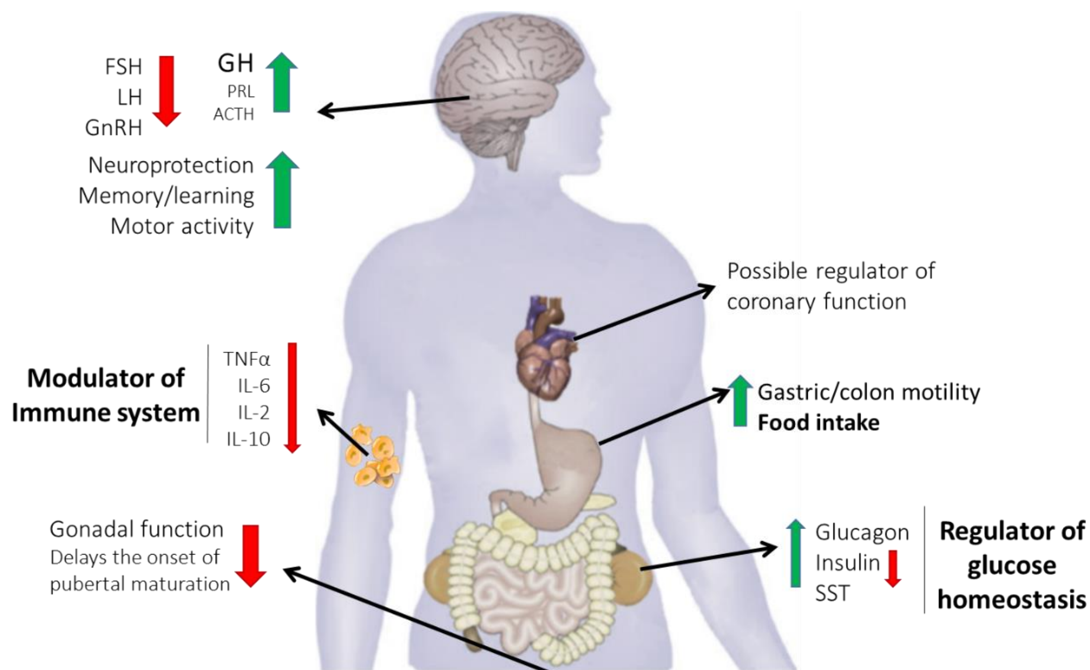


***Figure I5:*** *Summary of the main endocrine and non-endocrine functions of the ghrelin system in humans.*

Ghrelin binding and activation of its canonical GHSR-1a receptor is determined by a unique post-translational processing of the ghrelin peptide. Namely, ghrelin needs to be modified by the ghrelin-O-acyltransferase (GOAT) enzyme, by adding an octanoyl group at the serine-3 residue [201], and producing acylated-ghrelin, which is then able to signal through GHSR-1a [202]. However, this modification is not mandatory and unacylated-ghrelin can also be released. In addition, acyl-ghrelin can lose its acylation and become deacylated ghrelin [198]. Indeed, unacylated-ghrelin is present in the bloodstream at higher concentrations than acylated ghrelin, and has been described to exert a number of activities [203] despite being unable to interact with GHSR-1a.

*Splicing variants of the ghrelin system*

The classical, simplistic conception of a ghrelin system comprised by a single hormone, acylated ghrelin, and a unique receptor, GHSR-1a, has been definitely challenged during the last years by the discovery of novel, additional components, functions, and implications, which have increased remarkably the complexity and versatility of the ghrelin system. The additional members of the ghrelin system arise from both, co- and post-transcriptional modifications and from post-translational processing mechanisms, including alternative splicing variants. Indeed, the GHSR gene has been shown to encode for the full-length 7TMDs, canonical receptor mentioned above, GHSR-1a, which presents complete functionality, but also gives rise to an additional, truncated splicing variant with 5TMDs, named GHSR-1b, which arises from a splicing process of intron retention, and whose precise function is still to be elucidated [204]. In fact, the putative GHSR-1b ligand is unknown, and available studies present this splicing variant as a dominant-negative of its counterpart GHSR-1a by retaining it at the endoplasmic reticulum [205].

The case of the GHRL gene is even more complex. In humans, GHRL is a single-copy gene located on the short arm of chromosome 3, which was originally thought to be composed of four coding exons (exons 1–4) [206]. However, subsequent studies revealed the existence of several alternative upstream exons (exon -1, exon 0 and extended exon 1) that can act as alternative sites for transcription initiation [207]. The initially identified, native ghrelin peptide results from the proteolytic processing of a precursor peptide named pre-pro-ghrelin, a 117-aa long peptide in humans, whose sequence is highly conserved among their mammalian counterparts [205-207]. Human pre-pro-ghrelin contains a 23-aa signal peptide and a 94-aa segment called pro-ghrelin, which undergoes a proteolytic processing to generate the mature ghrelin peptide [206], and also an alternative but functional peptide named obestatin, which was initially considered as the antagonist hormone for ghrelin [208].

In addition, a growing body of evidence supports the existence of a number of alternative ghrelin gene-derived mRNA splice variants and peptides [16, 198, 207] (**Figure I6**). Some of those mRNA splicing variants encode peptides with minor changes

in their sequences as compared to native ghrelin, as it is the case of des-Gln14-ghrelin, which is identical to native ghrelin except for the deletion of one glutamine (Gln14) residue [206]. However, the ghrelin gene can also undergo more complex alternative splicing processes, such as exon skipping or intron retention. Indeed, an event of exon 3 skipping has been reported to generate a 91-aa peptide named Ex3-deleted ghrelin, which lacks the coding region for obestatin [208].

Along these lines, our group identified a novel ghrelin variant generated by retention of intron 1 (In1), which was consequently named **In1-ghrelin**. Owing to its molecular structure, In1-ghrelin shares the signal peptide and the initial portion of the peptide with native ghrelin, including the first 5-aa, which is the minimum sequence required for ghrelin acylation by GOAT, and for binding and activation of GHSR-1a [207]. However, the aa sequence of In1-ghrelin is subsequently altered by the retention of intron 1. In addition, In1-ghrelin pre-pro-peptide, similar to that observed for native pre-pro-ghrelin, also exhibits sites for putative protease action, suggesting that the full peptide could be processed to yield mature peptides whose precise chemical nature is still to be defined. Of note, a similar intron retention process had been identified, previously, in other mammalian species such as mice (generating an alternative splice variant named In2-ghrelin [210]) and a non-human primate model [17]. These data suggest that this new splicing variant might exert important physiological roles conserved among vertebrates.

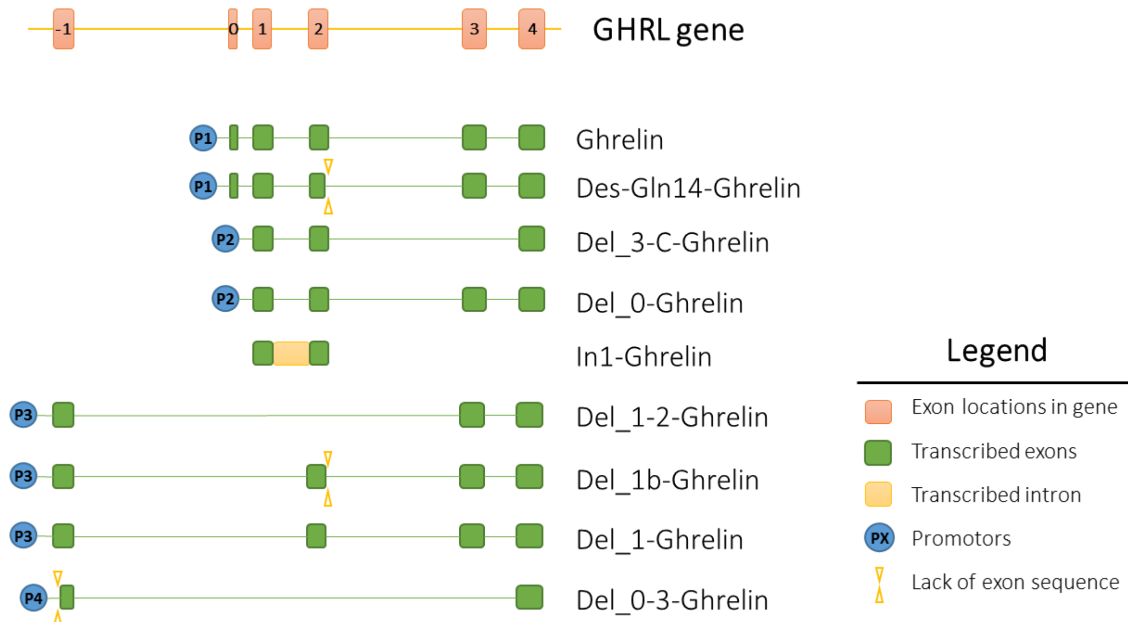


***Figure I6:*** *Splicing variants of the ghrelin gene described in the bibliography [16, 198, 207]*

#### 1.3.2.1. Ghrelin system in cancer

Due to the relevant role exerted by the ghrelin system in the regulation of a plethora of physiological processes, changes in this system have also been associated to the development and/or progression of a number of pathological conditions, including cancer. Indeed, the ghrelin system has been often associated to increased cell proliferation in several healthy and cancer tissues, as well as with the regulation of invasiveness, migration, metastasis, and apoptosis in various cell types [161, 211-216].

In particular, ghrelin actions in cancer are controversial, inasmuch as some reports correlated its presence with cancer malignancy features [161, 212], while others found ghrelin to be a good prognosis marker [217, 218]. Additionally, GHSR-1a has been found to be overexpressed in pituitary tumors, wherein it correlates with tumor size and invasiveness [219], and in endometrial cancer, where its inhibition decreases proliferation rate [220]. Similarly, overexpression of GHSR-1b has been found in different cancer types, including pituitary [221], prostate [222], pancreas [223] and lung [224]. Consistent with a relevant role of the components of the ghrelin system on the pathogenesis of different tumor types, the expression of several ghrelin gene-derived splicing variants has been found to be altered in certain tumoral pathologies. Specifically, although the precise functions of Ex3-deleted ghrelin variant remain uncertain, its expression is increased in human prostate and breast cancers [207], suggesting a putative role in these pathologies. In line with this, the recently identified In2c-ghrelin variant has also been shown to be expressed in prostate cancer cell lines [225]; however, the pathophysiological role of this variant is still unknown [225].

More recently, a series of studies on the In1-ghrelin variant identified by our group in 2011 has indicated that this variant is overexpressed in different tumoral pathologies such as breast cancer [17], and pituitary [15] and neuroendocrine tumors [16]. Interestingly, In1-ghrelin expression seems to be associate to tumor malignancy, for its mRNA levels tightly correlate with proliferation markers such as Ki-67 and cyclin D3 in breast cancer [17], and In1-ghrelin overexpression and/or treatment induced not only a higher cell viability and inhibited apoptotisis [15], but also stimulated hormone secretion from different pituitary and neuroendocrine tumors [15, 16]. Thus, when taken together, these data support the clear relevance of the In1-ghrelin splicing variant expression in tumoral processes, and suggest the necessity to implement additional studies to clarify the regulation, specific functions, and putative clinical implications of this splicing variant in relevant tumoral processes, such as breast cancer.

# Objetives

# 2. Objectives

The **general aim** of the present Thesis is to **identify basic molecular factors and mechanisms involved in the genesis and regulation of two novel, relevant splicing variants of the SST/CORT/ghrelin system, sst5TMD4 and In1-ghrelin, as well as to explore their role and putative clinical implications in breast cancer.**

In order to fulfill this general aim, we proposed the following **specific objectives**:

1) To establish molecular elements and mechanisms underlying the generation, regulation and functioning of **sst5TMD4**, and to ascertain its potential clinical implication in breast cancer development and progression, by exploring:

    a) The process of sst5TMD4 mRNA genesis, and the regulation of sst5TMD4 mRNA stability and translation. To this end we will evaluate:

        i) The putative involvement of SNPs, *de novo* mutations, and splicing factor binding sites in the regulation of ss5TMD4 mRNA expression

        ii) The role of miRNAs in the expression and degradation of sst5TMD4 mRNA.

    b) The consequences and clinical implications of sst5TMD4 presence in tumoral-associated processes, such as angiogenesis and cell dedifferentiation, as well as in the development and progression of breast cancer in human patients.

2) To investigate the molecular and cellular factors involved in the generation, regulation and functioning of the **In1-ghrelin splicing variant**, and its potential clinical implication in breast cancer development and progression, by exploring:

    a) The process of In1-ghrelin mRNA genesis and the regulation of In1-ghrelin mRNA stability and translation To this end we will evaluate::

        i) The putative involvement of SNPs, *de novo* mutations, and splicing factor binding sites in the regulation of In1-ghrelin mRNA expression

        ii) The role of lncRNAs in In1-ghrelin mRNA expression and degradation.

    b) The consequences and clinical implications of In1-ghrelin presence in tumoral-associated signaling pathways and processes, as well as in the development and progression of breast cancer in human patients.

# Materials & Methods

# 3. Materials & Methods

## 3.1. Human samples

In order to study different aspects of the transcriptional regulation and pathological implication of both sst5TMD4 and In1-ghrelin splicing variants, three different sets of human samples were used in this Thesis.

Firstly, a series of 127 **infiltrating ductal breast carcinomas** (IDCs) obtained from the archives of the Pathology Department of the MD Anderson Cancer Center (Madrid, Spain), classified as high grade tumors (G3). Patients underwent surgery between 2003 and 2004. Mean patient age at surgery was 54.9 years (range, 27-79 years). According to the TNM Classification staging, 48 of the tumors were stage I, 34 were stage II, and 35 were stage III-IV. Clinical data of the patients are shown in **Table M1**. Two different tumor areas from each sample were included into a tissue microarray (TMA) according to manufacturer's procedures. Histological and immunohistochemical studies were all carried out on formalin-fixed, paraffin-embedded tissue samples.

| | n (%)* |
|---|---|
| **sst5TMD4 protein expression (n=117):** | |
| Low | 61 (52.1) |
| High | 56 (47.9) |
| **sst5TMD4 mRNA expression (n=98):** | |
| Low | 44 (44.9) |
| High | 54 (55.1) |
| **In1-ghrelin mRNA expression (127):** | |
| Low-null | 63(49.6) |
| Medium | 32(25.2) |
| High | 32(25.2) |
| **CD34 protein expression (n=117):** | |
| Low | 81 (69.2) |
| High | 36 (30.8) |
| **Lymph node metastasis (n=117):** | |
| Negative | 56 (47.9) |
| Positive | 61 (52.1) |
| **Distant metastasis (n=117):** | |
| Negative | 82 (70.1) |
| Positive | 35 (29.9) |

*n(%), number of analysed cases and (percentage).

***Table M1.*** *Summary of clinical, pathological, immunohistochemical and molecular features of breast IDC Grade 3 samples.*

Secondly, a cohort of 47 **mammary gland biopsies**, both positive and negative for breast cancer, from women with age comprised between 31 and 77 years (mean age

49.19±3.5) was included in the study. Participants were recruited through the Mammary Gland Unit of Hospital Universitario Reina Sofia (HURS, Córdoba, Spain), with suspect of non-familial, sporadic breast cancer, considered within a group of common risk that were selected in the study after being image-diagnosed within grade 4-5 of the BIRADS (Breast Imaging Reporting and Data System).

Alternatively, in order to analyze the expression of different transcripts in normal, control human tissues, a **commercial panel of total RNA** from various human tissues was obtained from Clontech (Total Master Panel II and pituitary poly-A RNA; Palo Alto, CA), where each tissue sample is a pool of multiple individuals

These studies were performed following standard ethical procedures of the Spanish regulation (Ley de Investigación Orgánica Biomédica, 14 July 2007) and were approved by the ethic committee of MD Anderson Cancer Center (Madrid, Spain) and the HURS (Córdoba, Spain).

## 3.2. Molecular biology: Nucleic Acids

### 3.2.1. Nucleic acid extraction

#### 3.2.1.1. Genomic DNA and total RNA from human samples

Genomic DNA (gDNA) and total RNA from human samples were extracted using the "AllPrep DNA/RNA/Protein Mini Kit" (Qiagen, Madrid, Spain) following manufacturer's instructions and subsequently quantified with Nanodrop 2000 spectrophotometer (Thermo Scientific, Wilmington, NC, USA). Briefly, samples were homogenized with an IKA T25 Ultra-Turrax (Gemini BV laboratory, Apeldoorn, Nederland) in the recommended *RLT Buffer* that allow the specific purification of gDNA and total RNA. Then, the homogenized samples were passed through two columns that retain firstly the gDNA and then total RNA. gDNA and total RNA were eluted with RNase- and DNase-free water, respectively.

#### 3.2.1.2. miRNAs from human samples

Total miRNAs were extracted from the battery of mammary gland biopsies maintained in RNALater (Thermoscientific, Barcelona, Spain) using All-in-One Purification Kit (Norgen Biotech, Canada) and following manufacturer's protocol. Specifically, flow-through obtained from RNA purification step was mixed with 200 µl of Ethanol 100%, and passed through the *microRNA Enrichment Column* provided in the kit, washed with *Wash Solution* and eluted with 50 µl of *Elution Buffer*. Resultant eluted miRNAs were quantified with Nanodrop 2000 spectrophotometer.

#### 3.2.1.3. Total RNA from cell lines

Total RNA was extracted from MDA-MB-231 and MCF-7 cell lines using Trizol (Life Technologies, Barcelona, Spain), following the manufacturer's protocol as previously reported [17, 18]. Briefly, cells were incubated until confluence in 6-well-plates. Then, wells were washed with PBS and, subsequently, 1ml Trizol was added and collected with lysed cells in 1.5ml tubes. RNA isolation was carried out by adding chloroform, centrifugating, and collecting the aqueous phase. RNA was recovered and concentrated with 2-propanol precipitation and 70% ethanol washing steps. Finally, samples were dried and resuspended with 8 µl of DEPC-treated water. Subsequently, samples were treated with 1 µl (1 unit) of DNase (Promega, Barcelona, Spain) and incubated 30 min at 37ºC, stopping the reaction by adding a *Stop Solution* and incubating at 65C for 5 min. Total RNA concentration and purity were assessed using Nanodrop 2000 spectrophotometer.

### 3.2.2. Total RNA and miRNA retrotranscription

Retrotranscription of total RNA was carried out with the "cDNA First Strand Synthesis kit" (MRI Fermentas, Hanover, MD, USA) using random primers and following manufacturer's instructions. Briefly, 1 µg of total RNA from each sample was mixed with random hexamers and incubated at 65ºC for 5 min. Subsequently, the appropriate buffers, the dNTPs and the reverse-transcriptase were added and the mix incubated for 1 h at 42ºC, finishing with an incubation of 5 min at 70ºC.

miRNAs of interest were specifically retrotranscribed using specific primers developed by our group (for details regarding specific sequences see **Table M2**) using the "ThermoScript H-Reverse Transcriptase kit" (Thermo Scientific) and following manufacturer's instructions. Briefly, 100 ng of miRNA samples were mixed with specific primers for each miRNA of interest and dNTPs in sterile distilled water. Then samples were incubated at 65ºC for 5 min, adding the appropriate buffer and enzyme provided in the kit at a final volume of 20 µl. Prepared samples were incubated then at 50ºC for 60 min and, finally, reaction was ended by 5 min at 85ºC.

| Transcript | Accesion number | Sense | Tm | Antisense | Tm | Size (bp) |
|---|---|---|---|---|---|---|
| hsa-miR-939 | NR_030635.1 | CTGGGGAGCTGAGGCTCT | 60,66 | TCAGACACTGGGGAGCAGA | 60,56 | 64 |
| hsa-miR-346 | NR_029907.1 | GCATGCCTGCCTCTCTGTTG | 64,40 | TGCCCAGGCAGCTGCA | 65,35 | 61 |
| hsa-miR-339 | NR_029898.1 | TCCCTGTCCTCCAGGAGCT | 62,33 | TCTGTCGTCGAGGCGCT | 61,96 | 54 |
| hsa-miR-326 | NR_029891.1 | ATCTGTCTGTTGGGCTGGAG | 60,26 | TCGGGGCTGGAGGAAG | 60,89 | 82 |
| hsa-miR-189 | NR_029496.1 | CCGGTGCCTACTGAGCTG | 60,57 | TGCTGAACTGAGCCAGTGTG | 61,24 | 56 |

***Table M2****: Details of primers used to amplify each miRNA. Tm = Melting temperature; Size (bp) = number of nucleotides of the sequences amplified by each pair of primers.*

### 3.2.3. PCR and qPCR

Conventional PCR has been used throughout this work to validate qPCR primers, using "DreamTaq DNA Polymerase" (Thermoscientific), and to amplify the SST5 and ghrelin gene sequences for further sequencing. All conventional PCRs were carried out in a thermocycler "Supercycler Gradient Cycler" (Kyratec, Belgium).

Primers used during the present work for PCR and qPCR have been designed using the bioinformatics tool *Primer Blast* (http://www.ncbi.nlm.nih.gov/tools/primer-blast/), using as template the mRNA sequences from NCBI database depicted in **Table M3**. In order to standardize the methodology, basic requirements of the primers for qPCR were fixed in a Tm range of 59 to 61ºC, and an amplified sequence of 80 to 200pb. Additionally, when possible, each primer, sense and antisense, was designed in different exons to prevent genomic amplification. Primers for genomic sequencing used the same Tm range but the sequences amplified were approximately of 200-500pb (specific details are shown in **Table M3**). Candidate primers were further tested and *in silico* optimized with Primer3 online tool (http://bioinfo.ut.ee/primer3-0.4.0/) in order to maximize the specificity and avoid PCR efficiency complications. Selected primer sequences, were synthesized by "Integrated DNA technologies" (Madrid, Spain). Subsequently, primers were validated by conventional PCR using cDNAs from different cell lines as template. Bands obtained by electrophoresis in agarose gels were purified with "FavorPrep™ GEL/PCR Purification Kit" (Favorgen, Vienna, Austria) following manufacturer's instructions and sequenced using the genomic services of SCAI (Servicio Centralizado de Apoyo a la Investigación, University of Córdoba, Spain). Then sequences' specificity was compared with expected sequences for each set of primer.

Quantitative real-time PCR (qPCR) was used to evaluate cDNA derived from human samples or cells lines, where samples were run against a standard curve to estimate the absolute mRNA copy number. No-RT sample was used as a negative control. qPCR was performed using Brilliant III SYBR Green Master Mix in the Stratagene Mx3000p instrument (both from Agilent, La Jolla, CA, USA) as previously described [17, 18]. The thermal profile used was:

- Initial denaturation: 95°C 3 min
- 40 cycles:
  - Denaturation: 95°C 20 seconds
  - annealing/extension: 60°C 20 seconds
- Dissociation cycle (Melting curve): as a proof of single product amplification, analysis of lost of fluorescent signaling within a gradual raising temperature is used to analyze the existence of different products.

| Transcript | Accession number | Primers application | Sense | Tm | Antisense | Tm | Size (bp) |
|---|---|---|---|---|---|---|---|
| 18S | NR_003286.2 | qPCR | CCCATTCGAACGTCTGCCCTATC | 67.89 | TGCTGCCTTCCTTGGATGTGGTA | 68.02 | 136 |
| ANG1 | NM_001146.3 | qPCR | GACAGATGTTGAGACCCAGGTA | 59.07 | TCTCTAGCTTGTAGGTGGATAATGAA | 59.74 | 89 |
| ANG2 | NM_001147.2 | qPCR | GGATGGAGACAACGACAAATG | 60.37 | GGACCACATGCATCAAACC | 59.76 | 78 |
| β-Actin | NM_001101 | qPCR | ACTCTTCCAGCCTTCCTTCCT | 60.74 | CAGTGATCTCCTTCTGCATCCT | 60.79 | 176 |
| β-catenin | NM_001904.3 | qPCR | TTAAGCCTCTCGGTCTGTGG | 60.39 | CAAATACCCTCAGGGGAACA | 59.78 | 176 |
| CD34 | NM_001025109.1 | qPCR | CACCAATCTGACCTGAAAAAGC | 61.01 | AAATAGCCAGTGATGCCCAAG | 61.36 | 143 |
| CSF3 | NM_000759.3 | qPCR | CCTCCCCATCCCATGTATTT | 60.77 | TGGGAGGACAGGAGCTTTTT | 61.12 | 167 |
| EGF | NM_001963.4 | qPCR | CTGAAGGTACTCTCGCAGGAAA | 60.91 | CACTGAGACACCAGCATCCAC | 61.77 | 146 |
| GAPDH | NM_002046.5 | qPCR | AATCCCATCACCATCTTCCA | 60.13 | AAATGAGCCCCAGCCTTC | 60.16 | 122 |
| GHRL gene | GU942497.1 | Sequentiation | ATGCTCTGGCTGGACTTGG | 61.38 | GTTCATCCTCTGCCCCTTCT | 60.60 | 366 |
| Ghrelin | NM_016362.3 | qPCR | CACCAGAGAGTCCAGCAGAGA | 60.74 | CCGGACTTCCAGTTCATC | 56.40 | 215 |
| GHRLOS-1 | NR_004431.3 | qPCR | AGCGCCTCATCTCTTCCATT | 61.26 | CTCAGTGGCTGCCCTCCT | 61.56 | 238 |
| GHRLOS-2 | NR_024144.2 | qPCR | GCCTTCCATTCCCTCCAGTA | 61.35 | GACTGATTTTCCTGCACCACAG | 61.99 | 159 |
| GHRLOS-3 | NR_024145.2 | qPCR | CGCTTCTAAACTTAGAGAGAGGAGAGT T | 61.53 | TAGGCCAGGCCAGCAGTT | 61.91 | 158 |
| GHSR-1a | NM_198407.2 | qPCR | TGAAAATGCTGGCTGTAGTGG | 61.20 | AGGACAAAGGACACGAGGTTG | 61.48 | 168 |
| GHSR-1b | NM_004122.2 | qPCR | GGACCAGAACCACAAGCAAA | 61.08 | AGAGAGAAGGGAGAAGGCACA | 60.52 | 107 |
| HIF-1a | NM_001530.3 | qPCR | TTAGATTTTGGCAGCAACGAC | 60.26 | GGGTGAGGGGAGCATTACA | 60.88 | 87 |
| HIF-1b | NM_001668.3 | qPCR | ACTACTGCCAACCCCGAAAT | 60.74 | ATGGCTCCTCCACCTTGAAT | 60.85 | 98 |
| HPRT | NM_000194.2 | qPCR | CTGAGGATTTGGAAAGGGTGT | 60.35 | TAATCCAGCAGGTCAGCAAAG | 60.39 | 157 |
| IGFBP1 | NM_000596.2 | qPCR | GTTTAGCCAAGGCACAGGAG | 59.88 | TATCTGGCAGTTGGGGTCTC | 60.07 | 203 |
| In1-ghrelin | GU942497.1 | qPCR | TCTGGGCTTCAGTCTTCTCC | 59.53 | GCTTGGCTGGTGGCTTCTT | 62.79 | 132 |
| ITGB2 | NM_000211.4 | qPCR | ACTGATGACGGCTTCCATTT | 59.56 | GATGGGCTGGATGTTGTTTT | 59.80 | 171 |
| Jagged 1 (JAG1) | NM_000214.2 | qPCR | GTGCTACAACCGTGCCAGT | 59.77 | CTTCAGGTGTGTCGTTGGAA | 59.72 | 152 |
| MMP1 | NM_002421.3 | qPCR | CTGATATCGGGGCTTTGATG | 60.43 | GATGGGCTGGACAGGATTTT | 61.22 | 122 |
| MMP10 | NM_002425.2 | qPCR | TCGCAAGATGATGTGAATGG | 60.63 | TGATGGCATCGAAGGACAAA | 60.53 | 145 |
| SST5 gene | NC_000016.10 | Sequentiation | AGGAGCAGAGGACGGTCA | 59.47 | TGTCCTCACTGCTTGGATGT | 59.26 | 461 |
| sst5 | NM_001053.3 | qPCR | CTGGTGTTTGCGGGATGTT | 61.92 | GAAGCTCTGGCGGAAGTTGT | 61.86 | 183 |
| sst5TMD4 | DQ448304 | qPCR | TACCTGCAACCGTCTGCC | 60.84 | AGCCTGGGCCTTTCTCCT | 61.27 | 98 |
| TGF-β1 | NM_000660.5 | qPCR | CACGTGGAGCTGTACCAGAAA | 61.27 | CAACTCCGGTGACATCAAAAG | 60.53 | 112 |
| VEGFa | NM_001171623.1 | qPCR | TTAAACGAACGTACTTGCAGATG | 59.37 | GAGAGATCTGGTTCCCGAAA | 59.21 | 93 |

***Table M3:*** *Details of primers used for qPCR and sequencing experiments. Primer application = primary use of each set of primer; Tm = Temperature of melting; Size (bp) = length of the sequences amplified by each pair of primers.*

qPCR standard curves where generated from amplifications of specific products by conventional PCR using the designed primers and purifying the resulting bands as detailed above. Purified bands were quantified with Nanodrop 2000 spectrophotometer and serial dilutions of each template were generated in order to obtain 1, 10, $10^2$, $10^3$, $10^4$, $10^5$ and $10^6$ copies of synthetic template for each transcript.

Expression of a given mRNA in terms of copy number were adjusted by the expression of housekeeping genes to control for variations in the quantity of cDNA used and the efficiency of the reverse-transcription among of each sample. To this end, housekeeping gene expression levels in each sample was used to generate a

normalization factor (NF) by *Genorm* software [226]. It should be noted that, as previously reported [17, 18] and based on the stringent criteria to maximize specificity and efficiency, the qPCR technique, as applied, can be used to accurately quantify copy numbers for all human transcripts included in this study (for details regarding primer sequences and product sizes refer to **Table M3**).

### 3.2.4. Plasmid vectors

#### 3.2.4.1. Cloning sequences

pCDNA3.1 plasmid (Invitrogen, Madrid, Spain) was used as expression vector for both, previously and newly, cloned sequences. Specifically, sst5TMD4 and In1-ghrelin where previously cloned as reported in [11, 17] and validation of sequences integrity was assessed by sequencing with standard primers by the Genomic Services of the SCAI. On the other hand, a plasmid containing the ghrelin sequence was purchased from DF/HCC DNA Resource Core (Boston, MA, USA). In this case, the carrier plasmid was pCMV-Sport6 but, in order to standardize the subsequent studies, ghrelin sequence contained in this plasmid was subcloned into pCDNA3.1 vector as explained below.

#### 3.2.4.2. Restriction enzyme digestion

Ghrelin sequence subcloning in pCDNA3.1 was pursued by restriction enzymes-mediated extraction of the cloned sequence from pCMV-Sport6. To this end, EcoR-I and HindIII restriction enzymes (New England Biolab, Barcelona, Spain) were used in order to extract ghrelin sequence from the carrier plasmid. Specifically, 1 µg of both, ghrelin carrier and pCDNA3.1 empty plasmids, were incubated with 1 µl of both enzymes for 1 h at 37ºC. Ghrelin sequence was then isolated by agarose gel electrophoresis and specific ghrelin and linearized pCDNA3.1 bands were purified with the FavorPrep™ GEL/PCR Purification Kit (Favorgen).

#### 3.2.4.3. Plasmid and ghrelin sequence ligation

Ligation of ghrelin sequence with pCDNA3.1 was performed with T4 DNA Ligase (New England Biolab) following manufacturer's instruction. Specifically, 100 ng of digested plasmid were used in combination with purified ghrelin sequence in a molar ratio of 1:1 (pCDNA3.1:Ghrelin) and 1 µl of T4 DNA ligase. This mix was incubated overnight at 4ºC.

#### 3.2.4.4. Competent bacteria transformation

Chemically competent DH5α cells, prepared by Hanahan method [227], were used in heat shock transformation protocol. Specifically, 50 µl aliquots of thawed bacteria were exposed to 0,2% β-Mercaptoethanol and maintained on ice for 20 min. Then 100 ng of plasmid or 5 µl of ligation were gently added and incubated for another 10 min.

Heat shock was applied at 37ºC during 2 min followed by 3 min on ice. Subsequently, 0.5ml of *LB media* was added to bacteria containing tubes and were incubated in a 37ºC shaking incubator for 1 h. Then, 100 µl of bacteria cultures were seeded in LB-Agar plates with L-ampicillin (Sigma Aldrich, San Louis, MO, USA) at 100 µg/ml concentration an incubated at 37ºC for 13-15 h. Positive colonies were checked by PCR, amplifying the specific sequence cloned. Briefly, part of the bacterial colony was dissolved in distilled water and used as a template for PCR. The rest of the colony was grown in a new LB-Agar Plate with L-ampicillin in order to be used for plasmid amplification and purification.

3.2.4.5. Plasmid DNA amplification and purification

Positive colonies were grown overnight on *LB medium* with the selection antibiotic (L-ampicillin) and plasmids were extracted with "ISOLATE II Plasmid Mini Kit" from Biolane (Paris, France) following manufacturer's instructions. Purified plasmids were quantified with Nanodrop 2000 spectrophotometer.

## 3.3. Molecular Biology: Proteins

### 3.3.1. Peptides

Acylated ghrelin was purchased from a commercial supplier (SC1357, PolyPeptide Laboratories, Limhamn, Sweden). On the other hand, two different In1-ghrelin derived acylated peptides were synthesized in collaboration with Ipsen Biosciences (Cambridge, MA, USA) and CPC Scientific (Chinese Peptide Company, Hangzhou, China). Although the mature endogenous In1-ghrelin derived peptides have not yet been identified, pre-pro-In1-ghrelin precursor exhibits target sites for protein-convertases suggesting a further proteolytic processing. As previously reported [17], In1-ghrelin precursor processing could generate 40-aa or 19-aa long peptides (thus named heretofore In1-40: GSSFLSPEHQRVQVRPPHKAPHVVPALPLSNQLCDLEQQR and In1-19: GSSFLSPEHQRVQVRPPHK), which share with native ghrelin the initial 13-aa,

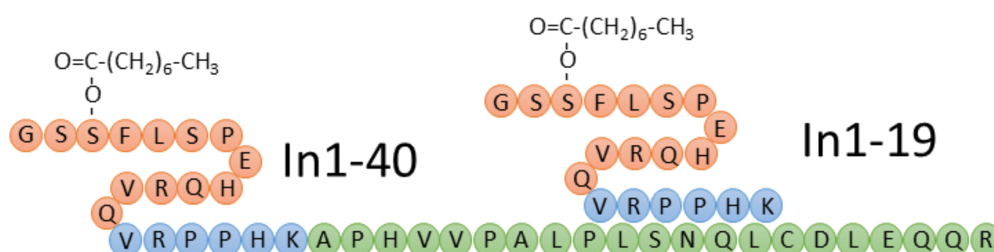


***Figure M1:*** *Representative image of putative acylated In1-ghrelin derived peptides (In1-40 and In1-19). Red aminoacids corresponds with those shared between ghrelin and In1-ghrelin, while blue and green aminoacids are only present in In1-ghrelin derived peptides, being green aminoacids those present only in In1-40 peptide.*

including the acylation site at Ser3. (**Figure M1**).

The acylated In1-19 and In1-40 peptides were synthesized using manual solid-phase peptide synthesis starting with Fmoc-Lys(Boc)-Wang resin on a 0.5 mmol scale. The resins were treated with DCM/DMF (151) for 1 h, followed by standard Fmoc single coupling cycles with 1.5 mmol amino acid and coupling agent. All amino acids were Fmoc-protected except for Ser 3, which was unprotected. All amino acids were activated with HBTU or DICGly1 and HATU. Octanoic acid was coupled to Ser3 using 235 mmol octanoic acid and HOBt, followed by 2310 mmol octanoic acid and HOBt. The peptides were then treated with a cocktail of TFA/EDT/ Thioanisole/Phenol/H2O (87.552.55552.5) for 150 min to remove the peptides from the resins. The peptides were confirmed by ESI MS and analytical RP-HPLC. The peptides were eluted with a gradient of Buffer B (0.09% TFA in 80% CH3CN/H2O) in aqueous 0.1% TFA. The peptides solubility was determined to be 1 mg/mL in water. Finally, peptide content was determined by AAA.

### 3.3.2. Protein extraction

Proteins from human and mouse tissues and cell lines were extracted with *SDS-DTT buffer* (62,5mM Tris-HCl, 2%SDS, 20% glicerol, 100mM DTT and 0,005% bromophenol blue) as follows.

#### 3.3.2.1. Tissues

Protein pellets obtained from "AllPrep DNA/RNA/Protein Mini Kit" (Qiagen Iberia S.L. Spain) were resuspended in 500 µl pre-warmed SDS-DTT buffer and disrupted by sonication. Finally, proteins were denaturalized by 5 min at 95ºC incubation.

#### 3.3.2.2. Cell lines

Cells were seeded in 6-wells plates and incubated at 37ºC, 5% $CO_2$ until confluence. Then, proteins were extracted by using 200 µl pre-warmed SDS-DTT buffer and denaturalized by boiling 5 min at 95ºC.

### 3.3.3. Western Blot

Proteins were separated by SDS-PAGE and transferred to nitrocellulose membranes (Millipore). Membranes were blocked with 5% non-fat dry milk in Tris-buffered saline with 0,05% Tween 20 and incubated O/N with the primary antibodies (detailed reference of antibodies utilized and concentrations are indicated in **Table M4**) at 4ºC, followed by 1-h incubation with the appropriate secondary antibodies. Proteins were developed using ECL-2 (GE Healthcare, City, UK) following the manufacturer´s instructions. Densitometric analysis of the bands was carried out with ImageJ software.

The relative p-ERK and p-Akt values were obtained from normalization of p-ERK or p-Akt values against the total ERK or Akt values, respectively.

| Antibody | Source | Technique | Dilution |
|---|---|---|---|
| Goat anti-human VEGF | R&D systems, Minneapolis (MN), USA | Western; IHC | 1:1000 |
| Rabbit anti-human ERK1/2 | Santa Cruz Biotech, Dallas (TX), USA | Western | 1:1000 |
| Rabbit anti-human p-ERK1/2 | Cell Signaling, Beverly (MA), USA | Western | 1:1000 |
| Rabbit anti-human AKT | Cell Signaling, Beverly (MA), USA | Western | 1:1000 |
| Rabbit anti-human p-AKT | Cell Signaling, Beverly (MA), USA | Western | 1:500 |
| Rabbit anti-human sst5TMD4 | Custom-made | Western; IHC | 1:1000 |
| Mouse anti-human CD34 Class II | Dako, Barcelona, Spain | IHC | 1:1000 |
| Goat anti-rabbit IgG HRP-linked | Cell Signaling, Beverly (MA), USA | Western | 1:2500 |
| Mouse anti-goat IgG HRP-linked | Santa Cruz Biotech, Dallas (TX), USA | Western | 1:1000 |
| Goat anti-rabbit biotin-conjugated | ThermoScientific Barcelona, Spain | IHC | 1:1000 |
| Rabbit anti-mouse biotin-conjugated | ThermoScientific Barcelona, Spain | IHC | 1:1000 |
| Donkey Anti-Goat IgG Alexa Fluor 488 | Abcam, Cambridge, UK | IHC | 1:1000 |

***Table M4**: Details regarding antibodies used in the different experimental approaches.*

### 3.3.4. Immunohistochemistry

sst5TMD4 and CD34 immunohistochemical staining of the human breast cancer samples was performed by the LSAB (Dako, Spain) method with a heat-induced antigen retrieval step. Briefly, sections were immersed in boiling 10mM sodium citrate at pH 6.0 for 3 min in a pressure cooker. The rabbit polyclonal antisera against human sst5TMD4 were previously described [11, 18]. The antibodies were used as 1:1000 dilution. The primary antibodies were omitted in the negative controls. sst5TMD4 staining was categorized as low or high expression compared to the average sst5TMD4 staining. CD34 (Clone QBEnd 10, Dako, Spain) staining was categorized as low or high with respect to normal mammary tissue.

## 3.4. Cellular Biology

### 3.4.1. Cell lines

In order to explore the functional role of In1-ghrelin and sst5TMD4, two breast cancer derived cell lines widely used in biomedical research were used. These cell lines were previously validated by analysis of STRs (GenePrint® 10 System, Promega, Barcelona, Spain) and checked for mycoplasma contamination by PCR as previously reported [228].

#### 3.4.1.1. MDA-MB-231

MDA-MB-231 (ATCC, VA, USA) represents a model of triple negative breast cancer-derived cell line with a highly malignant phenotype. We used this cell line as model of

poorly differentiated and aggressive tumor taking advantage of its endogenous In1-ghrelin expression to study the effects of overexpression and downregulation of this splicing variant. It was also used to explore some aspects of the pathological role of the truncated receptor sst5TMD4.

#### 3.4.1.2. MCF-7

MCF-7 (ATCC) represents a model of highly differentiated luminal-epithelial breast cancer derived cell line with a phenotype of lower malignancy than MDA-MB-231. We used this cell line as model of breast cancer to analyze the effects of both splicing variants.

### 3.4.2. Freezing/Thawing and maintenance of cells in culture

Cryotubes containing $10^6$ cells were thawed by fast warm in a water bath until ice disappeared, followed by resuspension of cells in pre-warmed complete medium. DMSO was removed by centrifugation and supernatant elimination. Then, cells were resuspended in the appropriate media and seeded in 75 $cm^2$ flasks.

MDA-MB-231 breast cancer cell line was cultured in DMEM with 4.5 g glucose (Lonza, Barcelona, Spain), 2 mM L-Glutamine (Sigma Aldrich, San Louis, MO, USA), 10 % FBS and 0,2 % Gentamicin/amphotericin B (Gibco, Barcelona, Spain) and maintained at 37ºC and 5 % $CO_2$.

MCF-7 cells were maintained in Dulbecco's Modified Eagle Medium with 1g glucose (Lonza), supplemented with 10 % fetal bovine serum, 1 % antibiotic-antimycotic and 2 mM L-glutamine, in a constant atmosphere with 37ºC and 5 % $CO_2$.

For freezing, aliquots of $10^6$ cells were resuspended in FBS with 10 % DMSO in cryotubes and maintained in a pre-warmed isopropanol bath that was stored O/N at -80ºC. Finally, the cryotubes were cryopreserved under liquid nitrogen conditions.

### 3.4.3. Plasmid, miRNA and siRNA transfection

MDA-MB-231 and MCF-7 cells were transfected with constructed plasmids using Lipofectamine-2000 (Gibco) following the manufacturer's instructions and, subsequently, stable transfectants were selected by addition of geneticin (G418) to the cell culture (500 µg/ml; Gibco). Briefly, 150.000 cells were seeded in 6-well plates and incubated for two days at 37ºC and 5 % $CO_2$. Then, cells media were withdrawn and 500 µl of Opti-MEM (Thermo Scientific, Wilmington, NC, USA) were added. In parallel, 1 µg of plasmids of interest (containing In1-ghrelin, ghrelin, sst5TMD4) empty vector were mixed with 3 µl of Lipofectamine-2000 in 100 µl Opti-MEM and incubated for 5 min. Next, transfection complexes were added to the cells and incubated at 37ºC and 5% $CO_2$ for 8 h. Finally, cell medium was replaced by adding 2ml of complemented

media. Success of transfections was validated by qPCR, comparing the expression of transcripts of interest with empty pCDNA3.1 transfected cells (mock). Transfections were stably maintained by adding 1% geneticin to the media, which selectively eliminates non-transfected cells.

Transfection of commercial hsa-miR-346 mimic and inhibitor miRNAs (Qiagen) in MCF-7 cells was carried out at a final concentration of 100 nM with Lipofectamine-2000 following the manufacturer's instructions (and as described above). All the experiments were implemented two days after transfection.

In1-ghrelin silencing was carried out with two custom-designed In1-ghrelin specific siRNA and reduction in mRNA expression and functional consequences were compared with a commercial scramble siRNA (Silencer Select Negative Control No. 1 siRNA; Life Technologies, Green Island, NY, USA). Specifically, MDA-MB-231 cells were transfected with these siRNAs (at a final concentration of 100 nM) using Lipofectamine RNAiMAX Reagent (Gibco) and following the manufacturer's instructions. siRNAs effectiveness was validated by qPCR. For all subsequently studies, cells were incubated 48 h in order to allow siRNAs to achieve the appropriate inhibition of In1-ghrelin expression.

### 3.4.4. Generation of monoclonal stably transfected cell lines

Given that In1-ghrelin, ghrelin, sst5TMD4 and mock transfected cells could have different levels of transfection and in order to minimize the variability in the transfection levels, we generated monoclonal stably-transfected cell lines of each construct. To this end, single clones of transfected cells were selected by limiting dilutions and validated by qPCR. Stably transfected and monoclonaly selected MDA-MB-231 or MCF-7 cells with empty pCDNA3.1 vector (mock) were used as a negative control.

### 3.4.5. Peptide treatments

Peptide treatments were applied at a final concentration of $10^{-7}$M for all treatments including Paclitaxel (Sigma-Aldrich); with the only exception of IGF-I (Sigma-Aldrich) that was used at a final concentration of $10^{-8}$M, following previous reports [15]. Peptides were added at the media at the beginning (0 h) of all experiments. For proliferation studies, peptide treatments were refreshed every 24 h in both Alamar blue and MTT experiments. For mammospheres formation assays, peptide treatments were refreshed every 48 h.

### 3.4.6. Gene expression microarray

Microarray experiments were performed using Human Whole Genome array V2 4*44K array G4845A (Agilent Technologies, Santa Clara, CA, USA). Four independent passages from stably-transfected sst5TMD4-pCDNA3.1 and empty-pCDNA3.1 vector, used as

control (mock), MFC-7 cells were used. RNA was labeled and array hybridized using the Low RNA Linear Amplification Kit and the In Situ Hybridization Kit Plus (Agilent technologies), respectively, following manufacturer's protocol. After hybridization and washing, the slides were scanned in an Axon GenePix Scanner (Axon Instruments) and analyzed using Feature Extraction Software 10.0 (Agilent technologies). RNA samples from independent sst5TMD4-stably transfected cells were labeled with Cy5-dUTP and equal concentrations of each RNA from mock control cells were labeled with Cy3-dUTP. Differentially expressed genes in sst5TMD4-transfected MCF-7 cells vs. control (mock) cells were identified with GEPAS (Gene Expression Pattern Analysis Suite, http://gepas3.bioinfo.cipf.es) selecting genes with a fold difference of at least two in all of the samples and standard deviation lower than 0.5. Functional enrichment analysis was performed using the FatiGO application (http://babelomics. bioinfo.cipf.es). Microarray raw data tables have been deposited in the Gene Expression Omnibus under the accession number of GSE85150.

### 3.4.7. ELISA measurements

VEGF concentration in the sst5TMD4- and mock-transfected MCF-7 culture media was determined using a commercial human VEGF ELISA kit (VEGF human ELISA kit; Cat. number: KHG0112; Life Technologies), following the manufacturer's instructions. The information regarding specificity, detectability and reproducibility for the assay can be accessed at the company website.

## 3.5. Functional assays

### 3.5.1. Proliferation studies

Cell proliferation was evaluated in stably-transfected or peptide-treated MDA-MB-231 and MCF-7 cells using Alamar-Blue reagent (Biosource International, New York, USA) and/or 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) (Sigma Aldrich) technique.

Alamar blue assays were performed as follows: 3.000 cells/well were plated in 96-well plates and, the day of measurement, cells were incubated for 4 h in 10% Alamar blue/serum free-DMEM. Then, Alamar blue reduction was measured in a FlexStation system plate reader (Molecular Devices), exciting at 560 nm and reading at 590 nm. Measurements were repeated during four consecutive days and medium was replaced by fresh medium immediately after each measurement.

Similarly, MTT technique was carried out plating the same number of cells. Each day, media of one plate were removed and 0.25% MTT reagent (Sigma Aldrich) resuspended in DPBS (Sigma Aldrich) was added to each well. Then, cells were incubated for 3 h at 37ºC and 5% $CO_2$, in order to allow MTT reduction. Finally, MTT was withdrawn from each well and cells and MTT crystals were lysed with an acid-SDS

solution (1 g SDS, 10 ml DMSO, 57.2 µl glacial acetic acid). MTT reduction was evaluated in a FlexStation system plate reader by measuring the absorbance of each well at 570nm, using empty wells as blank.

In all instances, cells were plated per quadruplicate and all assays were repeated a minimum of three times with independent cell preparations. Results are expressed as percentage vs. control (mock-transfected or vehicle-treated cells). Peptide treatments were added in the media. IGF-I and Paclitaxel were used as positive and negative controls, respectively.

In the case of siRNA transfected cells, after 2 days of transfection, cells were collected, counted and plated in 96-well plates following the protocol indicated above to assess proliferation rates. In1-ghrelin siRNAs effects were compared with scramble siRNA as a negative control.

### 3.5.2. Migration assays

The ability of mock, ghrelin and In1-ghrelin stably transfected MDA-MB-231 and MCF-7 cells to migrate was evaluated by wound healing technique as previously reported [18]. Briefly, cells were plated at sub-confluence in 12-well plates. Once cells reached confluence, a wound was made using a 200 µl sterile pipette tip. Then, cells were incubated for 24h in media complemented without FBS in order to minimize cell proliferation effects on wound recovery. Wound healing was calculated as the area of a rectangle centered in the picture 24h after the wound vs. the area of the rectangle just after doing the wound. Four or eight experiments (MCF-7 and MDA-MB-231 stably transfected cells, respectively) were performed in independent days, in which six random pictures along the wound were acquired per well. Results are expressed as percentage vs. control (mock transfected cells).

Similar to that indicated in proliferation assays, siRNA transfected cells were collected 2 days after the transfection, counted and plated. Subsequently, the same protocol used with overexpressing cells was applied.

In peptide treatment experiments, In1-ghrelin derived peptides, ghrelin, IGF-I and paclitaxel were added with the media after doing the wound in order to analyze the effects on MDA-MB-231 and MCF-7 cells, applying the same protocol described above.

### 3.5.3. Cell plasticity studies

Mesenchymal phenotype of MDA-MB-231 and MCF-7 transfected cells was analyzed by counting the percentage of mesenchymal- and epithelial-like cells as previously reported [18]. Mesenchymal phenotype cells were considered those whose morphology was slender, fusiform or fibroblast-like; while the epithelial-like phenotype cells were comprised by the remainder population. Specifically, three

different fields were counted from, at least, three independent cell preparations (in total, more than 1200 cells were counted for each transfection). Experiments were carried out by two independent observers in a blinded fashion.

### 3.5.4. Mammospheres generation assays

In order to analyze the percentage of tumor-initiating (TICs) or cancer stem-like (CSCs) cells in each cell line and the changes induced by plasmids transfections and peptide treatments, we implemented mammosphere formation assays as described by Frances L. Shaw [229]. These studies are based on the ability of cancer stem cells (CSCs) to survive and proliferate in low adherence plates forming **cell clusters** named tumorspheres. Each tumorsphere (or mammosphere in these cases) represents a colony formed from one CSC and, therefore, the number of mammospheres formed is a representation of the CSCs percentage in the cellular population.

Briefly, 6-well plates were treated with 1 ml of poly(2-hydroxyethyl methacrylate) (pHEMA), a chemical that blocks cell adhesion to the well surface, eluted in 98 % ethanol and incubated at 60ºC for two days in order to allow pHEMA to dry and appropriately cover the surface. Then, 6.000-10.000 cells (MDA-MB-231 and MCF-7 cell lines, respectively) were plated (in triplicate) in 6-well plates pretreated with pHEMA in 2ml DMEM/F12 medium without phenol red and supplemented with recombinant epithelial growth factor, B27 supplement and Gentamicin/Amphotericin B. After 7 days of incubation at 37ºC, 5 % $CO_2$, mammospheres were counted in an inverted microscope. Finally, mammospheres from each condition were collected in tubes and treated with trypsin for 5 min in order to disaggregate the mammospheres and be able to count the number of cells in each condition with a Neubauer chamber.

## 3.6. Xenografts studies

Animal maintenance and experiments were carried out following the European Regulation for Animal Care and under the approval of the University of Córdoba Research Ethics Committee. Six-week-old female athymic Swiss nu/nu mice (Charles River Laboratories, Barcelona, Spain) were subcutaneously grafted in the flank with $2x10^6$ mock- or sst5TMD4-stably transfected MCF-7 cells (n=4-5 mice per condition). Tumor growth was monitored weekly during 3 months. Each tumor was dissected and different pieces were snap-frozen for qPCR analysis or formalin-fixed and sectioned for histopathological examination after hematoxylin-eosin staining as previously described [18].

Xenografted tumors were processed for the detection of VEGF by immunohistochemistry using standard procedures [18]. Specifically, a tumoral piece was formalin-fixed and paraffin-embedded. After antigen-retrieval, sections were incubated with a specific anti-human VEGF antibody (AB-293-NA; R&D systems) and

the appropriate fluorescence-labeled secondary antibody (Donkey Anti-Goat IgG Alexa Fluor 488; ab150129; Abcam, Cambridge, UK). Signal intensity was determined using the ImageJ software [230].

Additionally, xenografted tumors were processed for the quantification of blood vessel density using a standard histopathological procedure [231]. Specifically, a formalin-fixed and paraffin-embedded tumor piece was stained with haematoxylin and eosin and vascular density was determined by counting the number of visible vessels per 20x objective tumor-full fields from a minimum of 5 fields per each of the 4-5 animals per genotype.

## 3.7. *In silico* studies

The bioinformatic analysis of SST5 and GHRL genes was carried out with the aim to unveil possible factors that could be involved in the regulation of the alternative splicing of both genes. To this end, we used different online tools in order to characterize the splicing factor target sites present in the sequence of SST5 and GHRL genes:

- **SpliceAid** (http://www.introni.it/splicing.html): A robust tool that uses a database of strictly experimentally assessed target RNA sequences in humans.
- **Human Splicing Finder** (http://www.umd.be/HSF3): An online tool that performs deep analysis in the context of splicing function, detecting canonical splice sites, branch points and enhancer and silencer sequences by statistical analysis. This software is able to estimate the global impact of different splicing factors sites in the splicing of a given exon and/or intron.

To further study the splicing factors more likely implicated in the alternative splicing of ghrelin gene, we compared the conserved motives between human and rodent sequences of In1-ghrelin and In2-ghrelin (human and mice splicing variants, respectively) by **Clustalw2** tool (http://www.ebi.ac.uk/Tools/msa/clustalw2/).

The *in silico* identification of specific target sites for known miRNAs in the sst5TMD4 sequence was performed by using two different on-line tools (miR-Base and RegRNA). The lists of putative regulatory miRNAs were integrated together and candidate miRNAs for subsequent experimental analysis were selected in accordance to pre-established criteria (identification by both tools, minimum free energy of the seed with the target mRNA and specificity for sst5TMD4 variant).

## 3.8. Statistical analysis

For *in vitro* experiments, data are expressed as mean ± SEM obtained from, at least, three separate, independent experiments carried out in different days and with different cell preparations. Data were evaluated for heterogeneity of variance using the Kolmogorov–Smirnov test. In studies in which only two experimental groups were

compared, the Student's t-test was used; however, studies with two or more experimental groups, statistical analysis was performed for analysis of variance (one-way ANOVA) followed by Dunnett's test for multiple comparisons.

For *in vivo* xenograft experiments, data are expressed as mean ± SEM obtained from n=4-5 animals. Data were evaluated for heterogeneity of variance using the Kolmogorov–Smirnov test. Statistical analysis was carried out using Student's t-test.

For the analysis of the impact of sst5TMD4 on human breast carcinoma, samples were categorized in low and high sst5TMD4 levels according to median sst5TMD4 expression levels. Differences in the expression of angiogenic markers between both groups were assessed by Student's t-test. Correlations between the expression of sst5TMD4 and angiogenic markers was assessed by Pearson's correlation test. Significant correlations between categorized sst5TMD4 mRNA/protein expression, CD34 IHC expression, presence of metastasis and disease-free survival were studied using Chi-square and Long-rank-p-value methods.

For the analysis of the impact of In1-ghrelin on human breast carcinoma, samples were categorized in low, moderate, and high In1-ghrelin levels according to quartile In1-ghrelin expression levels (0-50% as low, 50-75% as moderate and 75-100% as high). Significant correlations between categorized In1-ghrelin mRNA expression, presence of metastasis and disease-free survival were studied using Chi-square and Long-rank-p-value methods.

Finally, Chi-Square test was used to compare the differences between the genomic frequencies of the different haplotypes of sst5 gene.

Statistical analyses were carried out with GraphPad Prism 6 (Graphpad Software, La Jolla, CA, USA). P-values smaller than 0.05 were considered statistically significant (* $p < 0.05$, ** $p < 0.01$; *** $p < 0.001$).

# Results

# 4. Results

The results presented in this Thesis have been subdivided in two independent but parallel sections for each splicing variant of interest (sst5TMD4 and In1-ghrelin). Within each section, we have analyzed some of the processes that could be regulating the appearance of each splicing variant, sst5TMD4 and In1-ghrelin, and the functional effects of these variants on tumor-related features.

## 4.1. sst5TMD4

### 4.1.1. Genesis and regulation of sst5TMD4

Splicing process has been shown to be tightly regulated through several overlapping mechanisms [124]. In order to analyze the possible mechanisms involved in the transcriptional regulation of sst5TMD4, we implemented a multidisciplinary strategy to unveil the molecular regulatory systems involved in the appearance and/or regulation of this splicing variant.

As mentioned above, genes encoding human SST receptors lack introns, except for sst2. Therefore, the splicing process involved in the appearance of sst5TMD4 should implicate non-canonical mechanisms that promote the splicing of intron-lacking genes. Specifically, there is a number of mechanisms by which sequences lacking canonical introns can undergo splicing processes [93, 124]. Among the possible mechanisms that could be involved in the genesis and regulation of sst5TMD4 expression, we explored: 1) the existence of SNPs in the cryptic intron of the SST5 sequence; 2) splicing factors that could be involved in the differential expression between the native receptor and the truncated sst5TMD4 form; and 3) the existence of miRNAs that could be involved in the regulation of sst5TMD4 expression.

4.1.1.1. Single nucleotide polymorphisms (SNPs) or *de novo* mutation affecting splicing process

Single nucleotide polymorphisms or SNPs, considered as demographically described allelic variability characterized by a shift of one nucleotide in the genomic sequence, as well as tumor-associated *de novo* mutations, are factors that have been previously associated with the generation of alternative splicing variants [232] and cancer progression [233]. Indeed, these phenomena have been described to be associated to modifications of transcription factor target sequences or alterations in the binding sequences of spliceosome components, which could give rise to the appearance of new splicing variants by generating splicing enhancer motifs or new spliceosome recruiting sequences [54]. These genomic modifications can interfere with the normal splicing, inducing the use of alternative elements of the gene, both exons and introns,

which can misbalance the ratio among different splicing variants or even induce the appearance of novel transcripts [234].

Unusually, the SST5 gene belongs to a small group of genes that are composed of a single coding exon and, therefore, does not present canonical intron sequences within the CDS. Consequently, and although it could be assumed a lack of canonic recognition sites for the spliceosome machinery in this sequence, it is also possible that certain SNPs or *de novo* mutation could generate the appearance of novel canonic splicing associated sequences.

Bearing this idea in mind, and based on previous studies indicating a clear overexpression of sst5TMD4 in endocrine-related tumors, such as breast cancer [18], we explored the sequence of the SST5 gene in a battery of genomic DNA from tumoral and healthy breast samples with different relative expression of both, the full-length canonical sst5 and the alternative spliced sst5TMD4, trying to find a haplotype that could correlate with increased sst5TMD4 expression.

Accordingly, we designed and validated specific primers to amplify the 5' and 3' ends of the cryptic intron spliced to generate the sst5TMD4 variant (**Figure R1**), and sequenced the resultant PCRs. Surprisingly, we only found variations in the 3' side of the genomic sequence, where two previously reported SNPs were found among the samples analyzed. These SNPs were previously described in PubMed database:

- rs197055: There are 3 different forms described for this SNP (A/C/T). However, in our studies we only found C and A variants.
- rs12599155: There are 2 different forms described (C and T). We found both of them in ours samples.

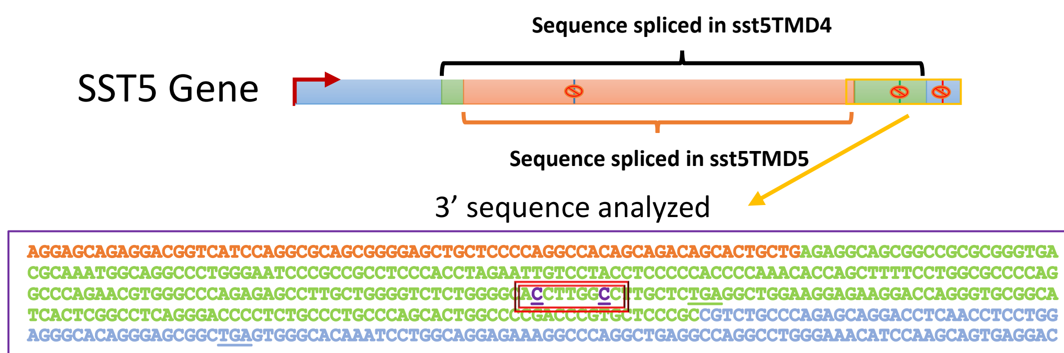


***Figure R1**: Representative scheme of SST5 gene highlighting the cryptic intron eliminated in sst5TMD4 and sst5TMD5. Sequenced 3' region is indicated by a yellow box. SNPs detected are located in the red box and indicated in purple.*

To analyze the putative implication of these SNPs in the regulation of SST5 splicing and, therefore, the generation of sst5TMD4, we compared the frequencies observed for each SNP between the different types of samples (healthy or tumoral) and between the tumoral samples with high or low sst5TMD4 expression (stratified according to the median expression level of sst5TMD4 in those samples) (Table R1).

| Sample | Type of sample | Copy number adjusted by b-act | | Genotype | |
|---|---|---|---|---|---|
| | | Full-length sst5 | sst5TMD4 | rs197055 | rs12599155 |
| 1 | Non-Tumoral | 0,00014920 | 0,00005037 | CC | CC |
| 2 | Non-Tumoral | 0,00006490 | 0,00009424 | | |
| 3 | Non-Tumoral | 0,00005073 | 0,00001687 | | |
| 4 | Non-Tumoral | 0,00002804 | 0,00002148 | | |
| 5 | Non-Tumoral | 0,00002051 | 0,00006526 | | |
| 6 | Non-Tumoral | 0,00001082 | 0,00008649 | | |
| 7 | Non-Tumoral | 0,00000960 | 0,00010762 | | |
| 8 | Non-Tumoral | 0,00000000 | 0,00009539 | | |
| 9 | Tumoral | 0,00038924 | 0,00000000 | | |
| 10 | Tumoral | 0,00009678 | 0,00001126 | | |
| 11 | Tumoral | 0,00007172 | 0,00003724 | | |
| 12 | Tumoral | 0,00004295 | 0,00011965 | | |
| 13 | Tumoral | 0,00002731 | 0,00001409 | | |
| 14 | Tumoral | 0,00001582 | 0,00015812 | | |
| 15 | Tumoral | 0,00000291 | 0,00001798 | | |
| 16 | Tumoral | 0,00000000 | 0,00104099 | | |
| 17 | Tumoral | 0,00000000 | 0,00021700 | | |
| 18 | Tumoral | 0,00000000 | 0,00005113 | | |
| 19 | Non-Tumoral | 0,00025930 | 0,00004285 | CC | CT |
| 20 | Non-Tumoral | 0,00020676 | 0,00019620 | | |
| 21 | Non-Tumoral | 0,00008139 | 0,00027384 | | |
| 22 | Non-Tumoral | 0,00007464 | 0,00003857 | | |
| 23 | Non-Tumoral | 0,00005244 | 0,00013085 | | |
| 24 | Non-Tumoral | 0,00002958 | 0,00006113 | | |
| 25 | Non-Tumoral | 0,00001540 | 0,00007653 | | |
| 26 | Non-Tumoral | 0,00000000 | 0,00006133 | | |
| 27 | Non-Tumoral | 0,00000000 | 0,00005646 | | |
| 28 | Tumoral | 0,00002823 | 0,00009872 | | |
| 29 | Tumoral | 0,00000000 | 0,00008203 | | |
| 30 | Non-Tumoral | 0,00003564 | 0,00011319 | CC | TT |
| 31 | Non-Tumoral | 0,00000829 | 0,00039630 | | |
| 32 | Non-Tumoral | 0,00000000 | 0,00006768 | | |
| 33 | Tumoral | 0,00027901 | 0,00010919 | | |
| 34 | Tumoral | 0,00003379 | 0,00000000 | | |
| 35 | Tumoral | 0,00000480 | 0,00004192 | | |
| 36 | Tumoral | 0,00000000 | 0,00006342 | | |
| 37 | Tumoral | 0,00085406 | 0,00000000 | CA | CC |
| 38 | Tumoral | 0,00000132 | 0,00000850 | | |
| 39 | Tumoral | 0,00004381 | 0,00008253 | CA | CT |
| 40 | Tumoral | 0,00003564 | 0,00011041 | | |
| 41 | Tumoral | 0,00000333 | 0,00003188 | | |
| 42 | Tumoral | 0,00000000 | 0,00001508 | | |
| 43 | Tumoral | 0,00029427 | 0,00067656 | CA | TT |
| 44 | Tumoral | 0,00000000 | 0,00037820 | AA | CT |
| 45 | Tumoral | 0,00119149 | 0,00095341 | AA | TT |

*Table R1: **List of breast samples used in SST5 sequencing studies.** Columns indicate the type of breast sample, the sst5TMD4 and full-length expression levels and the genotypes o both SNPs.*

Subsequently, the frequencies of rs197055 variants were analyzed by Chi-Square for trends analysis. As can be observed in **Table R2** and **Figure R1**, this SNP showed significant differences in the genotypic frequencies between control healthy (non-tumoral) and tumoral samples, being the C/A and A/A genotypes clearly associated to the tumoral samples. More interestingly, although this analysis did not indicate significant differences between the genotype of samples with high and low expression of sst5TMD4, these results showed that the homozygous AA genotype was only present in those tumoral samples with high sst5TMD4 expression (Figure R1). Remarkably, the presence and/or expression of the full-length sst5 variant was not associated with rs197055 SNP variability.

| Analysis of rs197055 | | Value | df | p-value |
|---|---|---|---|---|
| Tumoral vs. non-tumoral | Chi-Square for Trends<br>N of Valid Cases | 7.863<br>45 | 1 | 0.005 |
| Tumoral samples with low vs. high sst5TMD4 expression | Chi-Square for Trends<br>N of Valid Cases | 0.6461<br>25 | 1 | 0.421 |

***Table R2:*** *Chi-Square tests for trends analyzing differences in genotype frequencies of rs197055 SST5 SNP between tumoral and non-tumoral samples and between tumoral samples with high and low sst5TMD4 expression.*

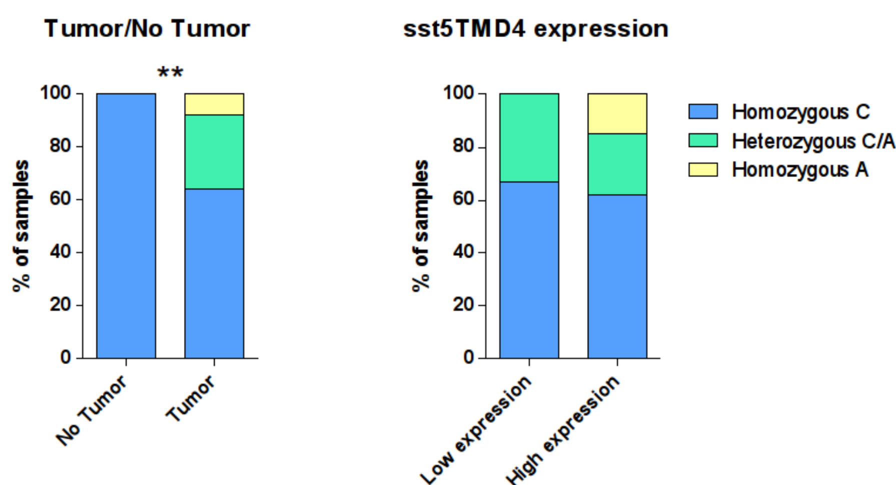


***Figure R1: Analysis of rs197055 frequencies.*** *Left graph shows the frequencies of the different genotypes in tumoral and non-tumoral samples. Right graph represents the frequencies within tumoral samples with high and low sst5TMD4 expression. Asterisks (**, p<0.01) indicate significant differences between categories.*

Conversely, the frequencies found for rs12599155 variants did not show any significant difference by Chi-square test for trends comparing tumoral vs. non-tumoral samples (**Table R3** and **Figure R2**). However, when analyzing the tumoral samples, rs12599155 variability seemed to be associated to the presence of sst5TMD4, in that tumoral samples with high or low sst5TMD4 expression showed a trend (p=0.066) to present different rs12599155 frequencies (**Table R3** and **Figure R2**). Indeed, tumoral samples

with high sst5TMD4 levels presented more frequently the C/T heterozygous and T/T homozygous genotypes, which could be associated to the presence of the truncated receptor.

| Analysis of rs12599155 | | Value | df | P value |
|---|---|---|---|---|
| Tumoral vs. non-tumoral | Chi-Square for Trends<br>N of Valid Cases | 0.0019<br>45 | 1 | 0.965 |
| Tumoral samples with low vs. high sst5TMD4 expression | Chi-Square for Trends<br>N of Valid Cases | 3.371<br>25 | 1 | 0.066 |

***Table R3** Chi-Square tests for trends analyzing differences in genotype frequencies of rs12599155 SST5 SNP between tumoral and non-tumoral samples and between tumoral samples with high and low sst5TMD4 expression*

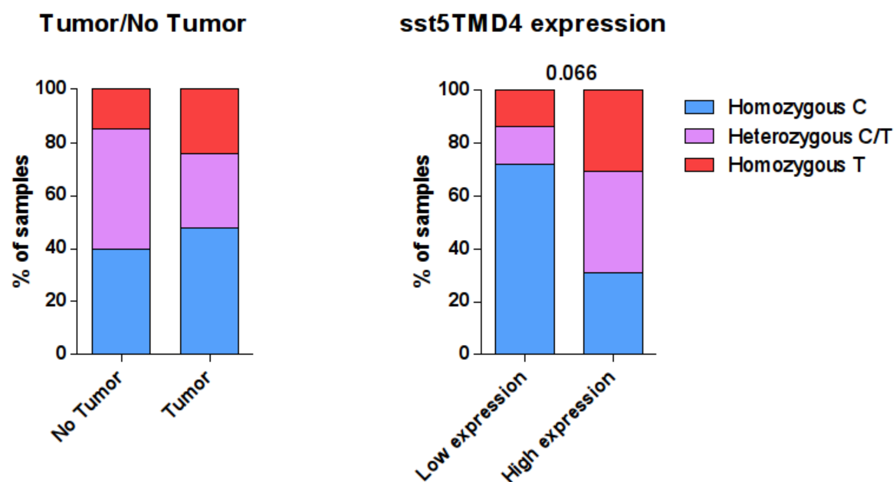


***Figure R2: Analysis of rs12599155 frequencies.** Left graph shows the frequencies of the different genotypes in tumoral and non-tumoral samples. Right graph represents the frequencies within tumoral samples with high and low sst5TMD4 expression.*

4.1.1.2. <u>Splicing factors involved in the splicing process</u>

Enhancer or silencer splicing factors are common regulators of splicing processes [81]. The binding of these proteins to a nascent mRNA regulates the splicing process triggered by the spliceosome. Therefore, depending on the splicing factors involved in the transcription/maturation process, the nascent transcript could generate a different splicing variant. In order to analyze the splicing factors that could play a role in the alternative splicing of the SST5 gene, and the consequent generation of sst5TMD4, we implemented a study of the splicing factor target sequences present in the SST5 cryptic intron, by using the online software SpliceAid and Human Splicing Finder.

As a result of the analysis with the SpliceAid software, we identified the presence of numerous splicing factor target sequences throughout the cryptic intron, including both exonic/intronic splicing enhancer (ESE and ISE) and exonic/intronic splicing silencer (ESS and ISS). In particular, this study revealed the presence of target

sequences for 38 different splicing factors, 17 enhancers and 21 silencers, illustratively represented in the **Figure R3** and detailed in **Table R4**.

| | Splicing Factor | Nº of target sites | Splicing Factor | Nº of target sites | |
|---|---|---|---|---|---|
| Enhancers | 9G8 (SRSF7) | 3 | CUG-BP1 | 2 | Silencers |
| | ETR-3 | 11 | DAZAP1 | 1 | |
| | Fox-1 | 1 | hnRNP A1 | 16 | |
| | Fox-2 | 1 | hnRNP A2/B1 | 2 | |
| | HTra2β1 (SRSF10) | 2 | hnRNP E2 | 2 | |
| | Nova-1 | 3 | hnRNP F | 58 | |
| | Nova-2 | 3 | hnRNP H1 | 71 | |
| | RBM4 | 1 | hnRNP H2 | 69 | |
| | SC35 (SRSF2) | 16 | hnRNP H3 | 61 | |
| | SF2/ASF (SRSF1) | 27 | hnRNP I (PTB) | 8 | |
| | SRp20 (SRSF3) | 11 | hnRNP K | 11 | |
| | SRp30c (SRSF9) | 31 | hnRNP P (TLS) | 13 | |
| | SRp40 (SRSF5) | 60 | KSRP | 6 | |
| | SRp55 (SRSF6) | 16 | MBNL1 | 36 | |
| | TIA-1 | 4 | nPTB | 1 | |
| | | | PSF | 3 | |
| | | | RBM5 | 4 | |
| | | | SF1 | 2 | |
| | | | TDP43 | 1 | |
| | | | ZRANB2 | 9 | |

***Table R4**: Splicing factor target sites identified within the cryptic intron of the SST5 gene by the SpliceAid software, and classified as enhancers or silencers, indicating the number of target sites found for each one.*

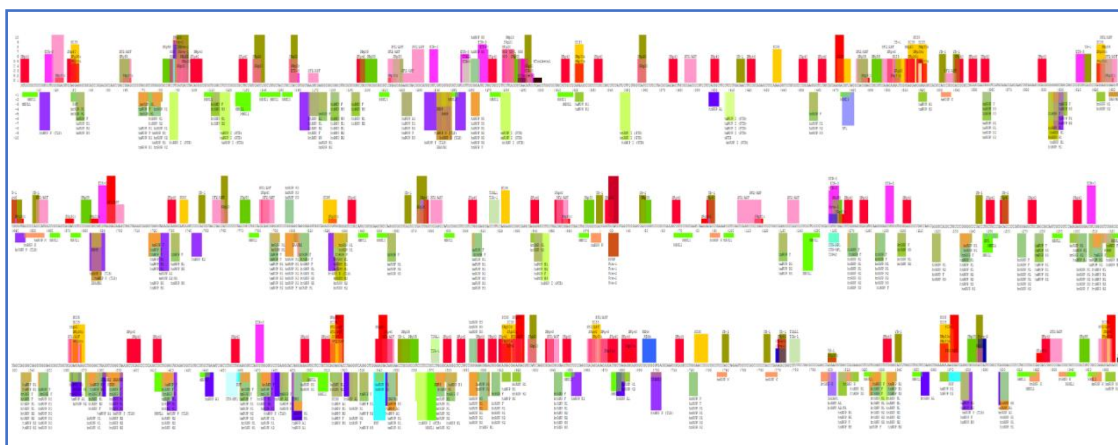

***Figure R3:** Distribution of splicing factors target sequences in the cryptic intron of SST5 gene generated by the analysis with SpliceAid software*

Thus, this *in silico* analysis revealed the presence of multiple binding sites (even more than 30) for numerous enhancer and silencer splicing factors (**Table R5**). Of note, several of these splicing factors are well known by their implication in the regulation of the alternative splicing of several genes (SRSF1, SRSF3, SRSF6, CUGBP, PTB, PSF, Nova-

1) and, hence, could represent crucial players in the regulation of SST5 alternative splicing.

Nevertheless, the global, final effect of the combined actions of these splicing factor target sequences would depend, ultimately, on the actual set of splicing factors present during the transcription process, and the interactions established among them. In spite of this, using the Human Splicing Finder tool, it is possible to estimate a relative measure of the overall splicing tendency, based on the number of splicing factor target sites along the cryptic sequence for each splicing factor and their relative strength. Of note, this approach demonstrated a high number and density of silencer splicing factor target sequences along the cryptic intron, as compared with the rest of the gene sequence (green bars in **Figure R4**), which could suggest a strong tendency to maintain this cryptic intron in the SST5 mature mRNA under normal, physiologically conditions. However, it can be also observed that the relative strength among ESE and ESS presents a strong peak of enhancer octamers at the end of the cryptic intron, which would suggest that these regions could be specially relevant for the elimination of the cryptic intron and the generation of sst5TMD4 mRNA.

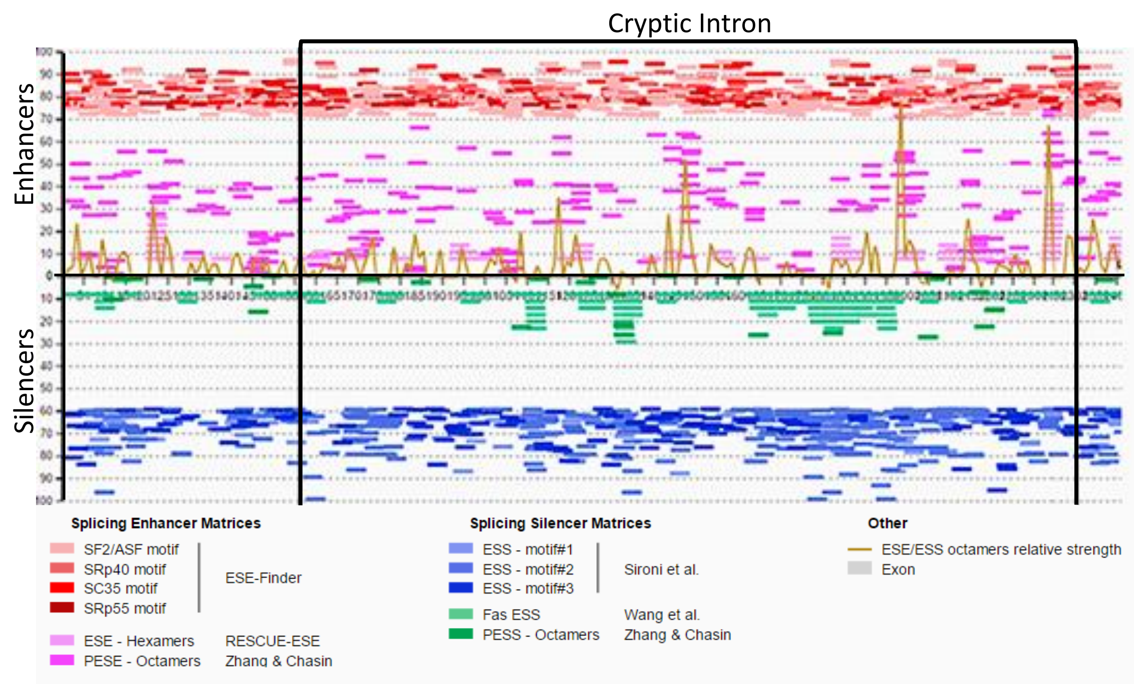


***Figure R4***: *Representative picture of the analysis of the presence of splicing factors target sequences in SST5 gene. Cryptic intron location is marked by two black vertical lines*

We further analyzed the putative effects of the SNPs identified in the previous section on the generation/elimination of novel ESS or ESE in the cryptic intron by applying a similar strategy. Results generated were somehow unexpected since, on one side, rs12599155, which exhibited a significant differential frequency between tumoral and non-tumoral samples, induced the disappearance of an exonic regulatory sequence (GCCTTG) that is predicted as an ESE [235]. On the other hand, the analyses of rs197055 SNP, which presents a near-significant trend for a differential frequency

between high and low sst5TMD4 expressors —differences that are also observed in the subset of tumoral samples— revealed that this nucleotide change induces the disappearance of a target sequence for the ESE YB-1 (**Figure R5**), a splicing factor that has been associated with the stimulation of exon inclusion, by enhancing the recruitment of U2AF to weak polypyrimidine tracts [236]. In addition, rs197055 variation generated a new weak branch point near to a stronger one.

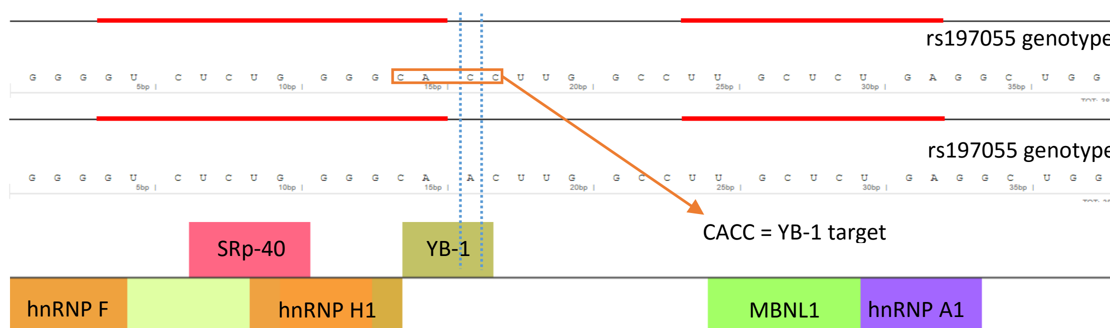


***Figure R5:*** *Elimination of YB-1 target sequence in rs197055 genotype A within SST5 cryptic intron*

### 4.1.1.3. Regulation of sst5TMD4 transcription and splicing by miRNAs

To investigate the putative regulation of sst5TMD4 generation by miRNAs, we first explored the miRNA target sequences present in the specific region of this alternative splicing variant by using two distinct web-based prediction tools (miRBase and RegRNA). This *in silico* approach revealed the existence of several putative miRNAs target sites in the sst5TMD4 sequence (**Table R5**). From the list of miRNAs identified herein as candidates to regulate the generation and/or stability of sst5TMD4, we selected a subset based on several criteria, including their identification by both tools, their specificity for sst5TMD4, or their MFE (minimum free energy, which estimates the stability of the miRNA-target mRNA duplex). Specifically, we finally selected hsa-miR-189, hsa-miR-326, hsa-miR-346, hsa-miR-339 and hsa-miR-939 for further analysis.

| miRNA | Location | Program used |
|---|---|---|
| hsa-miR-189 | sst5TMD4 3' UTR | RegRNA |
| hsa-miR-326 | Spanning | RegRNA |
| hsa-miR-339 | Spanning | miRBase & RegRNA |
| hsa-miR-346 | Spanning | miRBase |
| hsa-mir-512-5p | sst5TMD4 3' UTR | RegRNA |
| hsa-miR-665 | sst5TMD4 3' UTR | miRBase |
| hsa-miR-708 | sst5TMD4 3' UTR | miRBase |
| hsa-miR-939 | sst5TMD4 3' UTR | miRBase |

***Table R5****: miRNA target sites identified in sst5TMD4 sequence with miRBase and RegRNA online tools*

In order to explore the putative association of these miRNAs with sst5TMD4 expression, we measured by qPCR the expression of those selected miRNAs in a battery of 16 tumoral and healthy human mammary gland biopsies. From this analysis, we found that all miRNAs were present at similar levels, with hsa-miR-189 showing the higher level of expression and hsa-miR-939 being the less expressed. The analysis of the expression of each miRNA compared with those of sst5TMD4 showed that three of them correlated with sst5TMD4 expression, but did not show the same tendencies. Indeed, hsa-miR-346 showed a clearly negative correlation with sst5TMD4 (p=0.045 and $r^2$=-0.38); while hsa-miR-326 (p=0.005 and $r^2$=0.53) and hsa-miR-189 (p=0.084 and $r^2$=0.25) showed a positive correlation, indicating that their effects, if so, could not be synonymous (**Figure R6**).

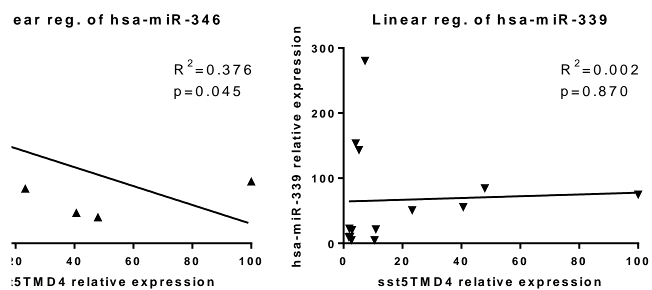


***Figure R6:*** *Correlation analysis, by linear regression, between sst5TMD4 mRNA expression levels and each miRNA (adjusted by $\beta$-Actin and RNU9, respectively) in breast cancer samples. (n=14)*

#### 4.1.1.3.1. Effect of hsa-miR-346 on sst5TMD4 expression

As the usual mode of action of a given miRNA is to bind to its mRNA target and induce its silencing or degradation through DICER complex [237], it would be expected that the relationship between the expression of a miRNA and that of its target mRNA would be inverse. Therefore, since among the miRNAs analyzed in the previous section only hsa-miR-346 presented a negative correlation with sst5TMD4, we mainly focused our attention on exploring the putative consequences of this miRNA on sst5TMD4 expression.

Firstly, to further explore the relationship between the selected miRNAs and sst5TMD4 expression, we analyzed the expression pattern of these miRNAs in the breast cancer cell line MCF-7. It is important to emphasize that we have previously reported that sst5TMD4 expression in MCF-7 is progressively lost through cell-passages [18], which provides a unique opportunity to explore the mechanisms involved in sst5TMD4 expression regulation. For this reason, the expression of the candidate miRNAs was first analyzed in early and late passages of MCF-7 cells. As shown in **Figure R7**, only hsa-miR-346 expression was significantly increased in MCF-7 cells from late passages as compared to early passages, thereby implying a negative correlation with sst5TMD4 expression, which could thus indicate a putative regulatory function of hsa-miR-346 on sst5TMD4 expression.

***Figure R7:*** *Relative expression levels of each miRNA at initial and final passages of MCF-7 cells, adjusted to the expression levels of the initial passages. Data represent mean ± SEM (n=3-6). Asterisks (**, p<0.01) indicate significant differences between initial and final passages miRNA expression levels.*

Base on these findings, we next studied the functional effects of hsa-miR-346 on sst5TMD4 expression. To this end, we used commercially available hsa-miR-346 mimic and inhibitor, which are chemically modified, double-stranded miRNA-like RNA designed to resemble the functionality of or inhibit, respectively, mature endogenous miRNA upon transfection. These hsa-miR-346 mimic and inhibitor were used to transfect the MCF-7 cell line and to determine sst5TMD4 expression at both RNA and protein levels, as compared with scrambled miRNA-transfected cells, which served as controls. As shown in **Figure R8**, mimic miRNA induced a clear decrease on sst5TMD4 protein levels, and a nearly-significant inhibition trend on sst5TMD4 mRNA levels, indicating that hsa-miR-346 can indeed regulate sst5TMD4 expression and/or stability.

***Figure R8****: sst5TMD4 mRNA and protein expression (n=8 and n=6, respectively) after hsa-miR-346 mimic and inhibitor transfection as compared with scrambled miRNA transfection in MCF-7 cell line. Data represent mean +/- SEM. Asterisks (**, p<0.01) indicate significant differences in sst5TMD4 expression levels compared with those of scramble-transfected MCF-7 cells.*

### 4.1.2. Functional role of sst5TMD4 in breast cancer

In previous studies, our group demonstrated that the sst5TMD4 splicing variant is overexpressed in human breast cancer samples compared with normal breast tissue [18]. Moreover, sst5TMD4 showed the ability to enhance the malignant phenotype in a cell line model (MCF-7), where overexpression of this splicing variant induced a higher rate of proliferation, migration, and invasion and, even, induced relevant changes related to epithelial to mesenchymal transition (EMT) [18].

In order to investigate new functional features of sst5TMD4, which were previously unexplored [18], and to further understand the molecular determinants underlying the alterations induced by sst5TMD4, we implemented new approaches, by exploring the changes in the gene expression pattern, and in specific signaling pathways induced by the overexpression of this variant in the breast cancer cell line MCF-7.

#### 4.1.2.1. Changes in gene expression pattern induced by sst5TMD4 overexpression

In order to explore the cellular and molecular mechanisms underlying the association between sst5TMD4 and breast cancer malignancy, a gene expression microarray was carried out using sst5TMD4 overexpressing MCF-7 cells, and compared with mock MCF-7 controls. This gene expression microarray revealed the existence of an elevated number of genes altered by sst5TMD4 overexpression (38% up-regulated and 62% down-regulated). Indeed, a software-driven functional-enrichment analysis of these data indicated that sst5TMD4 overexpression in MCF-7 altered the expression of numerous genes involved in several cellular processes such as epithelial-to-mesenchymal transition (EMT; 10% of altered genes with known function), cell growth (6%), cell metabolism (6%), signal transduction (13%) or, even, angiogenesis (13%) (**Figure R9**).

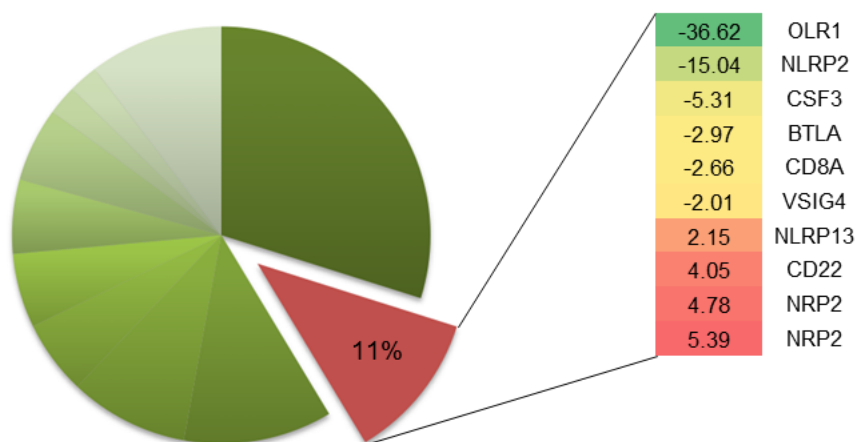


***Figure R9:*** *Representative picture of software-driven functional analysis of genes whose expression is altered by the presence of sst5TMD4 in MCF-7 cells by gene expression microarray (green = inhibition, red = overexpression).*

We have previously demonstrated that sst5TMD4 overexpression alters EMT and cell growth (among other cellular processes) in MCF-7 cell line [18]; thence, these new results served to reinforce those previous studies.

Interestingly, this approach also revealed a putative implication of sst5TMD4 in the angiogenic process, which had been previously unexplored. Indeed, supporting these results, a user-driven functional-enrichment analysis of the data generated unveiled that 31 out of 78 genes with known function (40%) significantly altered by the presence of sst5TMD4 were associated, somehow, to the angiogenic process (**Figure R10**).

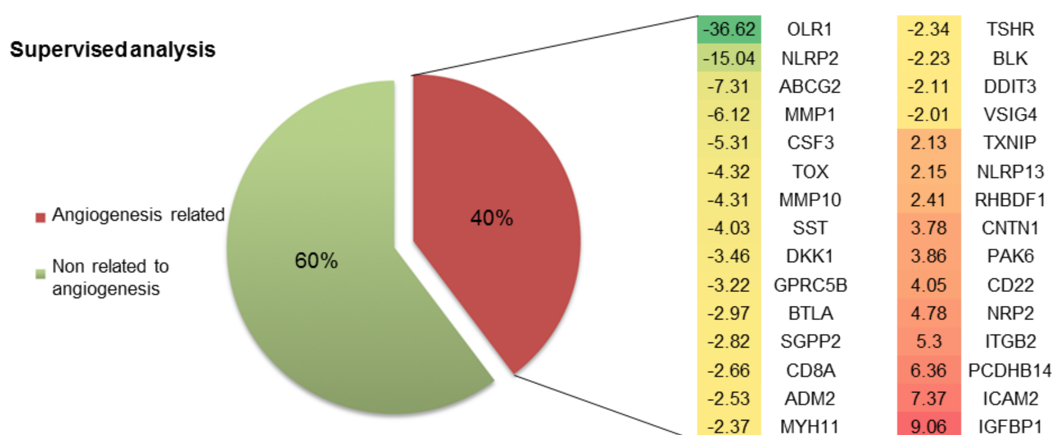


***Figure R10:*** *Representative picture of user-driven supervised functional analysis of genes whose expression is altered by the presence of sst5TMD4 in MCF-7 cells by gene expression microarray (green = inhibition, red = overexpression).*

Further validation of the genes found to be altered in the gene expression array by qPCR confirmed the alteration of several angiogenesis-related genes in sst5TMD4-overexpressing MCF-7 cells, including the overexpression of ITGB2 or IGFBP1 (**Figure R11**).

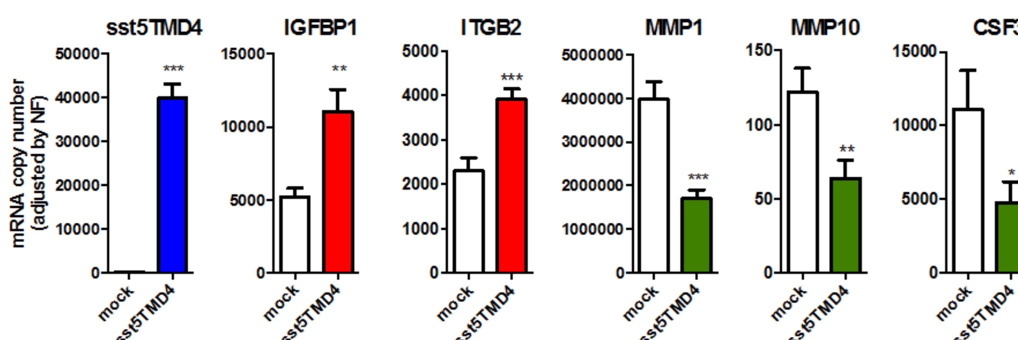


***Figure R11:*** *Validation of sst5TMD4 overexpression and examples of sst5TMD4-induced gene expression changes validated by additional qPCR in transfected cell lines. Data represent mean +/- SEM (n=14). Asterisks (*, p<0.05**, p<0.01; ***, p<0.001) indicate significant differences between mock and sst5TMD4 transfected cells.*

4.1.2.1.1. sst5TMD4 increased the production of proangiogenic factors in MCF-7 cells

Angiogenic process is driven by the expression and secretion of certain pro-angiogenic factors, wherein some of them are usually involved in cancer progression. Although multiple growth factors regulate the formation of new vessels, VEGF [238], EGF [239] and the angiopoietins Ang-1 and Ang-2 [240] are especially relevant. Therefore, we analyzed the expression of these relevant growth factors with angiogenic activity in sst5TMD4-expressing MCF-7 cells. In particular, sst5TMD4-transfected cells showed elevated levels of VEGF ($p<0.05$), EGF ($p<0.05$) and Ang1 ($p<0.001$), with no changes in Ang2 expression, compared with mock-transfected cells (**Figure R12**). However, this elevation in the expression of pro-angiogenic factors was not accompanied by an increase in the expression of hypoxia-induced genes such as HIF-1a or HIF-1b, whose elevation generally precedes those of the pro-angiogenic factors. Nevertheless, changes in protein expression or phosphorylation cannot be ruled out.

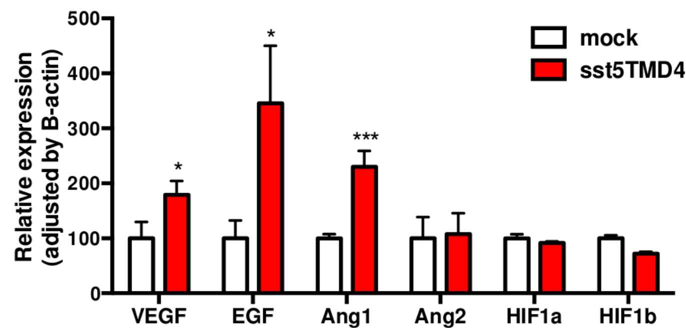


***Figure R12:*** *mRNA expression levels of different proangiogenic factors in sst5TMD4- and mock-transfected MCF-7 cells measured by qPCR. Data represent mean +/- SEM. Asterisks (*, p<0,05; ***, p<0.001) indicate significant differences in mRNA expression levels between mock and sst5TMD4 transfected cells.*

Consistently, MCF-7 cells at initial passages, which had high levels of sst5TMD4 expression, showed high expression levels of two of the most relevant angiogenic factors (VEGF and EGF) compared with MCF-7 from advanced passages (which had virtually lost sst5TMD4 expression) (**Figure R13**).

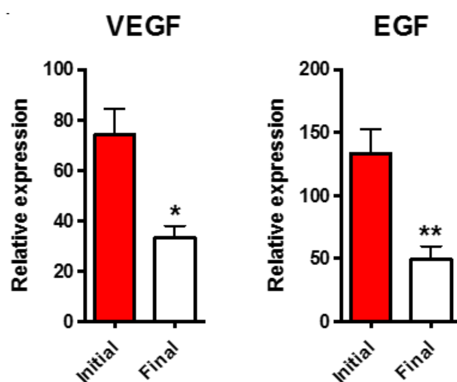


***Figure R13:*** *VEGF and EGF relative mRNA expression levels (adjusted by $\beta$-actin) in MCF-7 cells at initial and final passages. Data represent mean +/- SEM (n=3-6). Asterisks (*, p<0,05; **, p<0.01) indicate significant differences between initial and final passages mRNA expression levels of EGF and VEGF.*

In line with these results, ELISA measurements revealed higher levels of secreted VEGF in culture medium derived from sst5TMD4-transfected cells, as compared to media from mock controls (**Figure R14**), which further demonstrated the implication of sst5TMD4 in the higher expression and secretion of angiogenic factors in MCF-7 cells.

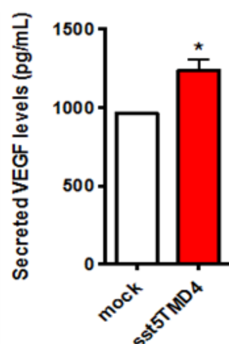


***Figure R14:*** *VEGF secretion measured by ELISA in sst5TMD4- and mock-transfected MCF-7 cells. Data represent mean +/- SEM (n=4). Asterisks (*, p<0,05) indicate significant differences in VEGF secretion between mock and sst5TMD4 transfected MCF-7 cells.*

4.1.2.1.2. sst5TMD4 increased *in vivo* VEGF production and angiogenic features

In order to confirm if there was a relevant association between the expression of the truncated receptor and the angiogenic process in a preclinical model, we analyzed the expression of pro-angiogenic factors in xenograft tumors previously generated by the inoculation of sst5TMD4-overexpressing MCF-7 cells [18]. Consistent with the results obtained *in vitro*, these *in vivo* xenograft tumors induced by sst5TMD4-transfected cells showed elevated VEGF and EGF mRNA expression levels (by qPCR) and increased VEGF protein (detected by western-blot and immunohistochemistry) compared to tumors generated by the inoculation of mock-transfected cells (**Figure R15**).

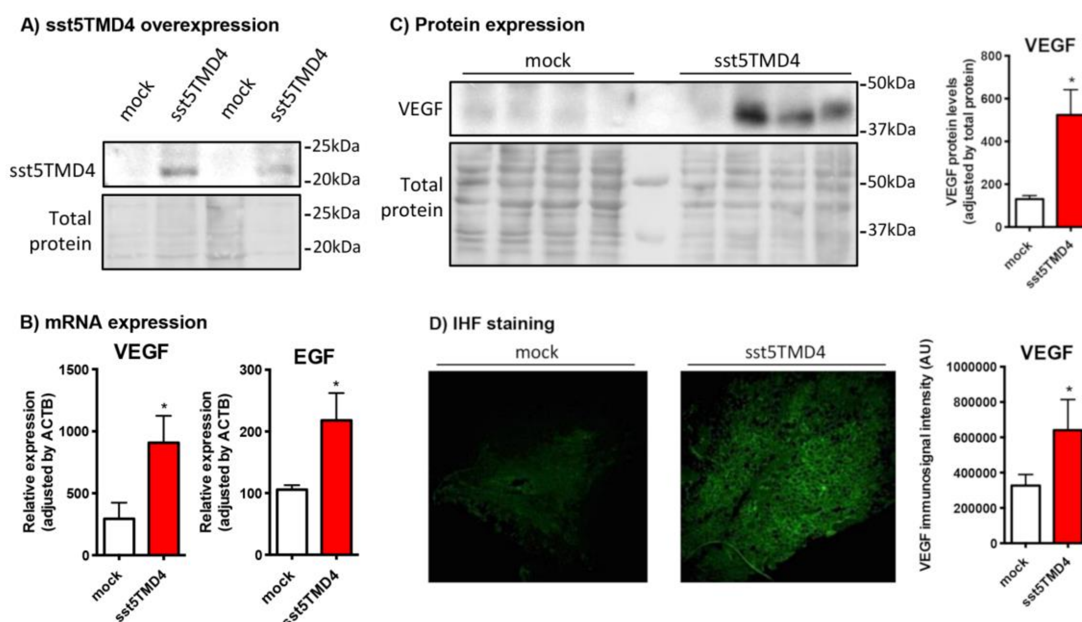


***Figure R15. A)*** *sst5TMD4 protein expression by western blotting in mock- and sst5TMD4-MCF-7 xenografted tumors.* ***B)*** *VEGF and EGF mRNA expression levels in xenografted tumors derived from MCF-7 cells overexpressing sst5TMD4.* ***C*** *and* ***D)*** *VEGF protein expression by western blot and IHC in mock- and sst5TMD4-MCF-7 xenografted tumors. Data represent mean ± SEM (n=4-5). Asterisks (*, p<0,05) indicate significant differences in protein levels between mock and sst5TMD4 transfected MCF-7 cells.*

In addition, according with these results, tumors induced by sst5TMD4-overexpressing MCF-7 cells exhibited a clearly distinct phenotype with a significantly increased number of blood vessels per field (p<0.05) (**Figure R16**)

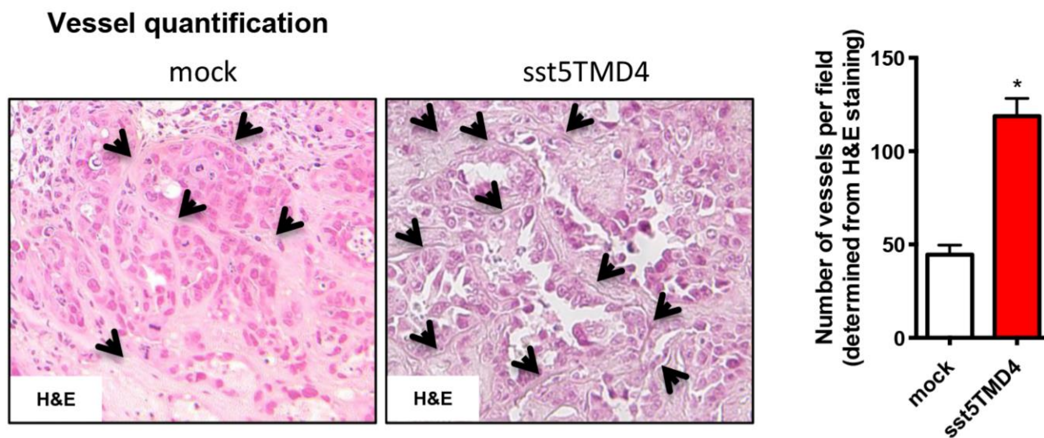


***Figure R16:*** *Representative images (x20) and quantification of straight blood vessels in xenografts derived from mock- and sst5TMD4-MCF-7 cells. Data represent mean ± SEM (n=4-5). Asterisks (*, p<0,05) indicate significant differences in the number of vessels per field between mock- and sst5TMD4-MCF-7 xenografted tumors.*

4.1.2.1.3. sst5TMD4 correlated with poor clinical outcome in breast cancer patients.

In order to explore the putative clinical consequences of sst5TMD4 presence, the expression of this truncated receptor and several tumoral markers was determined by qPCR in a battery of 127 grade 3 IDC tumors resected in 2003-04. Patients were categorized as low or high sst5TMD4 expression levels according to the median sst5TMD4 expression (**Figure R17**) in order to further analyze the putative association between sst5TMD4 presence and angiogenic markers and clinical data.

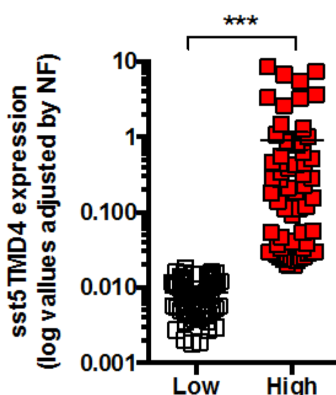


***Figure R17:*** *sst5TMD4 expression levels in breast cancer samples distributed in low and high sst5TMD4 expression. Asterisks (***, p<0.001) indicate significant differences in sst5TMD4 expression levels between samples with low and high sst5TMD4 expression.*

This analysis revealed that the group of breast carcinomas with a high sst5TMD4 expression presented higher mean expression levels of VEGF (p<0.05), Ang1 (p<0,01) and a clear tendency in CD34 (p=0.068). Moreover, there was an evident correlation between the expression levels of the truncated receptor sst5TMD4 and the expression of VEGF, EGF and Ang1 in the tumoral piece (**Figure R18**).

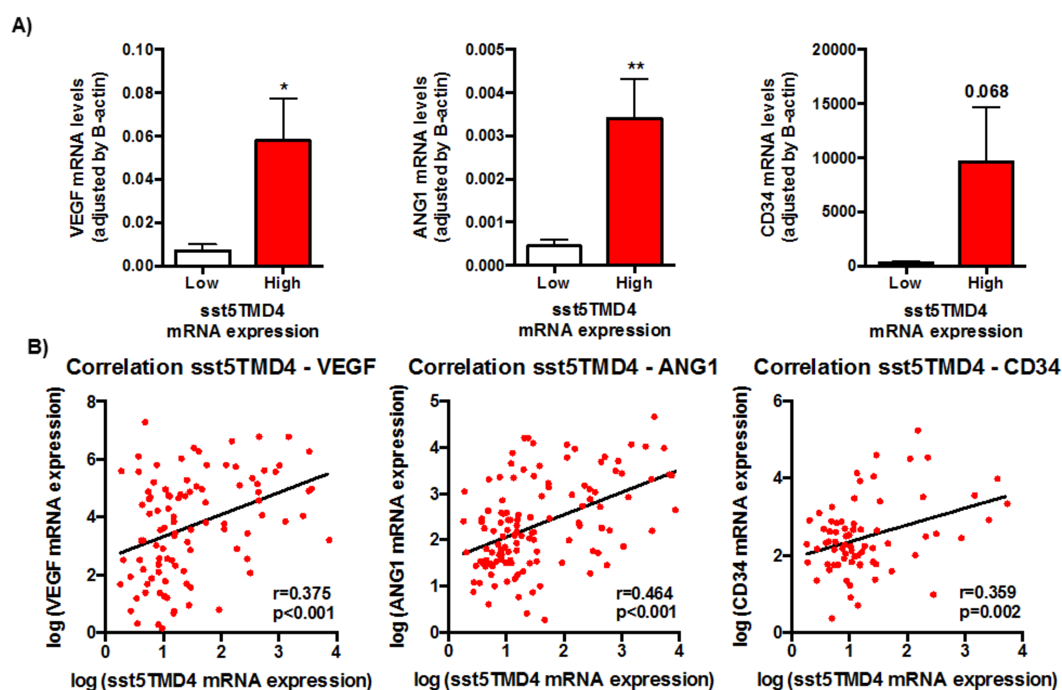


***Figure R18:*** *VEGF, Angiopoietin 1 (ANG1) and CD34 mRNA expression levels in samples with low and high sst5TMD4 levels* ***(A)*** *and their correlation with sst5TMD4 levels in breast cancer tumoral samples* ***(B)****. Data in* ***(A)*** *represent mean ± SEM. Asterisks (*, p<0,05; **, p<0.01) indicate significant differences in mRNA expression between low and high sst5TMD4 breast cancer patients.*

Of note, our data indicated that the presence of the truncated receptor sst5TMD4 was associated to a greater probability to develop metastasis, in that a higher proportion of the breast cancers that underwent lymphatic metastasis presented high sst5TMD4 expression (p=0.021), and a parallel non-significant trend was found for distant metastasis (p=0.092). Most importantly, sst5TMD4 expression was also associated to disease-free survival in breast cancer patients, in that patients with high sst5TMD4 expression exhibited a clearly lower disease-free survival (p=0.015) (**Figure R19**).

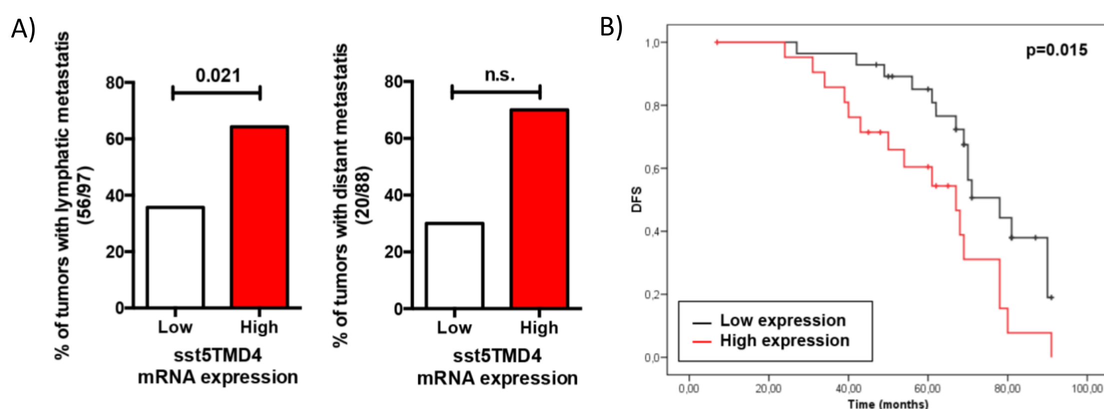


***Figure R19: A)*** *Association between the presence of sst5TMD4 and lymphatic and distant metastasis in breast carcinoma samples. Graphs, obtained from a frequency table, show the distribution of 117 grade 3 ductal breast carcinoma with low or high sst5TMD4 expression according to lymphatic and distant metastasis.* ***B)*** *Kaplan-Meier plots showing the association between increased sst5TMD4 mRNA and disease-free survival (DFS) in breast carcinoma series. Significant correlation was studied using a Chi-square and Long-rank-p-value methods.*

In order to further validate the association of the sst5TMD4 truncated receptor with breast cancer aggressiveness, sst5TMD4 presence by immunohistochemistry was determined in these breast cancer samples using a TMA as described above. Representative images of some samples are shown in **Figure R20A**. The presence and expression of the sst5TMD4 receptor at the mRNA and protein levels significantly correlated (p=0.043), confirming the validity of the measurements (**Figure R20B**).

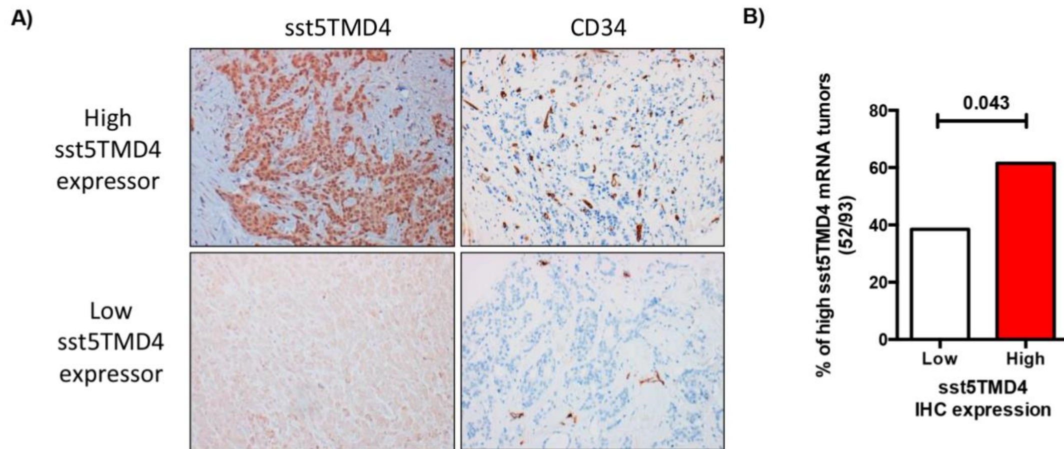


***Figure R20: A)*** *A TMA including the 127 breast carcinoma samples was employed to determine the presence of sst5TMD4 at the protein level by using an sst5TMD4 specific custom-made antibody and of the angiogenic marker CD34. Representative pictures (x20) of sst5TMD4 and CD34 staining in samples with low and high sst5TMD4 expression are depicted.* ***B)*** *Association between presence and expression of sst5TMD4 at protein and mRNA levels in the battery of breast carcinoma samples.*

Interestingly, higher expression of sst5TMD4 at the protein level was also associated with CD34 positive tumors [p<0.001], with lymphatic metastasis [p=0.035] and with disease-free survival (p=0.058) (**Figure R21**).

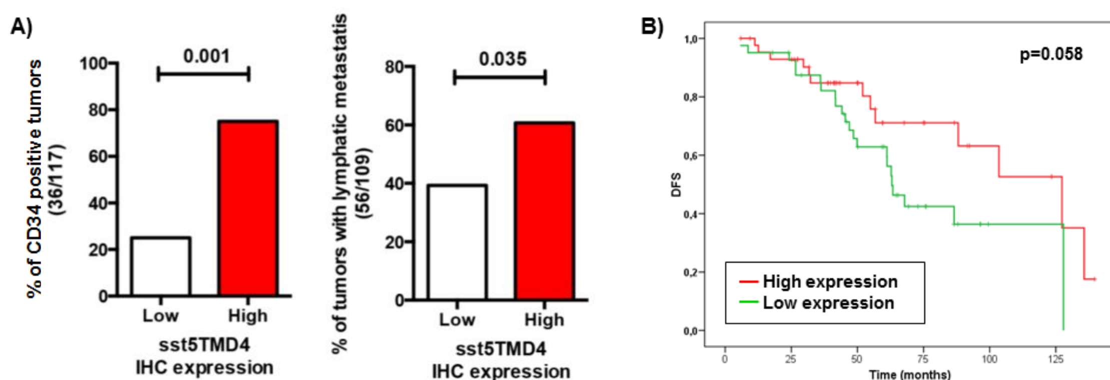


***Figure R21: A)*** *Association between the presence of sst5TMD4 and CD34 and lymphatic metastasis in breast carcinoma samples. Graphs, obtained from a frequency table, show the distribution of 117 grade 3 breast carcinoma with low or high sst5TMD4 protein levels according to CD34 staining and lymphatic metastasis.* ***B)*** *Kaplan-Meier plots showing the association between increased sst5TMD4 by IHC and disease-free survival (DFS) in breast carcinoma series. Significant correlation was studied using a Chi-square and Long-rank-p-value methods.*

### 4.1.2.2. <u>Effects of sst5TMD4 on cellular dedifferentiation in breast cancer cells</u>

A more profound analysis of the gene expression microarray data revealed that the expression of the truncated receptor sst5TMD4 could be associated to cellular dedifferentiation processes such as EMT, which has been already shown in sst5TMD4-overexpressing MCF-7 cells [18]; hence, this finding reinforced our previous data. However, we aimed to further explore the putative role of sst5TMD4 receptor on other dedifferentiation-associated processes with relevant translational consequences, such as the population of tumor-initiating cells (also referred to as "**cancer stem cells**" or **CSCs**) within breast cancer cell lines.

To achieve this aim, we analyzed the percentage of tumor-initiating cells by using a mammosphere-forming assay in MDA-MB-231 and MCF-7 sst5TMD4-transfected breast cancer cell lines, as compared to their respective mock control cells. As shown in **Figure R22**, overexpression of sst5TMD4 induced in both cell lines a clear increase in the number of mammospheres obtained, compared to mock control cells, indicating a relevant role of sst5TMD4 on the maintenance of tumor-initiating cells.

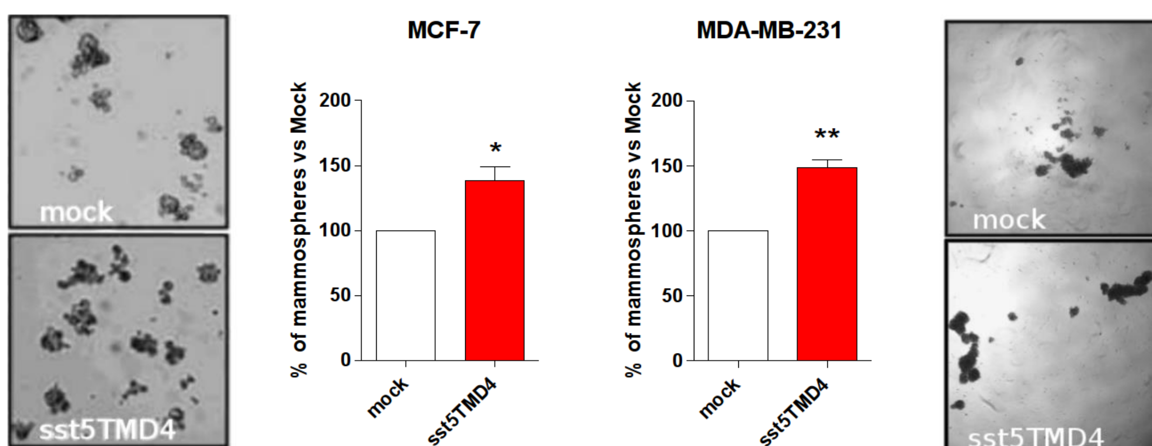


***Figure R22:*** *Percentage of mammospheres generated by MCF-7 (left) and MDA-MB-231 (right) cell lines transfected with sst5TMD4 as compare with mock transfected cells. Data represent mean ± SEM (n=5). Asterisks (*, p<0.05; **, p<0.01) indicate significant differences between mock and sst5TMD4 transfected cells*

In order to find the molecular determinants and signaling pathways underlying the association between sst5TMD4 overexpression and cellular dedifferentiation processes, we analyzed by qPCR the expression pattern of key component of dedifferentiation-related signaling pathways in both, MCF-7 and MDA-MB-231 breast cancer cell lines. This study revealed that sst5TMD4-overexpressing cells display a clear increase in the expression of Jag1 and β-Catenin, two key components of the Notch and Wnt signaling pathways (**Figure R23**). Similarly, we also found that sst5TMD4-overexpression in these cells induced a clear overexpression of TGF-β1 (**Figure R23**), which is a protein related not only with dedifferentiation processes, but also with several malignancy-associated functional characteristics [241, 242].

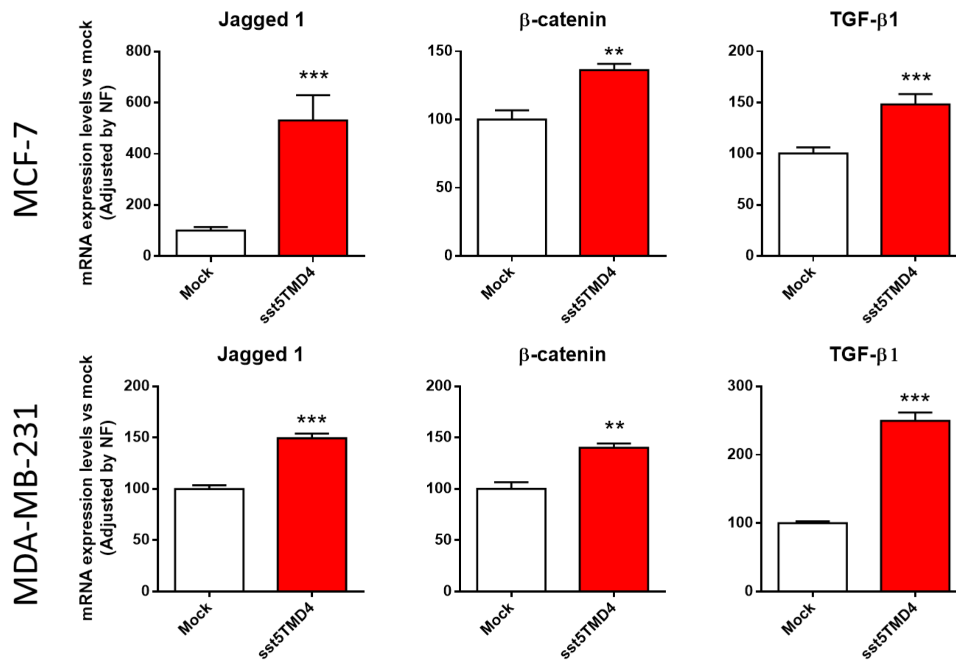


***Figure R23:*** *Jagged 1 (JAG1), β-catenin and TGF-β1 mRNA expression levels in mock- and sst5TMD4-overexpressing MCF-7 and MDA-MB-231. Expression levels are adjusted by genorm-generated normalization factors. Data represent mean ± SEM (n=10). Asterisks (**, p<0.01; ***, p<0.001) indicate significant differences between mock and sst5TMD4 transfected cells.*

## 4.2. In1-ghrelin

### 4.2.1. Genesis and regulation of In1-ghrelin splicing variant

The In1-ghrelin splicing variant is generated by a process of intron retention, in which intron 1 of the ghrelin gene, GHRL, is not spliced out, thus modifying the sequence of the final mature mRNA. However, the systems and/or mechanisms that could regulate the retention of the intron 1 in the mRNA, and the generation of In1-ghrelin variant, are still to be elucidated. Similar to the approach implemented in the case of sst5TMD4, we explored herein several aspects that could be involved in the genesis and regulation of In1-ghrelin. Specifically:

1. We analyzed the presence of certain SNPs or the *de novo* mutations that could be associated with a specific expression pattern of ghrelin or In1-ghrelin splicing variants.
2. We explored *in silico* the splicing factors that could be involved in the regulation of the splicing process that originates the expression of In1-ghrelin by exploring the presence of splicing factor target sites.
3. We investigated the putative regulatory role of the Ghrelin Opposite Strand gene (GHRLOS), which encodes for lncRNAs that overlap with the ghrelin and In1-ghrelin sequences.

#### 4.2.1.1. *Genomic changes associated to the regulation of In1-ghrelin splicing process*

To analyze the possible role of SNPs or *de novo* mutation on the expression of In1-ghrelin, we sequenced the genomic region comprising the intron 1 (**Figure R24**). Specifically, we studied the putative variations within the intron sequence in a battery of tumoral and healthy breast samples. Surprisingly, the analysis of the intron 1 sequence in these samples revealed the absence of any nucleotide alteration in this region. Thus, this initial study demonstrated the absence of SNPs or *de novo* mutations in the intron 1, which could have helped to explain the regulation of In1-ghrelin expression. These results were somehow unexpected, for an intron sequence is not predicted to be so remarkably well-preserved. Moreover, this discovery prompted the idea that the sequence of this intron may have been conserved throughout evolution, perhaps because its protein product could be playing a relevant function, which may explain the lack of the natural heterogeneity found in normal intron sequences. In any case, these results demonstrate that heterogeneity in the expression of In1-ghrelin variant is not determined by particular SNPs, at least in the population studied herein.

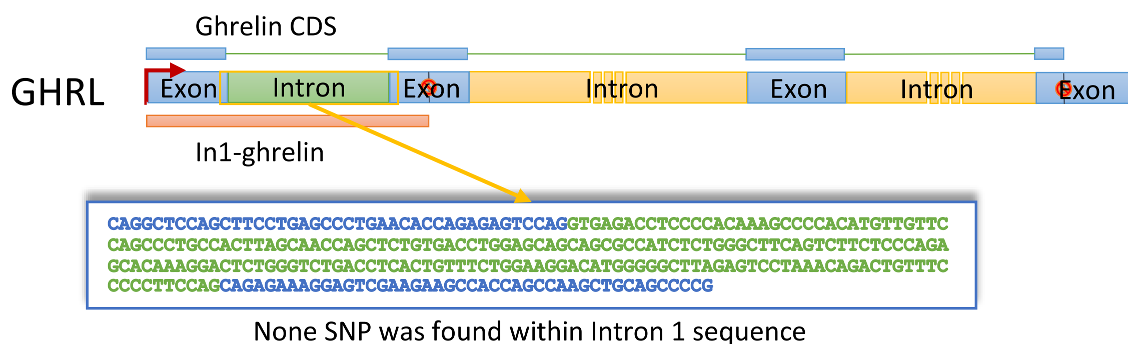


***Figure R24**: Representative scheme of GHRL gene. Sequenced region is indicated by a yellow box.*

#### 4.2.1.2. *Splicing factors involved in In1-ghrelin splicing* regulation

In order to identify splicing factors that could be involved in the generation of In1-ghrelin variant, we explored the presence of splicing factor target sites in the sequence of the intron 1, following a similar *in silico* approach to that shown in the case of sst5TMD4. The results obtained from the "SpliceAid" software revealed the existence of putative target sites for a balanced number of enhancer and silencer splicing factors. Specifically, this *in silico* analysis revealed the existence of putative target sites for 9 different enhancers (including SRSF1, SRSF3, SRSF6 or Tia-1) and for 13 different silencer splicing factors (such as PTB or KSRP). Interestingly, ESE and ESS target sequences were preferentially clustered at the central region of the intron sequence (**Figure R25** and **Table R6**), where several enhancer and silencer target sequences for different splicing factors are overlapping, which implies probable interactions among the different splicing factors.

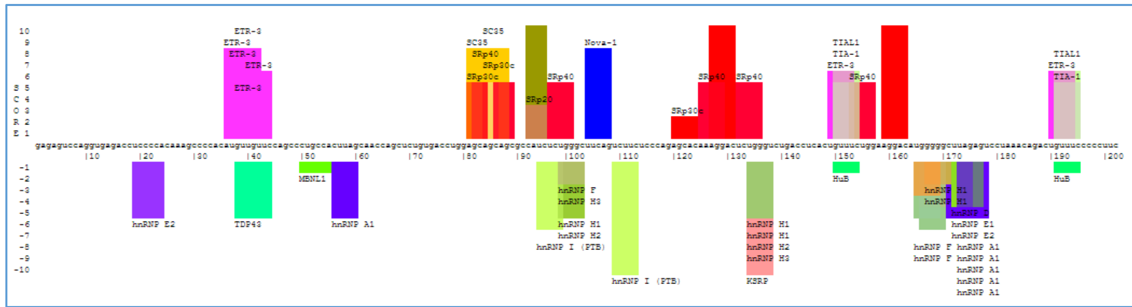


***Figure R25:*** *Representative picture of the splicing target sites predicted in the sequence of the intron 1 GHRL gene by SpliceAid.*

| | Splicing Factor | Number of targets sites | Splicing Factor | Number of targets sites | |
|---|---|---|---|---|---|
| Enhancers | ETR-3 | 3 | hnRNP A1 | 1 | Silencers |
| | Nova-1 | 1 | hnRNP D | 1 | |
| | SC35 (SRSF2) | 1 | hnRNP E1 | 1 | |
| | SRp20 (SRSF3) | 1 | hnRNP E2 | 2 | |
| | SRp30c (SRSF9) | 4 | hnRNP F | 2 | |
| | SRp40 (SRSF5) | 5 | hnRNP H1 | 3 | |
| | TIA-1 | 2 | hnRNP H2 | 3 | |
| | TIAL1 | 2 | hnRNP H3 | 3 | |
| | YB-1 | 1 | hnRNP I (PTB) | 2 | |
| | | | HuB | 1 | |
| | | | KSRP | 1 | |
| | | | MBNL1 | 1 | |
| | | | TDP43 | 1 | |

***Table R6****: Splicing factor target sites detected by SpliceAid software and classified as enhancers or silencers. The number of target sequences present within the intron 1 is also indicated.*

Moreover, analysis of In1-ghrelin complete sequence by “Human Splicing Finder” software showed that, while 3’ splice site present a strong signal for ESE, 5’ splice site, even presenting a positive ESE/ESS relative strength, showed less ESE or ISE sequences (**Figure R26**). These results suggest that the weakness of the 5’ splice site could help to promote intron 1 inclusion in the nascent transcripts.

To gain further insight, we also compared the sequences of the human and rodent ghrelin splicing variants (In1-ghrelin and In2-ghrelin, respectively) by using the bioinformatics tool *Clustalw2* in order to identify conserved sequences that could be leading to the specific regulation of the splicing variants in both species.

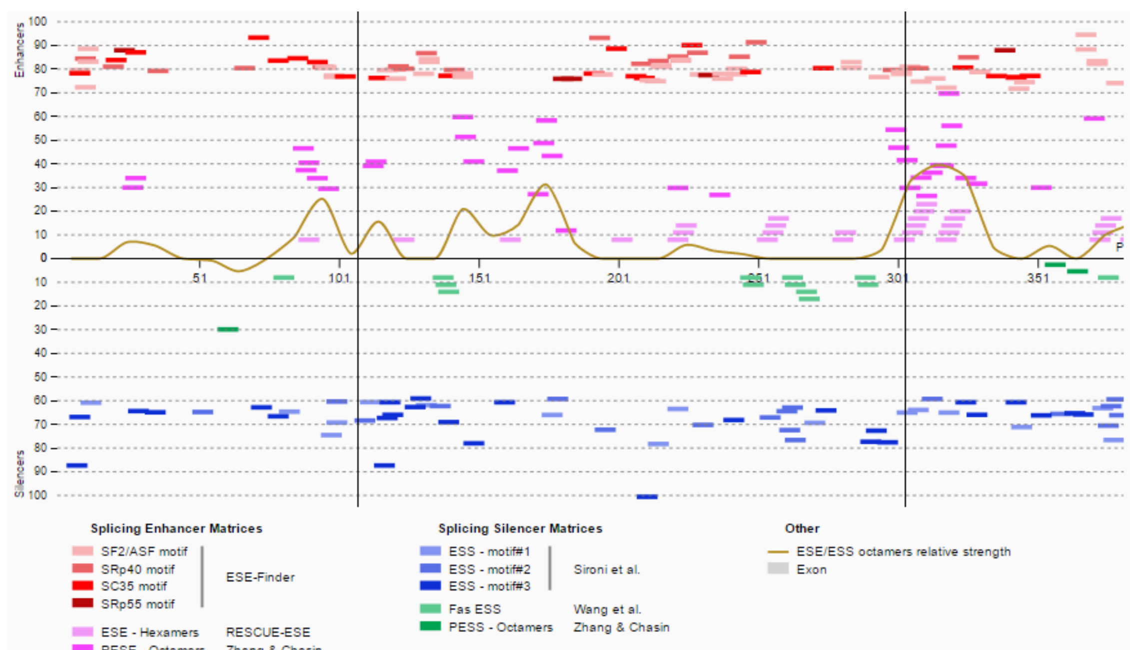


***Figure R26:** Representative picture of the In1-ghrelin transcript analyzed by the "Human Splicing Finder" software. The limits of the Intron 1 are marked with two vertical black lines.*

As illustrated in **Figure R27**, ghrelin exons are strongly conserved in both species (85% of homology). However, sequences of human and mouse retained introns are significantly different. Although this is a common hallmark of introns that are normally eliminated during the splicing processes, it is surprising in this case, inasmuch as this intron seems to be retained in both species in a physiological manner (it is amply expressed in healthy tissues in both species). Nevertheless, it is worth mentioning that there are several conserved regions that correspond with certain splicing factor target sites, including both enhancers and silencers (**Table R7**), which suggest that similar splicing factors could be regulating intron 1 retention in different species.

```
Human In1-ghrelin      ATGCCCTCCCCAGGGACCGTCTGCAGCCTCCTGCTCCTCGGCATGCTCTGGCTGGACTTGGCCATGGCAGGCTCCAGC
Mouse In2-ghrelin      ATGCTGTCTTCAGGCACCATCTGCAGTTTGCTGCTACTCAGCATGCTCTGGATGGACATGGCCATGGCAGGCTCCAGC
                       ****  **  **** *** *******  * ***** *** *********** ***** ********************

Human In1-ghrelin      TTCCTGAGCCCTGAACACCAGAGAGTCCAGGTGAGACC-----TCCCCACAAAGCCCCACATGTTGTTCCAGCCCTGC
Mouse In2-ghrelin      TTCCTGAGCCCAGAGCACCAGAAAGCCCAGGTCAGTCAGTCTGTCTCCCTAAGCCCCACA------------------
                       *********** ** ******* ** ****** ** *      ** **  **  *******

Human In1-ghrelin      CACTTAGCAACCAGCTCTGTGACCTGGAGCAGCAGCGCCATCTCTGGGCTTCAGTCTTCTCCCAGAGCACAAAGGACT
Mouse In2-ghrelin      ------------------------------------------------------TCTACCCGA--------------T
                                                                             *** * ** *             *

Human In1-ghrelin      CTGGGTCTGACCTCACTGTTTCTGGAAGGACATGGGGGCTTAGAGTCCTAAACAGACTGTTTCCCCCTTCCAGCAGAG
Mouse In2-ghrelin      CTGTGTGTG-------TGTGTGTGAGAGAGAGAGAAAG---AGAACCCT----------CTTTTCCTTTCCAGCAGAG
                       *** ** **       *** * **  **      *   *  ***  ***           **  ** ***********

Human In1-ghrelin      AAAGGAGTCGAAGAAGCCACCAGCCAAGCTGCAGCCCCGAGCTCTAGCAGGCTGGCTCCGCCCGGAAGATGGAGGTCA
Mouse In2-ghrelin      AAAGGAATCCAAGAAGCCACCAGCTAAACTGCAGCCACGAGCTCTGGAAGGCTGGCTCCACCCAGAGGACAGAGGACA
                       ****** ** ************** ** ******** ******** * *********** *** ** **  **** **
```

***Figure R27:** Human In1-ghrelin and mouse In2-ghrelin sequences alignment using the ClustalW2 software. The intron sequences are highlighted in blue and the stop codons in red.*

| Enhancers | Silencers |
|---|---|
| SRp40 (SRSF5) | hnRNP A1 |
| | hnRNP E2 |
| TIA-1 | hnRNP H1, H2, H3 |
| | hnRNP I (PTB) |
| TIAL-1 | Hub |
| | KSRP |

***Table R7:*** *Splicing factor target sequences present within the retained intron of human In1-ghrelin and mouse In2-ghrelin variants.*

4.2.1.3. <u>Regulation of In1-ghrelin transcription by GHRLOS</u>

A growing number of studies point out the crucial role of lncRNAs in the regulation of alternative splicing processes [243]. Particularly, it has been recently described the existence of a lncRNA in the antisense strand of GHRL gene, named GHRLOS. Specifically, GHRLOS presents six *in silico* predicted splicing variants, which could be involved in regulating the alternative splicing of ghrelin gene variants (**Figure R28**).

***Figure R28:*** *Representative illustration of GHRLOS and GHRL genes overlapping locations (adapted from Seim I. et al. BMC Mol Biol. 2008)*

In order to understand the analysis implemented herein, it is necessary to mention that the number of GHRLOS splicing variants described to date has been altered during the progression of this Thesis. Initially, four variants (GHRLOS-1 to GHRLOS-4) were described at the NCBI database; however, the GHRLOS-4 variant was later eliminated and three additional variants have been recently added to the database (GHRLOS-5 to GHRLOS-7). At the same time, the revised version of the GHRLOS variants included new exons at the 5' end of the three variants firstly described **(Figure R29)**. All of those variants present the exon 1, which overlaps with the last exon of ghrelin mRNA and, therefore, could be associated with its generation.

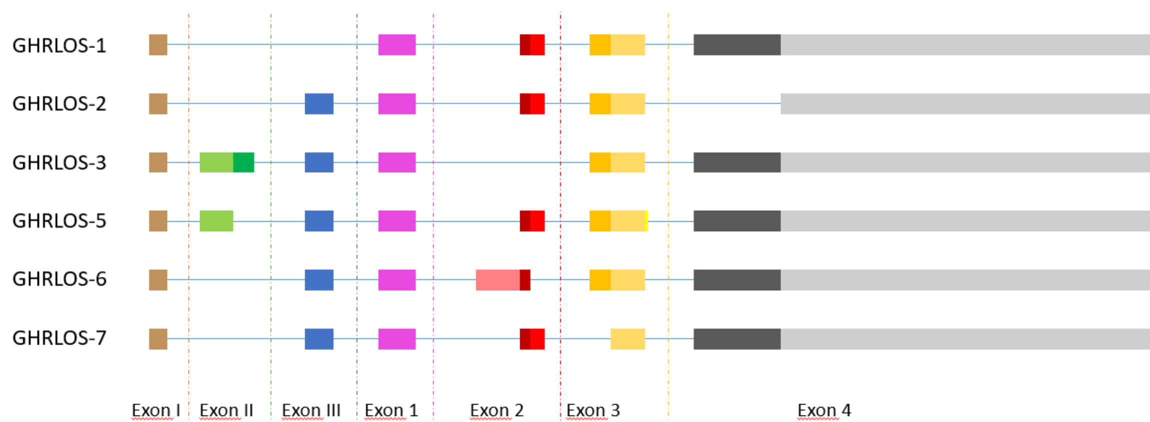


***Figure R29**: Distribution of exons in the GHRLOS splicing variants. Exons numbered by arabic numbers correspond to those firstly described while the exons numbered by latin numbers where described later.*

In order to explore the putative regulatory function of the GHRLOS splicing variants on the differential expression of the ghrelin gene variants, we decided to analyze the mRNA expression levels of the different GHRLOS splicing variants and to compare their expression with the expression pattern and mRNA abundance of ghrelin gene splicing variants.

Consequently, we designed different sets of specific primers for each GHRLOS variant and validated them in different cDNA samples. Surprisingly, we were unable to define a set of primer that could amplify the newly described GHRLOS variants (GHRLOS-5, -6 and -7), while we easily designed and validated primers for the initially described GHRLOS splicing variants (GHRLOS-1, -2 and -3). We tested all sets of primers in cDNA samples from breast cell lines and only primers for GHRLOS-1, GHRLOS-2 and GHRLOS-3 showed specific products of amplification. Moreover, we were unable to detect the newly added 5’ sequences of the GHRLOS-1, GHRLOS-2 and GHRLOS-3 variants by conventional PCR.

Taking these results into account, we decided to analyze the expression patterns of the splicing variants GHRLOS-1, GHRLOS-2 and GHRLOS-3 in a cDNA battery of human healthy tissues and to compare them with the expression patterns of ghrelin and In1-ghrelin. This analysis demonstrated that the three GHRLOS variants were widely expressed in normal healthy tissues, being GHRLOS-1 the predominant isoform and GHRLOS-3 the less expressed one (**Figure R30**). Interestingly, all three GHRLOS variants showed a similar expression pattern, exhibiting a high expression level in thymus, testis and lung (**Figure R30**).

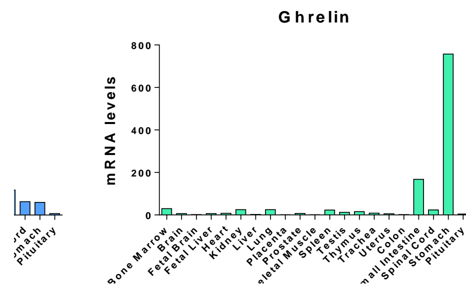


***Figure R30:*** *mRNA expression levels determined by qPCR of GHRL gene splicing variants (In1-ghrelin and ghrelin) and GHRLOS gene splicing variants (GHRLOS-1, GHRLOS-2 and GHRLOS-3) on 21 different human tissues (Bone marrow, brain, fetal brain, fetal liver, heart, kidney, liver, lung, placenta, prostate, skeletal muscle, spleen, testis, thymus, trachea, uterus, colon, small intestine, spinal cord, stomach and pituitary).*

Subsequently, we compared the expression pattern of the GHRLOS variants with those exhibited by ghrelin gene splicing variant. This demonstrated that the expression patterns of the GHRLOS variants were differentially correlated with the expression of ghrelin and In1-ghrelin. Specifically, the mRNA levels of the three GHRLOS variants analyzed, positively and significantly correlated with In1-ghrelin mRNA expression, but not with that of ghrelin mRNA (**Figure R31**), suggesting that these lncRNAs could be associated with In1-ghrelin mRNA expression, but not with ghrelin, in different tissues.

### 4.2.2. Role of In1-ghrelin in breast cancer

The role of In1-ghrelin splicing variant in several types of endocrine and endocrine-related tumors , such as pituitary or NETs, has been recently analyzed in detail by our group [15, 16]. Results from these studies indicate that In1-ghrelin is a pro-tumoral factor associated to increased malignancy, for its expression correlates with poor prognosis in these different types of tumors [15-17]. Likewise, we have developed an initial set of analysis on In1-ghrelin presence and functional effects in breast cancer [17]. However, the precise role of this splicing variant in this type of cancer is still to be completely unveiled. Therefore, in order to further expand our knowledge on the role of In1-ghrelin on breast cancer and to determine the functional and mechanistic consequences on In1-ghrelin expression, we conducted a battery of functional and molecular studies on breast cancer cell lines (MCF-7 and MDA-MB-231), by analyzing different malignancy-associated features commonly observed in tumor progression and comparing the effects of In1-ghrelin with those elicited by native ghrelin. In

addition, in order to define the clinical and pathological implications of In1-ghrelin overexpression, we explored the putative associations between In1-ghrelin expression and malignancy markers in a cohort of breast cancer patients.

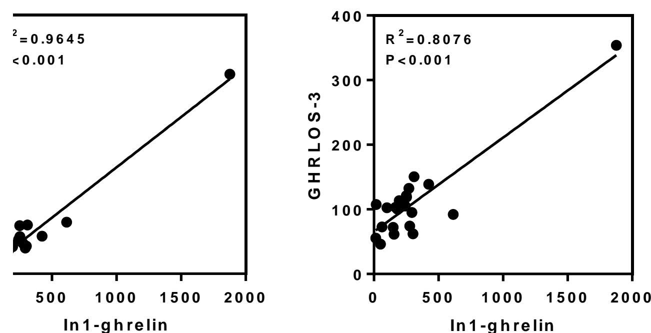


***Figure R31:*** *Correlations of mRNA expression levels (adjusted by $\beta$-actin expression) of ghrelin and In1-ghrelin with those of GHRLOS-1, GHRLOS-2 and GHRLOS-3 expression on 21 different human tissues.*

In particular, we designed a set of functional and mechanistic studies in which we challenged MDA-MB-231 and MCF-7 cells with In1-ghrelin (overexpressing In1-ghrelin or treating with In1-ghrelin derived peptides), as well as, similarly, with ghrelin. Furthermore, we also analyzed the functional consequences of downregulating endogenous In1-ghrelin on proliferation and migration capacity in the MDA-MB-231 cell line, which provides a suitable model for these assays due to their higher basal expression of In1-ghrelin compared to MCF-7 cells.

To pursue these goals, we generated stably-transfected monoclonal cell lines for ghrelin, In1-ghrelin, and empty pCDNA3.1 (mock), which served as a control. Validation of the overexpression approach by qPCR showed that both breast cancer cell lines exhibited similar efficiency in the overexpression of ghrelin, and also overxpressed similar In1-ghrelin levels, at least in terms of absolute copy numbers (**Figure R32**). In contrast, it also revealed that overexpression of ghrelin in both cell lines was markedly higher than that obtained in the case of In1-ghrelin, which could be related to the fact that basal levels of In1-ghrelin in non-transfected or mock cells were already several fold higher than those of ghrelin.

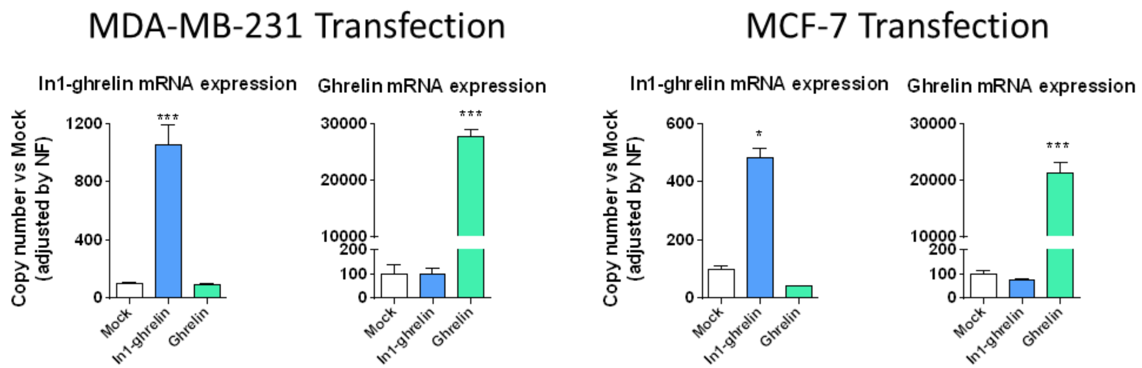


***Figure R32:*** *Validation of In1-ghrelin and ghrelin overexpression by qPCR in MDA-MB-231 and MCF-7 stably transfected cells. Copy number are normalized by a normalization factor (NF) and data are showed as percentage of mock cells (n=10 different cell preparations). Data represent mean +/- SEM. Asterisks (*, p<0.05; ***, p<0.001) indicate significant differences between mock and ghrelin or In1-ghrelin transfected cells.*

4.2.2.1. <u>Effects of In1-ghrelin on proliferation rate</u>

As a first approach, we analyzed cell survival/proliferation by using an Alamar blue assay in stably-transfected cell lines. This demonstrated an increased proliferation rate in both In1-ghrelin and ghrelin transfected cells, as compared with mock cells (**Figure R33**). Specifically, both In1-ghrelin-overexpressing MDA-MB-231 and MCF-7 cells exhibited significantly increased cell proliferation rates compared to mock cells, which is consistent with our previous results using transiently transfected MDA-MB-231 cells [17]. In addition, our data demonstrated that ghrelin overexpression increased cell proliferation after 72 h in MDA-MB-231, but not in MCF-7, cell line (**Figure R33**).

***Figure R33:*** *Proliferation/survival rates determined by Alamar blue in transfected MDA-MB-231 and MCF-7 cell lines with In1-ghrelin and ghrelin. Data represent mean ± SEM (n=6) Asterisks (*, p<0.05; **, p<0.01) indicate significant differences between mock and ghrelin or In1-ghrelin transfected cells.*

These results were further confirmed by means of an additional cell viability methodology, as is the case of MTT assay, which generated similar results (**Figure R34**).

***Figure R34:*** *Proliferation rates determined by MTT techniques in transfected MDA-MB-231 and MCF-7 cell lines (n=5-4, respectively). Data represent mean ± SEM. Asterisk (*, p<0.05) indicates significant differences between mock and ghrelin or In1-ghrelin transfected cells.*

Moreover, proliferation assays in response to In1-ghrelin and ghrelin peptides in both cell lines showed comparable results, in that In1-19, In1-40, and ghrelin increased proliferation in MDA-MB-231 cells compared to vehicle, but only In1-ghrelin derived peptides (In1-19 and In1-40) were able to stimulate proliferation in MCF-7 cells. (**Figure R35**).

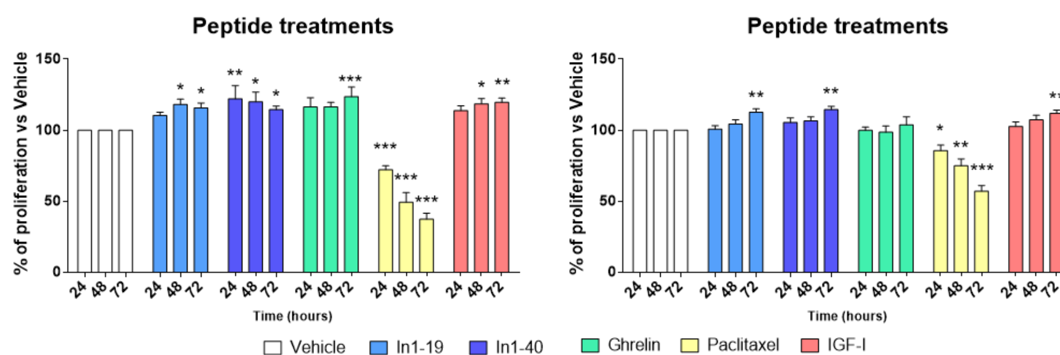


***Figure R35:*** *Alamar blue assay after treatment with In1-ghrelin and ghrelin peptides (n=8-7 in non-transfected MDA-MB-231 (left) and MCF-7 (right) cell lines, respectively). Asterisks (*, p<0.05; **, p<0.01; ***, p<0.001) indicate significant differences with vehicle treated cells.*

### 4.2.2.2. Effects of In1-ghrelin on migration capacity

Migration ability in response to In1-ghrelin was analyzed by wound healing assay in both cell lines. This approach demonstrated a significant increase in migration capacity of In1-ghrelin-overexpressing cells (**Figure R36**), while ghrelin-overexpressing cells did not show any significant change in migration capacity compared with controls (**Figure R36**). Similarly, non-transfected MDA-MB-231 and MCF-7 cells also exhibited higher

migration capacity when treated with the In1-ghrelin derived peptides (In1-19 and In1-40) compared to vehicle-treated cells (**Figure R36**). Interestingly, native ghrelin treatment only increased migration capacity in MDA-MB-231 cell line, but not in MCF-7 cells (**Figure R36**).

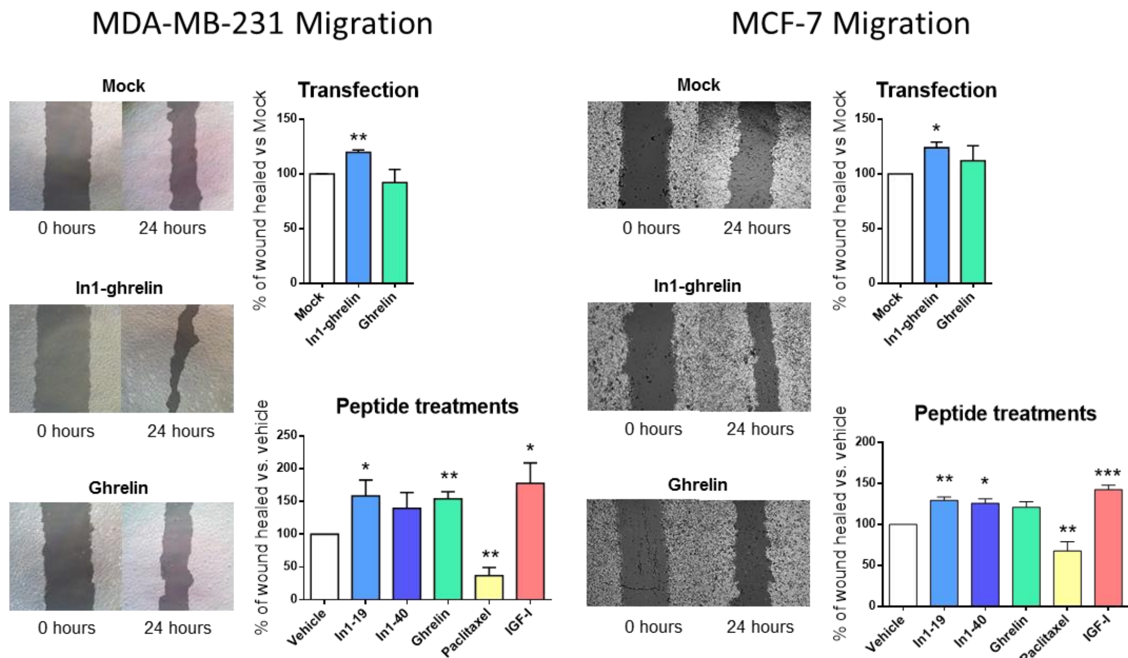


***Figure R36:*** *Migration ability was determined by wound healing technique in overexpressing (n=4-6 in MDA-MB-231 and MCF-7 cell lines, respectively) and treated (n=4-5 in MDA-MB-231 and MCF-7 cell lines, respectively) cells. Data are presented as percentage of migration vs. mock cells, and representative pictures at 0 and 24 h in overexpressing cells migration studies are depicted. Data represent mean ± SEM. Asterisks (*, p<0.05; **, p<0.01; ***, p<0.001) indicate significant differences with control cells.*

To explore the molecular basis for these actions, basal phosphorylation levels of two signaling pathways associated to cell proliferation and migration (e.g. MEK-ERK and PI3K-AKT), were analyzed by Western Blot. Interestingly, In1-ghrelin overexpression stimulated basal ERK phosphorylation, whereas ghrelin did not (**Figure R37**). In contrast, neither In1-ghrelin nor ghrelin altered basal AKT phosphorylation.

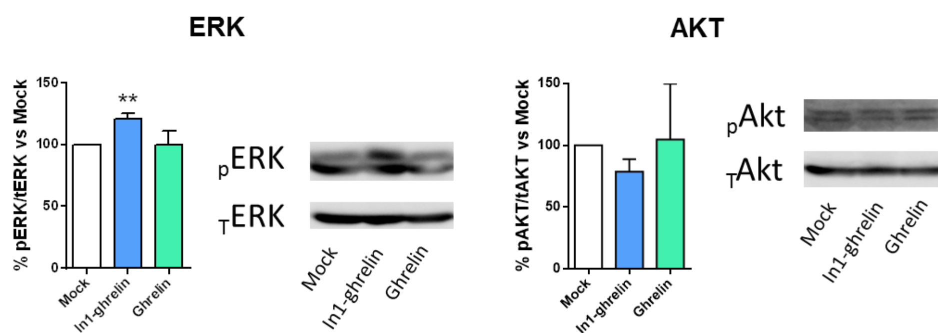


***Figure R37:*** *Basal ERK (n=5) and AKT (n=4) phosphorylation levels in In1-ghrelin and ghrelin transfected MDA-MB-231 cell lines compared with mock cells. Values represent means ± SEM. Asterisks (**, p<0.01) indicate significant differences between mock and ghrelin or In1-ghrelin transfected cells.*

#### 4.2.2.3. Effects of In1-ghrelin downregulation on proliferation rate and migration capacity

To obtain further experimental evidence of a direct relationship between In1-ghrelin overexpression and exacerbated malignant features of breast cancer MDA-MB-231 cells, we performed downregulation experiments to silence endogenous In1-ghrelin using specific siRNAs. In particular, two different In1-ghrelin specific siRNAs were successfully employed that reduced In1-ghrelin mRNA expression with respect to scramble-transfected cells, as observed in **Figure R38**.

***Figure R38:*** *In1-ghrelin mRNA expression levels in MDA-MB-231 cells after transfection of each siRNA (n=4). Values represent means ± SEM. Asterisks (***, p<0.001) indicate significant differences between scramble and In1-ghrelin siRNAs transfected cells.*

After transfection with these siRNAs, cells were used to analyze proliferation rate and migration capacity. Remarkably, we found opposite results to those obtained by In1-ghrelin overexpression, since In1-ghrelin silencing induced a significant decrease in both proliferation and migration rates compared with scramble-transfected cells (**Figure R39**). Of note, similar results were obtained with both siRNAs, in separate assays, thus reinforcing the idea that In1-ghrelin could play a relevant role in controlling the proliferation and migration of breast cancer cells.

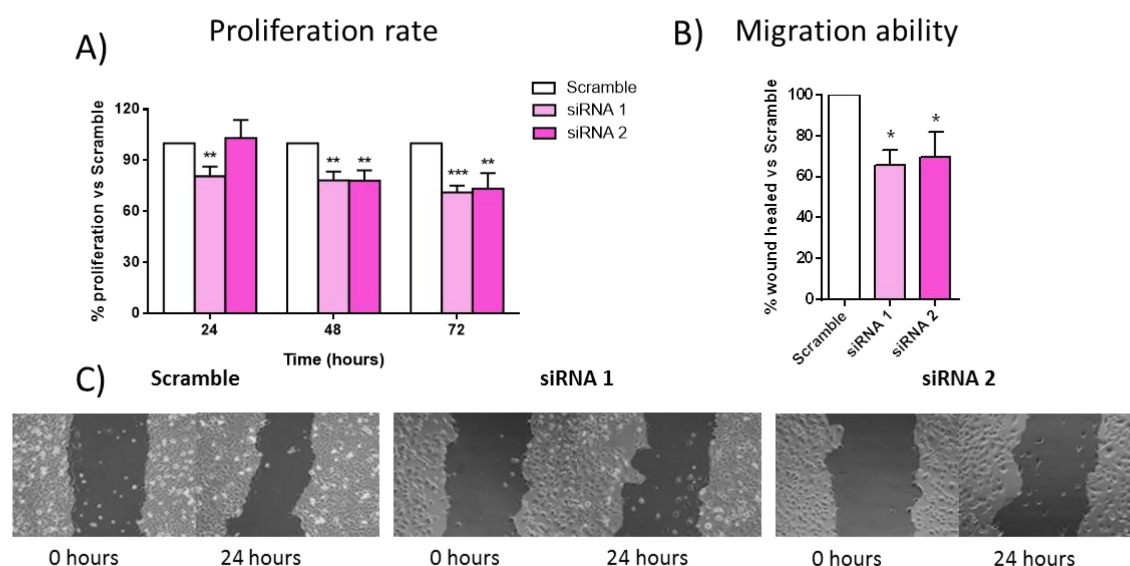


***Figure R39:*** ***A)*** *Proliferation rates in cells transfected with both siRNAs compared with scramble transfected cells (n=5-3 in siRNA-1 and siRNA-2, respectively).* ***B)*** *Migration ability analysis by wound healing technique (n=3-4 in siRNA-1 and siRNA-2, respectively).* ***C)*** *Representative images of transfected cells at 0 and 24 h. Values represent means ± SEM. Asterisks (*, p<0.05; **, p<0.01; ***, p<0.001) indicate significant differences between scramble and In1-ghrelin siRNAs transfected cells.*

#### 4.2.2.4. Effects of In1-ghrelin on cell dedifferentiation

Since higher rates of cell proliferation and migration are usually found in poorly differentiated cancers [244], we then sought to study the putative implication of In1-ghrelin, and ghrelin, on cell dedifferentiation hallmarks. Specifically, two characteristic features such as cell morphology (as a measure of cell plasticity) and mammosphere formation, were explored and compared with those found in control (mock) cells.

As illustrated in **Figure R40**, there was a clear increase in the proportion of cells with mesenchymal-like phenotype among In1-ghrelin overexpressing cells, compared with mock cells, in MDA-MB-231 cell cultures, but not in MCF-7 cells (data not shown). In contrast, no changes in cellular morphology were observed within native ghrelin-overexpressing cells.

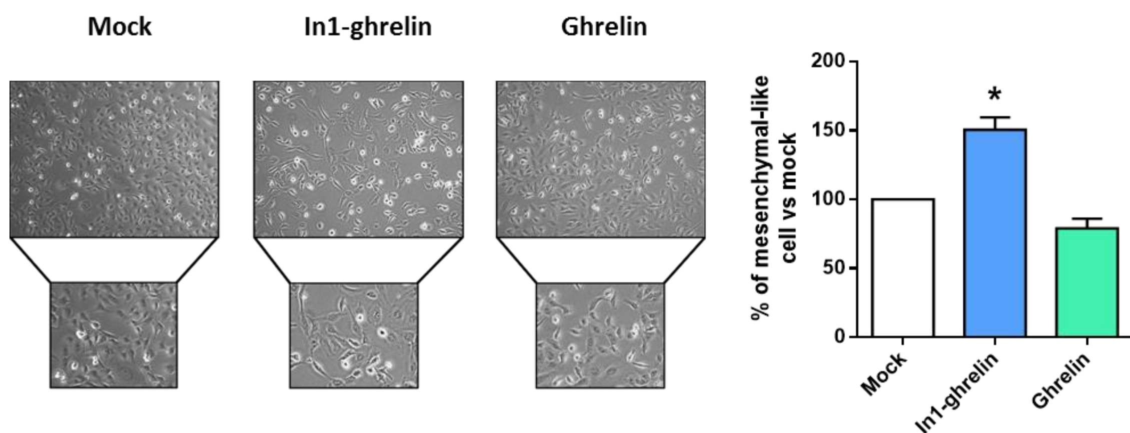


***Figure R40:*** *Percentage of mesenchymal-like cells in In1-ghrelin and ghrelin transfected MDA-MB-231 cells compared to mock controls (n=4). Values represent means ± SEM. * indicate significant differences (p<0.05) between mock and ghrelin or In1-ghrelin transfected cells.*

This cell line-dependent EMT induction of In1-ghrelin highlights the importance of the context where In1-ghrelin is exerting its functions, and suggests that additional elements should exist that contribute to modulate the function of this splicing variant in breast cancer malignancy.

To gain further insight into In1-ghrelin function, we next explored the abundance of CSC-like cells within the stably-transfected cell lines. CSCs have been in the focus of cancer research for the last two decades, and different studies have showed that the presence of these cells within the tumor population increases their malignancy as well as their resistance to chemotherapy [244, 245]. Thus, appearance of CSCs in a tumor seems to involve a dedifferentiation process from the epithelial-phenotype of the cells that comprise the original population. Inasmuch as the overexpression of In1-ghrelin, but not ghrelin, induced an increase in the number of mesenchymal-like cells within the MDA-MB-231 cell line, and since EMT is a process of dedifferentiation, it could be possible that the CSCs population was affected by the overexpression of In1-ghrelin.

Therefore, we implemented mammosphere-formation assays in MCF-7 and MDA-MB-231 cells as a measure of the CSC colony within the breast cancer cell line population.

The results depicted in **Figure R41** indicate that overexpression of In1-ghrelin, but not ghrelin, induced an increase in the proportion of cells able to generate mammospheres (thus, in the number of CSCs) within the cell population. Interestingly, this increase was especially marked in MCF-7 cells, where In1-ghrelin transfection induced a 3-fold increase in the number of mammospheres.

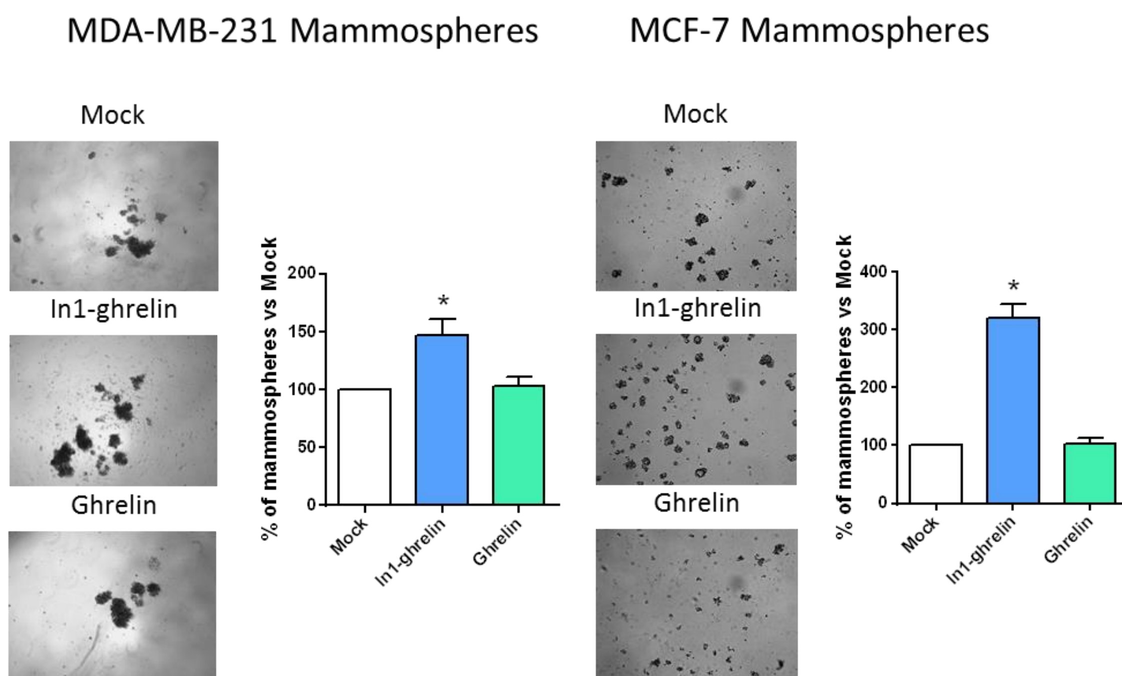


***Figure R41:*** *Percentage of mammospheres in MDA-MB-231 (n=5) and MCF-7 (n=3) cell lines transfected with In1-ghrelin, ghrelin compared to empty vector (mock) transfected cells. Values represent means ± SEM. Asterisks (*, p<0.05) indicate significant differences between mock and ghrelin or In1-ghrelin transfected cells.*

Moreover, mammospheres generated by the MDA-MB-231 cells transfected with In1-ghrelin seemed to be larger than those observed in ghrelin-transfected or mock cells. Actually, as shown in **Figure R42,** the number of cells in the mammospheres generated by In1-ghrelin-overexpressing MDA-MB-231 cells, was almost 3-fold higher than that in ghrelin-overexpressing or mock transfected cells.

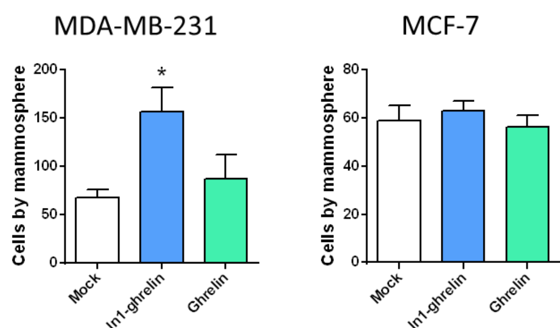


***Figure R42:*** *Average number of cells by mammosphere in each transfected cell line (n=3). Values represent means ± SEM. Asterisks (*, p<0.05) indicate significant differences between mock and ghrelin or In1-ghrelin transfected cells.*

Consistent with these results, cells treated during 7 days with In1-ghrelin derived peptides also exhibited an increased capacity to form mammospheres, as compared to vehicle- and ghrelin-treated cells (**Figure R43**).

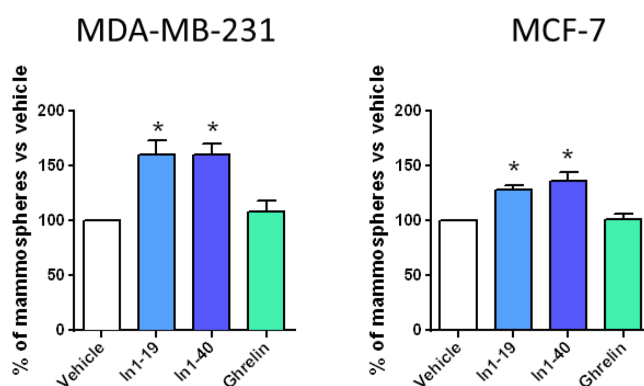


***Figure R43:*** *Percentage of mammospheres generated by MDA-MB-231 and MCF-7 cell lines treated with In1-ghrelin peptides (In1-19 and In1-40) or ghrelin peptide compared with vehicle -treated cells (n=4). Values represent means ± SEM. Asterisks (*, p<0.05) indicate significant differences between vehicle and ghrelin or In1-ghrelin peptides treated cells.*

### 4.2.2.5. Signaling pathways altered by In1-ghrelin overexpression

In order to explore the molecular basis of these In1-ghrelin-induced changes, the expression levels of key genes of three signaling pathways related with dedifferentiation processes were measured in the stably-transfected cell lines. First, we found that both ghrelin and In1-ghrelin overexpression increased TGF-β1 expression, which has been associated not only with EMT or CSCs, but also with higher proliferation and migration rates. In that TGF-β1 has been previously associated with both Notch and Wnt/β-catenin signaling pathways [246], we then measured JAG1, a Notch pathway ligand, and β-catenin (CTNNB1), an effector protein in the Wnt pathway, which is also tightly related to EMT [247, 248]. Indeed, as depicted in **Figure R44,** JAG1 showed elevated mRNA levels in In1-ghrelin- and also in ghrelin-transfected cells. Interestingly, β-catenin mRNA levels revealed a differential response to In1-ghrelin and ghrelin, as they were increased in both cell lines only when In1-ghrelin, but not when ghrelin was overexpressed (**Figure R44**).

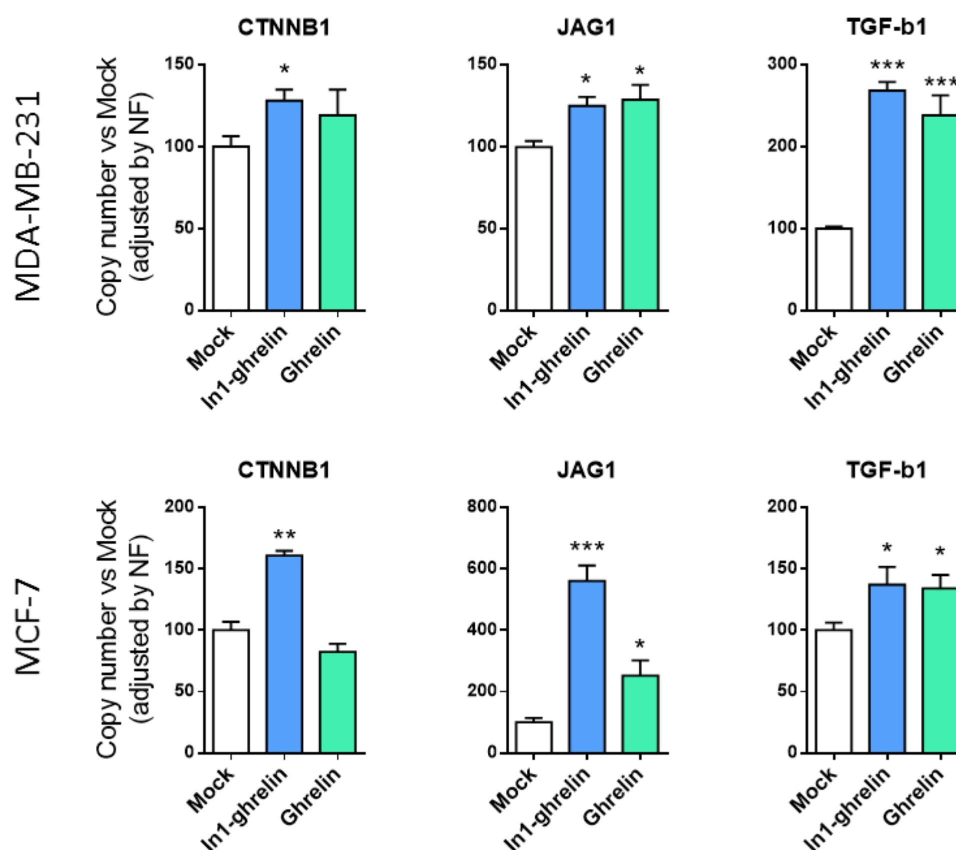


***Figure R44:*** *mRNA expression levels of β-Catenin (CTNNB1), Jagged1 (JAG1), and Tumor Growth Factor-β1 (TGF-β1) are showed as percentage of expression in mock cells (n=10 independent cell preparations). Values represent means ± SEM. Asterisks (*, p<0.05; **, p<0.01; ***, p<0.001) indicate significant differences between mock and ghrelin or In1-ghrelin transfected cells.*

Finally, we also explored the expression levels of the canonical receptor for native ghrelin, GHSR-1a, and its truncated form GHSR-1b, and found that both were absent or close to the detection limit (GHSR-1a: Ct = 37,0 – 37,1 in MDA-MB-231 and MCF-7 cell lines and GHSR-1b: Ct = 35,2 – 33,4 in MDA-MB-231 and MCF-7 cell lines, respectively).

#### 4.2.2.6. In1-ghrelin correlated with poor clinical outcome in breast cancer patients

To ascertain the potential clinical significance of In1-ghrelin expression in breast cancer patients, we analyzed the expression of In1-ghrelin variant in a cohort of 127 breast cancer samples, and explored the association of In1-ghrelin expression levels with malignancy-associated features, such as lymph-node metastasis and disease-free survival. To this end, samples were categorized among low, moderate and high In1-ghrelin levels according to quartile In1-ghrelin expression levels (0-50% as low, 50-75% as moderate and 75-100% as high) (**Figure R45**).

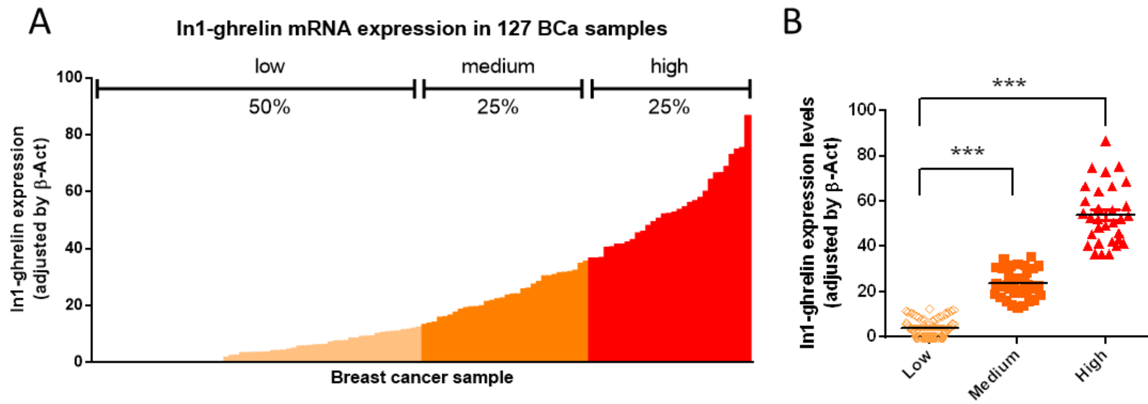


***Figure R45: Expression of In1-ghrelin variant in a battery of 127 breast cancer samples. A)*** *Categorization of breast cancer samples by In1-ghrelin mRNA expression levels. Samples within the fourth and third quartile (Q4 and Q3) were grouped as none/low In1-ghrelin expression, while samples within second and first quartiles were considered as medium and high In1-ghrelin expression, respectively.* ***B)*** *Average In1-ghrelin expression in the three categories. Asterisks (***, p<0.001) indicate significant differences between the groups compared.*

This analysis showed a strong, direct association (p=0.001) between In1-ghrelin expression and lymph-node metastasis, in which patients with low expression of In1-ghrelin presented less lymph-node invasion than those with high In1-ghrelin expression levels (**Figure R46**). Most strikingly, breast cancer patients with high expression levels of In1-ghrelin presented significantly (p<0.001) lower disease-free survival than those with low or moderate In1-ghrelin expression (**Figure R46**).

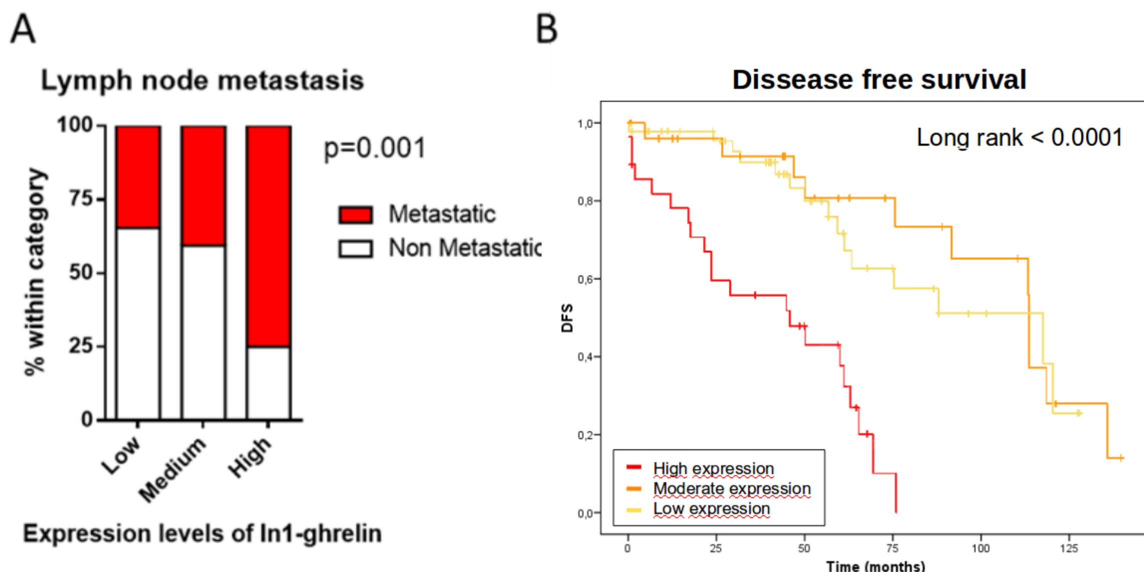


***Figure R46: Correlation of In1-ghrelin expression and clinical parameters in the cohort of breast cancer patients. A)*** *Association between the presence of In1-ghrelin and lymphatic metastasis in breast carcinoma samples.* ***B)*** *Kaplan-Meier plots showing the association of increased In1-ghrelin mRNA expression and disease-free survival in breast carcinoma series.*

# Discussion

# 5. Discussion

Cancer is nowadays a major public health problem worldwide, and the second leading cause of death in industrialized countries [1]. Unfortunately, cancer development and progression comprises an extremely complex, dynamic process, highly influenced by genetics, but also determined by external factors, such as metabolic and nutritional status, natural environment, and life style factors [3]. As a result, the natural history of cancers, their progression, biological and clinical features, as well as their response to treatment can be extremely variable and heterogeneous, even within a similar type of tumor. In fact, such tumoral heterogeneity can also be the result of the specific internal milieu in which the tumor develops and progresses, a notion that is particularly evident in the so-called endocrine-related cancers, in which tumor genesis and evolution is profoundly conditioned by metabolic-endocrine dysregulations [20, 21]. Among them, breast cancer is a major cancer type, both in terms of incidence (the most common cancer in women), but also in terms of mortality rate (representing the second leading cause of cancer-associated deaths) [1].

Breast cancer represents in itself an example of a highly complex and heterogeneous cancer type, in which the available tools for the diagnosis, prognosis and medical treatment are not as specific and effective as desired. In this scenario, the strong regulation by the endocrine milieu exhibited by breast cancer [244] prompted our group to explore the potential role of several hormonal a neuropeptide systems — particularly SST/CORT/sst and ghrelin systems— in the development and progression of these tumors, as well as in determining the suitability of certain members of these systems as novel biomarkers for the diagnosis, prognosis and/or putative therapeutic treatment of this endocrine-related tumor type [10-19].

Of note, despite their remarkable, intrinsic complexity, most cancers share a group of common hallmarks, which likely includes altered alternative splicing processes [6]. Indeed, aberrant splicing is gaining an unexpected relevance in this field, since recent studies point out that cancer heterogeneity, from that in the tumoral tissue to the disparate patient outcome and survival, can reside, at least in part, on genetic variations (such as splicing variants) present and/or originating in the primary tumor [7]. In support of this idea, our group has recently identified novel splicing variant of SST and ghrelin systems (the truncated receptor sst5TMD4 and the splicing variant In1-ghrelin), which are overexpressed in several, distinct tumors and cancer (e.g. pituitary, thyroid or NETs), wherein they are associated with enhanced malignant phenotypes [10-19]. However, the molecular mechanisms implicated in their generation and their precise role and clinical implications in breast cancer have not been completely unveiled yet, and therefore represent the main focus of this Thesis.

## 5.1. sst5TMD4

There is increasing interest in the regulation of the splicing processes associated to SST receptors (ssts), for they seem to represent relevant factors in cancer malignancy, as indicated by the fact that two non-canonical splicing variants of the sst5 gene (sst5TMD4 and sst5TMD5) have been found to be associated with the malignancy of tumoral cells from several cancer pathologies. In particular, sst5TMD4 expression has shown specially relevant associations with the progression of a subset of pituitary [11, 12], breast [18] and thyroid [10, 13] tumors, in which there is evidence for a clear relationship between sst5TMD4 expression and/or function and increased proliferation, migration and invasion abilities of tumoral cells. For this reason, it seems necessary to explore the processes underlying the altered splicing events occurring in SST5 gene that could promote the generation of this splicing variant, as well as the molecular mechanisms and the functional features associated to sst5TMD4 presence in tumoral pathologies, focusing on breast cancer.

### 5.1.1. SST5 splicing regulation

Alternative splicing of the SST5 gene represents a case of non-canonical splicing, in that the cryptic introns generating the two known alternatively spliced variants, sst5TMD4 and sst5TMD5, do not present the canonical splicing recognition sites at 5' (GT) and 3' (AG) intron-exon boundaries. In particular, the cryptic intron eliminated to generate sst5TMD4 presents a novel non-canonical splicing pair of sequences (5'GC-GC3´), which have not been described before; although the 5'GC splice site has been previously reported as a non-canonical donor splice sequence [35]. In order to unveil the putative mechanisms associated to the generation of this splicing variant, mainly in tumoral pathologies, we implemented a triple approach aimed at screening the most common regulatory mechanisms of the splicing process, including alteration in genomic sequence [249], dysregulation of splicing factors [100] or changes in the expression patterns of regulatory ncRNAs [68, 250].

#### Role of genomic alterations in sst5TMD4 expression

Genomic sequence alterations (also known as alterations in *cis*), such as certain SNPs or *de novo* mutations, can alter the sequence of the splicing sites or modify the actions of different *trans* elements, such as the splicing factors, by modifying their target sequences [48], and have been, therefore, found associated with changes in the expression pattern and in the appearance of splicing isoforms of different genes [249].

In order to explore this possibility in the case of sst5TMD4, we sequenced key regions of the cryptic intron (those regions close to the intron-exon boundaries) in a battery of normal and tumoral breast biopsies, which showed variable levels of the truncated sst5TMD4 and of the full-length sst5. Interestingly, the results obtained revealed the absence of *de novo* mutations in the regions sequenced from the tumoral samples,

thus minimizing the possible role of *de novo* mutations in the generation of the truncated sst5TMD4 variant.

This notwithstanding, we detected in our population the presence of two previously identified SNPs (rs197055 and rs12599155), which had not been associated earlier to any known pathological implication. Interestingly, our results suggest that demographic changes in the genomic sequence of SST5 could be associated with the presence of breast cancer, and that the C>A transition in the SNP rs197055 could be a significant risk factor for breast cancer, inasmuch as heterozygous C/A and A/A homozygous genotypes only appeared in tumoral samples. In addition, our results indicate that the A/A homozygous genotype was present in samples with high sst5TMD4 expression, suggesting that this polymorphism could be implicated in the generation of this splicing variant. Consistent with this idea, *in silico* studies revealed that the presence of this SNP could disrupt a target sequence for the splicing factor YB-1, and the consequent appearance of a weak branch point, which could be involved in the generation of the sst5TMD4. However, the actual implication of this C>A transition in the generation of the sst5TMD4 variant and its association with breast cancer still remains to be fully elucidated. On the other hand, although the frequencies observed in the SNP rs12599155 was not significantly altered in tumoral samples compared to control non-tumoral samples, the fact that the proportion of C/T and T/T genotypes tended to be higher in tumoral samples with high sst5TMD4 expression suggests that the C>T transition in this locus could be related to an increase in the expression levels of sst5TMD4. In addition, *in silico* studies found that this SNP induces the disappearance of a splicing enhancer sequence; however, this splicing factor target has been predicted only bioinformatically and the splicing factor that could bind to this sequence is unknown. Altogether, although it would be necessary to increase the number of samples sequenced in order to fully confirm the implication of rs197055 and rs197055 SNPs in the generation of the splicing variant sst5TMD4, these results provide suggestive evidence to propose that genomic alterations of the SST5 gene can play a relevant role of in breast cancer by influencing the expression of the truncated receptor sst5TMD4 and, thereby, tumoral features.

### Presence of splicing factor target sequences in SST5.

Splicing factors are the most studied splicing *trans* elements. They comprise a group of regulatory molecules that, by binding the nascent transcript, can modify the splicing process and, therefore, the splicing variants generated. Splicing factors are altered in many different pathologies [251] and their specific role in cancer is being intensely studied nowadays [100]. Splicing factors interact not only with the spliceosome or the nascent transcripts, but also among them, and the results from these complex interactions determine the final pattern of splicing variants expressed. Indeed, the interaction among different splicing factors substantially alters the splicing process [252] and the modification in the ratio of enhancer and silencer splicing factors drives to changes in the splicing pattern of several genes [253]. Consequently, the landscape of splicing factors present during the transcription of a particular gene represents a key

point in defining splicing regulation, being this presence regulated not only by the expression of the different splicing factors in the cell, but also by the cellular localization of each specific factor [254] and its functionality [255].

In this scenario, our *in silico* studies indicate that the cryptic intron excised to generate the sst5TMD4 variant encompasses a unique distribution of splicing factor target sequences, with robust clusters of sequences for silencer factors, followed by peaks of high strength for splicing enhancers, which are not present in the rest of the gene. The robust clusters of splicing silencers target sites could indicate a normal, physiological tendency to prevent the splicing of this intron, thus generating the full-length sst5. Whereas, the peaks of high strength for splicing enhancers could indicate that, under certain circumstances, the presence and/or activity of key splicing enhancers could play a predominant role, promoting the splicing of the intron and the generation of the truncated sst5TMD4 variant. Hence, these studies suggest that the ratio among SR proteins and hnRNPs could be on the basis of the regulation of the splicing of this cryptic intron, and are in agreement with the fact that sst5TMD4 is barely expressed in few healthy tissues [11].

Of particular interest seems the role of SRp40 (SRSF5), for it is the splicing enhancer with the higher presence in terms of number of target sequences in the cryptic sst5TMD4 intron, and has been previously reported to be overexpressed in breast cancer. Specifically, SRSF5 has been related in breast cancer with the skipping of an exon in CD44 that promotes the appearance of an oncogenic variant, but also with changes in the splicing pattern of c-Myc, an oncogene with anti-apoptotic effects [85, 109, 110]. Likewise, hnRNP H1 and H2, the silencer splicing factors with more target sequences within the cryptic intron, have been also involved, in breast cancer, with the dysregulation of the splicing of the angiogenesis related enzyme thymidine phosphorylase, which has been associated to chemotherapy resistance [256].

In summary, *in silico* studies on SST5 splicing strongly suggest the existence of a complex regulatory system controlling the alternative splicing of the cryptic SST5 intron, which must be further analyzed by *in vitro* studies focused on the identification of the precise set of splicing factors involved, their effects on the splicing of this cryptic intron, and the putative interactions relevant for the regulation of sst5TMD4 expression, as well as their ultimate implications for breast cancer.

### Regulatory role of miRNAs on sst5TMD4 mRNA stability

Another group of *trans* elements commonly associated to the regulation of the expression of different splicing variants are the miRNAs [68, 250]. This family of ncRNAs regulates the expression of different transcripts by blocking their translation or inducing the degradation of the mRNA [59]. These regulatory elements have been shown to be involved in several biological processes, including those related with cell proliferation, differentiation, invasion or apoptosis [257-259], which are, obviously, important processes for tumoral aggressiveness and malignancy. Actually, dysregulation in the miRNAs landscape is a common event in cancer malignancy [151].

As an initial approach to investigate the miRNAs with capacity to regulate sst5TMD4 expression, we performed an *in silico* study to identify putative target sequences within the specific sst5TMD4 sequence. This analysis indicated that sst5TMD4 mRNA presents specific targets for 5 different miRNAs: hsa-miR-189, hsa-miR-326, hsa-miR-339, hsa-miR-346 and hsa-miR-939.

Subsequently, in order to explore the putative implication of these miRNAs in the expression of sst5TMD4, we analyzed the expression of the different miRNAs that, by *in silico* prediction, showed possibilities of binding to sst5TMD4 mRNA in a battery of 16 breast samples (both healthy and tumoral) and correlated them with those of the truncated sst5TMD4 variant. Three of them, hsa-miR-326, hsa-miR-189 and hsa-miR-346, showed significant correlation with the expression levels of sst5TMD4, which suggest that they could somehow be involved in the regulation of the mRNA stability or expression levels of this splicing variant. Specifically, hsa-miR-326 and hsa-miR-189 exhibited a direct correlation with sst5TMD4 whereas, conversely, hsa-miR-346 expression levels showed a clear inverse correlation with those of the truncated sst5TMD4 variant.

The direct correlation of both hsa-miR-189 and hsa-miR-326 with the expression of sst5TMD4 was somewhat unexpected, since the usual mode of action of miRNAs is a downregulation of the target transcript expression and, thus, an inverse correlation was likely expected. However, it has been previously reported that certain miRNAs can act as activators of mRNA transcription and translation in some conditions [260]. hsa-miR-189 has been found downregulated in cutaneous malignant melanoma [261] and with enhanced protection to radiation in primary microvascular endothelial cells [262]; however, its role, if any, in other tumoral malignancies has not been described. Conversely, hsa-miR-326 is considered a tumor suppressor, which is involved in chemotherapy resistance in breast cancer cells [263], and regulation of apoptosis thought Notch signaling pathways in glioblastoma [264]; however, the regulatory network of hsa-miR-326 is highly complex and involves interaction among transcription factors and different miRNAs, which could modulate the precise role of this miRNA [265], thus suggesting the possibility that additional elements could be involved in the putative contribution of this miRNA to the control of sst5TMD4 expression.

On the other hand, the negative correlation of hsa-miR-346 and sst5TMD4 suggests a more plausible regulatory system, based in the classical model of miRNA function, wherein hsa-miR-346 could be inhibiting sst5TMD4 translation or even promoting its mRNA degradation. Consistent with these idea, the study of the expression of both transcripts in different passages of the breast cancer cell line MCF-7, where we previously reported that sst5TMD4 expression is progressively decreased along the passages [18], revealed that the expression of hsa-miR-346 is progressively increased, which further supports a functional association between hsa-miR-346 and sst5TMD4. In an attempt to confirm the direct action of hsa-miR-346 on sst5TMD4 expression, we evaluated the effects of a commercial mimic and inhibitor of hsa-miR-346 on the expression of sst5TMD4 in MCF-7 cells. Interestingly, hsa-miR-346 inhibitor did not

induce any change at sst5TMD4 mRNA or protein levels, which could be consistent with the fact that MCF-7 cells from initial passages express high levels of sst5TMD4 but low levels of hsa-miR-346. However, the use of the hsa-miR-346 mimic induced a slight reduction on sst5TMD4 mRNA expression and a significant decrease in sst5TMD4 protein levels, strongly suggesting a direct effect of hsa-miR-346 on the stability or translation rate of the sst5TMD4 mRNA and, consequently, on the protein expression of sst5TMD4.

Interestingly, hsa-miR-346 seems to regulate a number of diverse actions and play differential roles in tumoral pathologies. Specifically, hsa-miR-346 has been shown to be involved in the modulation of anti-inflammatory effects [266], the regulation of epithelial to mesenchymal transition [267], promotion of osteogenic dedifferentiation through Wnt/β-Catenin pathway [268], and it has even been implicated in chemotherapy resistance [269]. It seems to be a widely expressed and tightly regulated molecule, since hsa-miR-346 has been reported to be downregulated in head and neck cancer [270] and in central nervous system tumoral cell lines [271], but overexpressed in cutaneous squamous cell carcinoma [272], and in follicular thyroid carcinoma [273]. These findings suggest that the final effects of hsa-miR-346 should be regulated in a tissue- or even cell-dependent manner. In the context of breast cancer, hsa-miR-346 has found to be overexpressed in metastatic MDA-MB-231 cells compared to normal MDA-MB-231 cells [274], but, most interestingly, it is downregulated in chemotherapy resistant MCF-7 cells [275]. As it will be discussed later, overexpression of sst5TMD4 increases the number of mammospheres (and thus presumably cancer stem cells, CSCs, or tumor-initiating cells, TICs), in both MCF-7 and MDA-MB-231 cell lines, which has been associated with chemotherapy resistance [276], and would be, therefore, in agreement with the reduced expression of hsa-miR-346 observed in chemotherapy resistant MCF-7 cells [275].

When viewed together, our results suggest the existence of a diverse set of potential regulatory pathways that could act at multiple levels to exert an integrated control of the altered, pathophysiologic expression of sst5TMD4 in tumoral pathologies. Certainly, further studies are required to understand the precise role of each of the potential players described in this work, from SNPs to splicing factors, and miRNAs in the genesis and regulation of sst5TMD4 expression, and to define what are their possible interactions in the context of breast cancer and, perhaps, other tumoral pathologies. Nevertheless, the present results provide for the first time a reasonable basis to explore in more detail the regulatory system underlying the non-canonical splicing and transcript stability of SST5 gene products.

### 5.1.2. sst5TMD4 functional role in cancer

The aberrantly spliced variant of the SST5 gene, sst5TMD4, was initially identified in pituitary adenomas [11], where its expression, and thus, the presence of the resulting truncated receptor, was associated with increased aggressiveness [14], and poor response to somatostatin analogues therapy [277]. Further studies found similar

results in poorly differentiated thyroids tumors [13] and medullar thyroids carcinoma [10], where sst5TMD4 was found overexpressed and involved in the lack of response to SST analogues treatment, and also in breast cancer samples, where sst5TMD4 expression was correlated with worse prognosis, and its overexpression induced increased malignant features in the breast cancer derived cell line MCF-7 [18].

Bearing these data in mind, it seemed essential to analyze the cellular and molecular mechanisms altered by the presence of sst5TMD4, in order to identify the molecular basis underlying the pathological implications of this splice variant in breast cancer. In this regard, the results obtained from the gene expression microarray performed on sst5TMD4-overexpressing MCF-7 cells showed ample effects of sst5TMD4 overexpression on different functional features, such as cell growth, metabolism, EMT or angiogenesis. Some of these malignancy-associated features have been previously related to sst5TMD4 overexpression; however, the role of this truncated receptor in tumoral angiogenesis was completely unexplored hitherto. Interestingly, a more comprehensive, user-driven analysis of the genes altered on this array showed that up to 40% of them were related to the angiogenesis process, further supporting a putative contribution of sst5TMD4 to the angiogenic process.

Tumor-induced angiogenesis is a key factor in cancer progression [4]. Growing tumors progressively increase their need of nutrient and oxygen and, in conditions of insufficient vasculature, tumoral cells become hypoxic [278], displaying a “dormant phenotype”, where the tumor stops growing, maintaining an equilibrium between proliferation and death rates [279]. In these conditions, secretion of angiogenic factors that stimulate the sprouting of new vessels by tumoral cells can raise the growth rate of the tumors and, therefore, promote cancer progression and metastasis [4]. Regulation of angiogenesis is controlled through stimulatory and inhibitory pathways, and the balance of the different signals determine the correct progression of the angiogenesis process [280]. This balance is clearly dysregulated in tumors, showing an overexpression of many pro-angiogenic factors [281]. For this reason, it is not surprising the use angiogenic factors as therapeutic targets in several pathologies, being the blockade of the VEGF pathway the most established angiogenic-targeting in cancer treatment [282].

In this scenario, our results unveiled that the truncated somatostatin receptor sst5TMD4, which has been reported to be present in different tumor types [10, 11, 13, 14, 18, 195, 277], profoundly alters the expression pattern of several angiogenesis related genes in the breast cancer cell line MCF-7. To be more specific, sst5TMD4 increases the expression and/or secretion of pro-angiogenic factors such as VEGF, EGF and angiopoietins (Ang1, Ang2), which have been widely studied by their involvement in cancer progression [283-285]. Interestingly, increased sst5TMD4-induced VEGF secretion has been recently reported in medullar thyroid carcinoma cells [10], which nicely agrees with and extends the present findings to other relevant endocrine-related tumors. The pro-angiogenic potential of sst5TMD4 was further confirmed in

preclinical mouse models, where xenografted tumors overexpressing sst5TMD4 showed increased expression of angiogenic factors and also a higher number of blood vessels as compared with tumor generated by mock cells. More importantly, sst5TMD4 has been found to be expressed at moderate or high levels in a representative proportion of samples from a cohort of 127 grade 3 infiltrating ductal breast carcinoma tumors, which is consistent with results reported previously wherein sst5TMD4 was detected in 28% of 49 breast cancer samples from a closed random series of tumor breast cancer samples classified as poorly differentiated grade 3 (G3) tumors [18]. In this new and more ample series, sst5TMD4 presence/expression is associated to several angiogenic markers, such as VEGF, ANG1 and CD34, and with the capacity of the tumors to metastasize, mainly to lymph nodes. Nevertheless, these analyses also revealed a clinically-relevant finding, namely, that sst5TMD4 presence/expression is associated with lower disease-free survival of the patients, clearly reinforcing our notion of an involvement of the truncated receptor in breast cancer progression, and further suggesting the potential value of sst5TMD4 as a novel biomarker for breast cancer prognosis.

In this sense, it is worth noting that SST and its synthetic analogues have been shown to play crucial roles in the angiogenic process. Indeed, SST analogues were found to reduce vascular cell proliferation [286] and to prevent hypoxia-induced changes in VEGF/VEGFRs system in vascular cells [287], likely through the sst1 and/or sst4 receptors [287]. In addition, SST and its analogues seem to modulate the angiogenic process in several models of retinal angiogenesis, likely acting through the sst2 receptor [288, 289]. Thus, and although much less is known about their role in tumoral cells, it seems that SST and its analogues can reduce VEGF production from some types of tumoral cells, such as gliomas [290], gastric carcinomas [291], or pancreatic cancer [292-294], acting through the sst2 receptor subtype [292-294]. Consistent with a role of SST and its receptors in reducing the angiogenic process by acting at the endothelial cells and the tumoral cells levels [295], the use of SST analogues in clinical trials has revealed that SST could exert its anti-vasculogenesis effect by downregulating the serum VEGFs and, therefore, could be used as an important adjuvant to improve the survival of gastric cancer patients [291].

However, despite the fact that SST receptors are densely expressed in breast cancer samples compared with normal tissues [296], being the sst2 subtype the most frequently and abundantly expressed sst in tumor cells [297], the clinical studies reporting treatment of breast cancer patients with SSAs have only demonstrated a limited success [298]. In this context, we have recently demonstrated that the presence of the truncated receptor sst5TMD4 correlates with a worse prognosis in a group of breast cancer tumors and its overexpression is associated with increased malignant features, such as invasion and proliferation abilities (both in cell cultures and nude mice) in the breast cancer MCF-7 cell line [18]. This was likely mediated by

sst5TMD4-induced increase in phosphorylated ERK1/2 and Akt levels, which also led to a mesenchymal-like phenotype. At the same time, this study demonstrated that sst5TMD4 interacts (physically and functionally) with sst2, promoting the disruption of SST/sst2 inhibitory feedback. These previous results provide a plausible basis for the findings shown herein in that sst5TMD4 could be inducing VEGF expression/secretion through a direct [increasing phosphorylated ERK and Akt levels [299]] and/or an indirect mechanism [disrupting the inhibitory loop established between SST and the sst2 [287, 288, 273, 294]]. Surprisingly, the changes in VEGF expression/secretion observed in the MCF-7 cells were not accompanied by changes in the expression of HIF-1a and HIF-1b, which suggest that these factors are more likely regulated at the protein level (amount and/or phosphorylation status) or that sst5TMD4 increases VEGF expression through a HIF-independent mechanism [300].

Importantly, the stimulatory actions of sst5TMD4 on the production of pro-angiogenic factors from MCF-7 cells were accompanied by functional alterations in MCF-7 induced tumor xenografts. As we have reported previously, sst5TMD4 overexpressing MCF-7 cells induce the formation of larger xenograph tumors, with a more undifferentiated histologic phenotype [18]. The present study expands those previous observations by demonstrating that the presence of the sst5TMD4 receptor is also associated with a higher number of blood vessels in the tumor, which, additionally, confirms its relevant role in tumoral angiogenic processes. It is known that angiogenesis represents a crucial step for tumor growth and metastasis; however, tumor progression towards metastasis is a complex, multistage process, which is classically simplified as: local invasion, intravasation, survival in the circulation, extravasation, and colonization [301]. Tumor-induced angiogenesis promotes the formation of altered vessels, which resembles chaotic networks of tortuous endothelium lacking the normal hierarchical arrangement of artery-arteriole-capillary [302]. These altered tumor vessels facilitate tumoral cells intravasation and therefore metastasis [303]; yet, tumoral cells need to survive in bloodstream and extravasate to other tissue in order to complete the metastasis process. Nevertheless, merely about 0.2% of the tumor cells can effectively induce angiogenesis and eventually form metastases in distant organs. In this scenario, presence of CSCs has arisen as a marker of malignancy, for it is considered a fundamental driving force of tumor development, initiation of invasion, and metastasis [304]. In fact, circulating tumor cells (CTCs), that is, cells that have escaped from primary tumor to the bloodstream, can present CSC markers such as CD133, and their expression correlates with poor prognosis associated to increased metastatic potential [305, 306]. Interestingly, sst5TMD4 overexpression induced a clear enrichment in CSCs among the cell population in both, MCF-7 and MDA-MB-231 cell lines, as it is evidenced by the increase in the number of mammospheres generated *in vitro.* The ability of the truncated receptor to increase the percentage of CSCs may very well be driven by its capacity to induce the overexpression of TGF-β1, JAG1 and β-catenin, since alteration in TGF-β/BMP, Notch and Wnt/β-catenin signaling pathways have

been associated to the appearance of CSCs [306, 307]. Consistent with this idea, in this patient series, sst5TMD4 presence/expression levels were associated to several angiogenic markers and with the capacity of the tumors to metastasize, mainly to lymph nodes. Even more importantly, sst5TMD4 presence/expression was found to be associated with lower disease-free survival of the patients, which is in agreement with previous and present *in vitro* and *in vivo* studies, emphasizing the important involvement of the truncated receptor in cancer progression. The cellular mechanisms by which sst5TMD4 can exert this function are multiple, from increased proliferation and invasiveness to dedifferentiation through enhancement of EMT [18], or to stimulated growth of CSCs; but also, by affecting its environment through secretion of angiogenic factors, which could be involved in the increase of the number of endothelial vessels irrigating the tumors and probably facilitating metastasis.

Therefore, altogether, these data demonstrate that the mere presence of the truncated receptor sst5TMD4, which is overexpressed in a substantial proportion of breast cancer patients (herein and [18]), can directly and indirectly increase the capacity of breast cancer cells to produce pro-angiogenic factors, which, in turn, would promote the sprouting of new vessels, thus facilitating tumor growth and making them more prone to invade and metastasize. Consequently, the presence/expression of this truncated receptor should be considered a risk factor for breast cancer progression and, thus, these findings could be used to identify new molecular targets for diagnosis, prognosis or therapy in these tumors.

## 5.2. In1-Ghrelin

The ghrelin gene exhibits a complex structure and regulation, which leads to the generation of a number of alternative products exhibiting diverse and ample actions throughout the organism [201]. Indeed, this gene represents a paradigmatic example of the different, multilevel regulatory systems that have to act coordinately to maintain cell homeostasis, including, among others, the modulation of gene transcription, mRNA translation [207], post-translational protein modifications [308], and hormone secretion [309]. At the co-transcriptional level, the splicing process has arisen as an important mechanism, leading to the generation of several alternative splicing variants, originated from different splicing events such as the use of alternative splice site (Des-Gln14-Ghrelin), exon skipping (Exon3-deleted preproghrelin), use of an upstream exon (NM_001134944), and even intron retention (In1-ghrelin) [207]. Inasmuch as some of these splicing variants have been found to be involved in the development and progression of different cancers, like breast [17, 21] and prostate cancer [209, 225], and pituitary [15] or neuroendocrine tumors [16], the regulation of the splicing process of the ghrelin gene under these pathological conditions has emerged as a relevant task, in order to be able to use the expression of these splicing variants for the development of novel diagnostic, prognostic or therapeutic tools.

Among these splicing variants, In1-ghrelin seems to play a clinically relevant role, for it has been reported to be abundantly expressed in various cancer types [15-17], where it is associated to increased malignancy. However, the regulatory mechanisms underlying the splicing events that control its expression are still unknown. Moreover, although this splicing variant has been found to be overexpressed in several types of cancer [15-17], the functional consequences that the expression of In1-ghrelin entails are not completely elucidated. Particularly, In1-ghrelin has shown the ability to increase the secretion and cell viability of pituitary adenomas [15] and to increase proliferation in breast cancer cells [17]; however, more comprehensive studies are needed to understand the complex alteration that In1-ghrelin presence can be leading to in tumor pathologies.

### 5.2.1. Ghrelin gene splicing regulation

The factors involved in the regulation of the alternative splicing of the ghrelin gene, which could be thus responsible of the appearance of In1-ghrelin, can be classified (as in the case of the sst5TMD4 studies discussed above) in *cis* and *trans* elements. *Cis* elements are those intrinsic to the sequence of the gene, such as SNPs and *de novo* mutation, while trans elements are those that modify the splicing and/or transcription by interacting with the genome or the proteins involved in these processes.

#### Role of genomic alterations in In1-ghrelin expression

As initial approach, GHRL intron 1 sequence was analyzed in a battery of normal and tumoral breast samples in order to search for alterations that could be correlated with changes in In1-ghrelin and ghrelin mRNA expression. However, we did not find any nucleotide alteration along the 194 bases of the intron sequence, in spite of the fact that 26 low frequency SNPs have been described at NCBI within the intron sequence (none of them related with any pathology). This is, of course, a surprising and interesting result that could be explained, at least partially, by the small size of the intron; but it could also suggest the existence of key regulatory sequences within the intron, or, more intriguingly, could indicate a relevant physiological function of this splicing variant, which would have prevented the alteration of the intron 1 sequence during the evolutionary process. In this sense, it is also worth noting that In1-ghrelin is expressed at variable levels in certain normal tissues [17], reinforcing the idea of a physiological function for this splicing variant in several tissues. Interestingly, the process of intron retention in the ghrelin gene has been found to be conserved among several mammalian species, including mouse, where In1-ghrelin has an orthologous splicing variant named In2-ghrelin [210]. Comparison of the sequences of both variants revealed that, while exon sequences are greatly conserved between human and mice, intron sequences were markedly different, being human intron more than 2-fold larger than its mice orthologous. However, we found similar sequences at both, 5’ and 3’ intron extremes, which are the sequences usually recognize by the splicing regulators

during splicing process. This data reinforce the idea that the retention of the intron sequence could be similarly regulated by the splicing machinery in both species.

Involvement of splicing factors target sequences in the retention of the intron 1

As a second approach, we explored *in silico* the splicing factor target sequences present along the In1-ghrelin intron sequence. These studies revealed a balanced presence among enhancer and silencer target sites. Specifically, bioinformatic analysis indicated the existence of target sites for several splicing factors such as SRSF5 (SRp40), which is the splicing enhancer factor with more target sequences within the intron, and also one of the conserved splicing factors between human and mice, which suggest that SRSF5 could be specially relevant in this process. However we did not find literature associating this SR protein with intron retention. Alternatively, SRSF2 (SC35), which presents a target sequence within the intron, has been previously correlated with intron retention events [310]. On the other hand, hnRNP H1, 2 and 3, which also presented conserved target sites between human and mice, have been previously related with intron retention in the PABPN1 gene, associated with oculopharyngeal muscular dystrophy [311].

Additionally, to identify relevant splicing factors potentially involved in the regulation of the retention of the intron 1, we explored the splicing factor target sequences located at the regions conserved between mice and human introns, inasmuch as they could represent evolutionary conserved regulatory sequences [312]. This comparison revealed that the majority of sequences conserved corresponded to inhibitory sequences, which could indicate that the retention of the intron in the mature mRNA could be evolutionary conserved. Interestingly, among the splicing factor target sequences located at the conserved regions, several hnRNPs appeared that have been found associated to breast cancer such as hnRNP E1 [313], hnRNP H1 [314] or PTBP1 (hnRNP-I) [122]. Particularly, hnRNP E1 activation through TGF-$\beta$ mediated phosphorylation has been shown to trigger EMT in breast cancer [313]; while hnRNP H1 has been found overexpressed in breast cancer associated to HER-2 [314], where it has been associated, together with SRSF3 or SRp20 (which presents a target sequence in a non-conserved region of the intron 1) to the appearance of the oncogenic splicing variant Δ16HER2 [315]. Additionally, PTBP1 has been reported to act with both oncogenic [113,122, 125, 126] and tumor repressor [127, 128] activities. All these data highlight the complex network of interactions that could be underlying the dysregulation of the splicing process in tumoral pathologies, and also, and more importantly, establish a starting point to further analyze by *in vitro* studies the specific splicing factors implicated in the regulation of the inclusion of the intron 1 in the nascent transcript from the GHRL gene.

## Regulatory role of GHRLOS lncRNAs on In1-ghrelin mRNA stability

There is growing evidence supporting the involvement of lncRNAs in the regulation of transcription and splicing processes [310], especially those lncRNAs that comprise natural antisense transcripts (NATs) of the target gene [311,312]. Interestingly, the GHRL gene has been reported to display a NAT named **GHRLOS**, which could be involved in the regulation of the splicing process that generates the In1-ghrelin variant. Intriguingly, GHRLOS represents a recently identified NATs, whose structure, exon composition and functionality is still controversial. Initially, GHRLOS was described to be composed of 4 exons numbered as exon 1, exon 2, exon 3 and exon 4, which could be differentially assembled to generate 4 different splicing transcript variants (GHRLOS-1, GHRLOS-2, GHRLOS-3 and GHRLOS-4). However, subsequent revisions of the sequence and structure of this antisense gene included 3 novel upstream exon named as exon I, exon II and exon III. This revision of the gene also implicated the addition of three new splicing variants (from GHRLOS-5 to GHRLOS-7) and the elimination of one of them (GHRLOS-4) [313, 314]. However, our studies were only capable to confirm the existence of the three initially described variants (GHRLOS-1, GHRLOS-2 and GHRLOS-3) despite the numerous attempts to validate the remaining splicing variants.

Based on these results, we explored the putative correlations between the expression of In1-ghrelin variant and the three initially described GHRLOS splicing variants in a battery of normal tissue samples with variable expression of In1-ghrelin mRNA. Interestingly, these results demonstrated a clear positive correlation between In1-ghrelin and the three GHRLOS splicing variants (GHRLOS-1, GHRLOS-2 and GHRLOS-3). These data clearly suggests a putative role of these GHRLOS splicing variants in regulating the GHRL gene, by specifically controlling the retention of the intron 1 and, thereby, regulating In1-ghrelin expression. Of note, none of the GHRLOS variants examined displayed levels of expression that correlated with ghrelin expression, which is in line with a specific regulation of In1-ghrelin, but not ghrelin, by these variants. The mechanisms of action underlying this specificity are unknown, but may involve diverse functions, like chromatin remodeling, which alters the binding of splicing regulators [236], or, masking of splice sites, avoids the spliceosome binding to the sequence [74] (for an extensive review of NATs functions see [311]). Interestingly, all known GHRLOS splicing variants share the exon 4, which overlaps with the promoter sequence of the GHRL gene and would therefore enable a putative functional interaction between the GHRLOS variants and the GHRL promoter. In this scenario, since GHRLOS expression is associated to In1-ghrelin expression, it could be conceivable that modifications at the GHRL gene promoter activity may contribute to regulate the In1ghrelin splicing variant expression, similar to the changes in CD44 splicing induced by changes in its promoter activity described in [315]. However, unequivocal evidence that GHRLOS splicing variants regulate In1-ghrelin transcription must be further obtained in *in vitro* studies.

### 5.2.2. Role of In1-ghrelin in breast cancer malignancy

As mentioned above, In1-ghrelin is overexpressed in a number of tumoral pathologies including pituitary [15] and neuroendocrine tumors [16], as well as in breast cancer [17], wherein In1-ghrelin increases malignant features such as hormone secretion and cell viability or proliferation rate. Although these studies are promising and point out a relevant role of this splicing variant in an ample number of pathologies, its precise role in breast cancer cells and its putative clinical relevance in breast cancer patients remain poorly understood. Thus, to explore the functional implications of In1-ghrelin on breast cancer cells, we used two different model cell lines, MDA-MB-231 and MCF-7, which represent different tumor subtypes. Specifically, MDA-MB-231 cell line represents a highly malignant triple negative tumor, which presents a highly advanced mesenchymal phenotype [316], while MCF-7 is an epithelial-like tumor derived cell line with a low malignancy phenotype [317]. Interestingly, the results generated herein indicate that In1-ghrelin exerts similar actions (at the level of proliferation, migration, or number of CSCs) in both cell lines, which suggest that this splicing variant could play a general role in breast cancer malignancy independently of the breast cancer subtype or disease stage, which increases the significance of the results obtained.

In particular, both cell lines showed increased proliferation and migration rates in response to In1-ghrelin overexpression and treatment, which is in line with the results reported previously [17], where transient transfection of In1-ghrelin in MDA-MB-231 enhanced cell proliferation. Moreover, endogenous In1-ghrelin downregulation on MDA-MB-231 by siRNA treatment reduced both migration and proliferation rates, which provides a key proof-of-concept and further corroborates the results obtained after In1-ghrelin overexpression and treatments. The fact that these results also compare favorably with those published in pituitary and neuroendocrine tumor cells treated or transfected with In1-ghrelin [15, 16], comprise compelling evidence that In1-ghrelin can play a common, pathological role across different types of tumors.

At variance with In1-ghrelin, overexpression or treatment with native ghrelin showed variable results, rising proliferation rate in MDA-MB-231 but not in MCF-7 cells, and improving migration ability in MDA-MB-231 after pharmacological treatment, but not after overexpression, results that are also in line with those observed in pituitary tumors [15]. Variable effects of ghrelin on tumoral behavior have been already reported [152]. Indeed, ghrelin can exert proliferative effects in some cell lines but not in others [22], and can even decrease cell proliferation in some cases on MCF-7 cell line [318]. On the other hand, ghrelin effects in cell motility have been less studied in the context of breast cancer, as it has only been related with increased migration rate in canine breast cancer cell lines (CMT-W1 and CMT-W2) [319]. The differences observed herein, between transfection and treatment, are possibly due to differences in the effective concentration achieved in each set of experiments. In this line, previous studies [319] observed that, while treatment with low doses (1-10nM) of

ghrelin peptide induces increased migration, high doses (100nM), which are similar to those used in our studies, did not induce significant changes in this parameter.

Interestingly, In1-ghrelin overexpression also induced changes in the basal phosphorylation rate of the tumor malignancy-associated signaling pathway MEK/ERK [320]. Specifically, overexpression of In1-ghrelin, but not ghrelin, increased ERK phosphorylation, which has been associated with tumor proliferation and malignancy [321] and could help to explain the changes in cellular proliferation and migration observed herein. It is important to highlight that although both, ghrelin and In1-ghrelin, affected proliferation and migration, at least in MDA-MB-231 cell line, only In1-ghrelin overexpression increased the basal phosphorylation of ERK, suggesting that the effects of each splice variant can be exerted through different receptors and/or signaling cascades. Moreover, MDA-MB-231 and MCF-7 cell lines showed virtually negligible mRNA levels of the known ghrelin receptor GHSR-1a or its truncated variant GHSR-1b, thus implying that the effects exerted by ghrelin and In1-ghrelin must be triggered by different, likely unknown alternative receptors.

In addition, In1-ghrelin was also able to modulate cell dedifferentiation, a process related to tumor malignancy [322] that is linked with key aggressiveness features, such as proliferation, migration, invasion, metastasis, and chemotherapy resistance [323]. Two important elements of tumor cell dedifferentiation are EMT [324] and the maintenance of CSCs [325]. Of note, our study demonstrated that both processes could be enhanced by the presence of In1-ghrelin, but not ghrelin. Specifically, EMT implies the transformation of epithelial cells to cells with a mesenchymal phenotype, which facilitates migration and invasion abilities, improving proliferation rates [326]. Interestingly, In1-ghrelin overexpression increased mesenchymal-like cell percentage in MDA-MB-231 cell line, but not in MCF-7, which could imply cell line specific effects of this splicing variant in breast cancer. Remarkably, although MDA-MB-231 cell line exhibits several features of EMT, In1-ghrelin overexpression aggravated this phenotype. In contrast, in MCF-7 cells, which present a clear epithelial phenotype, In1-ghrelin overexpression did not induce significant morphological changes. Therefore, the effects of In1-ghrelin on breast cancer cell plasticity could be cell line-specific.

CSCs are not only present in breast cancer [327], but in virtually every type of cancer, including colon [328], brain [329], gastric [330], prostate [331], pancreatic [332] and hematopoietic tumors [333]. As mentioned above, CSCs are postulated to reside in the basis of tumor development, recurrence, and drug resistance [298]. Hence, the increase in the number of CSCs induced by In1-ghrelin, but not ghrelin overexpression and treatment in both MDA-MB-231 and MCF-7 cell lines indicates that In1-ghrelin could play a relevant, selective role in cell dedifferentiation in breast cancer. Furthermore, these results are consistent with, and can provide a mechanistic basis for the association found herein between high In1-ghrelin expression levels and increased

lymph node metastasis, and, most relevantly, with diminished disease-free survival in breast cancer patients.

Interestingly, while In1-ghrelin only modulated cell plasticity in MDA-MB-231, but not in MCF-7, cells, the overexpression of In1-ghrelin in MCF-7 cells appeared more effective in inducing mammosphere formation than in MDA-MB-231 cells, despite the fact that both cell lines showed similar levels of In1-ghrelin overexpression. These differences may relate to the fact that non-transfected MDA-MB-231 cells already presented appreciable basal expression levels of In1-ghrelin, whereas MCF-7 did not express In1-ghrelin at detectable levels. Thus, it is not unreasonable to suggest that endogenous In1-ghrelin expression in MDA-MB-231 could be exerting a basal stimulation of CSC population, which would somewhat conceal the effects of an overexpression of this splicing variant. Nevertheless, In1-ghrelin overexpressing MDA-MB-231 cells generated larger mammospheres (as measured by the number of cells that integrate each mammosphere), which suggest that In1-ghrelin overexpressing MDA-MB-231 CSCs could present increased proliferation rates.

Due to the noticeable overexpression of In1-ghrelin in breast cancer [17] and the marked functional effects of this peptide on tumor cell malignancy, elucidation of the signaling mechanisms underlying the actions of In1-ghrelin in breast cancer become imperative. The information obtained would be important, first, to better understand the role and relevance of In1-ghrelin in breast cancer, but it can also provide a valuable knowledge to identify biomarkers and design new therapeutic tools for the treatment of this pathology. Due to the marked effects found for In1-ghrelin in the dedifferentiation processes, we analyzed the expression levels of several candidate molecules, which are known to be involved in signaling pathways associated with cell plasticity and CSCs [334-336]. Particularly, TGF-β [301], Notch [337] and Wnt/β-catenin [338] signaling pathways seem to be tightly associated with both cell plasticity and CSCs, exhibiting a marked bidirectional crosstalk [239]. Interestingly, our present results demonstrate that both cell lines exhibited an altered expression of TGF-β and JAG1 in response to both ghrelin and In1-ghrelin peptides; whereas, in contrast, only In1-ghrelin overexpression induced changes in β-catenin expression levels, which could represent an In1-ghrelin specific signaling pathway that may mediate the In1-ghrelin specific effects observed in our studies. In addition, the alteration in mRNA expression levels induced by In1-ghrelin in both cell lines was clearly different: while TGF-β1 was strongly overexpressed in MDA-MB-231 cell line with modest, but significant, increases in JAG1 and β-catenin mRNA expression, MCF-7 cells transfected with In1-ghrelin showed a discrete increase in TGF-β1 expression but an almost 6-fold increased JAG1 expression, and a strong overexpression of β-catenin compared with that of MDA-MB-231 cells. The disparate relative strength of In1-ghrelin in the stimulation of this three signaling pathways in both cell lines could be associated with some of the differences observed in the response to In1-ghrelin. These data, together with the fact that ghrelin

overexpression induced dissimilar changes in the expression of the mentioned signaling pathways, suggest that In1-ghrelin triggers specific changes in TGF-β1, Notch and Wnt/β-catenin signaling pathways to modulate cell plasticity and CSCs biology, which could be breast cancer cell line-dependent.

Most importantly, as it was highlighted above, all these *in vitro* studies support the relevant association found between high In1-ghrelin expression levels and lymph node metastasis, where patient categorized as low or null expression of In1-ghrelin presented less metastasis than those included in the high-expressing category. Moreover, 10 years survival studies on these patients showed that high expression of In1-ghrelin correlated with diminished disease-free survival, which could be in line with the higher number of CSCs induced by In1-ghrelin in both tumoral cell lines, since CSCs have been associated with tumor relapse [339] and metastatic potential [340].

In summary, the present study provides compelling support for the possible implication of In1-ghrelin on breast cancer, where it could influence tumor progression, metastasis, and relapse in clinical patients. Moreover, this work demonstrates the ability of the In1-ghrelin splicing variant to enhance the malignancy features of two breast cancer derived cell lines (MDA-MB-231 and MCF-7), by increasing proliferation and migration rates, as well as higher capacity to form mammospheres in response to In1-ghrelin over-exposition (overexpression and/or peptide treatment). Moreover, our study unveils key signaling cascades which can serve in the mechanism of action of In1-ghrelin in this context: namely, the activation of the MAPK-ERK, Jag1/Notch, Wnt/β-catenin and/or TGF-β1 signaling pathways. Of note, these effects are different from those exerted by the canonical splicing variant of the ghrelin gene and, although further studies are needed, these results postulate In1-ghrelin as a promising marker/target in breast cancer diagnosis and treatment .

# Conclusions

# 6. Conclusions

1. **sst5TMD4 expression** is a tightly and intricately modulated process that likely involves a number of **regulatory layers**, wherein **genomic alterations**, **splicing factors**, **RNA-binding proteins,** and **miRNAs** can play crucial roles. In particular:

   **1.1.** Genomic variability within the SST5 cryptic intron originated by the SNPs rs195570 and rs12599155 is associated with breast cancer features and enhanced sst5TMD4 mRNA expression, respectively, and could therefore provide new genetic markers of tumoral malignancy.

   **1.2.** The pattern of splicing factor target sequences within the SST5 gene suggests a sophisticated regulation of the splicing of sst5TMD4 cryptic intron. The presence of clusters of target sequences for silencer splicing factors could imply an especially relevant role of the hnRNPs during sst5TMD4 generation.

   **1.3.** Regulation of sst5TMD4 mRNA stability by miRNAs is a plausible step in sst5TMD4 protein appearance. In particular, hsa-miR-346 is associated with the expression of sst5TMD4 in breast cancer samples and tightly modulates the expression levels of the sst5TMD4 truncated receptor *in vitro*.

2. **sst5TMD4** expression can increase breast cancer malignancy by **enhancing** of **angiogenesis** and **dedifferentiation** processes, and increasing the **metastatic potential** of tumoral cells. Specifically:

   **2.1.** sst5TMD4 exhibits a strong **angiogenic potential,** as it stimulates expression and secretion of pro-angiogenic factors (especially VEGF) in *in vitro* and *in vivo* models, increases vessel formation in xenografted tumors and directly associates with the expression of pro-angiogenic markers in human breast cancer samples.

   **2.2.** sst5TMD4 seems to influence the **metastatic potential** of breast cancer cells, as shown by its ability to modulate cell plasticity and CSCs population in breast cancer cell lines, likely through Jag1/Notch, Wnt/B-catenin and/or TGF-b signaling pathways. Indeed, sst5TMD4 expression positively correlates with lymph-node and distant metastasis in breast cancer patients and, ultimately, with the onset of disease-free survival, thus suggesting a relevant role of this splicing variant in the progression and malignancy of breast cancer.

3. **In1-ghrelin** expression is tightly modulated by several *trans* regulatory elements, wherein **splicing factor micro-environment** and expression of lncRNA NATs from **GHRLOS gene** could play a relevant role.

   **3.1.** The lncRNA NATs from **GHRLOS gene** could play a relevant role in controlling GHRL gene splicing processes and, therefore, in the generation of the different GHRL gene derived splicing variants, in that In1-ghrelin mRNA expression strongly correlates with the expression of, at least, three GHRLOS splicing variants, which do not correlate with native ghrelin expression.

   **3.2.** *In silico* analysis of intron 1 sequence and structure suggests that the small size of the intron, the relatively weakness of the 5' splice site and the splicing factors cellular environment could play a key role in the intron definition and, therefore, in the appearance of In1-ghrelin.

   **3.3.** On the contrary, In1-ghrelin expression does not seem to be relevantly determined by genomic alterations inasmuch as the GHRL gene intron 1 did not present any punctual mutation in the samples analyzed.

4. In1-**ghrelin** presence can **enhance** the **progression** and **malignancy** of **breast cancer** by acting **at several levels** of cancer pathology, which suggests a **relevant role** for this splice variant in breast cancer. Specifically:

   **4.1.** In1-ghrelin splicing variant modulates proliferation and migration abilities of breast cancer cell lines, likely through the activation of MEK/ERK pathway, as shown by overexpression, silencing, and peptide treatment experiments. Interestingly, these actions were different from those exerted by native ghrelin, indicating dissimilar roles of ghrelin gene derived products on breast cancer cells.

   **4.2.** In1-ghrelin increases plasticity and dedifferentiation of breast cancer cell lines probably by modulation of Jag1/Notch and Wnt/β-catenin signaling pathways. In line with this, In1-ghrelin expression correlates with lymph node metastasis and lower disease-free survival in breast cancer patients, which strongly suggests a prominent role of this splicing variant in the progression of breast cancer pathology. Interestingly, these effects are specific of this splicing variant since native ghrelin did not exhibit any effects on these processes.

***Global Corollary:***

From a general point of view, the studies implemented in the present Thesis regarding the presence and functional/pathological role of sst5TMD4 and In1-ghrelin splicing variants in breast cancer expand and reinforce the contention of the key role that the dysregulation of the splicing process represents in endocrine-related cancers, their development and progression. More specifically, the results presented herein demonstrate that both splicing variants (sst5TMD4 and In1-ghrelin) could represent promising prognostic and therapeutic target in breast cancer pathology, which invites to the implementation of future efforts to generate and examine novel approaches based on these splicing variants.

# Bibliography

# 7. Bibliography


1. Siegel RL, Miller KD, Jemal A, *Cancer statistics, 2016.* CA Cancer J Clin, 2016. 66(1): p. 7-30
2. Koutsogiannouli E, Papavassiliou AG, Papanikolaou NA, *Complexity in cancer biology: is systems biology the answer?* Cancer Med., 2013. 2(2): p. 164-77
3. Caiado F, Silva-Santos B, Norell H, *Intra-tumour heterogeneity - going beyond genetic.* FEBS J., 2016. 283(12): p. 2245-58.
4. Hanahan D, Weinberg RA, *Hallmarks of cancer: the next generation.* Cell, 2011. 144(5): p. 646-74
5. Hanahan D, Weinberg RA, *The hallmarks of cancer.* Cell, 2000. 100(1): p. 57-70.
6. Ladomery M, *Aberrant Alternative Splicing Is Another Hallmark of Cancer.* Int J Cell Biol, 2013. 2013: p. 463786.
7. Stratton MR, Campbell PJ, Futreal PA, *The cancer genome.* Nature, 2009. 458(7239): p. 719-24.
8. Koren S, Bentires-Alj M, *Breast Tumor Heterogeneity: Source of Fitness, Hurdle for Therapy.* Mol Cell, 2015.60(4): p. 537-46.
9. Brooks MD, Burness ML, Wicha MS, *Therapeutic Implications of Cellular Heterogeneity and Plasticity in Breast Cancer.* Cell Stem Cell. 2015. 17(3): p. 260-71
10. Mole D, Gentilin E, Ibáñez-Costa A, et al., *The expression of the truncated isoform of somatostatin receptor subtype 5 associates with aggressiveness in medullary thyroid carcinoma cells.* Endocrine, 2015.
11. Durán-Prado M, Gahete MD, Martínez-Fuentes AJ, et al., *Identification and characterization of two novel truncated but functional isoforms of the somatostatin receptor subtype 5 differentially present in pituitary tumors.* J Clin Endocrinol Metab, 2009. 94(7): p. 2634-43.
12. Durán-Prado M, Saveanu A, Luque RM, et al., *A potential inhibitory role for the new truncated variant of somatostatin receptor 5, sst5TMD4, in pituitary adenomas poorly responsive to somatostatin analogs.* J Clin Endocrinol Metab, 2010. 95(5): p. 2497-502.
13. Puig-Domingo M, Luque RM, Reverter JL, et al., *The truncated isoform of somatostatin receptor5 (sst5TMD4) is associated with poorly differentiated thyroid cancer.* PLoS One, 2014. 9(1): p. e85527.
14. Luque RM, Ibáñez-Costa A, Neto LV, et al., *Truncated somatostatin receptor variant sst5TMD4 confers aggressive features (proliferation, invasion and reduced octreotide response) to somatotropinomas.* Cancer Lett, 2015. 359(2): p. 299-306.
15. Ibáñez-Costa A, Gahete MD, Rivero-Cortés E, et al., *In1-ghrelin splicing variant is overexpressed in pituitary adenomas and increases their aggressive features.* Sci Rep, 2015. 5: p. 8714.
16. Luque RM, Sampedro-Núñez M, Gahete MD, et al., *In1-ghrelin, a splice variant of ghrelin gene, is associated with the evolution and aggressiveness of human neuroendocrine tumors: Evidence from clinical, cellular and molecular parameters.* Oncotarget, 2015. 6(23): p. 19619-33
17. Gahete MD, Córdoba-Chacón J, Hergueta-Redondo M, et al., *A novel human ghrelin variant (In1-ghrelin) and ghrelin-O-acyltransferase are overexpressed in breast cancer: potential pathophysiological relevance.* PLoS One, 2011. 6(8): p. e23302.
18. Durán-Prado M, Gahete MD, Hergueta-Redondo M, et al., *The new truncated somatostatin receptor variant sst5TMD4 is associated to poor prognosis in breast cancer and increases malignancy in MCF-7 cells.* Oncogene, 2012. 31(16): p. 2049-61.

19. Sampedro-Núñez M, Luque RM, Ramos-Levi AM, et al., *Presence of sst5TMD4, a truncated splice variant of the somatostatin receptor subtype 5, is associated to features of increased aggressiveness in pancreatic neuroendocrine tumors.* Oncotarget. 2016. 7(6): p. 6593-608
20. Henderson BE, Ross RK, Pike MC, et al., *Endogenous hormones as a major factor in human cancer.* Cancer Res, 1982. 42(8): p. 3232-9.
21. Belfiore A, Perks CM, *Grand challenges in cancer endocrinology: endocrine related cancers, an expanding concept.* Front Endocrinol (Lausanne), 2013. 4: p. 141.
22. Jeffery PL, Murray RE, Yeh AH, et al., *Expression and function of the ghrelin axis, including a novel preproghrelin isoform, in human breast cancer tissues and cell lines.* Endocr Relat Cancer, 2005. 12(4): p. 839-50.
23. Venter JC, Adams MD, Myers EW, et al., *The sequence of the human genome.* Science, 2001. 291(5507): p. 49.
24. Hillier LW, Coulson A, Murray JI, et al., *Genomics in C. elegans: so many genes, such a little worm*. Genome Res, 2005. 15(12): p. 1651-60.
25. Ruppert EE, Fox RS, Barnes RD, *Invertebrate Zoology: A Functional Evolutionary Approach*. 7th ed. Thomson-Brooks/Cole 2004. p. 989
26. Albertin W, Marullo P, Bely M, et al., *Linking post-translational modifications and variation of phenotypic traits*. Mol Cell Proteomics, 2013. 12(3): p. 720-35.
27. Murphy KP, *Janeway's Immunobiology*. 8th ed. Garland Science. 2012. p. 823.
28. Martínez-Montiel N, Rosas-Murrieta N, Martinez-Contreras R, *[Alternative splicing regulation: implications in cancer diagnosis and treatment]*. Med Clin (Barc), 2015. 144(7): p. 317-23.
29. Kornblihtt AR, de la Mata M, Fededa JP, et al., *Multiple links between transcription and splicing*. RNA, 2004. 10(10): p. 1489-98.
30. Grzybowska EA, *Human intronless genes: functional groups, associated diseases, evolution, and mRNA processing in absence of splicing*. Biochem Biophys Res Commun, 2012. 424(1): p. 1-6.
31. Papasaikas P, Valcarcel J*, The Spliceosome: The Ultimate RNA Chaperone and Sculptor*. Trends Biochem Sci, 2015.
32. Matera AG, Wang Z, *A day in the life of the spliceosome.* Nat Rev Mol Cell Biol, 2014. 15(2): p. 108-21.
33. Keren H, Lev-Maor G, Ast G, *Alternative splicing and evolution: diversification, exon definition and function.* Nat Rev Genet, 2010. 11(5): p. 345-55.
34. Baralle D, Baralle M, *Splicing in action: assessing disease causing sequence changes.* J Med Genet, 2005. 42(10): p. 737-48.
35. Burset M, Seledtsov IA, Solovyev VV, *Analysis of canonical and non-canonical splice sites in mammalian genomes.* Nucleic Acids Res, 2000. 28(21): p. 4364-75.
36. Sheth N, Roca X, Hastings ML, et al., *Comprehensive splice-site analysis using comparative genomics.* Nucleic Acids Res, 2006. 34(14): p. 3955-67.
37. Patel AA, Steitz JA, *Splicing double: insights from the second spliceosome.* Nat Rev Mol Cell Biol, 2003. 4(12): p. 960-70.
38. Twigg SR, Burns HD, Oldridge M, et al., *Conserved use of a non-canonical 5' splice site (/GA) in alternative splicing by fibroblast growth factor receptors 1, 2 and 3.* Hum Mol Genet, 1998. 7(4): p. 685-91.
39. Brackenridge S, Wilkie AO, Screaton GR, *Efficient use of a 'dead-end' GA 5' splice site in the human fibroblast growth factor receptor genes.* EMBO J, 2003. 22(7): p. 1620-31.

40. Quan F, Forte MA, *Two forms of Drosophila melanogaster Gs alpha are produced by alternate splicing involving an unusual splice site.* Mol Cell Biol, 1990. 10(3): p. 910-7.
41. Pollard AJ, Krainer AR, Robson SC, et al., *Alternative splicing of the adenylyl cyclase stimulatory G-protein G alpha(s) is regulated by SF2/ASF and heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1) and involves the use of an unusual TG 3'-splice Site.* J Biol Chem, 2002. 277(18): p. 15241-51.
42. Szafranski K, Schindler S, Taudien S, et al., *Violating the splicing rules: TG dinucleotides function as alternative 3' splice sites in U2-dependent introns.* Genome Biol, 2007. 8(8): p. R154.
43. Wahl MC, Will CL, Lührmann R, *The spliceosome: design principles of a dynamic RNP machine.* Cell, 2009. 136(4): p. 701-18.
44. Lee C, Namshin K, Meenakshi R, et al., *Massive expansions of Dscam splicing diversity via staggered homologous recombination during arthropod evolution.* RNA, 2010. 16(1): p. 91-105.
45. Wang ET, Rickard S, Shujun L, et al., *Alternative isoform regulation in human tissue transcriptomes.* Nature, 2008. 456(7221): p. 470-6.
46. Pan Q, Shai O, Lee LJ, et al., *Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing.* Nat Genet, 2008. 40(12): p. 1413-5.
47. Koren E, Lev-Maor G, Ast G, *The emergence of alternative 3' and 5' splice site exons from constitutive exons.* PLoS Comput Biol, 2007. 3(5): p. e95.
48. Cartegni L, Chew SL, Krainer AR, *Listening to silence and understanding nonsense: exonic mutations that affect splicing.* Nat Rev Genet, 2002. 3(4): p. 285-98.
49. Wang Z, Burge CB, *Splicing regulation: from a parts list of regulatory elements to an integrated splicing code.* RNA, 2008. 14(5): p. 802-13.
50. Giulietti M, Piva F, D´Antonio M, et al., *SpliceAid-F: a database of human splicing factors and their RNA-binding sites.* Nucleic Acids Res, 2013. 41(Database issue): p. D125-31.
51. Blencowe BJ, *Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases.* Trends Biochem Sci, 2000. 25(3): p. 106-10.
52. Wang Y, Ma M, Xiao X, et al., *Intronic splicing enhancers, cognate splicing factors and context-dependent regulation rules.* Nat Struct Mol Biol, 2012. 19(10): p. 1044-52.
53. Pozzoli U, Sironi M, *Silencers regulate both constitutive and alternative splicing events in mammals.* Cell Mol Life Sci, 2005. 62(14): p. 1579-604.
54. Silipo M, Gautrey H, Tyson-Capper A, *Deregulation of splicing factors and breast cancer development.* J Mol Cell Biol, 2015.
55. David CJ, Manley JL, *The RNA polymerase C-terminal domain: a new role in spliceosome assembly.* Transcription, 2011. 2(5): p. 221-5.
56. Zhong XY, Wang P, Han J, et al., *SR proteins in vertical integration of gene expression from transcription to RNA processing to translation.* Mol Cell, 2009. 35(1): p. 1-10.
57. Twyffels L, Gueydan C, Kruys V, *Shuttling SR proteins: more than splicing factors.* FEBS J, 2011. 278(18): p. 3246-55.
58. He Y, Brown MA, Rothnagel JA, et al., *Roles of heterogeneous nuclear ribonucleoproteins A and B in cell proliferation.* J Cell Sci, 2005. 118(Pt 14): p. 3173-83.
59. He L, Hannon GJ, *MicroRNAs: small RNAs with a big role in gene regulation.* Nat Rev Genet, 2004. 5(7): p. 522-31.
60. Quinn JJ, Chang HY, *Unique features of long non-coding RNA biogenesis and function.* Nat Rev Genet. 2016. 17(1): p. 47-62.

61. Fatica A, Bozzoni I, *Long non-coding RNAs: new players in cell differentiation and development.* Nat Rev Genet. 2014. 15(1): p. 7-21
62. Rodríguez A, Griffiths-Jones S, Ashurst JL, et al., *Identification of mammalian microRNA host genes and transcription units.* Genome Res, 2004. 14(10A): p. 1902-10.
63. Kim YK, Kim VN, *Processing of intronic microRNAs.* EMBO J, 2007. 26(3): p. 775-83.
64. Shomron N, Golan D, Hornstein E, *An evolutionary perspective of animal microRNAs and their targets.* J Biomed Biotechnol. 2009. 2009: p. 594738
65. Wahid F, Shehzad A, Khan T, et al., *MicroRNAs: synthesis, mechanism, function, and recent clinical trials.* Biochim Biophys Acta, 2010. 1803(11): p. 1231-43.
66. Valencia-Sánchez MA, Liu J, Hannon GJ, et al., *Control of translation and mRNA degradation by miRNAs and siRNAs.* Genes Dev, 2006. 20(5): p. 515-24.
67. Zhang X, Zuo X, Yang B, et al., *MicroRNA directly enhances mitochondrial translation during muscle differentiation.* Cell, 2014. 158(3): p. 607-19.
68. Kalsotra A, Wang K, Li PF, et al., *MicroRNAs coordinate an alternative splicing network during mouse postnatal heart development.* Genes Dev, 2010. 24(7): p. 653-8.
69. Bertone P, Stolc V, Royce TE, et al., *Global identification of human transcribed sequences with genome tiling arrays.* Science, 2004. 306(5705): p. 2242-6.
70. Guttman M, Amit I, Garber M, et al., *Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals.* Nature, 2009. 458(7235): p. 223-7.
71. Pelechano V, Steinmetz LM, *Gene regulation by antisense transcription.* Nat Rev Genet, 2013. 14(12): p. 880-93
72. He Y, Vogelstein B, Velculescu VE, et al., *The antisense transcriptomes of human cells.* Science, 2008. 322(5909): p. 1855-7.
73. Katayama S, Tomaru Y, Kasukawa T, et al., *Antisense transcription in the mammalian transcriptome.* Science, 2005. 309(5740): p. 1564-6.
74. Okada Y, Tashiro C, Numata K, et al., *Comparative expression analysis uncovers novel features of endogenous antisense transcription.* Hum Mol Genet, 2008. 17(11): p. 1631-40.
75. Beltran M, Puig I, Peña C, et al., *A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition.* Genes Dev, 2008. 22(6): p. 756-69.
76. Tripathi V, Ellis JD, Shen Z, et al., *The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation.* Mol Cell, 2010. 39(6): p. 925-38.
77. Sveen A, Kilpinen S, Ruusulehto A, et al., *Aberrant RNA splicing in cancer; expression changes and driver mutations of splicing factor genes.* Oncogene, 2015. 35(19): p. 2413-27
78. Barta A, Schumperli D, *Editorial on alternative splicing and disease.* RNA Biol, 2010. 7(4): p. 388-9.
79. Cooper TA, Wan L, Dreyfuss G, *RNA and disease.* Cell, 2009. 136(4): p. 777-93.
80. Dhir A, Buratti E, *Alternative splicing: role of pseudoexons in human disease and potential therapeutic strategies.* FEBS J, 2010. 277(4): p. 841-55.
81. Hallegger M, Llorian M, Smith CW, *Alternative splicing: global insights.* FEBS J, 2010. 277(4): p. 856-66.
82. Scotti MM, Swanson MS, *RNA mis-splicing in disease.* Nat Rev Genet. 2016. 17(1): p. 19-32
83. Chen J, Weiss WA, *Alternative splicing in cancer: implications for biology and therapy*. Oncogene. 2015. 34(1): p. 1-14.
84. Hu A, Fu XD, *Splicing oncogenes.* Nat Struct Mol Biol, 2007. 14(3): p. 174-5.

85. Dutertre M, Vagner S, Auboeuf D, *Alternative splicing and breast cancer.* RNA Biol, 2010. 7(4): p. 403-11
86. Miao L, Cao Z, Shen C, et al., *Cloning and functional identification of two novel BRCA1 splicing variants.* Biochemistry (Mosc), 2008. 73(11): p. 1214-23.
87. David CJl Manley JL, *Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged.* Genes Dev, 2010. 24(21): p. 2343-64.
88. Kim E, Goren A, Ast G, *Insights into the connection between cancer and alternative splicing.* Trends Genet, 2008. 24(1): p. 7-10.
89. He C, Zhou F, Zuo Z, et al., *A global view of cancer-specific transcript variants by subtractive transcriptome-wide analysis.* PLoS One, 2009. 4(3): p. e4732.
90. Ritchie W, Granjeaud S, Puthier D, et al., *Entropy measures quantify global splicing disorders in cancer.* PLoS Comput Biol, 2008. 4(3): p. e1000011.
91. Sette C, Ladomery M, Ghigna C, *Alternative splicing: role in cancer development and progression.* Int J Cell Biol, 2013. 2013: p. 421606.
92. Pagliarini V, Naro C, Sette C, *Splicing Regulation: A Molecular Device to Enhance Cancer Cell Adaptation.* Biomed Res Int, 2015. 2015: p. 543067.
93. Venables JP, *Aberrant and alternative splicing in cancer.* Cancer Res, 2004. 64(21): p. 7647-54.
94. Sebestyén E, Zawisza M, Eyras E, *Detection of recurrent alternative splicing switches in tumor samples reveals novel signatures of cancer*. Nucleic Acids Res. 2015. 43(3): p. 1345-56
95. Sebestyén E, Singh B, Miñana B, et al. *Large-scale analysis of genome and transcriptome alterations in multiple tumors unveils novel cancer-relevant splicing networks.* Genome Res. 2016. 26(6): p. 732-44.
96. Ward AJ, Cooper TA, *The pathobiology of splicing.* J Pathol, 2010. 220(2): p. 152-63.
97. Raponi M, Baralle D, *Alternative splicing: good and bad effects of translationally silent substitutions.* FEBS J, 2010. 277(4): p. 836-40.
98. Xing Y, Lee C, *Alternative splicing and RNA selection pressure--evolutionary consequences for eukaryotic genomes.* Nat Rev Genet, 2006. 7(7): p. 499-509.
99. Matlin AJ, Clark F, Smith CW, *Understanding alternative splicing: towards a cellular code.* Nat Rev Mol Cell Biol, 2005. 6(5): p. 386-98.
100. Grosso AR, Martins S, Carmo-Fonseca M, *The emerging role of splicing factors in cancer.* EMBO Rep, 2008. 9(11): p. 1087-93.
101. Karni R, de Stanchina E, Lowe SW, et al., *The gene encoding the splicing factor SF2/ASF is a proto-oncogene.* Nat Struct Mol Biol, 2007. 14(3): p. 185-93.
102. Ghigna C, Giordano S, Shen H, et al., *Cell motility is controlled by SF2/ASF through alternative splicing of the Ron protooncogene.* Mol Cell, 2005. 20(6): p. 881-90.
103. Xiao R, Sun Y, Ding JH, et al., *Splicing regulator SC35 is essential for genomic stability and cell proliferation during mammalian organogenesis.* Mol Cell Biol, 2007. 27(15): p. 5393-402.
104. Shi J, Hu Z Pabon K, et al., *Caffeine regulates alternative splicing in a subset of cancer-associated genes: a role for SC35.* Mol Cell Biol, 2008. 28(2): p. 883-95.
105. Jia R, Li C, McCoy JP, et al., *SRp20 is a proto-oncogene critical for cell proliferation and tumor induction and maintenance.* Int J Biol Sci, 2010. 6(7): p. 806-26.
106. Stickeler E, Kittrell F, Medina D, et al., *Stage-specific changes in SR splicing factors and alternative splicing in mammary tumorigenesis.* Oncogene, 1999. 18(24): p. 3574-82.
107. Tang Y, Horikawa I, Ajiro M, et al., *Downregulation of splicing factor SRSF3 induces p53beta, an alternatively spliced isoform of p53 that promotes cellular senescence.* Oncogene, 2013. 32(22): p. 2792-8.
108. He X, Ee PL, Coon JS, et al., *Alternative splicing of the multidrug resistance protein 1/ATP binding cassette transporter subfamily gene in ovarian cancer creates functional splice*

*variants and is associated with increased expression of the splicing factors PTB and SRp20.* Clin Cancer Res, 2004. 10(14): p. 4652-60.

109. Huang CS, Shen CY, Wang HW, et al., *Increased expression of SRp40 affecting CD44 splicing is associated with the clinical outcome of lymph node metastasis in human breast cancer.* Clin Chim Acta, 2007. 384(1-2): p. 69-74.
110. Gautrey HL, Tyson-Capper AJ, *Regulation of Mcl-1 by SRSF1 and SRSF5 in cancer cells.* PLoS One, 2012. 7(12): p. e51497.
111. Lal S, Allan A, Markovic D, et al., *Estrogen alters the splicing of type 1 corticotropin-releasing hormone receptor in breast cancer cells.* Sci Signal, 2013. 6(282): p. ra53.
112. Pind MT, Watson PH, *SR protein expression and CD44 splicing pattern in human breast tumours.* Breast Cancer Res Treat, 2003. 79(1): p. 75-82.
113. Ghigna C, Valacca C, Biamonti G, *Alternative splicing and tumor progression.* Curr Genomics, 2008. 9(8): p. 556-70.
114. Best A, James K, Dalgliesh C, et al., *Human Tra2 proteins jointly control a CHEK1 splicing switch among alternative and constitutive target exons.* Nat Commun, 2014. 5: p. 4760.
115. Watermann DO, Tang Y, Zur Hausen A, et al., *Splicing factor Tra2-beta1 is specifically induced in breast cancer and regulates alternative splicing of the CD44 gene.* Cancer Res, 2006. 66(9): p. 4774-80.
116. Han N, Li W, Zhang M, *The function of the RNA-binding protein hnRNP in cancer metastasis.* J Cancer Res Ther, 2013. 9 Suppl: p. S129-34.
117. Patry C, Bouchard L, Labrecque P, et al., *Small interfering RNA-mediated reduction in heterogeneous nuclear ribonucleoparticule A1/A2 proteins induces apoptosis in human cancer cells but not in normal mortal cell lines.* Cancer Res, 2003. 63(22): p. 7679-88.
118. Li S, Xu H, Din H, et al., *Identification of an aptamer targeting hnRNP A1 by tissue slide-based SELEX.* J Pathol, 2009. 218(3): p. 327-36.
119. Bonomi S, di Matteo A, Buratti E, et al., *HnRNP A1 controls a splicing regulatory circuit promoting mesenchymal-to-epithelial transition.* Nucleic Acids Res, 2013. 41(18): p. 8665-79.
120. Oh Hk, Lee E, Jang HN, et al., *hnRNP A1 contacts exon 5 to promote exon 6 inclusion of apoptotic Fas gene.* Apoptosis, 2013. 18(7): p. 825-35.
121. Jo OD, Martin J, Bernath A, et al., *Heterogeneous nuclear ribonucleoprotein A1 regulates cyclin D1 and c-myc internal ribosome entry site function through Akt signaling.* J Biol Chem, 2008. 283(34): p. 23274-87.
122. He X, Arslan AD, Ho TT, et al., *Involvement of polypyrimidine tract-binding protein (PTBP1) in maintaining breast cancer cell growth and malignant properties.* Oncogenesis, 2014. 3: p. e84.
123. Brinkman BM, *Splice variants as cancer biomarkers.* Clin Biochem, 2004. 37(7): p. 584-94.
124. Chen M, Manley JL, *Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches.* Nat Rev Mol Cell Biol, 2009. 10(11): p. 741-54.
125. David CJ, Chen M, Assanah M, et al., *HnRNP proteins controlled by c-Myc deregulate pyruvate kinase mRNA splicing in cancer.* Nature, 2010. 463(7279): p. 364-8.
126. Izquierdo JM, Majós N, Bonnal S, et al., *Regulation of Fas alternative splicing by antagonistic effects of TIA-1 and PTB on exon definition.* Mol Cell, 2005. 19(4): p. 475-84.
127. Cote J, Dupuis S, Wu JY, *Polypyrimidine track-binding protein binding downstream of caspase-2 alternative exon 9 represses its inclusion.* J Biol Chem, 2001. 276(11): p. 8535-43.
128. Baraniak AP, Chen JR, Garcia-Blanco MA, *Fox-2 mediates epithelial cell-specific fibroblast growth factor receptor 2 exon choice.* Mol Cell Biol, 2006. 26(4): p. 1209-22.
129. Mandal M, Vadlamudi R, Nguyen D, et al., *Growth factors regulate heterogeneous nuclear ribonucleoprotein K expression and function.* J Biol Chem, 2001. 276(13): p. 9699-704.

130. Zhou R, Shanas R, Nelson MA, et al., *Increased expression of the heterogeneous nuclear ribonucleoprotein K in pancreatic cancer and its association with the mutant p53.* Int J Cancer, 2010. 126(2): p. 395-404.
131. Venables JP, Koh CS, Froehlich U, et al., *Multiple and specific mRNA processing targets for the major human hnRNP proteins.* Mol Cell Biol, 2008. 28(19): p. 6033-43.
132. Tang F, Li W, Chen Y, et al., *Downregulation of hnRNP K by RNAi inhibits growth of human lung carcinoma cells.* Oncol Lett, 2014. 7(4): p. 1073-1077.
133. Paronetto MP, Cappellari M, Busà R, et al., *Alternative splicing of the cyclin D1 proto-oncogene is regulated by the RNA-binding protein Sam68.* Cancer Res, 2010. 70(1): p. 229-39.
134. Song L, Wang L, Li Y, et al., *Sam68 up-regulation correlates with, and its down-regulation inhibits, proliferation and tumourigenicity of breast cancer cells.* J Pathol, 2010. 222(3): p. 227-37.
135. Valacca C, Bonomi S, Buratti E, et al., *Sam68 regulates EMT through alternative splicing-activated nonsense-mediated mRNA decay of the SF2/ASF proto-oncogene.* J Cell Biol, 2010. 191(1): p. 87-99.
136. Janz M, Harbeck N, Dettmar P, et al., *Y-box factor YB-1 predicts drug resistance and patient outcome in breast cancer independent of clinically relevant tumor biologic factors HER2, uPA and PAI-1.* Int J Cancer, 2002. 97(3): p. 278-82.
137. Rubinstein DB, Stortchevoi A, Boosalis M, et al., *Overexpression of DNA-binding protein B gene product in breast cancer as detected by in vitro-generated combinatorial human immunoglobulin libraries.* Cancer Res, 2002. 62(17): p. 4985-91.
138. Swamynathan SK, Varma BR, Weber KT, et al., *Targeted disruption of one allele of the Y-box protein gene, Chk-YB-1b, in DT40 cells results in major defects in cell cycle.* Biochem Biophys Res Commun, 2002. 296(2): p. 451-7.
139. Davies AH, Barrett I, Pambid MR, et al., *YB-1 evokes susceptibility to cancer through cytokinesis failure, mitotic dysfunction and HER2 amplification.* Oncogene, 2011. 30(34): p. 3649-60.
140. Shibata T, Kan H, Murakami Y, et al., *Y-box binding protein-1 contributes to both HER2/ErbB2 expression and lapatinib sensitivity in human gastric cancer cells.* Mol Cancer Ther, 2013. 12(5): p. 737-46.
141. Mylona E, Melissaris S, Giannopoulou I, et al., *Y-box-binding protein 1 (YB1) in breast carcinomas: relation to aggressive tumor phenotype and identification of patients at high risk for relapse.* Eur J Surg Oncol, 2014. 40(3): p. 289-96.
142. Lapuk A, Marr H, Jakkula L, et al., *Exon-level microarray analyses identify alternative splicing programs in breast cancer.* Mol Cancer Res, 2010. 8(7): p. 961-74.
143. Norris JD, Fan D, Sherk A, et al., *A negative coregulator for the human ER.* Mol Endocrinol, 2002. 16(3): p. 459-68.
144. Zhang C, Zhang Z, Castle J, et al., *Defining the regulatory network of the tissue-specific splicing factors Fox-1 and Fox-2.* Genes Dev, 2008. 22(18): p. 2550-63.
145. Venables JP, Klinck R, Koh C, et al., *Cancer-associated regulation of alternative splicing.* Nat Struct Mol Biol, 2009. 16(6): p. 670-6.
146. Shkreta L, Bell B, Revil T, et al., *Cancer-Associated Perturbations in Alternative Pre-messenger RNA Splicing.* Cancer Treat Res, 2013. 158: p. 41-94.
147. Yeo GW, Coufal NG, Liang TY, et al., *An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells.* Nat Struct Mol Biol, 2009. 16(2): p. 130-7.
148. Rintala-Maki ND, Goard CA, Langdon CE, et al., *Expression of RBM5-related factors in primary breast tissue.* J Cell Biochem, 2007. 100(6): p. 1440-58.
149. Inoue A, Yamamoto N, Kimura M, et al. *RBM10 regulates alternative splicing*. FEBS Lett. 2014. 588(6): p. 942-7.

150. Wang K, Bacon ML, Tessier JJ, et al. *RBM10 modulates apoptosis and influences TNF-$\alpha$ gene expression*. J Cell Death. 2012. 5: p. 1-19
151. Sampath J, Long PR, Shepard RL, et al., *Human SPF45, a splicing factor, has limited expression in normal tissues, is overexpressed in many tumors, and can confer a multidrug-resistant phenotype to cells.* Am J Pathol, 2003. 163(5): p. 1781-90.
152. Perry WL 3rd, Shepard RL, Sampath J, et al., *Human splicing factor SPF45 (RBM17) confers broad multidrug resistance to anticancer drugs when overexpressed--a phenotype partially reversed by selective estrogen receptor modulators.* Cancer Res, 2005. 65(15): p. 6593-600.
153. Jansson MD, Lund AH, *MicroRNA and cancer.* Mol Oncol, 2012. 6(6): p. 590-610.
154. van Kouwenhove M, Kedde M, Agami R, *MicroRNA regulation by RNA-binding proteins and its implications for cancer.* Nat Rev Cancer, 2011. 11(9): p. 644-56.
155. Nie L, Wu HJ, Hsu JM, et al., *Long non-coding RNAs: versatile master regulators of gene expression and crucial players in cancer.* Am J Transl Res, 2012. 4(2): p. 127-50.
156. Oltean S, Bates DO, *Hallmarks of alternative splicing in cancer.* Oncogene, 2014. 33(46): p. 5311-8.
157. Chopin LK, Seim I, Walpole CM, et al., *The ghrelin axis--does it have an appetite for cancer progression?* Endocr Rev, 2012. 33(6): p. 849-91.
158. Chanclón B, Martínez-Fuentes AJ, Gracia-Navarro F, *Role of SST, CORT and ghrelin and its receptors at the endocrine pancreas.* Front Endocrinol (Lausanne), 2012. 3: p. 114.
159. Barbieri F, Bajetto A, Pattarozzi A, et al., *Peptide receptor targeting in cancer: the somatostatin paradigm.* Int J Pept, 2013. 2013: p. 926295.
160. Volante M, Rosas R, Allìa E, et al., *Somatostatin, cortistatin and their receptors in tumours.* Mol Cell Endocrinol, 2008. 286(1-2): p. 219-29.
161. Duxbury MS, Waseem T, Ito H, et al., *Ghrelin promotes pancreatic adenocarcinoma cellular proliferation and invasiveness.* Biochem Biophys Res Commun, 2003. 309(2): p. 464-8.
162. Corbetta S, Peracchi M, Cappiello V, et al., *Circulating ghrelin levels in patients with pancreatic and gastrointestinal neuroendocrine tumors: identification of one pancreatic ghrelinoma.* J Clin Endocrinol Metab, 2003. 88(7): p. 3117-20.
163. Tostivint H, Joly L, Lihrmann I, et al., *Comparative genomics provides evidence for close evolutionary relationships between the urotensin II and somatostatin gene families.* Proc Natl Acad Sci U S A, 2006. 103(7): p. 2237-42.
164. Brazeau P, Vale W, Burgus R, et al., *Hypothalamic polypeptide that inhibits the secretion of immunoreactive pituitary growth hormone.* Science, 1973. 179(4068): p. 77-9.
165. de Lecea L, Criado JR, Próspero-García O, et al., *A cortical neuropeptide with neuronal depressant and sleep-modulating properties.* Nature, 1996. 381: p. 242-245.
166. Fukusumi S, Kitada C, Takekawa S, et al., *Identification and characterization of a novel human cortistatin-like peptide.* Biochem Biophys Res Commun, 1997. 232(1): p. 157-63.
167. Tostivint H, Lihrmann I, Bucharles C, et al., *Occurrence of two somatostatin variants in the frog brain: characterization of the cDNAs, distribution of the mRNAs, and receptor-binding affinities of the peptides.* Proc Natl Acad Sci U S A, 1996. 93(22): p. 12605-10.
168. Gahete MD, Durán-Prado M, Luque RM, et al., *Are somatostatin and cortistatin two siblings in regulating endocrine secretions? In vitro work ahead.* Mol Cell Endocrinol, 2008. 286(1-2): p. 128-34.
169. Gahete MD, Córdoba-Chacón J, Durán-Prado M, et al., *Somatostatin and its receptors from fish to mammals.* Ann N Y Acad Sci, 2010. 1200: p. 43-52.
170. Patel YC, *Somatostatin and its receptor family.* Front Neuroendocrinol, 1999. 20(3): p. 157-98.
171. Ruscica M, Arvigo M, Steffani L, et al., *Somatostatin, somatostatin analogs and somatostatin receptor dynamics in the biology of cancer progression.* Curr Mol Med, 2013. 13(4): p. 555-71.

172. Schweitzer P, Madamba SG, Siggins GR, *The sleep-modulating peptide cortistatin augments the h-current in hippocampal neurons.* J Neurosci, 2003. 23(34): p. 10884-91.
173. González-Rey E, Chorny A, Robledo G, et al., *Cortistatin, a new antiinflammatory peptide with therapeutic effect on lethal endotoxemia.* J Exp Med, 2006. 203(3): p. 563-71.
174. de Lecea L, Castaño JP, *Cortistatin: not just another somatostatin analog.* Nat Clin Pract Endocrinol Metab, 2006. 2(7): p. 356-7.
175. Dalm VA, van Hagen PM, van Koetsveld PM, et al., *Cortistatin rather than somatostatin as a potential endogenous ligand for somatostatin receptors in the human immune system.* J Clin Endocrinol Metab, 2003. 88(1): p. 270-6.
176. Winsky-Sommerer R, Spier AD, Fabre V, et al., *Overexpression of the human beta-amyloid precursor protein downregulates cortistatin mRNA in PDAPP mice.* Brain Res, 2004. 1023(1): p. 157-62.
177. Córdoba-Chacón J, Gahete MD, Pozo-Salas AI, et al., *Cortistatin is not a somatostatin analogue but stimulates prolactin release and inhibits GH and ACTH in a gender-dependent fashion: potential role of ghrelin.* Endocrinology, 2011. 152(12): p. 4800-12.
178. González-Rey E, Chorny A, del Moral RG, et al., *Therapeutic effect of cortistatin on experimental arthritis by downregulating inflammatory and Th1 responses.* Ann Rheum Dis, 2007. 66(5): p. 582-8.
179. González-Rey E, Varela N, Sheibanie AF, et al., *Cortistatin, an antiinflammatory peptide with therapeutic action in inflammatory bowel disease.* Proc Natl Acad Sci U S A, 2006. 103(11): p. 4228-33.
180. Dalm VA, van Hagen PM, van Koetsveld PM, et al., *Expression of somatostatin, cortistatin, and somatostatin receptors in human monocytes, macrophages, and dendritic cells.* Am J Physiol Endocrinol Metab, 2003. 285(2): p. E344-53.
181. Broglio F, Koetsveld Pv Pv, Benso A, et al., *Ghrelin secretion is inhibited by either somatostatin or cortistatin in humans.* J Clin Endocrinol Metab, 2002. 87(10): p. 4829-32.
182. Ferone D, Boschetti M, Resmini E, et al., *Neuroendocrine-immune interactions: the role of cortistatin/somatostatin system.* Ann N Y Acad Sci, 2006. 1069: p. 129-44.
183. Criado JR, Li H, Jiang X, et al., *Structural and compositional determinants of cortistatin activity.* J Neurosci Res, 1999. 56(6): p. 611-9.
184. de Lecea L, *Cortistatin--functions in the central nervous system.* Mol Cell Endocrinol, 2008. 286(1-2): p. 88-95.
185. Moller LN, Stidsen CE, Hartmann B, et al., *Somatostatin receptors.* Biochim Biophys Acta, 2003. 1616(1): p. 1-84.
186. Moaeen-ud-Din M, Yang LG, *Evolutionary history of the somatostatin and somatostatin receptors.* J Genet, 2009. 88(1): p. 41-53.
187. Durán-Prado M, Malagón MM, Gracia-Navarro F, et al., *Dimerization of G protein-coupled receptors: new avenues for somatostatin receptor signalling, control and functioning.* Mol Cell Endocrinol, 2008. 286(1-2): p. 63-8.
188. Reubi JC, Waser B, Schaer JC, et al., *Somatostatin receptor sst1-sst5 expression in normal and neoplastic human tissues using receptor autoradiography with subtype-selective ligands.* Eur J Nucl Med, 2001. 28(7): p. 836-46.
189. Castaño JP, Sundin A, Maecke HR, et al., *Gastrointestinal neuroendocrine tumors (NETs): new diagnostic and therapeutic challenges.* Cancer Metastasis Rev, 2014. 33(1): p. 353-9.
190. Canzian F, McKay JD, Cleveland RJ, et al., *Genetic variation in the growth hormone synthesis pathway in relation to circulating insulin-like growth factor-I, insulin-like growth factor binding protein-3, and breast cancer risk: results from the European prospective investigation into cancer and nutrition study.* Cancer Epidemiol Biomarkers Prev, 2005. 14(10): p. 2316-25.
191. Hofland LJ, Lamberts SW, *Somatostatin receptor subtype expression in human tumors.* Ann Oncol, 2001. 12 Suppl 2: p. S31-6.

192. Lamberts SW, de Herder WW, Hofland LJ, *Somatostatin analogs in the diagnosis and treatment of cancer.* Trends Endocrinol Metab, 2002. 13(10): p. 451-7.
193. Córdoba-Chacón J, Gahete MD, Durán-Prado M, et al., *Identification and characterization of new functional truncated variants of somatostatin receptor subtype 5 in rodents.* Cell Mol Life Sci, 2010. 67(7): p. 1147-63.
194. Durán-Prado M, Gahete MD, Delgado-Niebla E, et al., *Truncated variants of pig somatostatin receptor subtype 5 (sst5) act as dominant-negative modulators for sst2-mediated signaling.* Am J Physiol Endocrinol Metab, 2012. 303(11): p. E1325-34.
195. Córdoba-Chacón J, Gahete MD, Durán-Prado M, et al., *Truncated somatostatin receptors as new players in somatostatin-cortistatin pathophysiology.* Ann N Y Acad Sci, 2011. 1220: p. 6-15.
196. Kojima M, Hosoda H, Date Y, et al., *Ghrelin is a growth-hormone-releasing acylated peptide from stomach.* Nature, 1999. 402(6762): p. 656-60.
197. Howard AD, Feighner SD, Cully DF, et al., *A receptor in pituitary and hypothalamus that functions in growth hormone release.* Science, 1996. 273(5277): p. 974-7.
198. Nishi Y, Yoh J, Hiejima H, et al., *Structures and molecular forms of the ghrelin-family peptides.* Peptides, 2011. 32(11): p. 2175-82.
199. Giordano R, Picu A, Broglio F, et al., *Ghrelin, hypothalamus-pituitary-adrenal (HPA) axis and Cushing's syndrome.* Pituitary, 2004. 7(4): p. 243-8.
200. Muller TD, Nogueiras R, Andermann ML, et al., *Ghrelin.* Mol Metab, 2015. 4(6): p. 437-60.
201. Gahete MD, Rincón-Fernández D, Villa-Osaba A, et al., *Ghrelin gene products, receptors, and GOAT enzyme: biological and pathophysiological insight.* J Endocrinol, 2014. 220(1): p. R1-24.
202. Heppner KM, Tong J, Kirchner H, et al., *The ghrelin O-acyltransferase-ghrelin system: a novel regulator of glucose metabolism.* Curr Opin Endocrinol Diabetes Obes, 2011. 18(1): p. 50-5.
203. Ukkola O, *Ghrelin and the metabolic balance.* J Endocrinol Invest, 2005. 28(9): p. 849-52.
204. Camina JP, *Cell biology of the ghrelin receptor.* J Neuroendocrinol, 2006. 18(1): p. 65-76.
205. Leung PK, Chow KB, Lau PN, et al., *The truncated ghrelin receptor polypeptide (GHS-R1b) acts as a dominant-negative mutant of the ghrelin receptor.* Cell Signal, 2007. 19(5): p. 1011-22.
206. Sato T, Nakamura Y, Shiimura Y, et al., *Structure, regulation and function of ghrelin.* J Biochem, 2012. 151(2): p. 119-28.
207. Seim I, Herington AC, Chopin LK, *New insights into the molecular complexity of the ghrelin gene locus.* Cytokine Growth Factor Rev, 2009. 20(4): p. 297-304.
208. Zhang JV, Ren PG, Avsian-Kretchmer O, et al., *Obestatin, a peptide encoded by the ghrelin gene, opposes ghrelin's effects on food intake.* Science, 2005. 310(5750): p. 996-9.
209. Yeh AH, Jeffery PL, Duncan RP, et al., *Ghrelin and a novel preproghrelin isoform are highly expressed in prostate cancer and ghrelin activates mitogen-activated protein kinase in prostate cancer.* Clin Cancer Res, 2005. 11(23): p. 8295-303.
210. Kineman RD, Gahete MD, Luque RM, *Identification of a mouse ghrelin gene transcript that contains intron 2 and is regulated in the pituitary and hypothalamus in response to metabolic stress.* J Mol Endocrinol, 2007. 38(5): p. 511-21.
211. De Vriese C, Grégoire F, De Neef P, et al., *Ghrelin is produced by the human erythroleukemic HEL cell line and involved in an autocrine pathway leading to cell proliferation.* Endocrinology, 2005. 146(3): p. 1514-22.
212. Jeffery PL, Herington AC, Chopin LK, *Expression and action of the growth hormone releasing peptide ghrelin and its receptor in prostate cancer cell lines.* J Endocrinol, 2002. 172(3): p. R7-11.

213. Maccarinelli G, Sibilia V, Torsello A, et al., *Ghrelin regulates proliferation and differentiation of osteoblastic cells.* J Endocrinol, 2005. 184(1): p. 249-56.
214. Majchrzak K, Szyszko K, Pawłowski KM, et al., *A role of ghrelin in cancerogenesis.* Pol J Vet Sci, 2012. 15(1): p. 189-97.
215. Pettersson I, Muccioli G, Granata R, et al., *Natural (ghrelin) and synthetic (hexarelin) GH secretagogues stimulate H9c2 cardiomyocyte cell proliferation.* J Endocrinol, 2002. 175(1): p. 201-9.
216. Yu H, Xu G, Fan X, *The effect of ghrelin on cell proliferation in small intestinal IEC-6 cells.* Biomed Pharmacother, 2013. 67(3): p. 235-9.
217. Gronberg M, Fjällskog ML, Jirström K, et al., *Expression of ghrelin is correlated to a favorable outcome in invasive breast cancer.* Acta Oncol, 2012. 51(3): p. 386-93.
218. Sever S, White DL, García JM, *Is there an effect of ghrelin/ghrelin analogs on cancer? A systematic review*. Endocr Relat Cancer, 2016. 23(9):R393-409.
219. Xu T, Ye F, Wang B, et al., *Elevation of growth hormone secretagogue receptor type 1a mRNA expression in human growth hormone-secreting pituitary adenoma harboring G protein alpha subunit mutation.* Neuro Endocrinol Lett, 2010. 31(1): p. 147-54.
220. Fung JN, Jeffery PL, Lee JD, et al., *Silencing of ghrelin receptor expression inhibits endometrial cancer cell growth in vitro and in vivo.* Am J Physiol Endocrinol Metab, 2013. 305(2): p. E305-13.
221. Korbonits M, Bustin SA, Kojima M, et al., *The expression of the growth hormone secretagogue receptor ligand ghrelin in normal and abnormal human pituitary and other neuroendocrine tumors.* J Clin Endocrinol Metab, 2001. 86(2): p. 881-7.
222. Cassoni P, Ghé C, Marrocco T, et al., *Expression of ghrelin and biological activity of specific receptors for ghrelin and des-acyl ghrelin in human prostate neoplasms and related cell lines.* Eur J Endocrinol, 2004. 150(2): p. 173-84.
223. Volante M, Allìa E, Gugliotta P, et al., *Expression of ghrelin and of the GH secretagogue receptor by pancreatic islet cells and related endocrine tumors.* J Clin Endocrinol Metab, 2002. 87(3): p. 1300-8.
224. Takahashi K, Furukawa C, Takano A, et al., *The neuromedin U-growth hormone secretagogue receptor 1b/neurotensin receptor 1 oncogenic signaling pathway as a therapeutic target for lung cancer.* Cancer Res, 2006. 66(19): p. 9408-19.
225. Seim I, Lubik AA, Lehman ML, et al., *Cloning of a novel insulin-regulated ghrelin transcript in prostate cancer.* J Mol Endocrinol, 2013. 50(2): p. 179-91.
226. Vandesompele J, De Preter K, Pattyn F, et al., *Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes.*Genome Biol, 2002. 3(7): RESEARCH0034
227. Sambrook J, Russell DW, *The Hanahan Method for Preparation and Transformation of Competent E. coli: High-efficiency Transformation.* CSH Protoc. 2006. 2006(1).
228. Dvorakova H, Valicek L, Reichelova M, *Detection of mycoplasma contamination in cell cultures and bovine sera.* International journal for veterinary and biomedical science, 2005. 50(6): p. 262-268.
229. Shaw FL, Harrison H, Spence K, et al., *A detailed mammosphere assay protocol for the quantification of breast stem cell activity.* J Mammary Gland Biol Neoplasia, 2012. 17(2): p. 111-7.
230. Schneider CA, Rasband WS, Eliceiri KW, *NIH Image to ImageJ: 25 years of image analysis.* Nat Methods, 2012. 9(7): p. 671-5.
231. Casanovas O, Hicklin DJ, Bergers G, et al., *Drug resistance by evasion of antiangiogenic targeting of VEGF signaling in late-stage pancreatic islet tumors.* Cancer Cell, 2005. 8(4): p. 299-309.
232. Kurmangaliyev YZ, Sutormin RA, Naumenko SA, et al., *Functional implications of splicing polymorphisms in the human genome.* Hum Mol Genet, 2013. 22(17): p. 3449-59.

233. Shiovitz S, Korde LA, *Genetics of breast cancer: a topic in evolution.* Ann Oncol, 2015. 26(7): p. 1291-9.
234. Kwan T, Benovoy D, Dias C, et al., *Genome-wide analysis of transcript isoform variation in humans.* Nat Genet, 2008. 40(2): p. 225-31.
235. Gelfman S, Burstein D, Penn O, et al., *Changes in exon-intron structure during vertebrate evolution affect the splicing pattern of exons.* Genome Res, 2012. 22(1): p. 35-50.
236. Wei WJ, Mu SR, Heiner M, et al., *YB-1 binds to CAUC motifs and stimulates exon inclusion by enhancing the recruitment of U2AF to weak polypyrimidine tracts.* Nucleic Acids Res, 2012. 40(17): p. 8622-36.
237. Hammond SM, *An overview of microRNAs.* Adv Drug Deliv Rev, 2015. 87: p. 3-14.
238. Gerhardt H, Golding M, Fruttiger M, et al., *VEGF guides angiogenic sprouting utilizing endothelial tip cell filopodia.* J Cell Biol, 2003. 161(6): p. 1163-77.
239. Dunn IF, Heese O, Black PM, *Growth factors in glioma angiogenesis: FGFs, PDGF, EGF, and TGFs.* J Neurooncol, 2000. 50(1-2): p. 121-37.
240. Fagiani E, Christofori G, *Angiopoietins in angiogenesis.* Cancer Lett, 2013. 328(1): p. 18-26.
241. Roberts A, Mishra L, *Role of TGF-β in stem cells and cancer.* Oncogene, 2005. 24(37): p. 5667-5667.
242. Imamura T, Hikita A, Inoue Y, *The roles of TGF-beta signaling in carcinogenesis and breast cancer metastasis.* Breast Cancer, 2012. 19(2): p. 118-24.
243. González I, Munita R, Agirre E, et al., *A lncRNA regulates alternative splicing via establishment of a splicing-specific chromatin signature.* Nat Struct Mol Biol, 2015. 22(5): p. 370-6.
244. Owens TW, Naylor MJ, *Breast cancer stem cells.* Front Physiol, 2013. 4: p. 225.
245. Pinto CA, Widodo E, Waltham M, et al., *Breast cancer stem cells and epithelial mesenchymal plasticity - Implications for chemoresistance.* Cancer Lett, 2013. 341(1): p. 56-62.
246. Guo X, Wang XF, *Signaling cross-talk between TGF-beta/BMP and other pathways.* Cell Res, 2009. 19(1): p. 71-88.
247. Zavadil J, Cermak L, Soto-Nieves N, et al., *Integration of TGF-beta/Smad and Jagged1/Notch signalling in epithelial-to-mesenchymal transition.* EMBO J, 2004. 23(5): p. 1155-65.
248. Spaderna S, Schmalhofer O, Hlubek F, et al., *Epithelial-mesenchymal and mesenchymal-epithelial transitions during cancer progression.* Verh Dtsch Ges Pathol, 2007. 91: p. 21-8.
249. Coe BP, Witherspoon K, Rosenfeld JA, et al., *Refining analyses of copy number variation identifies specific genes associated with developmental delay.* Nat Genet, 2014. 46(10): p. 1063-71.
250. Boutz PL, Chawla G, Stoilov P, et al., *MicroRNAs regulate the expression of the alternative splicing factor nPTB during muscle development.* Genes Dev, 2007. 21(1): p. 71-84.
251. Wang GS, Cooper TA, *Splicing in disease: disruption of the splicing code and the decoding machinery.* Nat Rev Genet, 2007. 8(10): p. 749-61.
252. Eperon IC, Makarova OV, Mayeda A, et al., *Selection of alternative 5' splice sites: role of U1 snRNP and models for the antagonistic effects of SF2/ASF and hnRNP A1.* Mol Cell Biol, 2000. 20(22): p. 8303-18.
253. Grosso AR, Gomes AQ, Barbosa-Morais NL, et al., *Tissue-specific splicing factor gene expression signatures.* Nucleic Acids Res, 2008. 36(15): p. 4823-32.
254. Misteli T, *Different site, different splice.* Nat Cell Biol, 2000. 2(6): p. E98-E100.
255. da Silva MR, Moreira GA, Gonçalves da Silva RA, et al., *Splicing Regulators and Their Roles in Cancer Biology and Therapy.* Biomed Res Int, 2015. 2015: p. 150514.

256. Stark M, Bram EE, Akerman M, et al., *Heterogeneous nuclear ribonucleoprotein H1/H2-dependent unsplicing of thymidine phosphorylase results in anticancer drug resistance.* J Biol Chem, 2011. 286(5): p. 3741-54.
257. Liu Z, Xu Y, Long J, et al., *microRNA-218 suppresses the proliferation, invasion and promotes apoptosis of pancreatic cancer cells by targeting HMGB1.* Chin J Cancer Res, 2015. 27(3): p. 247-57.
258. Chen T, Xu C1, Chen J1, et al., *MicroRNA-203 inhibits cellular proliferation and invasion by targeting Bmi1 in non-small cell lung cancer.* Oncol Lett, 2015. 9(6): p. 2639-2646.
259. Lin L, Liang H, Wang Y, et al., *microRNA-141 inhibits cell proliferation and invasion and promotes apoptosis by targeting hepatocyte nuclear factor-3beta in hepatocellular carcinoma cells.* BMC Cancer, 2014. 14: p. 879
260. Vasudevan S, Tong Y, Steitz JA, *Switching from repression to activation: microRNAs can up-regulate translation.* Science, 2007. 318(5858): p. 1931-4.
261. Sand M, Skrygan M, Sand D, et al., *Comparative microarray analysis of microRNA expression profiles in primary cutaneous malignant melanoma, cutaneous malignant melanoma metastases, and benign melanocytic nevi.* Cell Tissue Res, 2013. 351(1): p. 85-98.
262. Wagner-Ecker M, Schwager C, Wirkner U, et al., *MicroRNA expression after ionizing radiation in human endothelial cells.* Radiat Oncol, 2010. 5: p. 25.
263. Liang Z, Wu H, Xia J, et al., *Involvement of miR-326 in chemotherapy resistance of breast cancer through modulating expression of multidrug resistance-associated protein 1.* Biochem Pharmacol, 2010. 79(6): p. 817-24.
264. Kefas B, Comeau L, Floyd DH, et al., *The neuronal microRNA miR-326 acts in a feedback loop with notch and has therapeutic potential against brain tumors.* J Neurosci, 2009. 29(48): p. 15161-8.
265. Ye S, Yang L, Zhao X, et al., *Bioinformatics method to predict two regulation mechanism: TF-miRNA-mRNA and lncRNA-miRNA-mRNA in pancreatic cancer.* Cell Biochem Biophys, 2014. 70(3): p. 1849-58.
266. Semaan N, Frenzel L, Alsaleh G, et al., *miR-346 controls release of TNF-alpha protein and stability of its mRNA in rheumatoid arthritis via tristetraprolin stabilization.* PLoS One, 2011. 6(5): p. e19827.
267. Xiao J, Liu D, Jiao W, et al., *Effects of microRNA-346 on epithelial-mesenchymal transition in mouse podocytes.* Gene, 2015. 560(2): p. 195-9.
268. Wang Q, Cai J, Cai XH, et al., *miR-346 regulates osteogenic differentiation of human bone marrow-derived mesenchymal stem cells by targeting the Wnt/beta-catenin pathway.* PLoS One, 2013. 8(9): p. e72266.
269. Park YT, Jeong JY, Lee MJ, et al., *MicroRNAs overexpressed in ovarian ALDH1-positive cells are associated with chemoresistance.* J Ovarian Res, 2013. 6(1): p. 18.
270. Tran N, McLean T, Zhang X, et al., *MicroRNA expression profiles in head and neck cancer cell lines.* Biochem Biophys Res Commun, 2007. 358(1): p. 12-7.
271. Gaur A, Jewell DA, Liang Y, et al., *Characterization of microRNA expression levels and their biological correlates in human cancer cell lines.* Cancer Res, 2007. 67(6): p. 2456-68.
272. Chen B, Pan W, Lin X, et al., *MicroRNA-346 functions as an oncogene in cutaneous squamous cell carcinoma.* Tumour Biol. 2016. 37(2): p. 2765-71.
273. Weber F, Teresi RE, Broelsch CE, et al., *A limited set of human MicroRNA is deregulated in follicular thyroid carcinoma.* J Clin Endocrinol Metab. 2006. 91(9): p. 3584-91.
274. Patel JB, Appaiah HN, Burnett RM, et al., *Control of EVI-1 oncogene expression in metastatic breast cancer cells through microRNA miR-22.* Oncogene, 2011. 30(11): p. 1290-301.
275. Xin F, Li M, Balch C, et al., *Computational analysis of microRNA profiles and their target genes suggests significant involvement in breast cancer antiestrogen resistance.* Bioinformatics, 2009. 25(4): p. 430-4.

276. Vidal SJ, Rodríguez-Bravo V, Galsky M, et al., *Targeting cancer stem cells to suppress acquired chemotherapy resistance.* Oncogene, 2014. 33(36): p. 4451-63.
277. Durán-Prado M, Saveanu A, Luque RM, et al., *A potential inhibitory role for the new truncated variant of somatostatin receptor 5, sst5TMD4, in pituitary adenomas poorly responsive to somatostatin analogs.* J Clin Endocrinol Metab, 2010. 95(5): p. 2497-502.
278. Vander Heiden MG, Cantley LC, Thompson CB, *Understanding the Warburg effect: the metabolic requirements of cell proliferation.* Science, 2009. 324(5930): p. 1029-33.
279. Folkman J, Kalluri R, *Cancer without disease.* Nature, 2004. 427(6977): p. 787.
280. Weis SM, Cheresh DA, *Tumor angiogenesis: molecular pathways and therapeutic targets.* Nat Med, 2011. 17(11): p. 1359-70.
281. Carmeliet P, Jain RK, *Angiogenesis in cancer and other diseases.* Nature, 2000. 407(6801): p. 249-57.
282. Kubota Y, *Tumor angiogenesis and anti-angiogenic therapy.* Keio J Med, 2012. 61(2): p. 47-56.
283. van Cruijsen H, Giaccone G, Hoekman K, *Epidermal growth factor receptor and angiogenesis: Opportunities for combined anticancer strategies.* Int J Cancer, 2005. 117(6): p. 883-8.
284. Khosravi Shahi P, Fernández Pineda I, *Tumoral angiogenesis: review of the literature.* Cancer Invest, 2008. 26(1): p. 104-8.
285. Cascone T, Heymach JV, *Targeting the angiopoietin/Tie2 pathway: cutting tumor vessels with a double-edged sword?* J Clin Oncol, 2012. 30(4): p. 441-4.
286. Danesi R, del Tacca M, *The effects of the somatostatin analog octreotide on angiogenesis in vitro.* Metabolism, 1996. 45(8 Suppl 1): p. 49-50.
287. Dal Monte M, Martini D, Ristori C, et al., *Hypoxia effects on proangiogenic factors in human umbilical vein endothelial cells: functional role of the peptide somatostatin.* Naunyn Schmiedebergs Arch Pharmacol, 2011. 383(6): p. 593-612.
288. Dal Monte M, Cammalleri M, Martini D, et al., *Antiangiogenic role of somatostatin receptor 2 in a model of hypoxia-induced neovascularization in the retina: results from transgenic mice.* Invest Ophthalmol Vis Sci, 2007. 48(8): p. 3480-9.
289. Dal Monte M, Petrucci C, Cozzi A, et al., *Somatostatin inhibits potassium-evoked glutamate release by activation of the sst(2) somatostatin receptor in the mouse retina.* Naunyn Schmiedebergs Arch Pharmacol, 2003. 367(2): p. 188-92.
290. Mentlein R, Eichler O, Forstreuter F, et al., *Somatostatin inhibits the production of vascular endothelial growth factor in human glioma cells.* Int J Cancer, 2001. 92(4): p. 545-50.
291. Zhao B, Yang P, Yang J, et al., *A randomized trial of somatostatin to regulate the VEGFs/VEGFRs in patients with gastric cancer.* Hepatogastroenterology, 2011. 58(109): p. 1425-30.
292. Kumar M, Liu ZR, Thapa L, et al., *Mechanisms of inhibition of growth of human pancreatic carcinoma implanted in nude mice by somatostatin receptor subtype 2.* Pancreas, 2004. 29(2): p. 141-51.
293. Kumar M, Liu ZR, Thapa L, et al., *Anti-angiogenic effects of somatostatin receptor subtype 2 on human pancreatic cancer xenografts.* Carcinogenesis, 2004. 25(11): p. 2075-81.
294. Kumar M, Liu ZR, Thapa L, et al., *Antiangiogenic effect of somatostatin receptor subtype 2 on pancreatic cancer cell line: Inhibition of vascular endothelial growth factor and matrix metalloproteinase-2 expression in vitro.* World J Gastroenterol, 2004. 10(3): p. 393-9.
295. Florio T, *Molecular mechanisms of the antiproliferative activity of somatostatin receptors (SSTRs) in neuroendocrine tumors.* Front Biosci, 2008. 13: p. 822-40.
296. van Eijck CH, Krenning EP, Bootsma A, et al., *Somatostatin-receptor scintigraphy in primary breast cancer.* Lancet, 1994. 343(8898): p. 640-3.

297. Orlando C, Raggi CC, Bianchi S, et al., *Measurement of somatostatin receptor subtype 2 mRNA in breast cancer and corresponding normal tissue.* Endocr Relat Cancer, 2004. 11(2): p. 323-32.
298. Watt HL, Kharmate G, Kumar U, *Biology of somatostatin in breast cancer.* Mol Cell Endocrinol, 2008. 286(1-2): p. 251-61.
299. Banerjee S, Dowsett M, Ashworth A, et al., *Mechanisms of disease: angiogenesis and the management of breast cancer.* Nat Clin Pract Oncol, 2007. 4(9): p. 536-50.
300. Arany Z, Foo SY, Ma Y, et al., *HIF-independent regulation of VEGF and angiogenesis by the transcriptional coactivator PGC-1alpha.* Nature, 2008. 451(7181): p. 1008-12.
301. Nguyen DX, Bos PD, Massague J, *Metastasis: from dissemination to organ-specific colonization.* Nat Rev Cancer, 2009. 9(4): p. 274-84.
302. Dudley AC, *Tumor endothelial cells.* Cold Spring Harb Perspect Med, 2012. 2(3): p. a006536.
303. Shayan R, Achen MG, Stacker SA, *Lymphatic vessels in cancer metastasis: bridging the gaps.* Carcinogenesis, 2006. 27(9): p. 1729-38.
304. Shiozawa Y, Nie B, Pienta KJ, et al., *Cancer stem cells and their role in metastasis.* Pharmacol Ther, 2013. 138(2): p. 285-93.
305. Pantel K, Alix-Panabieres C, *The clinical significance of circulating tumor cells.* Nat Clin Pract Oncol, 2007. 4(2): p. 62-3.
306. Hou JM, Krebs MG, Lancashire L, et al., *Clinical significance and molecular characteristics of circulating tumor cells and circulating tumor microemboli in patients with small-cell lung cancer.* J Clin Oncol, 2012. 30(5): p. 525-32.
307. Watabe T, Miyazono K, *Roles of TGF-beta family signaling in stem cell renewal and differentiation.* Cell Res, 2009. 19(1): p. 103-15.
308. Hosoda H, Kojima M, Mizushima T, et al., *Structural divergence of human ghrelin. Identification of multiple ghrelin-derived molecules produced by post-translational processing.* J Biol Chem, 2003. 278(1): p. 64-70.
309. Camiña JP, Carreira MC, Micic D, et al., *Regulation of ghrelin secretion and action.* Endocrine, 2003. 22(1): p. 5-12.
310. Walsh CM, Suchanek AL, Cyphert TJ, et al., *Serine arginine splicing factor 3 is involved in enhanced splicing of glucose-6-phosphate dehydrogenase RNA in response to nutrients and hormones in liver.* J Biol Chem, 2013. 288(4): p. 2816-28.
311. Bergeron D, Pal G, Beaulieu YB, et al., *Regulated Intron Retention and Nuclear Pre-mRNA Decay Contribute to PABPN1 Autoregulation.* Mol Cell Biol, 2015. 35(14): p. 2503-17.
312. Yeo GW, Van Nostrand EL, Liang TY, *Discovery and analysis of evolutionarily conserved intronic splicing regulatory elements.* PLoS Genet, 2007. 3(5): p. e85.
313. Chaudhury A, Hussey GS, Ray PS, et al., *TGF-beta-mediated phosphorylation of hnRNP E1 induces EMT via transcript-selective translational induction of Dab2 and ILEI.* Nat Cell Biol, 2010. 12(3): p. 286-93.
314. Zhang D, Tai LK, Wong LL, et al., *Proteomic characterization of differentially expressed proteins in breast cancer: Expression of hnRNP H1, RKIP and GRP78 is strongly associated with HER-2/neu status.* Proteomics Clin Appl, 2008. 2(1): p. 99-107.
315. Gautrey H, Jackson C, Dittrich AL, et al., *SRSF3 and hnRNP H1 regulate a splicing hotspot of HER2 in breast cancer cells.* RNA Biol, 2015. 12(10): p. 1139-51.
316. Mercer TR, Dinger ME, Mattick JS, *Long non-coding RNAs: insights into functions.* Nat Rev Genet, 2009. 10(3): p. 155-9.
317. Wight M, Werner A, *The functions of natural antisense transcripts.* Essays Biochem, 2013. 54: p. 91-101.
318. Morrissy AS, Griffith M, Marra MA, *Extensive relationship between antisense transcription and alternative splicing in the human genome.* Genome Res, 2011. 21(8): p. 1203-12.

319. Seim I, Collet C, Herington AC, et al., *Revised genomic structure of the human ghrelin gene and identification of novel exons, alternative splice variants and natural antisense transcripts.* BMC Genomics, 2007. 8: p. 298.
320. Seim I, Carter SL, Herington AC, et al., *Complex organisation and structure of the ghrelin antisense strand gene GHRLOS, a candidate non-coding RNA gene.* BMC Mol Biol, 2008. 9: p. 95.
321. Hofmann M, Rudy W, Günthert U, et al., *A link between ras and metastatic behavior of tumor cells: ras induces CD44 promoter activity and leads to low-level expression of metastasis-specific variants of CD44 in CREF cells.* Cancer Res, 1993. 53(7): p. 1516-21.
322. Brinkley BR, Beall PT, Wible LJ, et al., *Variations in cell form and cytoskeleton in human breast carcinoma cells in vitro.* Cancer Res, 1980. 40(9): p. 3118-29.
323. Soule HD, Vazguez J, Long A, et al., *A human cell line from a pleural effusion derived from a breast carcinoma.* J Natl Cancer Inst, 1973. 51(5): p. 1409-16.
324. Cassoni P, Papotti M, Ghè C, et al., *Identification, characterization, and biological activity of specific receptors for natural (ghrelin) and synthetic growth hormone secretagogues and analogs in human breast carcinomas and cell lines.* J Clin Endocrinol Metab, 2001. 86(4): p. 1738-45.
325. Majchrzak K, Pawłowski KM, Orzechowska EJ, et al., *A role of ghrelin in canine mammary carcinoma cells proliferation, apoptosis and migration.* BMC Vet Res, 2012. 8: p. 170.
326. Dhillon AS, Hagan S, Rath O, et al., *MAP kinase signalling pathways in cancer.* Oncogene, 2007. 26(22): p. 3279-90.
327. Ciuffreda L, Incani UC, Steelman LS, et al., *Signaling intermediates (MAPK and PI3K) as therapeutic targets in NSCLC.* Curr Pharm Des, 2014. 20(24): p. 3944-57.
328. Jogi A, Vaapil M, Johansson M, et al., *Cancer cell differentiation heterogeneity and aggressive behavior in solid tumors.* Ups J Med Sci, 2012. 117(2): p. 217-24.
329. Ye X, Weinberg RA, *Epithelial-Mesenchymal Plasticity: A Central Regulator of Cancer Progression.* Trends Cell Biol, 2015.
330. Pasquier J, Abu-Kaoud N, Al Thani H, et al., *Epithelial to Mesenchymal Transition in a Clinical Perspective.* J Oncol, 2015. 2015: p. 792182.
331. Islam F, Qiao B, Smith RA, et al., *Cancer stem cell: fundamental experimental pathological concepts and updates.* Exp Mol Pathol, 2015. 98(2): p. 184-91.
332. Klingbeil P, Isacke CM, *The 'alternative' EMT switch.* Breast Cancer Res, 2011. 13(4): p. 313.
333. May CD, Sphyris N, Evans KW, et al., *Epithelial-mesenchymal transition and cancer stem cells: a dangerously dynamic duo in breast cancer progression.* Breast Cancer Res, 2011. 13(1): p. 202.
334. O'Brien CA, Pollett A, Gallinger S, et al., *A human colon cancer cell capable of initiating tumour growth in immunodeficient mice.* Nature, 2007. 445(7123): p. 106-10.
335. Singh SK, Hawkins C, Clarke ID, et al., *Identification of human brain tumour initiating cells.* Nature, 2004. 432(7015): p. 396-401.
336. Zhao Y, Feng F, Zhou NY, *Stem cells in gastric cancer.* World J Gastroenterol, 2015. 21(1): p. 112-23.
337. Collins AT, Berry PA, Hyde C, et al., *Prospective identification of tumorigenic prostate cancer stem cells.* Cancer Res, 2005. 65(23): p. 10946-51.
338. Li C, Lee CJ, Simeone DM, *Identification of human pancreatic cancer stem cells.* Methods Mol Biol, 2009. 568: p. 161-73.
339. Lapidot T, Sirard C, Vormoor J, et al., *A cell initiating human acute myeloid leukaemia after transplantation into SCID mice.* Nature, 1994. 367(6464): p. 645-8.
340. Espinoza I, Miele L, *Deadly crosstalk: Notch signaling at the intersection of EMT and cancer stem cells.* Cancer Lett, 2013. 341(1): p. 41-5.
341. Kotiyal S, Bhattacharya S, *Breast cancer stem cells, EMT and therapeutic targets.* Biochem Biophys Res Commun, 2014. 453(1): p. 112-6.

342. Noseda M, McLean G, Niessen K, et al., *Notch activation results in phenotypic and functional changes consistent with endothelial-to-mesenchymal transformation.* Circ Res, 2004. 94(7): p. 910-7.
343. Pannuti A, Foreman K, Rizzo P, et al., *Targeting Notch to target cancer stem cells.* Clin Cancer Res, 2010. 16(12): p. 3141-52.
344. Bandapalli OR, Dihlmann S, Helwa R, et al., *Transcriptional activation of the beta-catenin gene at the invasion front of colorectal liver metastases.* J Pathol, 2009. 218(3): p. 370-9.
345. Yu Y, Ramena G, Elble RC, *The role of cancer stem cells in relapse of solid tumors.* Front Biosci (Elite Ed), 2012. 4: p. 1528-41.
346. Velasco-Velázquez MA, Popov VM, Lisanti MP, et al., *The role of breast cancer stem cells in metastasis and therapeutic implications.* Am J Pathol, 2011. 179(1): p. 2-11.