Universidade de Aveiro Departamento de Química
Ano 2011

Miguel Costa Ribeiro

Análise protéomica de células mamárias em processo de diferenciação

Proteomic analysis of mammary cells throughout differentiation

Universidade de Aveiro Departamento de Química
Ano 2011

Miguel Costa Ribeiro

Análise protéomica de células mamárias em processo de diferenciação

Proteomic analysis of mammary cell secretoma throughout differentiation

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Bioquímica, realizada sob a orientação científica da Doutora Luísa Helguero, investigadora auxiliar do Departamento de Química da Universidade de Aveiro e co-orientação científica do Doutor Rui Vitorino, investigador auxiliar no Departamento de Química da Universidade de Aveiro.

Dedico este trabalho à minha família…

O júri

Presidente

Prof. Doutor Pedro Miguel Dimas Neves Domingues
Professor auxiliar do Departamento de Química da Universidade de Aveiro

Prof. Doutora Margarida Sâncio da Cruz Fardilha
Professora auxiliar do Departamento de Ciências da Saúde da Universidade de Aveiro

Doutora Luisa Alejandro Helguero Shepherd
Investigadora auxiliar do Departamento de Química da Universidade de Aveiro

Doutor Rui Miguel Pinheiro Vitorino
Investigador auxiliar do Departamento de Química da Universidade de Aveiro

Agradecimentos

Aos meus orientadores Doutora Luisa Helguero e Doutor Rui Vitorino toda a disponibilidade e conhecimentos transmitidos sem os quais a realização deste trabalho não seria possível.

Aos companheiros de mestrado e todo o grupo de trabalho pelo óptimo ambiente e por toda a ajuda disponibilizada. Guardarei com carinho todos os bons momentos passados.

Aos amigos que sempre me apoiaram.

Palavras-chave Proteómica, Glândula mamária, Diferenciação, Linha celular HC11, Receptor de Estrogénio β

Resumo

A sinalização paracrina e autocrina desempenha um papel importante na proliferação, diferenciação e apoptose das células epiteliais mamárias. Apesar de este facto ser globalmente aceite os mecanismos pelos quais estes processos ocorrem ainda não são totalmente compreendidos. Em particular, a acção dos estrogénios através do receptor de estrogénio alpha (ERα) e beta (ERβ) afectam profundamente a proliferação e diferenciação; no entanto o papel do ERβ nestes processos é ainda pouco conhecido.

Com o objectivo de melhor compreender os mecanismos de sinalização envolvidos na diferenciação das células epiteliais mamárias fez-se a recolha do meio condicionado (CM) de células HC11 nas 3 etapas de diferenciação: stem-cell like (SCL), pre-differenciadas (PD) e funcionalmente diferenciadas (D). Em seguida, as proteínas do CM foram separadas com recurso a SDS-PAGE e 2D-PAGE e identificadas por MALDI-TOF/TOF. Os nossos resultados mostram que células em SCL segregam factores que possuem um papel pró-angiogenico/invasor e de sobrevivência, bem como outros factores que desempenham a função oposta. Estes resultados estão de acordo com o estado activo de proliferação e migração observado nestas células. Com o inicio da diferenciação (células em PD e D) ocorreu uma mudança nas proteínas mais secretadas, sendo possível observar um aumento do nível de proteínas que regulam a adesão e apoptose.

Para estudar os mecanismos regulados pelo ERβ, o CM de células não diferenciadas, bem como de células a entrar em diferenciação (-EGF) e totalmente diferenciadas (D) foi recolhido após a activação do ERβ por ligandos. As proteínas presentes no CM foram novamente separadas por SDS-PAGE e 2D-PAGE e identificadas por MALDI-TOF/TOF. Os nossos resultados mostraram uma diferença nos níveis de proteínas segregadas em SCL+EGF quando comparado com SCL-EGF, indicando que as vias metabóicas que mantêm as células numa condição indiferenciada como o EGF também comunicam com a sinalização activada por ERβ. Além disso, em concordância com SCL-EGF, as células em D, as proteínas reguladas pelo ERβ estavam relacionadas com adesão e apoptose indicando que o estado celular de diferenciação modula a sinalização do ERβ; sendo que o já identificado papel de supressor tumoral do ERβ é desencadeado apenas quando as células começam a diferenciar.

keywords Proteomic, Mammary gland, Differentiation, HC11 cell line, Estrogen receptor β

abstract

Paracrine and autocrine signalling play an important role in proliferation, differentiation and apoptosis of mammary epithelial cells. Despite this fact being accepted unanimously, the mechanisms by which these processes occur are not fully understood. In particular, estrogens acting through estrogen receptors alpha (ERα) and beta (ERβ) profoundly affect proliferation and differentiation; however, knowledge about ERβ role in these processes is still unclear.

With the goal to best understand the signalling mechanisms involved in mammary epithelial cells differentiation, we collected the conditioned media (CM) of HC11 cells in the three stages of differentiation: stem-cell like (SCL), pre-differentiated (PD) and functionally differentiated (D). Next, the proteins in the CM were separated using SDS-PAGE and 2D-PAGE and identified by MALDI-TOF/TOF. Our results showed that SCL cells secrete factors with pro-angiogenic/invasion and survival roles as well as factors with the opposite functions. These results were in agreement with the controlled active proliferation and migration observed in these cells. With the beginning of differentiation (PD and D cells) a change in the most secreted proteins occurred, being possible to observe a raise in the levels of proteins that regulate adhesion and apoptosis.

To study the mechanisms regulated by ERβ, the CM of SCL cells either kept undifferentiated (+EGF), entering differentiation (-EGF) or fully differentiated (D) was collected following ERβ ligand activation. Proteins present in the CM were once again separated by SDS-PAGE and 2D-PAGE and identified by MALDI-TOF/TOF. Our results showed a difference in the levels of proteins secreted in SCL+EGF compared to SCL-EGF, indicating that pathways that maintain the undifferentiated SCL stage such as EGF also cross talk with ERβ signalling. Moreover, similar to SCL-EGF, in D cells, proteins regulated by ERβ were cell adhesion and apoptosis related proteins indicating that the cellular state of differentiation modulates ERβ signalling; being that the previously identified tumour suppressor role of ERβ is triggered only with the beginning of differentiation.

Index

Figures Index.....3

Tables Index.....4

I - Introduction.....5

1. The Mammary gland…….....6

    1.1. Hormonal regulation of mammary gland development…….....8

2. Cell adhesion and differentiation in the mammary gland.....12

3. Estrogen receptors …….....13

4. Breast cancer.....15

    4.1 Breast cancer with ER expression.....…….....16

5. EGF/EGFR signalling.....17

6. Proteomic approach to breast cancer study.....18

    6.1. Matrix-Assisted Laser Desorption Ionization (MALDI).....19

    6.2. Time of Flight (TOF) Analyzer.....20

    6.3. Tandem Mass Spectrometry (MS/MS).....21

    6.4. HC11 cells.....22

7. Aim…….....24

II - Methods and material…….....25

III – Results and Discussion.....29

1. Proteins indentified in the mammary epithelial cell secretome throughout the differentiation program.....30

2. Protein expression in undifferentiated mammary epithelial cells under estrogenic stimulation with or without EGFR activation.....41

3. Protein expression in differentiated mammary epithelial cells with or without ERβ activation.....48

IV - Conclusion.....50

V - References.....54

VI – Annexes.....61

Figures Index

Figure 1- Schematic representation of the mammary gland. Source: www.cancer.org.....6

Figure 2 – Development stages of the mammary gland. Source: Robinson, G., 2005.....7

Figure 3 – Model of the ductal tree growth. Source: Brisken, C., 2007…….....9

Figure 4 - Model of alveolar differentiation. Source: Brisken, C., 2007……..11

Figure 5 – Comparison between ERα and ERβ domains. Source: Weihua, Z., et al., 2003.....13

Figure 6 – Scheme of transcription activation by ERs. A) Direct binding to promoter region ERE; B) Indirect binding to the promoter region AP-1. Source: Schiff, R. and Osborne, C.K. 2005.....14

Figure 7 – Origin hypothesis for the several breast cancer subtypes. A) Cell of origin model; B) Tumor subtype–specific transforming event model. Source: Polyak, K., 2007.....15

Figure 8 – Signalling network dependent of EGFR activation. Source: Singh, A; 2005.....18

Figure 9 - HC11 cell differentiation program. A) Culture conditions used to achieve each differentiation stage. B) Expression of stem-cell like marker (Melk) and differentiation marker (beta-casein) achieved under the indicated growth conditions. Source: adapted from Williams, 2009 and Aydogdu (upublished).....20

Figure 10 – MALDI ionization process. Source: El-Aneed, A.; Cohen, A.; Banoub, J. 2009.....21

Figure 11 – Scheme of a conventional TOF analyzer (A) and a TOF with reflectron (B). Source: El-Aneed, A.; Cohen, A.; Banoub, J. 2009.....22

Figure 12 – Representative scheme of tandem mass spectrometry (MS/MS). Source: El-Aneed, A.; Cohen, A.; Banoub, J. 2009.....23

Figure 13 - SDS gel comparing protein profile between SCL, PD and D. The numbers indicate proteins differentially identified.....31

Figure 14 - 2D-PAGE comparing protein profile between A) SCL, B) PD and C) D.....32

Figure 15 - Venn diagram of the identified proteins and their distribution.....35

Figure 16 - Pie charts with Biological processes of identified proteins A) SCL B) Pre-differentiated C) Differentiated.....36

Figure 17 - SDS gel comparing protein profile between SCL +EGF (CTR and DPN) and -EGF (CTR and DPN).....42

Figure 18 - 2D gel comparing protein profile between SCL cells A) +EGF CTR; B) +EGF DPN; C) –EGF CTR and D) –EGF DPN.....43

Figure 19 - Pie charts with Biological processe of identified proteins A) Proteins overexpressed in +EGF +DPN B) Proteins underexpressed in +EGF +DPN C) Proteins overexpressed in -EGF DPN and D) Proteins underexpressed in -EGF DPN.....46

Figure 20 – SDS gel comparing protein profile of differentiated cells with or without DPN.....48

Tables Index

Table 1 – Proteins identified from the secretome of mammary epithelial cells with a C.I.%>95%.....33

Table 2 – Proteins bellowing to the 2 major proteins classes.....37

Table 3 – Protein variation on SCL and PD when compared to D stage.....37

Table 4 – Proteins up and downregulated in +EGF (CTR and DPN) and -EGF (CTR and DPN).....44

Table 5 – Proteins upregulated on differenciated cells with different treatments.....49

I – Introduction

1. The mammary gland

All female mammals have a mammary gland with the main purpose to feed their infants in the earlier stages or their lives. The mammary gland is a very specialized secretory gland that when fully developed is composed of the alveolus grouped into lobules, where milk is produced, and ducts, the canals that transport milk from lobules to the nipple. Alveoli and ducts are highly differentiated epithelial structures surrounded by fatty tissue (Fig. 1). The mammary gland develops after birth and undergoes several development stages to reach its final mature form and function. Five different stages in mammary gland development can be identified [1]: embryonic, pre-pubertal, pubertal, sexually mature adult, pregnancy and lactation (Fig. 2).

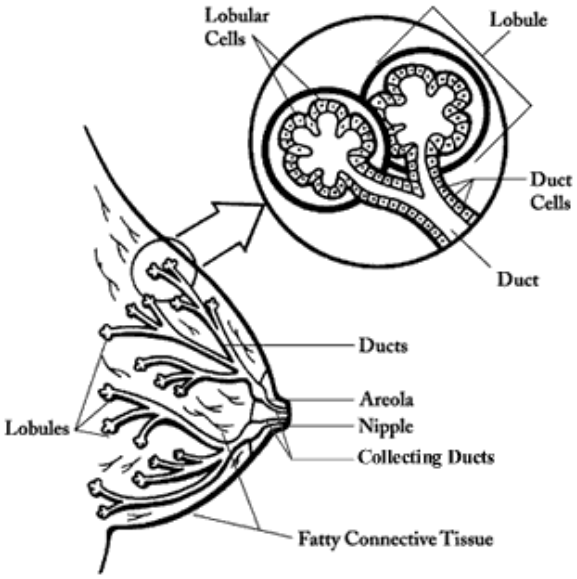


Figure 1- Schematic representation of the mammary gland. Source: www.cancer.org

During the embryonic phase, mammary placodes are formed in the ventral zone of the embryo where the nipples will grow. The placodes give rise to a rudimentary ductal system that remains unaltered until puberty. Therefore, throughout pre-puberty, the mammary gland maintains a growth rate similar to the rest of the body [2-4].

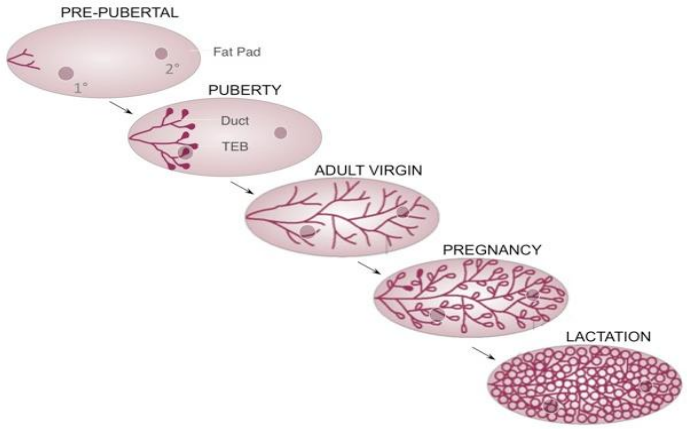


Figure 2– Development stages of the mammary gland. Source: Robinson, G., 2005. TEB: Terminal end bud.

It is with the beginning of puberty, the establishment of a menstrual cycle and dramatic increase in ovarian estrogen (E2) and progesterone (Pg) levels that the majority of changes in the mammary gland occur. Ductal ramification resumes, permitting the formation of a more complex network with the formation of secondary ducts occupying a large amount of fat area, like branches of a tree. Pregnancy, with an increase of Pg and prolactin (Prl), initiates the last stage of maturation with tertiary ductal ramification and alveolar differentiation, thus preparing the gland for lactation. Finally, after weaning, and in response to the lack of suction stimuli, alveoli collapse, cells are removed by apoptosis and the mammary gland returns to a state similar to before pregnancy and ready to undergo a new proliferation-differentiation-death cycle (Fig. 2). It is postulated that the regenerative capacity of the mammary gland is possible thanks to mammary stem cells (SC) and although such cell has not been identified, several progenitor cells capable of originating mammary gland structures such as alveoli and ducts, have been isolated [5-7].

1.1. Hormonal regulation of mammary gland development

Hormonal regulation of mammary gland development can be divided into 2 processes: Growth of the ductal tree and alveolar differentiation.

I. Growth of the ductal tree: The first step in mammary gland development is elongation of the ducts. Ductal elongation is the result of proliferation in the tips of ducts, known as terminal end buds (TEBs). Ductal proliferation is induced by E2 and occurs through a paracrine mechanism, where binding of E2 to estrogen receptor alpha (ERα) activates secretion of amphiregulin (Areg) which then activates proliferation of neighbouring epithelial cells (paracrine mechanism) [8]. For this process to occur, induction of insulin-like growth factor 1 (IGF1) by growth hormone (GH) in the stroma is also necessary [9-10].

A normal ductal ramification requires the presence of estrogen (E2) and ERα while estrogen receptor beta (ERβ) is not fundamental [11]. Studies carried out in ovariectomized mice (no E2) or ERα null mice (ERα-/-), showed that administration of epidermal growth factor (EGF) leads to a normal mammary development [12]. Considering these facts and knowing that administration of E2 leads to EGF receptor (EGFR) activation by Areg, it seems clear that EGFR pathway is important for ductal growth [12-13]. The enzyme in charge of releasing Areg to make it available for the paracrine effects is a member of desintegrin and metalloproteinase family (ADAM) called ADAM-17 [2, 14]. Once EGFR is activated, it guides the release of fibroblast growth factors (FGFs) particularly FGF2 and FGF7 to mediate proliferation and TGF-β to inhibit growth of certain areas of the duct, thus influencing the architecture of the ductal branches [15-16].

One model that tries to explain these mechanisms suggest that the luminal sensor cells (expressing ERα) upon stimulation by E2 secrete membrane bound Areg, which is liberated by ADAM17 promoting proliferation. Stem cells are induced to proliferate creating progenitor cells that can originate myoepthelial and epithelial cells as well as luminal cells. In this process FGF/FGFR and EGF/EGFR have an important role. TEB growth and consequent fat pad invasion is only possible through extracellular matrix (ECM) degradation by metalloproteinases (MMPs) secreted by activated stromal cells (Fig. 3).

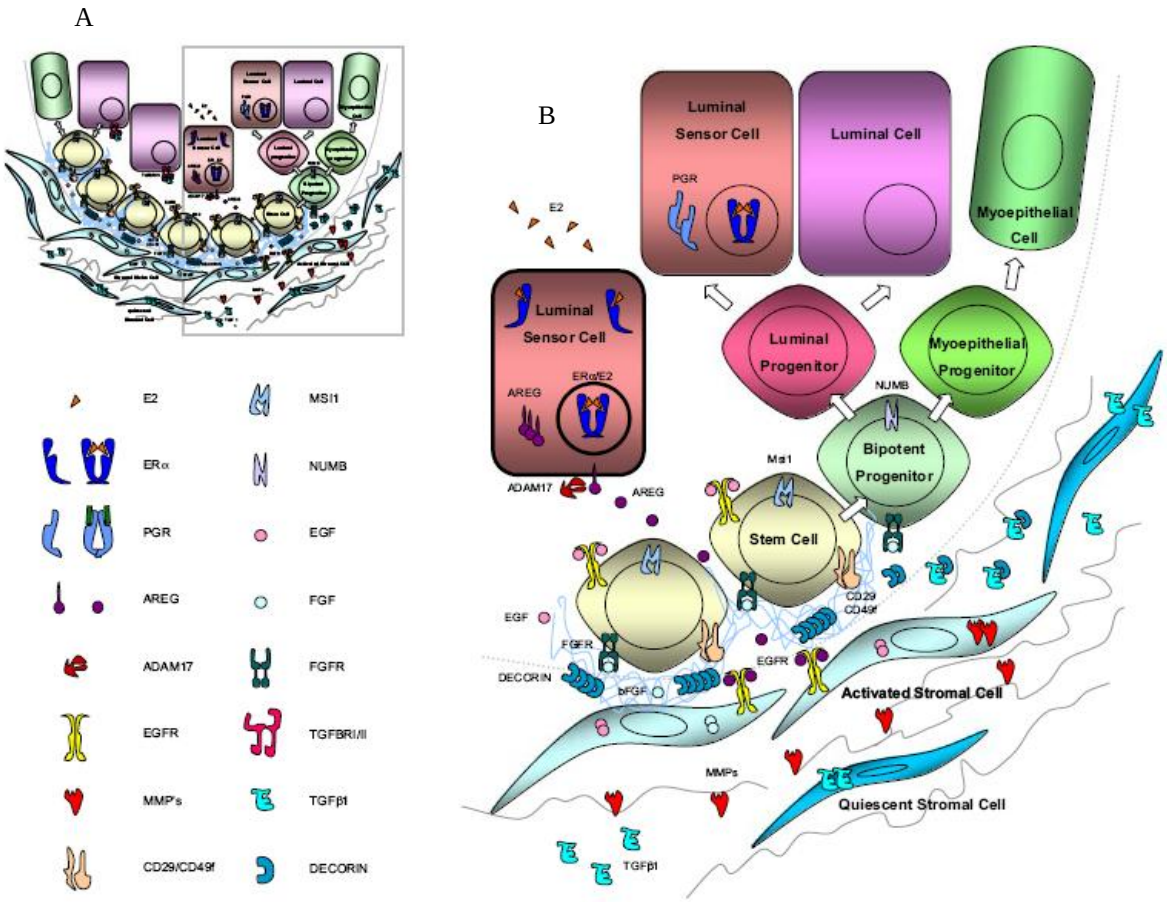


Figure 3– Model of the ductal tree growth. Source: Brisken, C., 2007. A) TEB B) Detail of TEB

EGFR is a tyrosine kinase receptor that participates in several signalling pathways associated with proliferation and cell survival. When activated by one of its ligands it dimerizes with another EGFR or with a related receptor (ErbB2, ErbB3 or ErbB4) [17]. Within different EGFR ligands, Areg is the only one upregulated in puberty (by E2) and it was identified as the only EGFR ligand necessary for mammary maturation [13-14].

Invasion of the mammary fat pad by epithelial cells is necessary for ramification and growth of the ducts. An increase of ECM-degrading proteases activity leads to ramification in kidneys, lungs, salivary glands pancreas and mammary gland. EGFR activation also leads to the release of proteases that degrade ECM allowing cell invasion of the surrounding fat tissue [18]. Within the ECM degradating group of enzymes, the most important are MMPs which are zinc dependent enzymes capable of degradating ECM proteins such as collagens and laminins. However some studies have indicated that a high raise in their activity could lead to ductal hyperplasia and even to tumorogenesis [18]. An important role in bifurcation of the ducts is played by collagenases, which degradate

collagen type I preventing bifurcation. On the contrary, inhibition of collagenases leads to collagen type I deposition facilitating bifurcation [18]. MMPs also affect ramification because they can interfere with ligand-receptor communication causing proliferation or apoptosis [13, 18-19]; ECM degradation may produce fragments with proliferative and morphogenic effect [7, 20]; and degradate cell adhesion molecules facilitating cell detachment and invasion [21-22].

Related to MMPs it is important to mention their inhibitors, tissue inhibitors of metalloproteinases (TIMPs) that in most cases stop ramification because they bind to MMPs and inhibit invasion [18].

II. Alveolar differentiation: When females reach their sexual maturity and establish regular menstrual cycles the ovaries release regular amounts of Pg. In pregnancy this function is sustained by the placenta. With the development of pregnancy the pituitary releases Prl. Both Pg, and later on Prl have an important role in alveolar differentiation [23].

Pg activates progesterone receptor (PR) and induces further ramification and alveolar differentiation. Pg effects on the mammary gland occur through paracrine/autocrine mechanisms, where activation of PR leads to secretion of Wnt 4 and receptor activator of NFκB ligand (RANKL) to induce proliferation and survival of alveolar cells, respectively [5, 24].

Alveolar differentiation is completed reaching the end of pregnancy and sustained throughout lactation. Prl and its receptor PrlR are also implicated in the expression of several proteins including RANKL. This protein and its receptor (RANK) participate in alveolar maturation and can also inhibit apoptosis [25].

Another factor involved in these mechanisms is Igf2 that is thought to participate in alveogenesis. Heregulin 1 (HRG1, an EGF family member) that activates ErbB receptors mediates alveolar differentiation and proliferation but can be replaced by other factors [26-27].

Brisken *et al.* [2], suggest that in alveolar differentiation stem cells originate bipotent progenitors from which the myoepithelial and luminal lineages will be derived. In this process FGF/FGFR and alternatively, EGF/EGFR are involved. A luminal sensor cell is stimulated by Pg leading to WNT and RANKL secretion which, in turn stimulate

proliferation and survival in neighbour cells. Functional differentiation of alveolar cells is achieved in the fully differentiated epithelium by Prl/PrlR (fig. 4).

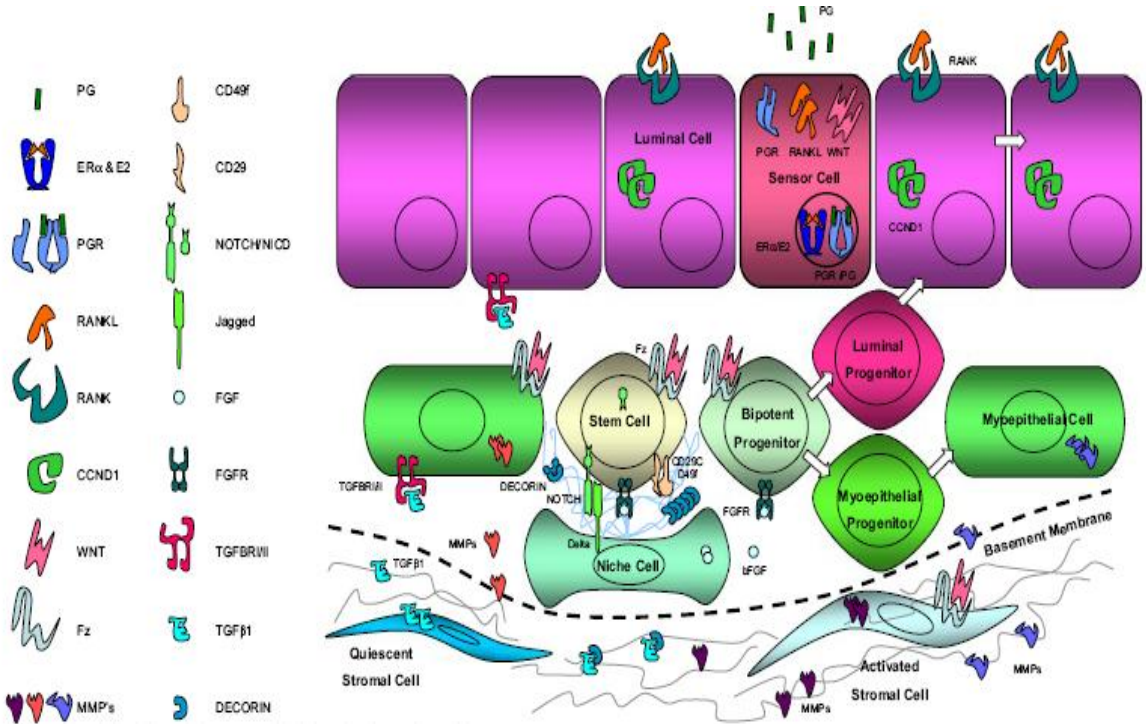


Figure 4- Model of alveolar differentiation. Source: Brisken, C., 2007.

Besides regulating paracrine pathways described previously Prl ligation to its receptor initiates an intracellular response. After Prl binding, PrlR dimerization happens and janus 2 kinase (JAK2) activation occurs, this in turn leads to signal transducer and activator of transcription 5 a (STAT5a) phosphorylation. STAT5a allows the expression of some proteins that constitute milk such as β-casein as well as suppressors of cytokine signalling (SOCS) to promote survival [27-29].

Although there is no role assigned to ERα in the differentiation process of alveoli, ERβ has a role in maintaining the epithelium through control of cell adhesion proteins such as cadherins and integrins. In fact, mammary glands from pregnant and lactating ERβ null mice (ERβ -/-) where able to produce milk. However, they showed alterations in the alveolar epithelial cells, and the basal membrane, indicating that ERβ participates in alveolar differentiation [30].

2. Cell adhesion and differentiation in the mammary gland

The current notion of the ECM considers that rather than act only as a barrier and substrate to cells, it also communicates and interacts with them. This communication can occur due to the connection of certain molecules present in the ECM to adhesion proteins such as integrins and modulation of cadherins. Alterations in cell adhesion mechanisms can lead to carcinogenesis [31]. Cadherins are a superfamily of transmembrane adhesion molecules, which participate in cellular cohesion, apoptosis inhibition and cellular signalling. Within this family it is important to highlight E-cadherins, proteins that influence morphology and polarization of epithelial cells. Thus, it was proved that loss of E-cadherin is correlated to carcinogenesis and metastasis progression in breast cancer. E-cadherins interact with catenins through their citoplasmatic domain and form bonds with the cytoskeleton called adherens junctions [32-33].

Integrins are transmembrane proteins that tie cells to ECM. The connection is once again made to the cytoskeleton of the cell through interaction of integrins cytosolic domains with a varying group of proteins [31]. Several studies have demonstrated that integrins have an active role in ductal ramification namely, α2 integrin in mice and α3 integrin in human cells. The different combinations of ligands (ECM proteins) and integrin receptors will lead to a stronger or weaker binding to cytoskeleton, activation of different signalling cascades and ultimately to the level of adhesion/invasion of cells. Within the integrin ligands the most relevant are laminins collagens and fibronectins. Collagen binding to integrin receptors besides mediating ductal ramification also participates in the production of basal membrane that surrounds the ducts [34-36].

Although integrins and cadherins are important in morphogenesis of mammary ducts and alveoli they are not the only receptors that participate in this process. It is important to refer the participation of other receptors such as β 1,4 glactosyltransferase and dystroglycan [37-38]. Moreover, expression and activity of E-cadherin as well as several integrins is regulated by E2 and EGFR ligands [33, 39-40], and their availability in the cell membrane is dependent of expression levels, as well as turnover rate and endocytosis [33].

3. Estrogen receptors

There are two estrogen receptors (ERs) that mediate estrogen action, ERα and ERβ. These receptors more than isoforms, are two distinct proteins codified by genes in different chromosomes. ERs are ligand activated transcription factors, members of the steroid/retinoid receptor family and like other members of this family are constituted by six functional and structural domains (A to F) [11, 41]. On the N'-terminal, the A/B domain functions as a transcriptional activator. Domain C or DNA binding domain (DBD) is responsible for mediating binding to Estrogen Response Elements (EREs) on the DNA. Between ERα and ERβ, there is a 96% homology on the DBD, suggesting that both ERs bind with the same affinity to EREs. The D domain is a region not well characterized. Ligand binding occurs in E/F Domain or Ligand Binding Domain (LBD). ERα and ERβ exhibit a 53% homology in the LBD (Fig. 5) which could indicate that although they bind the endogenous hormone E2 with the same affinity, they have different affinities for other ligands [42-43]. Two of the compounds that are used as selective agonists of ERα and ERβ are 4,4',4''- (4-propyl-(1H)-pyrazole-1,3,5-triyl)trisphenol (PPT) which is 410-fold more selective for ERα over ERβ and 2,3-bis (4-hydroxy-phenyl)-propionitrile (DPN), 72-fold more selective for ERβ rather than ERα [44-45]. ER selective agonists and antagonists are important tools to help elucidate ER role in mammary development and have been proposed as possible alternative therapy for breast cancer [46]. In addition, ERs have two activation functions (AFs) AF-1 in the DBD which is a consensus phosphorylation site activated in the absence of ligand, and AF-2 which is activated in a ligand dependent manner.

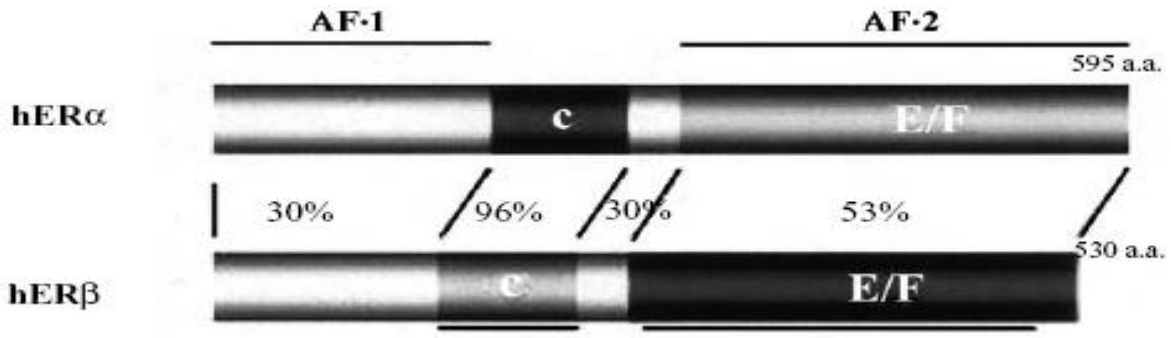


Figure 5– Comparison between human ERα and ERβ domains. Source: Weihua, Z., et al., 2003.

Estrogen and ERs regulate several cellular processes through distinct pathways. Ligand activation of ERs leads do receptor dimerization and direct binding to DNA at the promoter region ERE (Fig. 6A), modulating gene transcription. However, ER activation can interact with transcription factors such as AP-1 (Fig. 6B) or SP-1 and as a result, modulate gene expression indirectly. ERα and ERβ can present different/opposite effects on AP-1 or SP-1 regulated promoters. ER-ligand binding may also cause rapid effects through non genomic mechanisms which are not well understood [47]. In addition, ER activation can also be ligand independent as for example, growth factor signalling (i.e. EGF/EGFR) may lead to ER phosphorylation in AF-1. This mechanism is thought to be responsible for hormone independent growth of ER positive breast cancers.

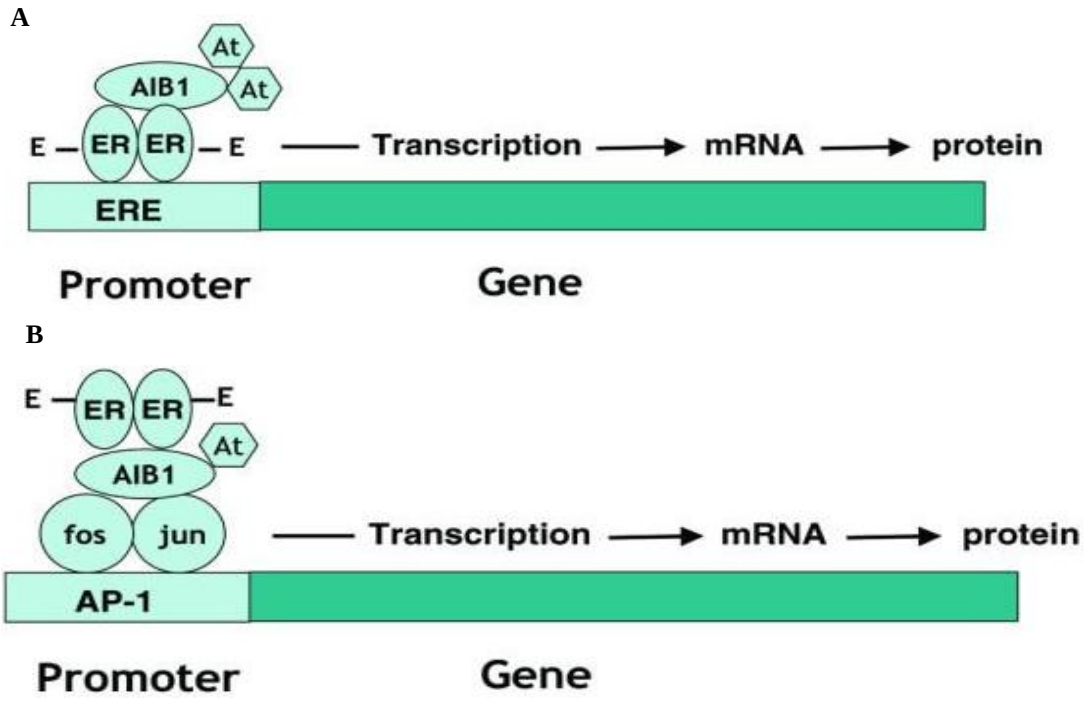


Figure 6– Scheme of transcription activation by ERs. A) Direct binding to promoter region ERE; B) Indirect binding to the promoter region AP-1. Source: Schiff, R. and Osborne, C.K. 2005.

The two ERs are differently expressed in the different body tissues. ERα is predominant in the uterus, mammary gland, testis, pituitary while ERβ is mainly encountered in the prostate and ovaries. In the mammary gland, it is possible to see a different distribution of both ERs. ERα is found in the ductal epithelium and ERβ is expressed in the ductal and lobular epithelium and also in the stroma [1, 11].

4. Breast Cancer

Of all the different types of cancer, breast cancer has the second highest incidence rate and the second highest death rate in western women [48]. Normally, in breast cancer, changes that transform normal cells in cancer cells occur in the alveolar or ductal epithelium and are associated to genetic or epigenetic alterations in normal signalling pathways that regulate proliferation, survival and differentiation [49]. In rare cases, this transformation can occur in fat tissue.

Despite all efforts to understand the mechanisms involved in this pathology there is still a lot to unravel and this is due to the heterogeneity of the disease. There are two hypotheses, not necessarily exclusive that aim to explain breast cancer heterogeneity [50]. One of the theories considers that each breast cancer subtype initiates in different types of cells (cell of origin model; Fig. 7A). The second theory refers that the cell of origin can be the same to different subtypes and the transformation is due to a series of genetic or epigenetic events (subtype transformation event model; Fig. 7B). Within these hypotheses, the existence of transformed stem cells (cancer stems cells) with capacity to differentiate into different epithelial subtypes is the most accepted nowadays and serves also to explain the plasticity of cancers that allows them to respond to environmental pressure and acquire resistance to therapy [50-51].

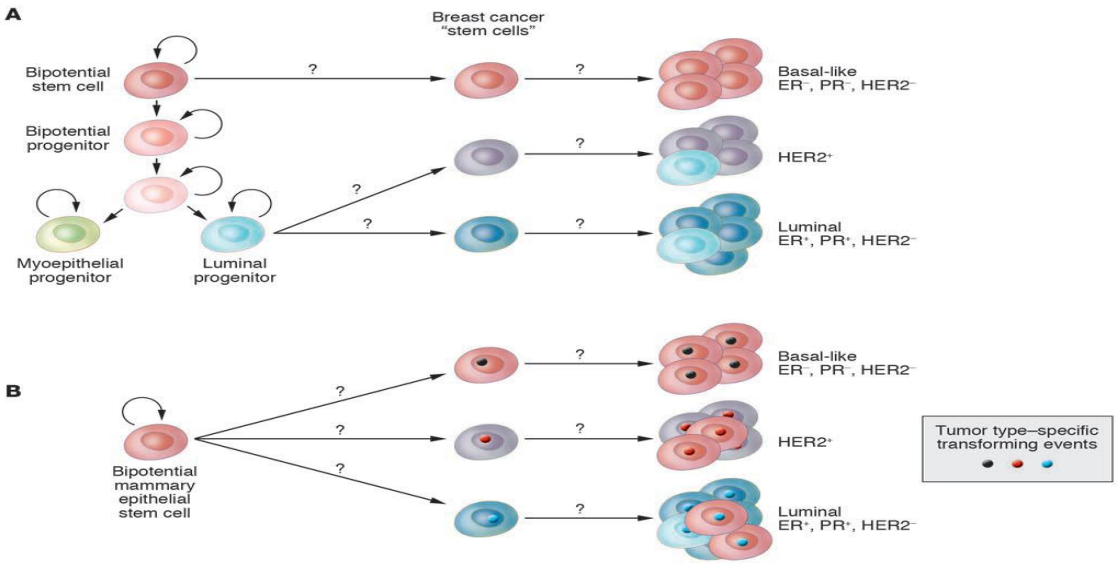


Figure 7– Origin hypothesis for the several breast cancer subtypes. A) Cell of origin model; B) Tumor subtype–specific transforming event model. Source: Polyak, K., 2007.

Through genetic profile analyses it was possible to group mammary tumours into five molecular subtypes that are transversal to all ethnicity [52-53]: luminal A and luminal B (both expressing ER), $HER2^{+}/ER^{-}$, basal-like (ER, PR and HER2 negative), and normal breast-like. Such classification is useful as in general, treatment of patients is administered according to the type of receptors expressed (i.e. ER antagonists, HER2 inactivating, antibodies). Luminal type can have expression of ER, PR and varying degrees of HER2. Luminal A tumours have high expression of ERα and PR and best prognosis; while basal-like tumours are the less differentiated, share expression of genes with mammary stem cell markers and have the worse prognosis and lower survival rate [54]. Due to the relatively novel discovery of ERβ, its potential as predictive factor is still under investigation.

4.1. Breast cancer and ER expression

Stimulation of normal mammary epithelium with E2 leads to increased proliferation and studies have demonstrated a correlation between breast cancer in humans and a long exposure to E2. Hence, E2 exposure has been regarded as a risk factor in breast cancer [11, 55]. This relation can be observed for example, by the fact that 70% of breast cancers express ERs and breast cancer occurs in women never before puberty and that women with a long fertile period present a higher risk [56]. Women with high E2 blood levels are also a risk group [57]. On the other hand, women with early menopause due to natural causes or castration have a lower risk of developing breast cancer [56].

The role of ERα inducing proliferation in breast cancer is well established and expression of this receptor is presently the only factor considered to indicate patient treatment with ER antagonists like tamoxifen or aromatase inhibitors to reduce E2 body levels. However, most cancers will develop resistance to this therapy and there is a need to find new treatments [55]. Studies in cell lines showed that ERβ inhibits proliferation and induces apoptosis and is important in maintaining differentiation [40, 58]. Therefore, it has been proposed that treatment with ERβ agonists could be used as an alternative therapy [42, 46]. However, there is still controversy regarding the role of ERβ in breast cancer and more studies *in vivo* need to be carried out [59]. Recently, we found that the pro-apoptotic

effect of ERβ both in mammary non-tumorigenic and breast cancer cells is reverted by activation of EGFR (Cotrim Z, in preparation). This suggests that cooperation of ERβ and EGFR pathways increase survival and growth, suggesting that the cellular context regulates ERβ activity.

5. EGF/EGFR signalling

The ErbB family of receptor tyrosine kinases is composed of 4 distinct members, EGFR (or HER1), HER2, HER3 and HER4. The four proteins present an extracellular ligand-binding domain, a transmembrane region and an intracellular tyrosine Kinase-containing domain [17, 60]. The intracellular domain is highly conserved within the receptors (except HER3), but the extracellular ligand-binding domain is less conserved between the receptors, suggesting that they have different affinity for different ligands [61-62].

EGFR can be activated by a variety of ligands, such as EGF, transforming growth factor α (TGFα), Areg, heparin-binding EGF, or betacellulin [63-64]. Upon ligand binding, receptor homo or heterodimerization occurs followed by internalization and subsequent autophosphorylation of tyrosine kinase domain that serves as binding site for the recruitment of signal transducers and other molecules. This activation, in turn, triggers several downstream signalling such as phospholipase C-γ (PLCγ), Ras-Raf MAP Kinase, STAT and PI3 kinase – Akt pathways (Fig. 8) [62-64].

Several studies have demonstrated that ErbB family members are important for the development of many organs such as central nervous system, mammary gland, lungs or heart [65-67], since it regulate multiple biologic processes, such as gene expression, cellular proliferation, angiogenesis, and inhibition of apoptosis [63-64]. Activation of EGFR was also shown to promote tumour development, motility, adhesion and metastasis [68-69] and HER2 is overexpressed in approximately 30% of breast cancers and is associated with poor prognosis [70]. In addition, activation of MAP kinase and PI3 kinase pathways leads to ERα ligand independent activation [71].

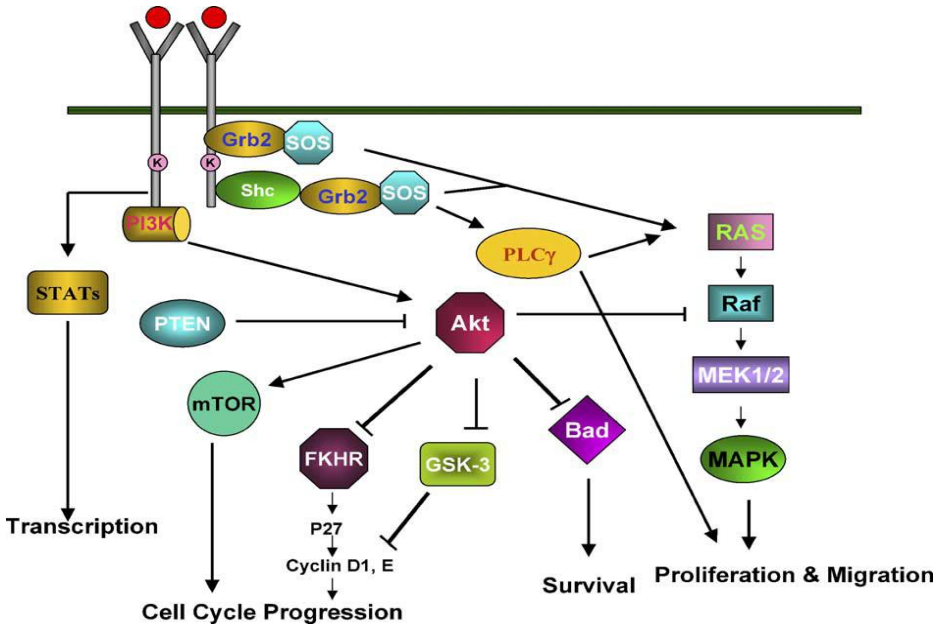


Figure 8– Signalling network dependent of EGFR activation. Source: Singh, A; 2005.

6. Proteomic approach to breast cancer study

Proteomics is the large-scale study of proteins using several techniques. These techniques allow the identification of proteins present in complex mixtures as is the case of the cell secretome. It also permits the study of protein-protein interactions and post-translational modifications (PTM). In this manner it is possible to use a proteomic approach to identify biomolecular markers and better understand the molecular pathways that regulate cell differentiation and that if altered may lead to breast cancer initiation and progression [72-75]. Ultimately, this knowledge will allow earlier identification of the pathology and the development of new drugs and therapies.

These studies can be performed with a variety of samples such as plasma, mammary gland tissue, nipple fluid or cell lines in culture. The results of these studies have demonstrated numerous differences on protein expression between tumorogenic and normal samples. Examples of relevant identifications are presented below.

Within proteomic methodologies, we can distinguish between mass spectrometry-based tools (MS-based tools) and Non-based-MS tools. The Non-based-MS methodologies include mainly antibody reactivity requiring a previous biological knowledge of the proteins to be tested in bodily and tissue samples. On the other hand the MS-based

methodologies do not require this kind of knowledge permitting identification and characterization of a large number of proteins. Within MS-based approaches, and facing the advances in instrumentation, sample analysis can be performed by techniques catalogued in gel-based and non-gel based [73, 76].

Gel-based proteomics include an initial separation step by gel electrophoresis (e.g. SDS-polyacrylamide gel electrophoresis (SDS-PAGE) or 2D-polyacrylamide gel electrophoresis (2D-PAGE)). For instance, in SDS-PAGE, the protein separation is done only by molecular weight, and in traditional 2D-PAGE proteins are separated in two dimensions being the first dimension related to the isoelectric point (pI) and the second dimension related to their molecular weight. Although 2D-PAGE allows a better protein separation and is the most used technique it has several disadvantages such as poor reproducibility, low sensitivity for identification of proteins in low concentrations and hydrophobic proteins and it is more expensive and time consuming than SDS-PAGE. However, the two dimensions allow separation of bands that in SDS-PAGE migrate together. After protein gel separation, bands/spots are cut from the gel, digested with a protease (typically trypsin) and peptides analysed by MS using peptide mass fingerprint. Matrix-assisted laser desorption ionization (MALDI-MS) and electrospray ionization (ESI-MS) are the most common sources to analyse the digested peptides. These sources can be coupled to different types of analyzers such as time-of-flight (TOF) quadrupole, ion-trap. In peptide mass fingerprint, protein identification is obtained through the comparison of experimental mass-to-charge ratio (m/z) of the peptides with theoretical m/z of the peptides stored in databases. Identification can also be done by fragment mass fingerprint. In this case, tandem MS spectra is used to retrieve the best peptide sequence based on fragment ion for further comparison with theoretical m/z values existing in database [74, 77].

6.1. Matrix-Assisted Laser Desorption Ionization (MALDI)

MALDI is a soft ionization technique utilized in MS, mainly for the analyses of biomolecules such as proteins, DNA or lipids. In this technique the sample is first dissolved in a suitable solvent and then co-crystallized with an appropriate matrix being spotted on a MALDI plate and air-dried or dried by nitrogen gas. The plate is put on the

spectrometer where a laser beam (typically a nitrogen laser) hits the co-crystallized sample leading to absorption of the laser energy by the matrix and consequent desorption and ionization of the analytes in the sample, mainly as single charged species (Fig. 10).

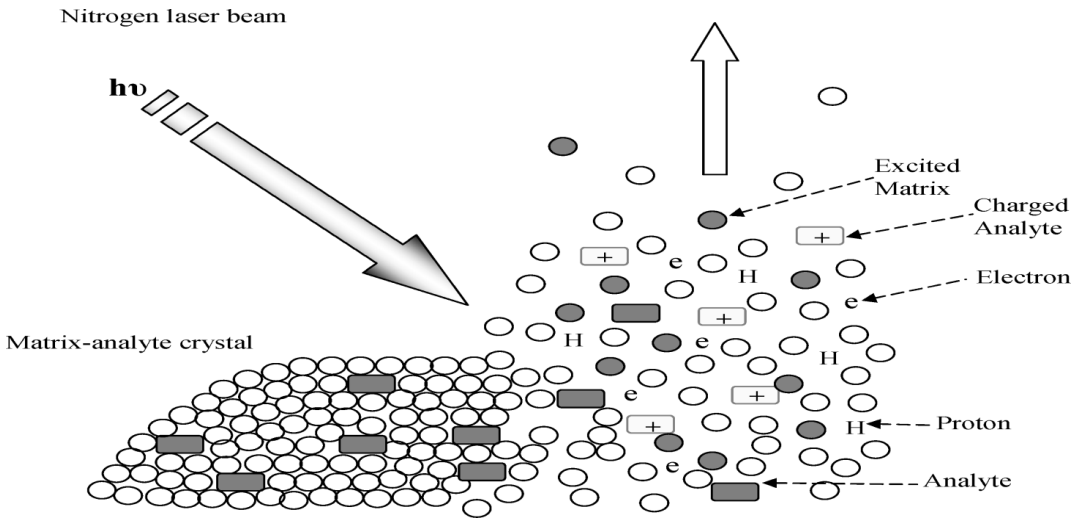


Figure 9 - MALDI ionization process. Source: El-Aneed, A.; Cohen, A.; Banoub, J. 2009.

Although, in the beginning MALDI was operated under vacuum it is now possible to operate at atmospheric pressure (AP) MALDI, reducing the cost of this technology and being easier to operate. The choice of the matrix depends not only on the sample to be analysed but also on the conditions under which the ionization is made. Normally, on the positive mode, acid matrices are used and on negative mode basic matrices are used [78-81].

6.2. Time of Flight (TOF) Analyzer

In a second step, ions enter a TOF analyzer and the m/z is calculated based on the time the ionized ion takes to travel along the tube (variable length) until it reaches the detector. The principle is that if two ions are created at the same time, have the same charge and are accelerated with the same energy but have different weights, the lighter ion will reach the detector earlier. So, the m/z is inversely proportional to the time of flight.

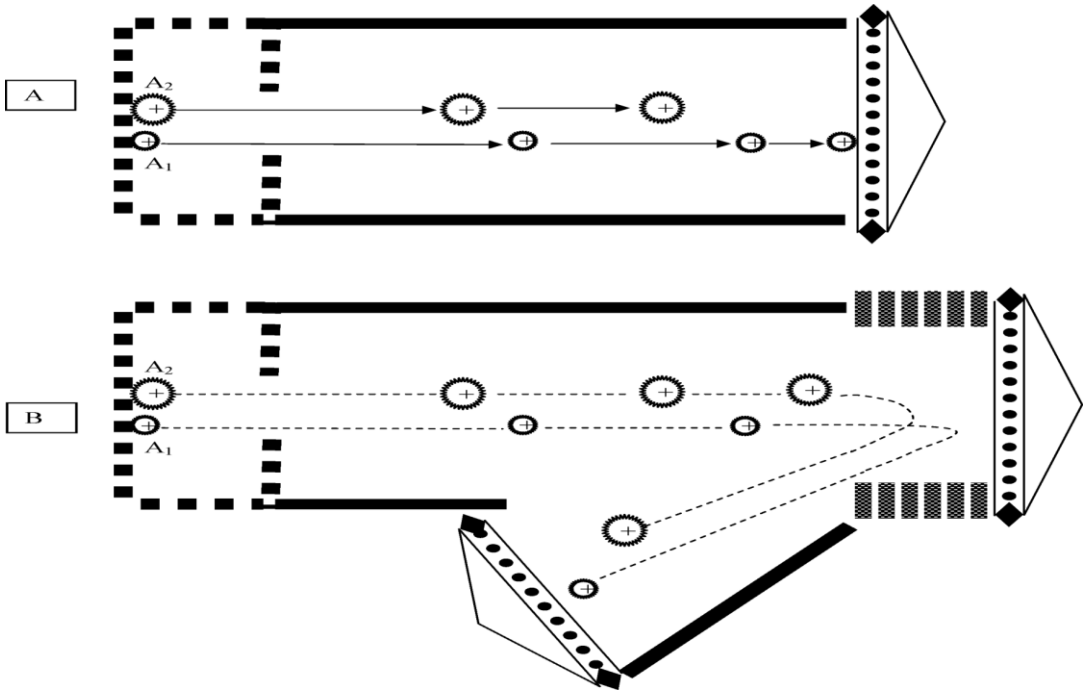


Figure 10 - Scheme of a conventional TOF analyzer (A) and a TOF with reflectron (B). Source: El-Aneed, A.; Cohen, A.; Banoub, J. 2009.

Comparing this kind of analyzer with others it has several advantages: all formed ions reach the detector and a very high mass range can be detected. However, the ions that enter the TOF have different kinetic energies affecting resolution and ion m/z measurement. To overcome this obstacle a reflectron (electrostatic ion mirror) was created. In a reflectron-TOF ions with higher kinetic energies penetrate deeper on the mirror and are expelled on a higher velocity, improving resolution (Fig. 11). As this factor increases the path ions perform, it also permits more accurate m/z measurements. This kind of analyzer is also capable of tandem MS (MS/MS) [82-84].

6.3. Tandem Mass Spectrometry (MS/MS)

Although first order spectrums permit obtaining of rigorous m/z measures of the fragmented ions, they give little information about molecular structure and most frequent

and favourable molecular fragmentation. Thus, the protein identification is more reliable if done from a MS/MS spectrum since it allows almost complete peptide sequencing.

On tandem mass spectrometry (MS/MS) the selected fragmented ions obtained in the first mass analyzer (MS1) are isolated and subjected to a posterior fragmentation through collision (normally with a gas) in a collision cell, being separated in a second analyzer (MS2) before reaching the detector (Fig. 12).

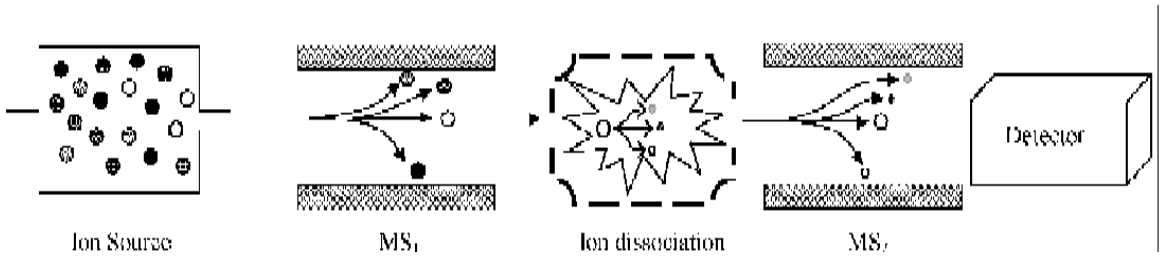


Figure 11 – Representative scheme of tandem mass spectrometry (MS/MS). Source: El-Aneed, A.; Cohen, A.; Banoub, J. 2009.

This type of analyzes can be made recurring to devices such as quadrupoles and TOF. MS/MS enhanced with these analyzers is called tandem-in-space because the analysis is performed by different mass analyzers. On the other hand, in tandem-in-time all ions are expelled from the trapping instrument except the selected ion. Within these instruments we can refer ITQ and FT-MS that can perform multiple MS experiments [85-87].

6.4. HC11 cells

HC11 cell line has been widely used to model mammary epithelial differentiation. This cell line was obtained from the mammary gland of BALB/c mice in the middle of pregnancy. HC11 cells retain progenitor cells characteristics as they can be induced to differentiate (express milk proteins) in vitro [88] and can originate ductal and alveolar structures when implanted in vivo [89]. Comparison of the transcriptome of undifferentiated stage of HC11 cells to public breast cancer cell data showed that they share expression of markers with other stem cells and basal-like breast cancer cells [90].

Since we do not exactly know if HC11 cells are "pure" stem cells or progenitor cells, we refer to them as stem cell like (SCL). HC11 cells are routinely kept in the undifferentiated, self-renewal and pluripotency stage by addition of EGF and induced to differentiate by withdrawal of EGF and addition of lactogenic hormones (dexamethasone and prolactin). For these reasons HC11 cell line presents several properties specific of normal mammary gland function, making them a relevant model to study markers and pathways in mammary gland development. A Schematic representation of the differentiation protocol of HC11 cells is presented in Fig. 9.

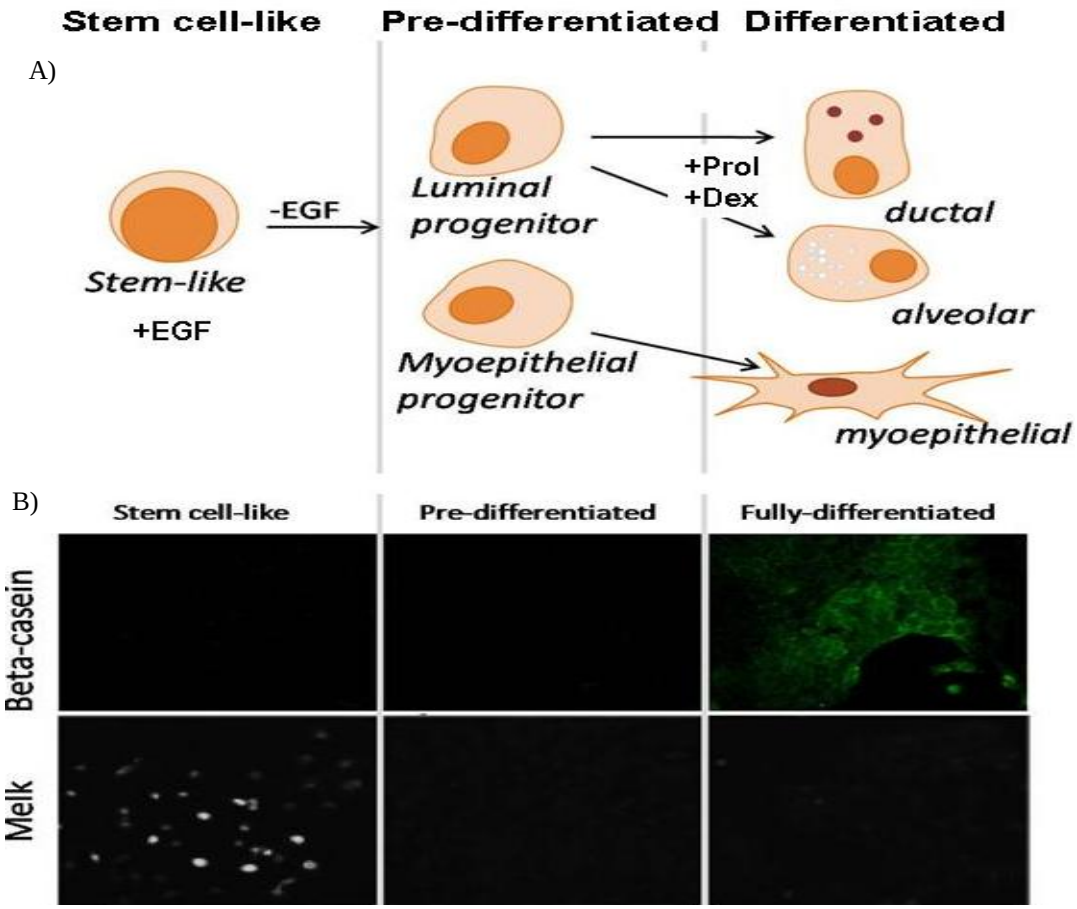


Figure 12 – HC11 cell differentiation program. A) Culture conditions used to achieve each differentiation stage. B) Expression of stem-cell like marker (Melk) and differentiation marker (beta-casein) achieved under the indicated growth conditions. Source: adapted from Williams, 2009 and Aydogdu (upublished).

7. Aim

Paracrine and autocrine signalling play important roles in mammary epithelial cell proliferation, differentiation and apoptosis. Current knowledge supports the idea that breast cancer is the result of genetic or epigenetic alterations in mammary stem or progenitor cells [91] which remain quiescent until induced to expand. Further, breast cancer was the first tissue from which a tumour initiating cell subpopulation was identified [92]. Long-term exposure to estrogens and activation of ERα has been associated to breast cancer development [55]. However, little is known about how ERβ activation influences differentiation and cancer development, but current knowledge indicates it has a role in regulating adhesion.

The main goal of this work was to identify novel secreted proteins which may serve as markers for different cellular states and provides a starting point for identification of novel pathways that regulate differentiation and that may become altered in breast cancer. It is also our objective to analyze the effect of ERβ ligand activation in the secretome of SCL cells and differentiated cells and identify possible mechanisms regulated by ERβ.

Therefore, this work was divided into three specific aims:

1. To use gel based MS to identify expression of secreted proteins in mammary epithelial cells throughout their differentiation program.
2. To use gel based MS to identify expression of secreted proteins in undifferentiated (SCL) mammary epithelial cells under ERβ stimulation.
3. To use gel based MS to identify expression of secreted proteins in differentiated mammary epithelial cells under ERβ stimulation.

II - Materials and Methods

In order to simplify the text reading, all solutions recipes used on the experimental approach can be consulted in annexe 2.

Cell culture

HC11 cells were routinely grown in complete medium. Collection of conditioned medium (CM) corresponding to each differentiation stage was done as follows. Actively proliferating, stem cell like (SCL) stage: once cells reached 60-70% confluence the medium was changed to SFM medium + EGF and after 24h incubation, the CM was collected. Pre-differentiated (-EGF 48h) stage: once cells reached confluence the medium was changed to pre-differentiation medium for 24h. Next, medium was changed for SFM –EGF and cells incubated for 24h more before the CM was collected. Differentiated (Dif) stage: once cells reached confluence the medium was changed to pre-differentiation medium for 48h. Then, medium was changed for differentiation medium and cells incubated for 48h. Next, medium was changed for SFM –EGF with addition of prolactin and dexamethasone and CM was collected after 24h. When ER ligand DPN was used, it was added to the indicated SFM for each differentiation stage and cells were incubated for 24h before CM was collected.

Protein precipitation and quantification

The collected CM was centrifuged at 10000g for 15 min at 4ºC to remove cellular debris. Protein precipitation was carried out using the trichloroacetic acid (TCA) /acetone method. Briefly, the samples were incubated 2h at 4ºC with 1/10 acetic acid and 1/10 TCA 100%. After the incubation, samples were centrifuged 15 min, 18000g, 4ºC. The resulting pellet was washed with ice cold acetone and subjected to a new centrifugation cycle after which the pellet was solubilised in solution 1 and proteins separated by SDS-PAGE. Pellets corresponding to samples resolved by 2D electrophoresis were solubilised in solution 2.

Protein quantification was performed using the DC assay or RC/DC assay (BioRad) for samples solubilised in solution 1 or solution 2 , respectively.

SDS-PAGE

Protein separation in a Tris-Glicine SDS-PAGE system was based on the protocol developed by Laemmli (1970) [93] where 10% and 6.25% gels were used. Equal amount of protein was used (20-30µg, depending on the gel). Proteins were solubilised with 1Vol of Laemmli loading buffer 5X every 4Vol of sample and proteins were denaturalized by boiling at 100ºC for 5 min. Then, samples were loaded in the gel and separated with 200V for 1h. The electrophoresis was ended once the front of the run reached the end of the gel.

2D electrophoresis

The samples separated by 2D electrophoresis were adjusted with rehydration buffer to a final volume of 250µl. Then, samples were placed in the holders and the 13cm 3-10 NL pH gradient Drystrips were laid on the top of the sample with the gel side down. The strips were covered with Drystrip cover fluid and the holders were placed on the IPG strip rehydrating chamber. Rehydration was carried out overnight (16.5h) at 20ºC at 50mA.

After isoelectric focusing (IEF), the excess of Drystrip cover fluid was removed and the strips were equilibrated for SDS-PAGE in 2.5ml of equilibrium buffer for 15 min with gentle agitation. The strips were washed with running buffer and placed in top of a 10% gel. Agar solution was applied on top of the strips to seal the junction and to make a well for the molecular weight marker. SDS-PAGE was carried out at 200V, 20ºC until bromophenol blue reaches the bottom of the gel.

Gel staining with Colloidal Coomassie (G250)

After protein separation (SDS-PAGE or 2D electrophoresis) proteins were fixed for 1h in fixing solution. Next they were rinsed overnight in Colloidal Coomassie. Following staining, the gels were placed in distaining solution, till background noise was removed. Gel images were acquired on Molecular Imager Gel Doc XR+ System (BioRad) and

analysed with QuantityOne version 4.6.3 (BioRad) in the case of 1D gels or PDQuest version 8.0.1 (BioRad) for 2D gels.

Protein identification

Bands/spots of interest were cut from the gel and rinsed in 50µl Buf1 for 30 min, then 50µl of ACN was added for additional 30 min, after which, it was also removed. These two steps were performed twice. The bands/spots were, once again, rinsed in 50µl ACN for 10 min. Next, the bands/spots were incubated with trypsin, overnight at 37ºC. The supernatants were collected and storage and the digests were rinsed with formic acid for 30 min and two times with a 1:1 solution of FA and ACN. All the supernatants were collected and evaporated in a SpeedVac and ressuspended in 10µl 50% ACN, 0.1% FA. Next, 0.5µl of the peptide mixture was spotted on the MALDI plate and covered with a 50% ACN, 0.1% TFA acid and 5mg/ml α-cyano-4-hydroxycinnamic acid.

The spots where analysed by MALDI-TOF, where the MS was performed with 900 shots randomly distributed while in MS/MS was acquired 6 peaks with S-N>30. Peptide masses obtained were exported using GPS software (v.2.0 Applied Biosystems) for searching the peptide mass fingerprints and MS/MS data. Searches were performed against the SwissProt (11/01/11) under *rodentia,* as taxonomic category and the following parameters: (i) two missed cleavages by trypsin; (ii) mass tolerance of precursor ions 25ppm and product ions 0.3Da; (iii) carboxymethylated cysteines fixed modification; and (iv) oxidation of methionine as variable modification.

III – Results and Discussion

1. Proteins indentified in the mammary epithelial cell secretome throughout the differentiation program.

Differentiation of mammary epithelial cells is dependent on autocrine and paracrine signalling. With the goal of better understanding these mechanisms, HC11 cells were cultured to obtain three stages of mammary epithelial cell differentiation:

a) Stem-cell like (SCL), when cells are capable of self-renewal and pluripotency (it was obtained by culturing with EGF (+EGF));

b) Pre-differentiation (PD) stage, when cells enter the differentiation process but the process can be partially reversed (it was obtained by withdrawal of EGF (-EGF for 24-48h));

c) Differentiation (D) stage, when cells express milk proteins (+Prl + Dex -EGF).

To better understand the protein profile of mammary epithelial cells as they transit the differentiation program, conditioned media (CM) of SCL, PD and D stages was analyzed using a gel based MS approach. The first step was to analyze the proteins in the CM of each differentiation stage using SDS-PAGE. A representative image of a SDS-PAGE (10%) showing a comparison between the protein profiles of the 3 conditions, SCL, PD and D is shown in Fig. 14. All bands were cut and identified by MALDI-TOF/TOF and presented in Table 1. Those bands differentially detected are indicated with numbers which are also referred to in Table 1.

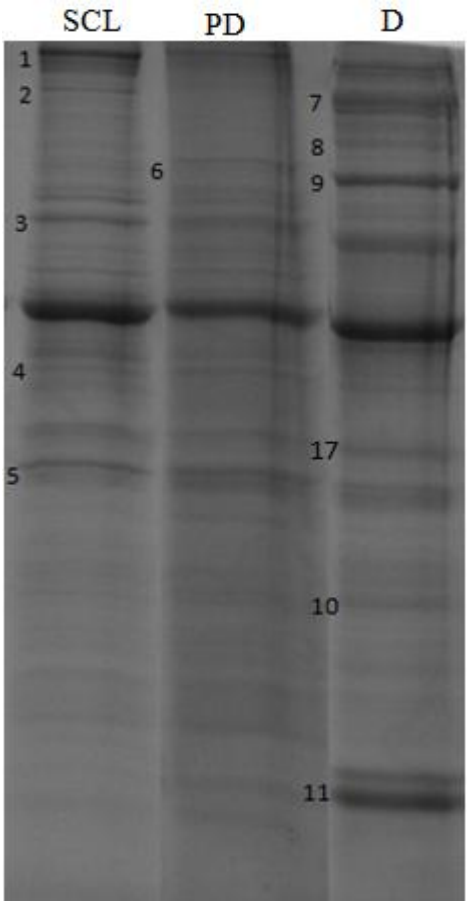


Figure 13 – SDS gel comparing protein profile between SCL, PD and D. The numbers indicate proteins differentially identified. Identifications are reported in Table 1.

Some tecnhical dificulties were encountered using SDS-PAGE such as contamination with serum albumin due to the high concentration levels even in the absence of bovine fetal serum. In addition, some of the proteins have similar molecular weight impossibiliting their identification and there was a high number of proteins present in the sample that due to their lower concentration were not possible to identify.

So, the next step was to perform a 2D-PAGE separation to try to overcome some of the dificulties presented to us when using SDS-PAGE. The protein profiles of the 3 conditions are shown in Fig. 14 and the proteins identified are also listed in Table 1.

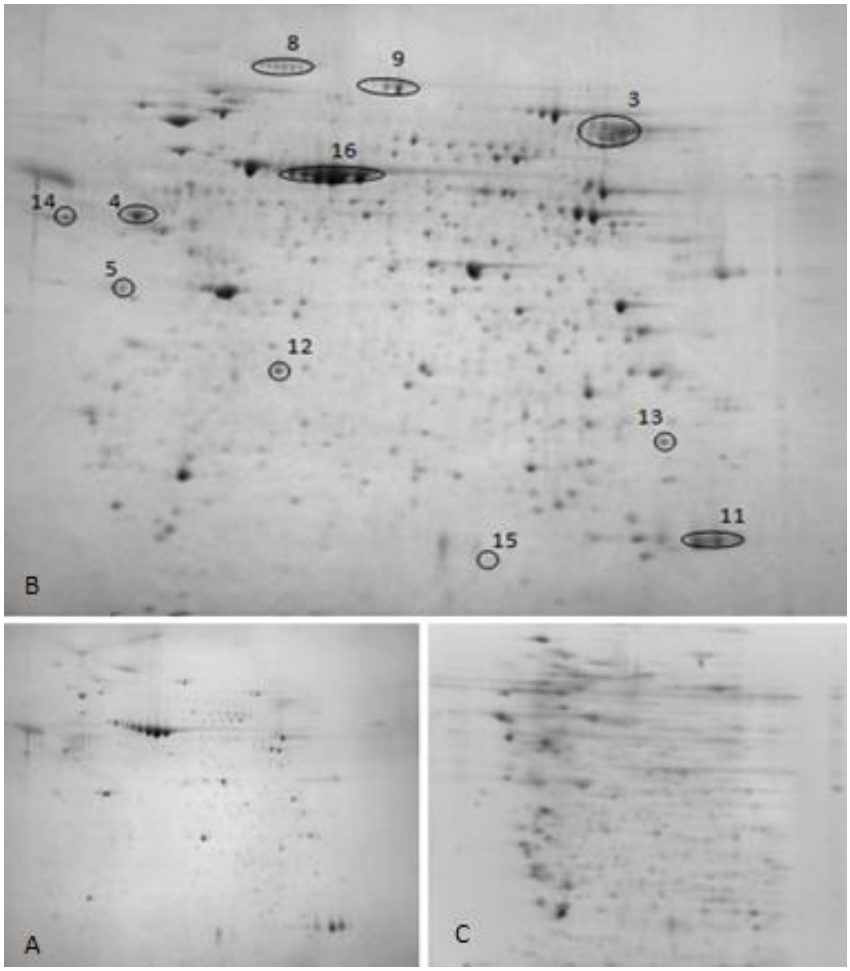


Figure 14 – 2D-PAGE comparing protein profile between A) SCL, B) PD and C) D.
The numbers indicate proteins differentially identified. Identifications are reported in Table 1

A total of 75 (SCL), 123 (PD) and 80 (D) positive identifications were obtained by 1D and 2D. Since this work pretends to study the secretome of cells in each stage, we only analyzed proteins that are known to be extracellular or are present in the membrane or lysosome/exosome compartments. However, tables with all proteins identified are presented in the annexes section. The number of secreted or potentially extracellular proteins was 20 (SCL), 28 (PD) and 24 (D).

A combined identification by SDS-PAGE or 2D-PAGE permitted the identification of a total of 40 different proteins within the 3 classes. Table 1 shows the proteins secreted in the 3 experimental conditions giving also information about the accession number, protein MW, pI and peptide count. In addition, Fig. 15, shows a Venn diagram for easier visualisation of the results.

Table 1 – Proteins identified from the secretome of mammary epithelial cells with a C.I.%>95%.

#	SCL	PD	D	Protein name (gene name)	Acession number	Protein MW (Da)	Protein pI	Peptide count
21		X		14-3-3 protein sigma (Sfn)	O70456	27696	4,75	8
22	X	X		6-phosphogluconate dehydrogenase, decarboxylating (Pgd)	Q9DCD0	53213	6,81	3
17	X	X	X	Alpha-enolase (Eno1)	P17182	47111	6,37	9
20	X		X	Annexin A2 (Anxa2)	P07356	38652	7,55	15
12	X	X		Annexin A3 (Anxa3)	O35639	36349	5,33	11
			X	Annexin A5 (Anxa5)	P48036	35730	4,83	19
		X		Annexin A8 (Anxa8)	O35640	36820	5,56	2
23		X		Basement membrane-specific heparan sulfate proteoglycan core protein (Hspg2)	Q05793	398039	5,88	2
			X	Cadherin-1 (Cdh1)	P09803	98195	4,69	5
14		X	X	Calreticulin (Calr)	P14211	47965	4,33	16
6		X		Calsyntenin-1 (Clstn1)	Q9EPL2	108831	4,82	11
			X	Carboxypeptidase E (Cpe)	Q00493	53222	5,07	10
26		X	X	Cathepsin D (Ctsd)	P18242	44925	6,71	7
8	X	X	X	Ceruloplasmin (Cp)	Q61147	121074	5,53	16
	X			Coagulation factor VIII (F8)	Q06194	265981	6,9	17
7		X	X	Collagen alpha-1(II) chain (Col2a1)	P28481	141886	6,58	32
18	X			Collagen alpha-1(VI) chain (Col6a1)	Q04857	108422	5,2	9
		X		Collagen alpha-1(XI) chain (Col11a1)	Q61245	180853	5,09	7
9	X	X	X	Complement C3 (C3)	P01027	186365	6,39	22
27			X	Connective tissue growth factor (Ctgf)	P29268	37768	8,22	2
		X		Cystatin-C (Cst3)	P21460	15521	9,18	9
			X	EF-hand domain-containing protein D2 (Efhd2)	Q9D8Y0	26775	5,01	9
	X	X		Extracellular superoxide dismutase [Cu-Zn] (Sod3)	O09164	27375	6,36	3
24	X	X		Fascin (Fscn1)	P85845	54457	6,29	3
1	X	X	X	Fibronectin (Fn1)	P11276	272319	5,39	35
13	X	X	X	Galectin-3 (Lgals3)	P16110	27498	8,46	6
10		X	X	Insulin-like growth factor-binding protein 5 (Igfbp5)	Q07079	30353	8,48	5
19		X		Liprin-alpha-3 (Ppfia3)	P60469	116208	5,72	4
15		X		Metalloproteinase inhibitor 2 (Timp2)	P25785	24312	7,45	5
11	X	X	X	Neutrophil gelatinase-associated lipocalin (Lcn2)	P11672	22861	8,96	6
3	X	X	X	Periostin (Postn)	Q62009	93085	7,27	25
			X	Prostate and testis expressed protein 3 (Pate3)	B3GLJ3	11652	9,14	5
4	X	X	X	Protein disulfide-isomerase (P4hb)	P09103	57108	4,79	18

	X	X		Protein disulfide-isomerase A3 (Pdia3)	P27773	56643	5,88	5
	X	X	X	Protein disulfide-isomerase A6 (Pdia6)	Q922R8	48143	5	6
			X	Renin receptor (Atp6ap2)	Q9CYN9	39067	5,36	6
25	X	X		Serpin B5 (Serpinb5)	P70124	42085	5,55	7
16	X	X	X	Serum albumin (Alb)	P07724	68648	5,75	8
5	X	X	X	SPARC (Sparc)	P07214	34428	4,77	11
2	X		X	Thrombospondin-1 (Thbs1)	P35441	129564	4,72	11

#: numbers presented in the gel images.
X: identified in the indicated stage.
OBS: Some identifications are not shown in the represented gels but were presented in others.

Eleven proteins were detected in all differentiation stages: Alpha-enolase, Ceruloplasmin, Complement C3, Fibronectin, Galectin-3, Neutrophil gelatinase-associated Lipocalin, Periostin, Protein disulfide-isomerase, Protein disulfide-isomerase A6, Serum Albumin and SPARC. Other proteins were only detected in 1 of the classes. Coagulation factor VIII and Collagen alpha-1(VI) chain were only identified in SCL stage. Twelve proteins were identified when cells entered the differentiation program (-EGF): 14-3-3 protein sigma, Anexin A8, Calreticulin, Calsystenin, Cathepsin D, Collagen alpha 1 (II and XI) chains, Cystatin-C, Basement membrane-specific heparan sulfate proteoglycan core protein, Insulin-like growth factor-binding protein 5, Liprin-alpha-3 and Metalloproteinase inhibitor 2. Since entering the differentiation program is induced by withdrawal of EGF, it is likely that this growth factor either regulates transcription, stability and/or secretion of the above listed proteins.

Functional differentiation induced by addition of Prol and Dex resulted in identification of 7 proteins characteristic of D stage: Anexin A5, Cadherin 1, Carboxypeptidase E, connective tissue growth factor, EF-hand domain-containing protein D2, Prostate and testis expressed protein 3 and Renin receptor. Unfortunately, due to technical problems, all spots present in D stage could not be analyzed. Therefore, proteins present in SCL and PD but not identified in D cannot be considered as differentially regulated in our analysis and more than the seven listed proteins may be characteristic of D stage.

Previously, it was shown that Insulin-like growth factor-binding protein 5 is expressed in differentiated HC11 cells [94]. In this analysis, Insulin-like growth factor-binding protein 5 was found in lower levels in SCL stage compared to D stage. In addition ceruloplasmin as well as complement C3 which were both upregulated in D (Table 3) have

been identified in breast milk [95-96]. Thus, indicating that in these experimental conditions, HC11 cells recapitulate the mammary epithelial cell differentiation program.

A previous transcriptomic study [90] showed different expression patterns in SCL comparatively to D stage. Several genes corresponding to proteins identified in this work, were found underexpressed in D when compared to SCL including Thrombospondin-1, Annexin A8 and Fascin. On the contrary, Cathepsin D, Collagen alpha-1(II) chain and Ceruloplasmin were upregulated.

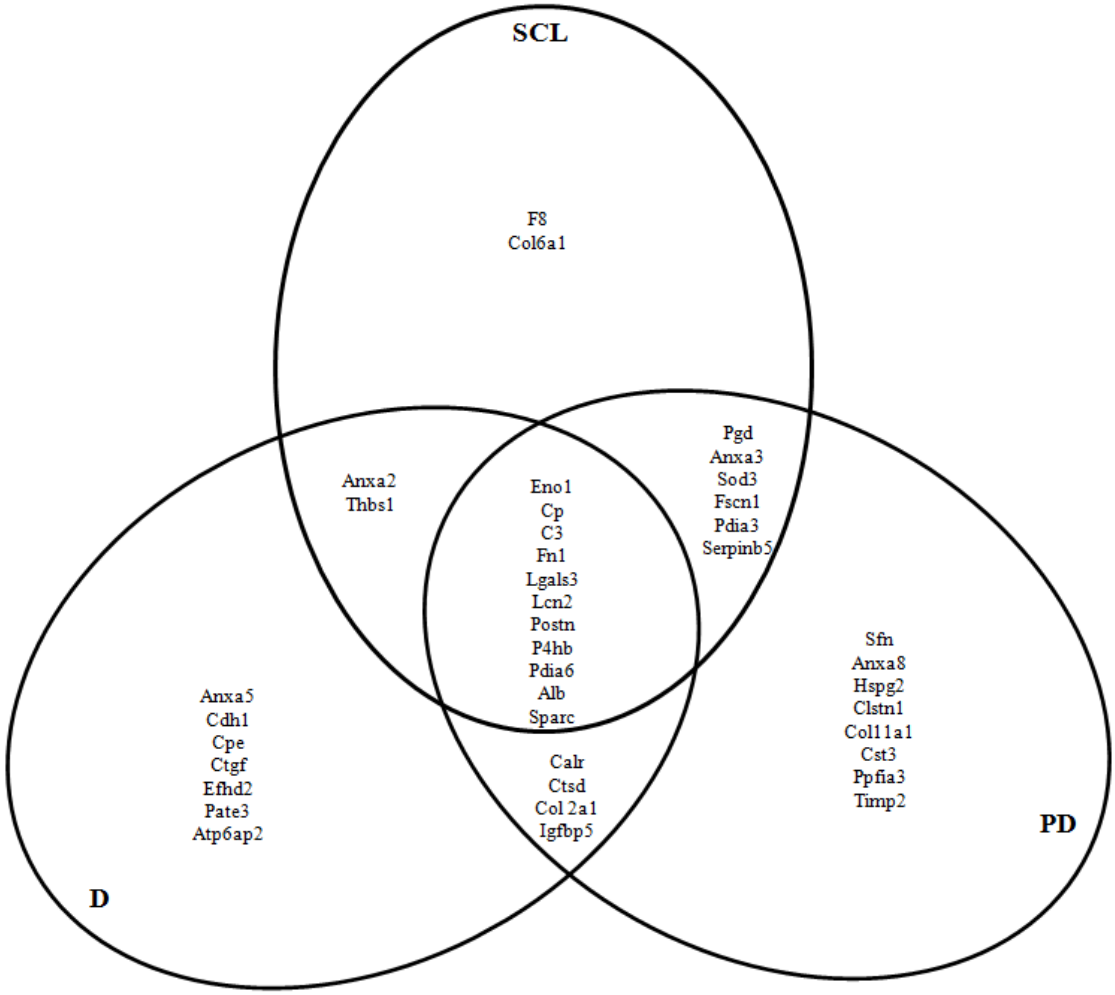


Figure 15– Venn diagram of the identified proteins and their distribution.

With information gathered from the proteins identified in each condition, pie charts with biological processes were built (Fig.16). It is possible to observe that the 3 top represented biological processes were the same in all the conditions: Cell communication (52,6% in SCL; 46,4% in PD; 42,9% in D), cell adhesion (30,0% in SCL; 33,3% in PD;

25,0% in D) and metabolic process (63,2% in SCL; 57,1% in PD; 39,3% in D). Another important biological process is transport, which was represented by 42,1% in SCL; 32,1% in PD and 25,0% in D. On the bottom down, apoptosis (1,3% in SCL; 1,1% in PD; 1,2% in D) was one of the less represented biological process as well as cell cycle (1,1% in PD; 2,4% in D) which was represented in PD and D conditions and not present in SCL.

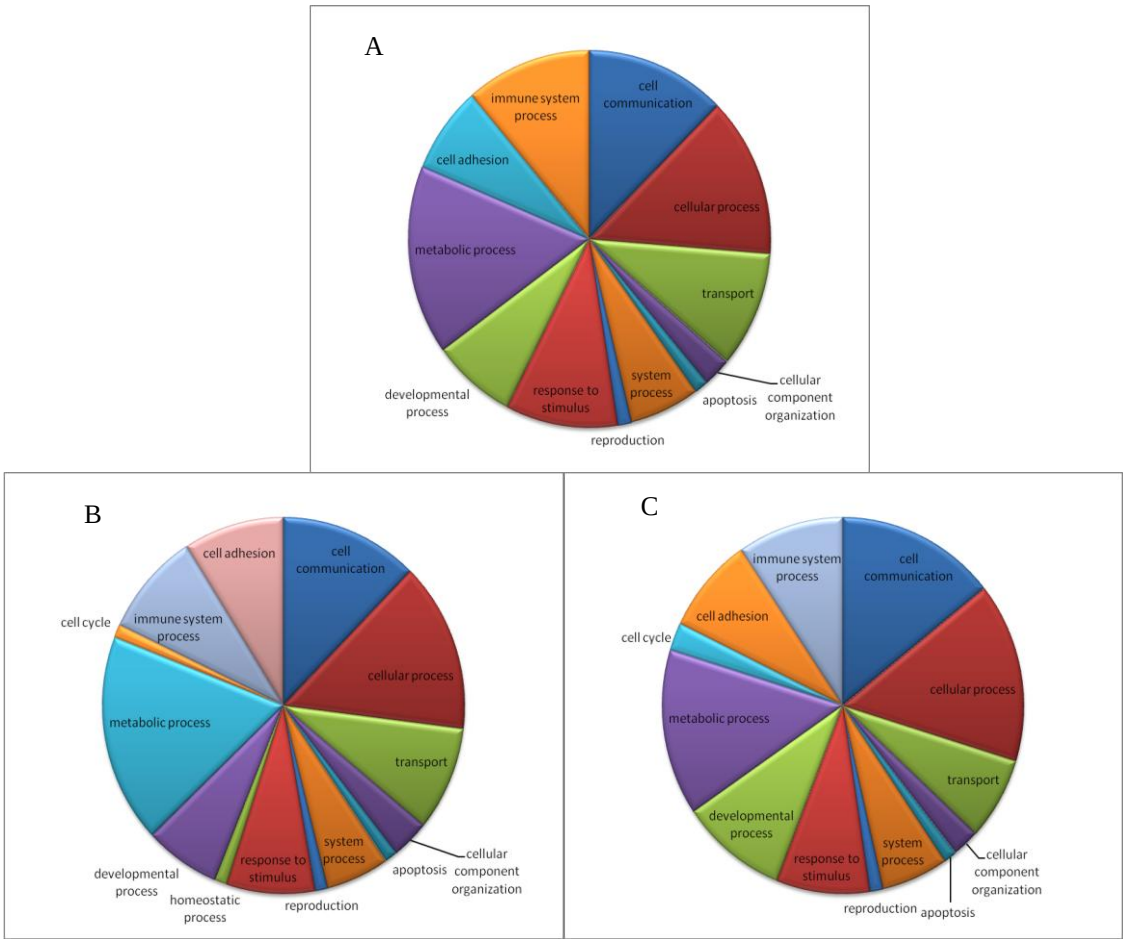


Figure 16– Pie charts with Biological processes of identified proteins A) SCL B) Pre-differentiated C) Differentiated

Information about molecular function of the proteins present in each condition was also analysed. The most represented functions were: binding (63% in SCL, 57% in PD; 54% in D) and Catalytic activity (42% in SCL, 29% in PD; 25% in D).

Relatively to the protein classes identified and according to Pantherdb.org database, we could identify 2 major classes in the 3 conditions (Table 2): signalling molecules (42,1% in SCL; 21,4% in PD; 25,0% in D) and transfer/carrier proteins (21,1% in SCL; 14,3% in PD; 14,3% in D).

Table 2 – Proteins belonging to the 2 major proteins classes.

Protein class	Proteins identified
signalling molecules	Ctgf; C3; Col6a1; Fn1; Lgals3; F8; Col11a1; Postn; Thbs1; Sparc
transfer/carrier proteins	Anxa8; Lcn2; Anxa2; Anxa3; Anxa5; Alb

Several metabolic pathways were identified; however, most of them were only represented by 1 protein. Integrin signalling pathway is the only one that in the 3 conditions showed more than 1 protein identified. Since levels of Collagen type VI, II, XI and Fibronectin activate different integrin dimmers and the levels of these ECM proteins varied between the differentiation stages. This finding is in agreement with the idea that integrin signalling regulates cell differentiation [97]. In SCL stage, 2 proteins participate in p53 pathway: Thrombospondin-1 and Serpin B5, which are upregulated by p53 and in turn, inhibit migration and angiogenesis [98-100].

The intensity of the bands/spots separeted by 1D or 2D-PAGE was measured and normalized, permiting a comparison between the 3 conditions. On table 3, the proteins that are only over or under represented in SCL or PD compared to D stage are shown. This should allow us to have a better understanding of the process of differentiation.

Table 3 – Protein variation on SCL and PD when compared to D stage.

Protein name	Variation (mean±SD)	Protein name	Variation (mean±SD)
Up only in SCL			
Thrombospondin-1	1,68±0,40	Protein disulfide-isomerase	1,81±0,40
Alpha-enolase	1,60±0,10	Fibronectin	2,87±0,20
SPARC	1,23±0,03		
Up only in PD			
Insulin-like growth factor-binding protein 5	1,31±0,06		
Down only in SCL			
Down only in PD			
Periostin	0,51±0,16	Collagen alpha-1(II) chain	0,13±0,08
Ceruloplasmin	0,48±0,25	Complement C3	0,16±0,03
Neutrophil gelatinase-associated lipocalin	0,05±0,57		

Comparing SCL with PD and D we found 5 upregulated proteins and 1 protein that was only detected in SCL (Table 1 and 3). Five of the upregulated proteins are involved in cell communication: Thrombospondin-1, Fibronectin, Collagen type VI, Coagulation factor

VIII and SPARC. They act as receptor ligands and are able to modulate cell-cell or cell-matrix adhesion. Thrombospondin-1 is also implicated in inhibition of angiogenesis and tumour growth as well as regulation of MMP-9 (facilitates angiogenesis and tumour invasion) activity an activation of vascular endothelial growth factor (VEGF) signalling [101-102]. In addition, SPARC was also found in higher levels in SCL compared to PD and D. SPARC is a glycoprotein that interacts with ECM proteins to promote cell adhesion to the matrix. It also plays an important role in tissue remodelling, angiogenesis and tumorigenesis [103-105]. In breast cancer SPARC can promote invasion, motility and activity of matrix metalloproteinases [106]. SPARC is overexpressed in breast tumour tissues when compared to normal breast tissues and is also highly expressed in aggressive breast cancer cell lines [75, 107]. Fibronectin is capable of modulating cell motility, adhesion and shape and is thought to inhibit angiogenesis, tumour growth and metastasis [108]. Thus, it seems that SCL cells have both pro-angiogenic/invasion factors (SPARC) as well as factors that counteract that activity (Thrombosponding and fibronectin) which may explain the regulated "normal" behaviour of these cells.

Coagulation factor VIII was only identified in SCL stage. Interestingly, mammary glands from transgenic rabbits expressing human recombinant Coagulation factor VIII showed decreased levels of apoptosis and higher number of mitochondria [109], indicative of a pro-survival effect as well as higher metabolism exerted in SCL stage by this factor.

Protein disulfide-isomerase and alpha-enolase were also upregulated in SCL being the second involved in glicolysis and is thought to participate in growth control and tolerance to hypoxia [75]. Alpha-enolase was found overexpressed in breast cancer [110].

Calreticulin, Cathepsin D and Insulin-like growth factor-binding protein 5 were identified in PD and D but not in SCL. Insulin-like growth factor-binding protein 5 has been shown to inhibit the growth promoting effects of the IGF growth factor family on cell culture and promote apoptosis [111-112], Calreticulin is a Ca^{2+} - binding chaperone that promotes the folding of synthesized proteins [113] and is thought to affect Fibronectin production and, hence, cell adhesion [114]. It was pointed as a possible prostate and bladder cancer biomarker [115-116].

Therefore, it seems that when cells enter the differentiation program, a reduction of Fibronectin levels may be related to higher levels of Calreticulin.

Cathepsin D is involved in enzymatic degradation in lysosomes and also activates some proteins precursors in specialized cells [117-118]. Using mass spectrometry this protein was found to be more abundant in aggressive tumours stimulating metastasis and promoting angiogenesis [75]. It also induces fibroblasts to produce extracellular matrix proteases [119-120]. Elevated Cathepsin D levels are correlated with higher incidence of metastasis and lower survival rates [121-122]. However, this lysosomal aspartic protease is present in mammary tissue and milk in various molecular forms and is secreted in an active form by Prl stimulation [123], is itself involved in cleavage of basally located Prl [124] and its activity is inhibited by Serpin B5 [125]. This is an interesting observation, since in the secretome of HC11 cell differentiation program Cathepsin D is not present in SCL stage, when Serpin B5 is found in higher levels, they are both present when cells begin to differentiate (PD) and only Cathepsin D is found in D stage, when the action of Prl would be needed.

Several proteins were found only in D stage: Annexin A5, Cadherin-1, Carboxypeptidase E, Connective tissue growth factor, EF-hand domain-containing protein D2, Prostate and testis expressed protein 3 and Renin receptor. Of these proteins, Cadherin-1 is involved in mechanisms regulating cell-cell adhesion, mobility and proliferation of epithelial cells and has a potent invasion suppressor role [126-127]. However, E-cadherin is a transmembrane protein which is prone to extracellular protease cleavage, a mechanism used by cancer cells to promote migration. Since this is not likely to occur in D stage, and E-cadherin is highly expressed in this stage [90], one possibility is that protein fragments resulted from apoptotic bodies present in the CM. Connective tissue growth factor mediates cell adhesion, is the most upregulated gene in HC11 D stage [90] and was recently identified as upregulated by Dex and highly expressed in pregnant and lactating mouse mammary glands [128]. Connective Tissue Growth Factor can function both as growth factor and, if associated to ECM, as cell adhesion molecule blocking anoikis [129]. EF-hand domain-containing protein D2 is capable of regulating apoptosis through BCL2L1 expression [130].

Besides the mentioned proteins, D condition presented an upregulation of Complement C3, Ceruloplasmin, Collagen alpha-1(II) chain and Neutrophil gelatinase-associated lipocalin. It has been proposed that Ceruloplasmin transports iron into malignant cells and it is upregulated in human breast cancer cell lines, but also in milk [131-132]. Although the

exact function of Neutrophil gelatinase-associated lipocalin is not yet established, some authors consider that it could be involved in protecting MMP-9 from degradation, enhancing its enzymatic activity and facilitating angiogenesis and tumor growth [133-134] or that it could be correlated with apoptosis and cell survival [135]. Further, Neutrophil gelatinase-associated lipocalin has been identified in bovine colostrum [136] and is also expressed in lumen of ducts close to breast carcinomas (a sign it is secreted) it is associated to poor prognosis [137] and may be a transporter and survival factor [135]. Interestingly, and similar to these findings, well- to moderately differentiated pancreas cancer (PaCa) cells express high levels of Neutrophil gelatinase-associated lipocalin but moderately to poorly differentiated PaCa cells express undetectable levels of this protein which may have a role as suppressor of invasion in PaCa cells [138].

Therefore, factors found in CM from D contribute to inhibition of invasion and survival.

2. Protein expression in undifferentiated mammary epithelial cells under estrogenic stimulation with or without EGFR activation.

Estrogen receptors are widely expressed in the mammary gland. ERα is majorly detected in ductal epithelial cells and its expression is reduced as cells differentiate during pregnancy and lactation. On the other hand, ERβ expression is maintained constant throughout the reproductive cycle [11]. The mammary response to estrogens is also variable throughout the reproductive cycle. Estrogens (through ERα), induce proliferation in puberty but in the pregnant mammary gland, E2 does not induce more proliferation or expression of E2 regulated genes such as PR [5, 139], indicating that in this reproductive stage, the mammary gland is not responsive to E2. On the other hand, ERβ has been shown to positively influence cell differentiation, especially during lactation [6]. Differential response to estrogens as well as the fact that even if cells express ERs, not all cells respond in the same way, may be due to activation of signalling pathways such as MAP kinase and PI3-kinase that influence the differentiation stage but also ER transcriptional activity.

Mammary stem and progenitor cells appear to be the most likely target of carcinogens, and long term estrogen exposure has a carcinogenic effect on the mammary epithelium. In this context, it could be of particular interest because the EGF/EGFR pathway is known to promote ERα transcriptional activity [12, 140] and proliferation of the mammary epithelium and is needed for maintenance of SCL phenotype. There is no information regarding the effect of EGF on endogenous ERβ. However, EGF can also stimulate transcription of overexpressed ERβ. Since ERβ was shown to regulate apoptosis, cell adhesion, migration and inhibit proliferation [6, 58] and is expressed throughout the whole reproductive cycle of the mammary gland, a detailed study of how EGF/EGFR signalling affects ERβ function is needed.

Therefore, this part of the study will compare the secretome of SCL cells (+EGF) and SCL cells as they begin the differentiation program (-EGF), when treated with the ERβ agonist DPN. For this purpose, we used a similar experimental approach as outlined in section 1. First, proteins form the CM were separated in a SDS-PAGE (10%) and secondly in a 2D-PAGE. Then bands/spots were cut, digested and analyzed by MALDI-TOF/MS/MS. Analysis was carried out only with proteins that are either secreted or can potentially be found in the extracellular space.

Fig. 17 shows the protein profile from SCL +EGF and -EGF cells simultaneously treated with DPN for 24h.

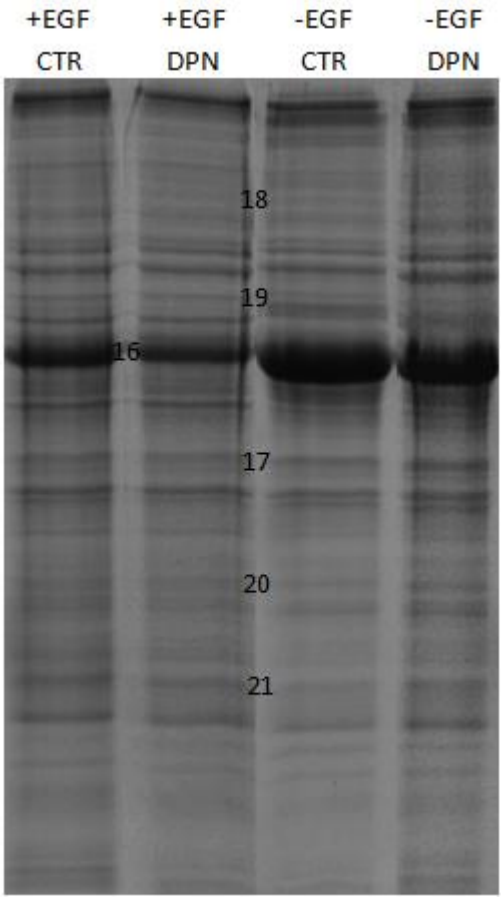


Figure 17 – SDS gel comparing protein profile between SCL +EGF (CTR and DPN) and -EGF (CTR and DPN). The numbers indicate some of the identified proteins which are reported in Table 1.

A 2D-PAGE separation was carried out with the 4 conditions mentioned above and the result is presented in Fig. 18. A larger number of spots were detected in the -EGF condition when compared to +EGF condition (panels C versus A, respectively). A similar observation was made when the same samples were separated by SDS-PAGE (annexe 2).

However, when density of all bands was calculated, we observed no difference which indicates that there was no error in protein quantification.

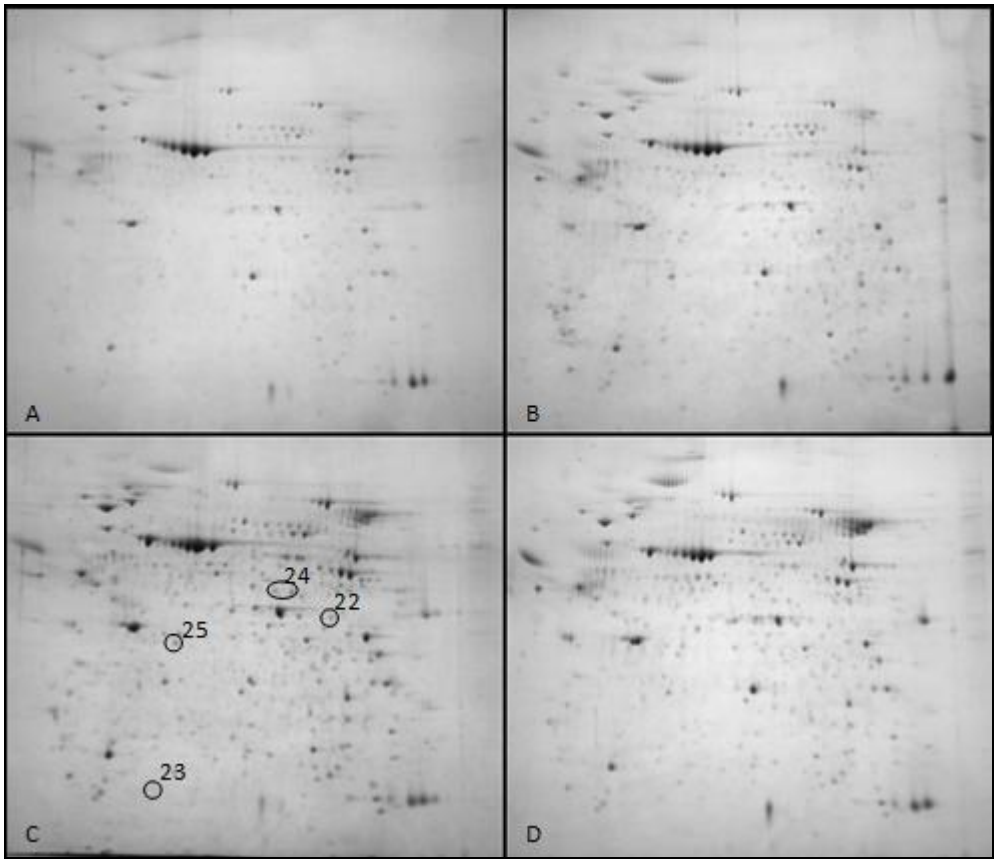


Figure 18 - 2D gel comparing protein profile between SCL cells A) +EGF CTR; B) +EGF DPN; C) –EGF CTR and D) –EGF DPN. Identifications are indicated by the numbers and are reported in Table 1.

The density values of the bands/spots were measured and the variation for each one was calculated and presented in Table 4.

A total of fifteen proteins were found regulated by DPN. Independently of EGF presence: DPN upregulated Ceruloplasmin, Extracellular superoxide dismutase [Cu-Zn], Protein disulfide-isomerase and Protein disulfide-isomerase A6 as well as SPARC; while Protein disulfide-isomerase A3 was downregulated by DPN. Fibronectin was downregulated by DPN only in cells + EGF and Basement membrane-specific heparan sulfate proteoglycan core protein, Calreticulin, Calsystenin, Collagen alpha-1 (VI) chain and Alpha-enolase were upregulated by DPN in cells -EGF.

Interestingly, Serpin B5 and Fascin were upregulated by DPN in cells +EGF, but downregulated in cells -EGF. On the other hand, Complement C3, Periostin, Galectin-3 and Neutrophil gelatinase-associated Lipocalin were downregulated by DPN in cells +EGF, but upregulated in cells –EGF.

Table 4 – Proteins up and downregulated in +EGF (CTR and DPN) and -EGF (CTR and DPN).

Protein name	Variation (mean±SD) *	Protein name	Variation (mean±SD) *
Up in +EGF DPN compared to +EGF CTR			
Ceruloplasmin	1,52	Protein disulfide-isomerase	1,49±0,13
Cystatin-C	1,13	Protein disulfide-isomerase A6	1,47
SPARC	1,51	Serpin B5	1,60
Fascin	1,18		
Down in +EGF DPN compared to -EGF CTR			
Complement C3	0,13	Neutrophil gelatinase-associated lipocalin	0,11
Fibronectin	0,70±0,07	Periostin	0,82
Galectin-3	0,56	Protein disulfide-isomerase A3	0,74
Up in -EGF DPN compared to -EGF CTR			
Basement membrane-specific heparan sulfate proteoglycan core protein	1,68	Extracellular superoxide dismutase [Cu-Zn]	1,24±0,13
Calreticulin	1,56	Galectin-3	1,16
Calsyntenin-1	1,41±0,06	Neutrophil gelatinase-associated lipocalin	1,53
Ceruloplasmin	1,75	Periostin	1,65
Collagen alpha-1(VI) chain	1,14±0,01	Protein disulfide-isomerase	1,29
Complement C3	1,47	Protein disulfide-isomerase A6	1,12
Alpha-enolase	1,21±0,05	SPARC	1,74
Down in -EGF DPN compared to -EGF CTR			
6-phosphogluconate dehydrogenase, decarboxylating	0,78	Protein disulfide-isomerase A3	0,83
Fascin	0,75	Serpin B5	0,65

***SD only if more than one measure was available.**

With the information from table 4, pie charts of biological processes were built (Fig. 19). The most represented processes of proteins upregulated in +EGF DPN were metabolic process (83,3%) and cell communication (33,3%). Proteins downregulated in +EGF DPN besides participating in metabolic processes (50%) and cell communication (66,7%) are also involved in immune system response (66,7%) and cell adhesion (50%). Taking into account the molecular function of the proteins over or under regulated in +EGF DPN, the 3 most represented functions were binding (61,5% in up and downregulated), catalytic activity (50% in up and 33,3% in downregulated) and enzyme regulator activity (50% in up and 16,7% in downregulated). When we separated the proteins according to their classes, we found that in +EGF +DPN condition, enzyme modulator (33,3%) and isomerase (33,3%) are the major classes of proteins in high levels, while proteins found in low levels in +EGF DPN are mostly signalling molecules (66,7%).

Proteins upregulated in -EGF condition treated with DPN participate in several processes such as metabolic processes (53,0% of the proteins), cell communication (53,0%) and cell adhesion (38,5%). Relatively to the molecular functions of those proteins the most represented are binding (61,5%) catalytic activity (38,5%) and enzyme regulator activity (15,4%). The proteins identified in high levels in –EGF DPN can be aggregated into classes being the most represented signalling molecules (38,5%) and isomerase (15,4%).

Relatively to the proteins found in low levels in –EGF DPN we cannot make this type of discussion due to the low number of identifications.

One interesting finding was the clear effect of DPN in the induction of developmental processes (Basement membrane-specific heparan sulfate proteoglycan core protein, Ceruloplasmin and Periostin) and apoptosis processes (Galectin-3) in SCL cells beginning to differentiate (-EGF). On the contrary, SCL cells +EGF with DPN treatment have shown reduced developmental process (only represented by ceruloplasmin) and absence of apoptosis. In addition, cell communication was more represented in –EGF+DPN (Complement C3, Basement membrane-specific heparan sulfate proteoglycan core protein, Ceruloplasmin, SPARC, Periostin and Neutrophil gelatinase-associated lipocalin) than in +EGF+DPN (only SPARC and Ceruloplasmin). When looking at the regulation of these proteins it is evident that those representing the above mentioned processes in -EGF+DPN are those proteins upregulated in this condition but downregulated in +EGF+DPN. Therefore, these results indicate that in HC11 cells, activation of ERβ will influence cell signalling differently depending on the differentiation stage of the cell. Further, it is evident that the tumor supressor role of ERβ occurs in cells on their way to differentiation and not in SCL cells, an observation suported by our unpublished data (Cotrim, Z in preparation).

It was also the goal of this work to study the pathways in each the proteins found however, all the pathways identified were only represented by one protein.

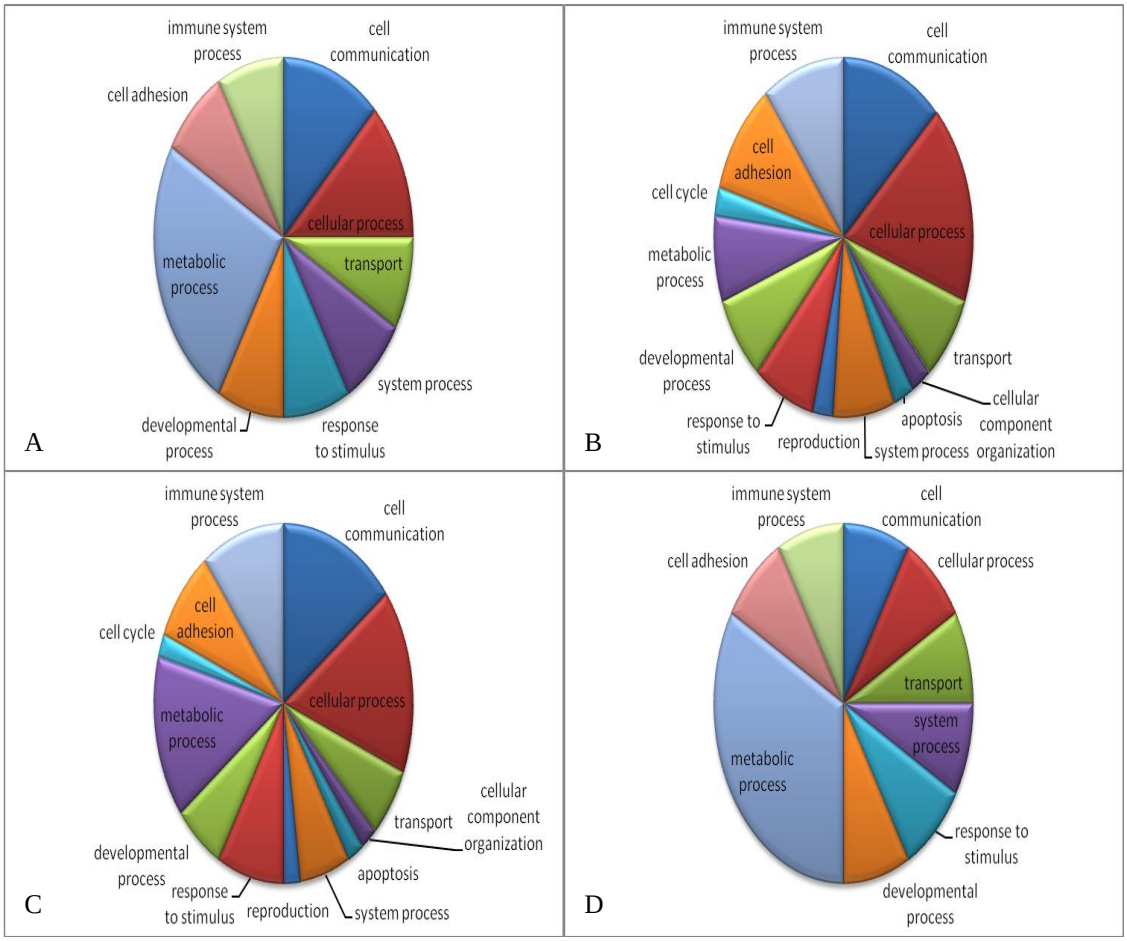


Figure 19 - Pie charts with Biological processe of identified proteins A) Overexpressed in +EGF +DPN B) Underexpressed in +EGF +DPN C) Overexpressed in -EGF DPN and D) Underexpressed in -EGF DPN.

Among proteins equally upregulated by DPN with or without EGF, Protein disulfide-isomerase was shown to bind intracellular E2 and through its release to modulate ER activity. In addition, Protein disulfide-isomerase can physically interact with ER acting somewhat as a chaperone [141-142]. Therefore, it is possible that increased ERβ signalling requires higher levels of Protein disulfide-isomerase. In adition, 2D electrophoresis of bovine smooth muscle cells identified Pdia as downregulated by E2 [143] which indicates that ERs regulate Protein disulfide-isomerase levels.

SOD3 has a role in protecting the cells from radical oxygen species (ROSs). E2 has a protective role on the endothelium and the beneficial effects of E2 in the circulatory system have been known for years [46]. In addition, ovariectomy of rats (depletion of endogenous E2) lead to SOD3 downregulation and increased oxidative stress, this effect can be reversed by administration of E2. Our results indicate that in the mammary epithelium, ERβ can protect from oxidative stress by induction of SOD3.

The extracellular matrix protein SPARC binds calcium, a co-factor needed for cadherin and integrin activation. Hence, SPARC regulates cell adhesion. Altered expression has been reported in breast cancer but the role of this potential tumour biomarker remains controversial. However, both Fibronectin and SPARC are present in an extracellular matrix cluster associated to good prognosis and disease free survival [144]. These results, indicate that in a non cancerous environment, activation of ERβ increases SPARC levels and may be one mechanism by which ERβ regulates cell adhesion.

Fibronectin expression is also associated to good prognosis in breast cancer [144]. The fact that it is only regulated by DPN (down) in SCL cells with EGF supports our finding that the tumour suppressor role of ERβ is lost when cells are grown with EGF (Cotrim C, in preparation).

Periostin, Galectin-3 and Neutrophil gelatinase-associated lipocalin were downregulated by DPN in cells +EGF and upregulated in cells –EGF. Periostin binds to heparin and seems to play a role in cell adhesion. It may also be involved in extracellular matrix mineralization [145-146]. Besides, Periostin mRNA was shown to be upregulated by ERβ in primary human periodontal ligament cells [147]. Galectin-3 is a lectin that is involved in normal and pathological processes. It may exhibit anti- or pro-apoptotic activity depending on the nature of the stimulus. [148]. Altered expression of galectin-3 has been described in many cancers, and studies of cancer cell lines have implicated this lectin in various aspects of the tumorigenic cascade. However, galectin-3 is strongly up-regulated during mammary gland involution [149] and may induce integrin beta1 endocytosis [150].

In summary, ERβ regulates expression of proteins involved in regulation of adhesion and apoptosis. However, since different regulation was observed in SCL (+EGF) cells compared to cells entering differentiation (-EGF), the cellular context seems to drive ER activity.

3. **Protein expression in differentiated mammary epithelial cells with or without ERβ activation.**

Hormone treatment in lactating mammary glands (i.e. birth control pills in humans, E2 treatment in cows) induces regression and reduced milk synthesis. This occurs through blockade of JAK/STAT5 signalling [151] and induction of cell death. Since mammary gland involution (regression) depends on detachment of alveolar cells from the ECM (anoikis) and during lactation ERβ regulates expression of several adhesion proteins, we aimed to study if ERβ influences secretion of factors which can modulate adhesion. To accomplish this goal differentiated cells were treated with DPN and the CM proteins separated in a SDS gel (10%) (Fig. 20).

Figure 20- SDS gel comparing protein profile of differentiated cells with or without DPN.
Identifications report to table 1.

The density variations of all the secreted proteins found in differentiated cells were calculated and regulated proteins are shown in Table 6.

DPN upregulated key proteins involved in mammary epithelial cell differentiation were: Connective tissue growth factor and Insulin-like growth factor-binding protein 5 as well as Cathepsin D, Collage alpha 1 (II) chain and Fibronectin, while it downregulated Neutrophil gelatinase-associated lipocalin. All these proteins clearly cluster within the extracellular matrix proteins which indicates that in the lactating epithelium, ERβ regulation of cell adhesion molecules like integrin alpha 6 and 2 as well as E-cadherin [30] may be a result of a complex regulation between, binding of these proteins to their substrate as well as protein stability.

Table 5 – Proteins upregulated on differenciated cells with different treatments.

Protein name	Variation (Mean±SD)*	Protein name	Variation (Mean±SD)*
Upregulated by DPN			
Cathepsin D	1,16	Fibronectin	1,33±0,17
Collagen alpha-1(II) chain	1,24±0,13	Insulin-like growth factor-binding protein 5	1,31
Connective tissue growth factor	1,35		
Downregulated by DPN			
Neutrophil gelatinase-associated lipocalin	0,46		

***SD only if more than one measure was available.**

Finally, Complement C3, Periostin and SPARC, three proteins upregulated by DPN in cells entering differentiation (-EGF) (Table 4), were not regulated in D cells. Therefore, and even thought D cells are obtained in –EGF medium, these results clearly show that ERβ activity is regulated by the differentiation stage of the cell.

IV – Conclusion

With the goal to identify secreted proteins which can potentially regulate mammary epithelial cells as they transit the differentiation program, the secretome of SCL, PD and D stages was analyzed using a gel based MS approach.

This analysis showed in SCL stage an increase of proteins such as SPARC, Coagulation factor VIII, Fibronectin, Thrombospondin-1 and Serpin B5.

SPARC is known to promote angiogenesis and invasion in breast cancer [105], being highly expressed in more aggressive breast cancer cell lines [107]. Besides, Coagulation factor VIII was only identified in this condition and was showed to inhibit apoptosis and increase metabolism [109]. These facts suggest that in SCL condition, secreted proteins have pro-angiogenic/invasion as well as pro-survival roles. These mechanisms are important for the maintenance of the proliferative state of this condition. However, Fibronectin, Thrombospondin-1 and Serpin B5 were also found in higher levels. These proteins are thought to inhibit angiogenesis, invasion and tumour growth [98, 100]. In fact, both Thrombospondin-1 and Serpin B5 participate in p53 pathways, being upregulated by the tumour suppressor p53 and in turn, inhibit migration and angiogenesis.

Thus, it seems that in SCL cells factors that enhance angiogenesis, invasion and inhibit apoptosis are co-secreted with factors that have the opposite effects. Is this crosstalk of signals that could explain the regulation of SCL state, that implicates a need to proliferate and migrate and at the same time the necessity to control these processes.

When entering differentiation (-EGF), we encountered several proteins upregulated such as Insulin-like growth factor-binding protein 5, Calreticulin, Cathepsin D, Connective tissue growth factor and Neutrophil gelatinase-associated lipocalin.

Insulin-like growth factor-binding protein 5 has been showed to inhibit the growth promoting effects of the IGF growth factor family on cell culture and promote apoptosis [111-112]. Calreticulin is thought to affect Fibronectin production and, hence, cell adhesion [114]. Although Cathepsin D was found to be more abundant in aggressive tumours stimulating metastasis and promoting angiogenesis it is also induced by Prl stimulation [123] and is itself involved in cleavage of Prl [124]; thus contributing to a positive feedback loop during lactation. Cathepsin D activity is inhibited by Serpin B5 [125]. This could explain why Cathepsin D is not present in SCL stage, when Serpin B5 is found in higher levels.

Connective tissue growth factor was detected in high levels in D stage both at the mRNA [90] and extracellular protein levels. It has a role as cell adhesion molecule blocking anoikis and therefore it is a pro/survival factor [129].

The exact function of Neutrophil gelatinase-associated lipocalin is not yet established, some authors consider that it could facilitate angiogenesis and tumour growth [133-134] while others consider that it could be correlated with apoptosis and cell survival [135]. However, and in agreement with these findings, Neutrophil gelatinase-associated lipocalin is only detected in differentiated PaCa cells but absent in poorly differentiated PaCa cells where it may have a role as suppressor of invasion [138].

Therefore, data obtained as cells differentiate point to pathways that inhibit invasion and increase survival, two fundamental processes in differentiation.

In the attempt to study how ERβ function is affected in different cellular states (i.e. influence with EGF/EGFR signalling or lactogenic hormone treatment) it was firstly compared the secretome of SCL cells (+EGF) and SCL cells as they begin the differentiation program (-EGF), when treated with the ERβ agonist DPN. Second, we compared the secretome of D cells treated with or without DPN.

These results showed that in cells beginning to differentiate (-EGF) developmental processes (Basement membrane-specific heparan sulfate proteoglycan core protein, Ceruloplasmin and Periostin) and apoptosis processes (Galectin-3) are upregulated by ERβ when compared to SCL (+EGF). Therefore, it seems also clear that the tumour suppressor role of ERβ is only triggered on cells beginning to differentiate (-EGF). This indicates that in HC11 cells, ERβ effects are influenced by EGF/EGFR signalling.

Independently of EGF prescence, ERβ activation leads to upregulation of SPARC which promotes migration and, in SCL +EGF, Fibronectin which is known to inhibit migration was downregulated. This indicates that in cells treated with DPN, EGF favours migratory signals which may be counteracted by Fibronectin in the absence of EGF; thus evidencing the tumour supressor role of ERβ which seems once again to be triggered when cells begin to differentiate (-EGF).

In functionally differentiated cells, ERβ activation lead to upregulation of Connective tissue growth factor, Insulin-like growth factor-binding protein 5, Cathepsin D, Collage alpha 1 (II) chain and Fibronectin. All these proteins are known to be present in differentiated cells and influence cell adhesion and/or apoptosis. Therefore, these results indicate that ERβ effects are dependent on the cell differentiation stage and that the clinically required tumour suppressor role of ERβ is only evidenced when SCL cells enter the differentiation program or are already fully differentiated.

V – Bibliographic References

1. Couse, J.F. and K.S. Korach, *Estrogen receptor null mice: what have we learned and where will they lead us?* Endocr Rev, 1999. **20**(3): p. 358-417.
2. Brisken, C. and S. Duss, *Stem cells and the stem cell niche in the breast: an integrated hormonal and developmental perspective.* Stem Cell Rev, 2007. **3**(2): p. 147-56.
3. Hens, J.R. and J.J. Wysolmerski, *Key stages of mammary gland development: molecular mechanisms involved in the formation of the embryonic mammary gland.* Breast Cancer Res, 2005. **7**(5): p. 220-4.
4. Foley, J., et al., *Parathyroid hormone-related protein maintains mammary epithelial fate and triggers nipple skin differentiation during embryonic breast development.* Development, 2001. **128**(4): p. 513-25.
5. Sternlicht, M.D., et al., *Hormonal and local control of mammary branching morphogenesis.* Differentiation, 2006. **74**(7): p. 365-81.
6. Brisken, C. and R.D. Rajaram, *Alveolar and lactogenic differentiation.* J Mammary Gland Biol Neoplasia, 2006. **11**(3-4): p. 239-48.
7. Giannelli, G., et al., *Induction of cell migration by matrix metalloprotease-2 cleavage of laminin-5.* Science, 1997. **277**(5323): p. 225-8.
8. Mallepell, S., et al., *Paracrine signaling through the epithelial estrogen receptor alpha is required for proliferation and morphogenesis in the mammary gland.* Proc Natl Acad Sci U S A, 2006. **103**(7): p. 2196-201.
9. Kleinberg, D.L., M. Feldman, and W. Ruan, *IGF-I: an essential factor in terminal end bud formation and ductal morphogenesis.* J Mammary Gland Biol Neoplasia, 2000. **5**(1): p. 7-17.
10. Brisken, C. and B. O'Malley, *Hormone action in the mammary gland.* Cold Spring Harb Perspect Biol, 2010. **2**(12): p. a003178.
11. Weihua, Z., et al., *Update on estrogen signaling.* FEBS Lett, 2003. **546**(1): p. 17-24.
12. Kenney, N.J., et al., *Effect of exogenous epidermal-like growth factors on mammary gland development and differentiation in the estrogen receptor-alpha knockout (ERKO) mouse.* Breast Cancer Res Treat, 2003. **79**(2): p. 161-73.
13. Luetteke, N.C., et al., *Targeted inactivation of the EGF and amphiregulin genes reveals distinct roles for EGF receptor ligands in mouse mammary gland development.* Development, 1999. **126**(12): p. 2739-50.
14. Sternlicht, M.D., et al., *Mammary ductal morphogenesis requires paracrine activation of stromal EGFR via ADAM17-dependent shedding of epithelial amphiregulin.* Development, 2005. **132**(17): p. 3923-33.
15. Andrechek, E.R., D. White, and W.J. Muller, *Targeted disruption of ErbB2/Neu in the mammary epithelium results in impaired ductal outgrowth.* Oncogene, 2005. **24**(5): p. 932-7.
16. Daniel, C.W., S. Robinson, and G.B. Silberstein, *The role of TGF-beta in patterning and growth of the mammary ductal tree.* J Mammary Gland Biol Neoplasia, 1996. **1**(4): p. 331-41.
17. Olayioye, M.A., et al., *The ErbB signaling network: receptor heterodimerization in development and cancer.* EMBO J, 2000. **19**(13): p. 3159-67.
18. Fata, J.E., Z. Werb, and M.J. Bissell, *Regulation of mammary gland branching morphogenesis by the extracellular matrix and its remodeling enzymes.* Breast Cancer Res, 2004. **6**(1): p. 1-11.
19. Stupack, D.G., et al., *Apoptosis of adherent cells by recruitment of caspase-8 to unligated integrins.* J Cell Biol, 2001. **155**(3): p. 459-70.
20. Sinha, S., et al., *Cellular and extracellular biology of the latent transforming growth factor-beta binding proteins.* Matrix Biol, 1998. **17**(8-9): p. 529-45.
21. Ito, K., et al., *Calcium influx triggers the sequential proteolysis of extracellular and cytoplasmic domains of E-cadherin, leading to loss of beta-catenin from cell-cell contacts.* Oncogene, 1999. **18**(50): p. 7080-90.
22. Noe, V., et al., *Release of an invasion promoter E-cadherin fragment by matrilysin and stromelysin-1.* J Cell Sci, 2001. **114**(Pt 1): p. 111-118.
23. Soyal, S., et al., *Progesterone's role in mammary gland development and tumorigenesis as disclosed by experimental mouse genetics.* Breast Cancer Res, 2002. **4**(5): p. 191-6.
24. Brisken, C., et al., *A paracrine role for the epithelial progesterone receptor in mammary gland development.* Proc Natl Acad Sci U S A, 1998. **95**(9): p. 5076-81.
25. Tanos, T. and C. Brisken, *High hopes for RANKL: will the mouse model live up to its promise?* Breast Cancer Res, 2011. **13**(1): p. 302.
26. Coleman, S., G.B. Silberstein, and C.W. Daniel, *Ductal morphogenesis in the mouse mammary gland: evidence supporting a role for epidermal growth factor.* Dev Biol, 1988. **127**(2): p. 304-15.

27. Teglund, S., et al., *Stat5a and Stat5b proteins have essential and nonessential, or redundant, roles in cytokine responses.* Cell, 1998. **93**(5): p. 841-50.
28. Watson, C.J. and T.G. Burdon, *Prolactin signal transduction mechanisms in the mammary gland: the role of the Jak/Stat pathway.* Rev Reprod, 1996. **1**(1): p. 1-5.
29. Harris, J., et al., *Socs2 and elf5 mediate prolactin-induced mammary gland development.* Mol Endocrinol, 2006. **20**(5): p. 1177-87.
30. Forster, C., et al., *Involvement of estrogen receptor beta in terminal differentiation of mammary gland epithelium.* Proc Natl Acad Sci U S A, 2002. **99**(24): p. 15578-83.
31. Muschler, J. and C.H. Streuli, *Cell-matrix interactions in mammary gland development and breast cancer.* Cold Spring Harb Perspect Biol, 2010. **2**(10): p. a003202.
32. Foty, R.A. and M.S. Steinberg, *Cadherin-mediated cell-cell adhesion and tissue segregation in relation to malignancy.* Int J Dev Biol, 2004. **48**(5-6): p. 397-409.
33. Bryant, D.M. and J.L. Stow, *The ins and outs of E-cadherin trafficking.* Trends Cell Biol, 2004. **14**(8): p. 427-34.
34. Stahl, S., S. Weitzman, and J.C. Jones, *The role of laminin-5 and its receptors in mammary epithelial cell branching morphogenesis.* J Cell Sci, 1997. **110 (Pt 1)**: p. 55-63.
35. Berdichevsky, F., et al., *Branching morphogenesis of human mammary epithelial cells in collagen gels.* J Cell Sci, 1994. **107 (Pt 12)**: p. 3557-68.
36. Alford, D., et al., *Integrin-matrix interactions affect the form of the structures developing from human mammary epithelial cells in collagen or fibrin gels.* J Cell Sci, 1998. **111 (Pt 4)**: p. 521-32.
37. Hathaway, H.J. and B.D. Shur, *Mammary gland morphogenesis is inhibited in transgenic mice that overexpress cell surface beta1,4-galactosyltransferase.* Development, 1996. **122**(9): p. 2859-72.
38. Muschler, J., et al., *A role for dystroglycan in epithelial polarization: loss of function in breast tumor cells.* Cancer Res, 2002. **62**(23): p. 7102-9.
39. Helguero, L.A., et al., *Different roles of estrogen receptors alpha and beta in the regulation of E-cadherin protein levels in a mouse mammary epithelial cell line.* Cancer Res, 2008. **68**(21): p. 8695-704.
40. Lindberg, K., et al., *Expression of estrogen receptor beta increases integrin alpha1 and integrin beta1 levels and enhances adhesion of breast cancer cells.* Journal of Cellular Physiology, 2010. **222**(1): p. 156-167.
41. Mangelsdorf, D.J., et al., *The nuclear receptor superfamily: the second decade.* Cell, 1995. **83**(6): p. 835-9.
42. Nilsson, S. and K.F. Koehler, *Oestrogen receptors and selective oestrogen receptor modulators: molecular and cellular pharmacology.* Basic Clin Pharmacol Toxicol, 2005. **96**(1): p. 15-25.
43. Koehler, K.F., et al., *Reflections on the discovery and significance of estrogen receptor beta.* Endocr Rev, 2005. **26**(3): p. 465-78.
44. Meyers, M.J., et al., *Estrogen receptor-beta potency-selective ligands: structure-activity relationship studies of diarylpropionitriles and their acetylene and polar analogues.* J Med Chem, 2001. **44**(24): p. 4230-51.
45. Stauffer, S.R., et al., *Pyrazole ligands: structure-affinity/activity relationships and estrogen receptor-alpha-selective agonists.* J Med Chem, 2000. **43**(26): p. 4934-47.
46. Nilsson, S. and J.A. Gustafsson, *Estrogen receptors: therapies targeted to receptor subtypes.* Clin Pharmacol Ther, 2011. **89**(1): p. 44-55.
47. Heldring, N., et al., *Estrogen receptors: how do they signal and what are their targets.* Physiol Rev, 2007. **87**(3): p. 905-31.
48. Kamangar, F., G.M. Dores, and W.F. Anderson, *Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world.* J Clin Oncol, 2006. **24**(14): p. 2137-50.
49. Huang, T.H. and M. Esteller, *Chromatin remodeling in mammary gland differentiation and breast tumorigenesis.* Cold Spring Harb Perspect Biol, 2010. **2**(9): p. a004515.
50. Polyak, K., *Breast cancer: origins and evolution.* J Clin Invest, 2007. **117**(11): p. 3155-63.
51. Visvader, J.E., *Cells of origin in cancer.* Nature, 2011. **469**(7330): p. 314-22.
52. Sorlie, T., et al., *Distinct molecular mechanisms underlying clinically relevant subtypes of breast cancer: gene expression analyses across three different platforms.* BMC Genomics, 2006. **7**: p. 127.
53. Yu, K., et al., *Conservation of breast cancer molecular subtypes and transcriptional patterns of tumor progression across distinct ethnic populations.* Clin Cancer Res, 2004. **10**(16): p. 5508-17.
54. Whitfield, M.L., et al., *Identification of genes periodically expressed in the human cell cycle and their expression in tumors.* Mol Biol Cell, 2002. **13**(6): p. 1977-2000.

55. Musgrove, E.A. and R.L. Sutherland, *Biological determinants of endocrine resistance in breast cancer.* Nat Rev Cancer, 2009. **9**(9): p. 631-43.
56. Feigelson, H.S. and B.E. Henderson, *Estrogens and breast cancer.* Carcinogenesis, 1996. **17**(11): p. 2279-84.
57. Miller, W.R. and R.M. Sharpe, *Environmental oestrogens and human reproductive cancers.* Endocr Relat Cancer, 1998. **5**(2): p. 69-96.
58. Helguero, L.A., et al., *Estrogen receptors alfa (ERalpha) and beta (ERbeta) differentially regulate proliferation and apoptosis of the normal murine mammary epithelial cell line HC11.* Oncogene, 2005. **24**(44): p. 6605-16.
59. Speirs, V. and A.M. Shaaban, *Role of ERbeta in Clinical Breast Cancer.* Cancer Treat Res, 2009. **147**: p. 1-20.
60. Yarden, Y., *The EGFR family and its ligands in human cancer. signalling mechanisms and therapeutic opportunities.* Eur J Cancer, 2001. **37 Suppl 4**: p. S3-8.
61. Yarden, Y. and M.X. Sliwkowski, *Untangling the ErbB signalling network.* Nat Rev Mol Cell Biol, 2001. **2**(2): p. 127-37.
62. Normanno, N., et al., *Epidermal growth factor receptor (EGFR) signaling in cancer.* Gene, 2006. **366**(1): p. 2-16.
63. Herbst, R.S., *Review of epidermal growth factor receptor biology.* Int J Radiat Oncol Biol Phys, 2004. **59**(2 Suppl): p. 21-6.
64. Singh, A.B. and R.C. Harris, *Autocrine, paracrine and juxtacrine signaling by EGFR ligands.* Cell Signal, 2005. **17**(10): p. 1183-93.
65. Miettinen, P.J., et al., *Epithelial immaturity and multiorgan failure in mice lacking epidermal growth factor receptor.* Nature, 1995. **376**(6538): p. 337-41.
66. Sibilia, M. and E.F. Wagner, *Strain-dependent epithelial defects in mice lacking the EGF receptor.* Science, 1995. **269**(5221): p. 234-8.
67. Sibilia, M., et al., *A strain-independent postnatal neurodegeneration in mice lacking the EGF receptor.* EMBO J, 1998. **17**(3): p. 719-31.
68. Chan, T.O., S.E. Rittenhouse, and P.N. Tsichlis, *AKT/PKB and other D3 phosphoinositide-regulated kinases: kinase activation by phosphoinositide-dependent phosphorylation.* Annu Rev Biochem, 1999. **68**: p. 965-1014.
69. Engebraaten, O., et al., *Effects of EGF, bFGF, NGF and PDGF(bb) on cell proliferative, migratory and invasive capacities of human brain-tumour biopsies in vitro.* Int J Cancer, 1993. **53**(2): p. 209-14.
70. Slamon, D.J., et al., *Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene.* Science, 1987. **235**(4785): p. 177-82.
71. Levin, E.R., *Bidirectional signaling between the estrogen receptor and the epidermal growth factor receptor.* Mol Endocrinol, 2003. **17**(3): p. 309-17.
72. Xue, H., B. Lu, and M. Lai, *The cancer secretome: a reservoir of biomarkers.* J Transl Med, 2008. **6**: p. 52.
73. Hondermarck, H., *Breast cancer: when proteomics challenges biological complexity.* Mol Cell Proteomics, 2003. **2**(5): p. 281-91.
74. Bertucci, F. and A. Goncalves, *Clinical proteomics and breast cancer: strategies for diagnostic and therapeutic biomarker discovery.* Future Oncol, 2008. **4**(2): p. 271-87.
75. Mbeunkui, F., et al., *Identification of differentially secreted biomarkers using LC-MS/MS in isogenic cell lines representing a progression of breast cancer.* J Proteome Res, 2007. **6**(8): p. 2993-3002.
76. Goncalves, A. and F. Bertucci, *Clinical application of proteomics in breast cancer: state of the art and perspectives.* Med Princ Pract, 2011. **20**(1): p. 4-18.
77. El Yazidi-Belkoura, I., et al., *Proteomics of breast cancer: outcomes and prospects.* Technol Cancer Res Treat, 2002. **1**(4): p. 287-96.
78. Gut, I.G., *DNA analysis by MALDI-TOF mass spectrometry.* Hum Mutat, 2004. **23**(5): p. 437-41.
79. Choi, M.Y., et al., *Mass spectrometry based proteomic analysis of human stem cells: a brief review.* Exp Mol Med, 2007. **39**(6): p. 690-5.
80. Cohen, A.M., et al., *Quantification of Greenland halibut serum vitellogenin: a trip from the deep sea to the mass spectrometer.* Rapid Commun Mass Spectrom, 2009. **23**(7): p. 1049-60.
81. Parker, C.E., et al., *Mass-spectrometry-based clinical proteomics--a review and prospective.* Analyst, 2010. **135**(8): p. 1830-8.

82. Brown, R.S. and J.J. Lennon, *Mass resolution improvement by incorporation of pulsed ion extraction in a matrix-assisted laser desorption/ionization linear time-of-flight mass spectrometer.* Anal Chem, 1995. **67**(13): p. 1998-2003.
83. Kaufmann, R., B. Spengler, and F. Lutzenkirchen, *Mass spectrometric sequencing of linear peptides by product-ion analysis in a reflectron time-of-flight mass spectrometer using matrix-assisted laser desorption ionization.* Rapid Commun Mass Spectrom, 1993. **7**(10): p. 902-10.
84. Merchant, M. and S.R. Weinberger, *Recent advancements in surface-enhanced laser desorption/ionization-time of flight-mass spectrometry.* Electrophoresis, 2000. **21**(6): p. 1164-77.
85. Chernushevich, I.V., A.V. Loboda, and B.A. Thomson, *An introduction to quadrupole-time-of-flight mass spectrometry.* J Mass Spectrom, 2001. **36**(8): p. 849-65.
86. Shukla, A.K. and J.H. Futrell, *Tandem mass spectrometry: dissociation of ions by collisional activation.* J Mass Spectrom, 2000. **35**(9): p. 1069-90.
87. Vestal, M.L. and J.M. Campbell, *Tandem time-of-flight mass spectrometry.* Methods Enzymol, 2005. **402**: p. 79-108.
88. Ball, R.K., et al., *Prolactin regulation of beta-casein gene expression and of a cytosolic 120-kd protein in a cloned mouse mammary epithelial cell line.* EMBO J, 1988. **7**(7): p. 2089-95.
89. Humphreys, R.C. and J.M. Rosen, *Stably transfected HC11 cells provide an in vitro and in vivo model system for studying Wnt gene function.* Cell Growth Differ, 1997. **8**(8): p. 839-49.
90. Williams, C., et al., *Gene expression in murine mammary epithelial stem cell-like cells shows similarities to human breast cancer gene expression.* Breast Cancer Res, 2009. **11**(3): p. R26.
91. Visvader, J.E. and G.J. Lindeman, *Stem cells and cancer - the promise and puzzles.* Mol Oncol, 2010. **4**(5): p. 369-72.
92. Al-Hajj, M., et al., *Prospective identification of tumorigenic breast cancer cells.* Proc Natl Acad Sci U S A, 2003. **100**(7): p. 3983-8.
93. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4.* Nature, 1970. **227**(5259): p. 680-5.
94. James, P.L., et al., *A highly conserved insulin-like growth factor-binding protein (IGFBP-5) is expressed during myoblast differentiation.* J Biol Chem, 1993. **268**(30): p. 22305-12.
95. Puchkova, L.V., et al., *Comparative analysis of the molecular heterogeneity of ceruloplasmin from human blood and breast milk.* Biochemistry (Mosc), 1997. **62**(8): p. 928-30.
96. Kassim, O.O., et al., *Immunoprotective factors in breast milk and sera of mother-infant pairs.* Trop Geogr Med, 1986. **38**(4): p. 362-6.
97. Klinowska, T.C., et al., *Epithelial development and differentiation in the mammary gland is not dependent on alpha 3 or alpha 6 integrin subunits.* Dev Biol, 2001. **233**(2): p. 449-67.
98. Martin-Manso, G., et al., *sFRP-1 binds via its netrin-related motif to the N-module of thrombospondin-1 and blocks thrombospondin-1 stimulation of MDA-MB-231 breast carcinoma cell adhesion and migration.* Arch Biochem Biophys, 2011. **509**(2): p. 147-56.
99. Grillo, J., et al., *Regulation of the Angiogenesis Inhibitor Thrombospondin-1 by the Breast Cancer Susceptibility Gene-1 (BRCA1).* Breast J, 2011.
100. Zhang, M., et al., *Maspin plays an important role in mammary gland development.* Dev Biol, 1999. **215**(2): p. 278-287.
101. Rodriguez-Manzaneque, J.C., et al., *Thrombospondin-1 suppresses spontaneous tumor growth and inhibits activation of matrix metalloproteinase-9 and mobilization of vascular endothelial growth factor.* Proc Natl Acad Sci U S A, 2001. **98**(22): p. 12485-90.
102. Jimenez, B., et al., *Signals leading to apoptosis-dependent inhibition of neovascularization by thrombospondin-1.* Nat Med, 2000. **6**(1): p. 41-8.
103. Francki, A., et al., *SPARC regulates cell cycle progression in mesangial cells via its inhibition of IGF-dependent signaling.* J Cell Biochem, 2003. **88**(4): p. 802-811.
104. Schiemann, B.J., J.R. Neil, and W.P. Schiemann, *SPARC inhibits epithelial cell proliferation in part through stimulation of the transforming growth factor-beta-signaling system.* Mol Biol Cell, 2003. **14**(10): p. 3977-3988.
105. Puolakkainen, P.A., et al., *Enhanced growth of pancreatic tumors in SPARC-null mice is associated with decreased deposition of extracellular matrix and reduced tumor cell apoptosis.* Mol Cancer Res, 2004. **2**(4): p. 215-224.
106. Jacob, K., et al., *Osteonectin promotes prostate cancer cell migration and invasion: a possible mechanism for metastasis to bone.* Cancer Res, 1999. **59**(17): p. 4453-7.

107. Watkins, G., et al., *Increased levels of SPARC (osteonectin) in human breast cancer tissues and its association with clinical outcomes.* Prostaglandins Leukot Essent Fatty Acids, 2005. **72**(4): p. 267-72.
108. Corbett, S.A., et al., *Covalent cross-linking of fibronectin to fibrin is required for maximal cell adhesion to a fibronectin-fibrin matrix.* J Biol Chem, 1997. **272**(40): p. 24999-5005.
109. Chrenek, P., et al., *Characteristics of Rabbit Transgenic Mammary Gland Expressing Recombinant Human Factor VIII.* Anatomia, Histologia, Embryologia, 2009. **38**(1): p. 85-88.
110. Espana, L., et al., *Bcl-x(L)-mediated changes in metabolic pathways of breast cancer cells: from survival in the blood stream to organ-specific metastasis.* Am J Pathol, 2005. **167**(4): p. 1125-37.
111. Tonner, E., et al., *Insulin-like growth factor binding protein-5 (IGFBP-5) induces premature cell death in the mammary glands of transgenic mice.* Development, 2002. **129**(19): p. 4547-57.
112. Butt, A.J., et al., *Insulin-like growth factor-binding protein-5 inhibits the growth of human breast cancer cells in vitro and in vivo.* J Biol Chem, 2003. **278**(32): p. 29676-85.
113. Gelebart, P., M. Opas, and M. Michalak, *Calreticulin, a Ca2+-binding chaperone of the endoplasmic reticulum.* Int J Biochem Cell Biol, 2005. **37**(2): p. 260-6.
114. Papp, S., M.P. Fadel, and M. Opas, *Dissecting focal adhesions in cells differentially expressing calreticulin: a microscopy study.* Biol Cell, 2007. **99**(7): p. 389-402.
115. Zhu, N. and Z. Wang, *Calreticulin expression is associated with androgen regulation of the sensitivity to calcium ionophore-induced apoptosis in LNCaP prostate cancer cells.* Cancer Res, 1999. **59**(8): p. 1896-902.
116. Kageyama, S., et al., *Identification by proteomic analysis of calreticulin as a marker for bladder cancer and evaluation of the diagnostic accuracy of its detection in urine.* Clin Chem, 2004. **50**(5): p. 857-66.
117. Barrett, A.J., *Cathepsin D. Purification of isoenzymes from human and chicken liver.* Biochem J, 1970. **117**(3): p. 601-7.
118. Diment, S., K.J. Martin, and P.D. Stahl, *Cleavage of parathyroid hormone in macrophage endosomes illustrates a novel pathway for intracellular processing of proteins.* J Biol Chem, 1989. **264**(23): p. 13403-6.
119. Basset, P., et al., *A novel metalloproteinase gene specifically expressed in stromal cells of breast carcinomas.* Nature, 1990. **348**(6303): p. 699-704.
120. Uria, J.A., et al., *Regulation of collagenase-3 expression in human breast carcinomas is mediated by stromal-epithelial cell interactions.* Cancer Res, 1997. **57**(21): p. 4882-8.
121. Duffy, M.J., et al., *Cathepsin D concentration in breast cancer cytosols: correlation with disease-free interval and overall survival.* Clin Chem, 1992. **38**(10): p. 2114-6.
122. Westley, B.R. and F.E. May, *Prognostic value of cathepsin D in breast cancer.* Br J Cancer, 1999. **79**(2): p. 189-90.
123. Castino, R., et al., *Prolactin promotes the secretion of active cathepsin D at the basal side of rat mammary acini.* Endocrinology, 2008. **149**(8): p. 4095-105.
124. Lkhider, M., et al., *Cathepsin D released by lactating rat mammary epithelial cells is involved in prolactin cleavage under physiological conditions.* J Cell Sci, 2004. **117**(Pt 21): p. 5155-64.
125. Khalkhali-Ellis, Z. and M.J. Hendrix, *Elucidating the function of secreted maspin: inhibiting cathepsin D-mediated matrix degradation.* Cancer Res, 2007. **67**(8): p. 3535-9.
126. Christofori, G. and H. Semb, *The role of the cell-adhesion molecule E-cadherin as a tumour-suppressor gene.* Trends Biochem Sci, 1999. **24**(2): p. 73-6.
127. Vleminckx, K. and R. Kemler, *Cadherins and tissue formation: integrating adhesion and signaling.* Bioessays, 1999. **21**(3): p. 211-20.
128. Wang, W., et al., *Glucocorticoid induced expression of connective tissue growth factor contributes to lactogenic differentiation of mouse mammary epithelial cells.* Journal of Cellular Physiology, 2008. **214**(1): p. 38-46.
129. Morrison, B. and M.L. Cutler, *The contribution of adhesion signaling to lactogenesis.* J Cell Commun Signal, 2010. **4**(3): p. 131-9.
130. Avramidou, A., et al., *The novel adaptor protein Swiprosin-1 enhances BCR signals and contributes to BCR-induced apoptosis.* Cell Death Differ, 2007. **14**(11): p. 1936-47.
131. Hough, C.D., et al., *Coordinately up-regulated genes in ovarian cancer.* Cancer Res, 2001. **61**(10): p. 3869-76.
132. Kluger, H.M., et al., *cDNA microarray analysis of invasive and tumorigenic phenotypes in a breast cancer model.* Lab Invest, 2004. **84**(3): p. 320-331.

133. Provatopoulou, X., et al., *Circulating levels of matrix metalloproteinase-9 (MMP-9), neutrophil gelatinase-associated lipocalin (NGAL) and their complex MMP-9/NGAL in breast cancer disease.* BMC Cancer, 2009. **9**: p. 390.
134. Fernandez, C.A., et al., *The matrix metalloproteinase-9/neutrophil gelatinase-associated lipocalin complex plays a role in breast tumor growth and is present in the urine of breast cancer patients.* Clin Cancer Res, 2005. **11**(15): p. 5390-5.
135. Tong, Z., et al., *Neutrophil gelatinase-associated lipocalin as a survival factor.* Biochem J, 2005. **391**(Pt 2): p. 441-8.
136. van Veen, H.A., et al., *Characterization of Bovine Neutrophil Gelatinase-Associated Lipocalin.* Journal of dairy science, 2006. **89**(9): p. 3400-3407.
137. Stoesz, S.P., et al., *Heterogeneous expression of the lipocalin NGAL in primary breast cancers.* International Journal of Cancer, 1998. **79**(6): p. 565-572.
138. Tong, Z., et al., *Neutrophil gelatinase-associated lipocalin: a novel suppressor of invasion and angiogenesis in pancreatic cancer.* Cancer Res, 2008. **68**(15): p. 6100-8.
139. Curtis Hewitt, S., J.F. Couse, and K.S. Korach, *Estrogen receptor transcription and transactivation: Estrogen receptor knockout mice: what their phenotypes reveal about mechanisms of estrogen action.* Breast Cancer Res, 2000. **2**(5): p. 345-52.
140. Sebastian, J., et al., *Activation and function of the epidermal growth factor receptor and erbB-2 during mammary gland morphogenesis.* Cell Growth Differ, 1998. **9**(9): p. 777-85.
141. Schultz-Norton, J.R., et al., *Protein Disulfide Isomerase Serves as a Molecular Chaperone to Maintain Estrogen Receptor α Structure and Function.* Molecular Endocrinology, 2006. **20**(9): p. 1982-1995.
142. Fu, X., P. Wang, and B.T. Zhu, *Protein disulfide isomerase is a multifunctional regulator of estrogenic status in target cells.* J Steroid Biochem Mol Biol, 2008. **112**(1-3): p. 127-37.
143. Molero, L., et al., *Proteomic approach to identify changes in protein expression modified by 17beta-oestradiol in bovine vascular smooth muscle cells.* Clin Sci (Lond), 2005. **109**(5): p. 457-63.
144. Helleman, J., et al., *Association of an extracellular matrix gene cluster with breast cancer prognosis and endocrine therapy response.* Clin Cancer Res, 2008. **14**(17): p. 5555-64.
145. Horiuchi, K., et al., *Identification and characterization of a novel protein, periostin, with restricted expression to periosteum and periodontal ligament and increased expression by transforming growth factor beta.* J Bone Miner Res, 1999. **14**(7): p. 1239-49.
146. Sugiura, T., et al., *Expression and characterization of murine osteoblast-specific factor 2 (OSF-2) in a baculovirus expression system.* Protein Expr Purif, 1995. **6**(3): p. 305-11.
147. Mamalis, A., et al., *Oestrogen regulates proliferation, osteoblastic differentiation, collagen synthesis and periostin gene expression in human periodontal ligament cells through oestrogen receptor beta.* Archives of oral biology, 2011. **56**(5): p. 446-455.
148. Mazurek, N., et al., *Phosphorylated galectin-3 mediates tumor necrosis factor-related apoptosis-inducing ligand signaling by regulating phosphatase and tensin homologue deleted on chromosome 10 in human breast carcinoma cells.* J Biol Chem, 2007. **282**(29): p. 21337-48.
149. Mengwasser, J., F.-T. Liu, and J.P. Sleeman, *Galectin-3 is strongly up-regulated in nonapoptosing mammary epithelial cells during rat mammary gland involution.* Glycobiology, 2002. **12**(2): p. 129-134.
150. Furtak, V., F. Hatcher, and J. Ochieng, *Galectin-3 Mediates the Endocytosis of [beta]-1 Integrins by Breast Carcinoma Cells.* Biochemical and Biophysical Research Communications, 2001. **289**(4): p. 845-850.
151. Faulds, M.H., et al., *Estrogen receptor functional activity changes during differentiation of mammary epithelial cells.* Mol Endocrinol, 2004. **18**(2): p. 412-21.

VI – Annexes

Annexe 1

Solutions

Cell culture

Complete medium:
RPMI 1640,
L-glutamine
10% FBS
50µl/ml gentamycin
10ng/ml EGF
5µl/ml insulin

Serum free medium (SFM):
RPMI 1640 (phenol red free)
L-glutamine
50µl/ml gentamycin
5µl/ml insulin
10ng/ml EGF (optional)

Pre-differentiation medium:
RPMI 1640,
L-glutamine
2 % FBS
50µl/ml gentamycin
5µl/ml insulin

Differentiation medium:
RPMI 1640,
L-glutamine
2 % FBS
50µl/ml gentamycin
5µl/ml insulin

100 μM dexamethasone
conc prolactin

ER ligand treatment: stock solutions of 1 μM E2, 10 μM PPT and μM DPN were prepared in absolute ethanol and for treatment, they were diluted 1/1000 in serum free medium. Cells were treated for 24 h

Protein precipitation and quantification

Solution 1:
PBS
0.1% Tween 20
1% SDS
1mM DTT

Solution 2:
8M Urea
2M Tiourea
1% CHAPS
10mg/ml DTT

SDS-PAGE

Running buffer 10X (before being used was diluted 1:10):
0.25M Tris
1.92M Glycine
1% SDS
Laemmli loading buffer (10mL):
4mL SDS 10%
2.5mL Tris 0.5M pH 6.8
1.5mL Glycerol
10.0mg Bromophenol Blue

2.0mL Mercaptoetanol

2D electrophoresis

Rehydration buffer:
8M Urea
2M Tiourea
1% CHAPS
12.9mM DTT
0.1% Pharmalyte 3-10
0.01% Bromophenol Blue

Equilibrium buffer (200ml):
6.7ml electrophoresis buffer ph 8.8
72.1g Urea
69ml glycerol 87%
4g SDS
Complete with distilled water

0.5% Agar solution (100ml):
0.5g Agar
100ml running buffer 1X
bromofenol blue

Gel staining with Colloidal Coomassie (G250)

Fixing solution:
40% (v/v) methanol
10% (v/v) acetic acid

Colloidal coomassie:
0.12% (w/v) coomassie blue G250

20% methanol

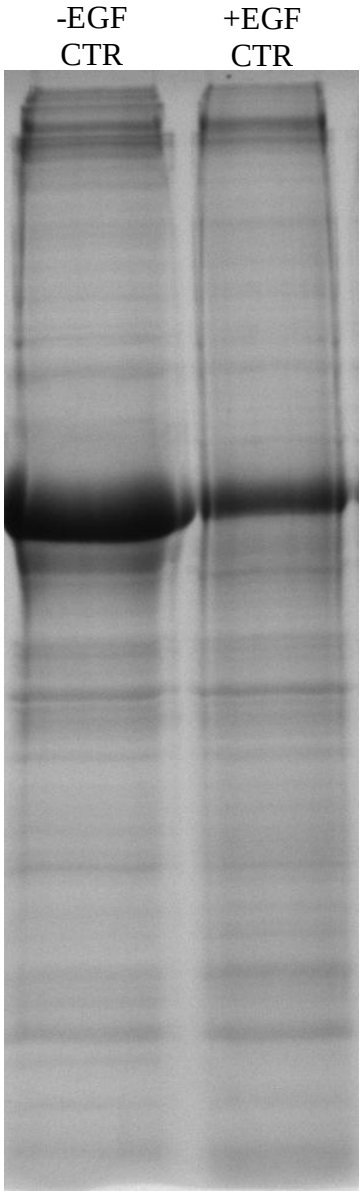
Destaining solution:
25% methanol

MALDI-TOF/TOF analysis

Buf1: 50mM Ammonium bicarbonate
FA: 10% (v/v) Formic acid
Trypsin (v511, Promega): 1 vial solubilised in 2ml Buf1

Annexe 2

Comparison between –EGF control and +EGF control appear to show a higher intensity in –EGF control. However, when intensities are measure the 2 conditions show a similar value, -EGF CTR (39449 CNT*mm^2) +EGF CTR (37825 CNT*mm^2). This effect was observed in other 1D electrophoresis.



Annexe 3

Table 1- Proteins expressed in Proliferation (SCL +EGF) identify by SDS-PAGE MALDI-TOF/TOF

Nº identifications	Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
2	14-3-3 protein epsilon	P62259	29155	4,63	10	100,00	mitochondrion	cell cycle
3	14-3-3 protein zeta/delta	P63102	27754	4,73	13	100,00	Cytoplasm	cell cycle
1	78 kDa glucose-regulated protein	P20029	72377	5,07	20	100,00	Cytoplasm	response to stimulus
5	Actin, cytoplasmic 2	P63259	41766	5,31	11	100,00	Cytoplasm	cellular component organization
1	Alpha-actinin-1	Q7TPR4	103004	5,25	19	100,00	Cytoplasm	cellular component organization
1	Alpha-actinin-4	P57780	104911	5,25	19	100,00	Cytoplasm	cellular component organization
3	Alpha-enolase	P17182	47111	6,37	8	100,00	Secreted	metabolic process
1	Annexin A2	P07356	38652	7,55	11	100,00	Secreted	cell comunnication
3	Beta-actin-like protein 2	Q8BFZ3	41977	5,3	8	100,00	Cytoplasm	cellular component organization
1	Coagulation factor VIII	Q06194	265981	6,9	17	96,47	Extracellular matrix	cell comunnication
3	Collagen alpha-1(VI) chain	Q04857	108422	5,2	9	100,00	Extracellular matrix	cell comunnication

Table 2- Proteins expressed in Proliferation (SCL +EGF) identify by SDS-PAGE MALDI-TOF/TOF (Cont.)

Nº identifications	Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
1	Elongation factor 1-alpha 1	P10126	50082	9,1	7	99,97	Cytoplasm	cellular process
4	Elongation factor 2	P05197	95223	6,41	13	100,00	Cytoplasm	cellular process
1	Extracellular superoxide dismutase [Cu-Zn]	O09164	27375	6,36	3	99,24	Secreted	
1	Fatty acid synthase	P19096	272257	6,13	15	100,00	Mitochondrion	metabolic process
2	Fibronectin	P11276	272319	5,39	30	100,00	Secreted	cell comunnication
1	Fructose-bisphosphate aldolase A	P05064	39331	8,31	11	100,00	flagellum	metabolic process
1	Heat shock cognate 71 kDa protein	P63018	70827	5,37	19	100,00	Cytoplasm	response to stimulus
1	Heat shock protein HSP 90-alpha	P07901	84735	4,93	9	100,00	Cytoplasm	response to stimulus
3	Heat shock protein HSP 90-beta	P11499	83273	4,97	16	100,00	Cytoplasm	response to stimulus
1	Homeobox protein Rhox5	P52651	23031	5,92	9	95,89	Nucleus	
1	Malate dehydrogenase, cytoplasmic	P14152	36488	6,16	3	96,28	Cytoplasm	metabolic process

Table 3- Proteins expressed in Proliferation (SCL +EGF) identify by SDS-PAGE MALDI-TOF/TOF (Cont.)

Nº identifications	Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
1	Peroxiredoxin-1	P35700	22162	8,26	7	100,00	Cytoplasm	metabolic process
1	Phosphoglycerate kinase 1 *	P09411	44522	8,02	9	100,00	Cytoplasm	metabolic process
1	Proteasome subunit beta type-1	O09061	26355	7,67	4	100,00	Cytoplasm	metabolic process
3	Protein disulfide-isomerase	P09103	57108	4,79	13	100,00	Cell membrane	metabolic process
5	Pyruvate kinase isozymes M1/M2	P52480	57808	7,18	16	100,00	Cytoplasm	metabolic process
1	Rho GDP-dissociation inhibitor 1	Q99PT1	23393	5,12	10	100,00	cytoplasm	cellular process
6	Serum albumin	P07724	68648	5,75	8	100,00	Secreted	transport
1	Thrombospondin-1	P35441	129564	4,72	8	99,67	Extracellular matrix	response to stimulus
1	Tubulin alpha-1C chain	P68373	49877	4,96	5	100,00	Cytoplasm	cellular component organization

Table 2 - Proteins expressed in Pre-diffentiation 48h identify by SDS-PAGE MALDI-TOF/TOF

Nº identifications	Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
1	Calsyntenin-1	Q9EPL2	108831	4,82	4	99,97	Extracellular matrix	cell comunnication
1	Fibronectin	P11276	272319	5,39	35	100,00	Extracellular matrix	cell comunnication
1	Metalloproteinase inhibitor 2	P25785	24312	7,45	5	100,00	Extracellular matrix	metabolic process
1	Neutrophil gelatinase-associated lipocalin	P11672	22861	8,96	2	100,00	Secreted	cell comunnication
1	Periostin	Q62009	93085	7,27	8	100,00	Extracellular matrix	cell comunnication
1	Protein disulfide-isomerase	P09103	57108			100,00	Cell membrane	metabolic process
1	Pyruvate kinase isozymes M1/M2	P52480	57808	7,18	6	100,00	Cytoplasm	metabolic process

Table 3 - Proteins expressed in Diffentiation identify by SDS-PAGE MALDI-TOF/TOF

Nº identifications	Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
1	78 kDa glucose-regulated protein	P20029	72377	5,07	11	99,86	Endoplasmic reticulum	response to stimulus
1	Actin, alpha skeletal muscle	P68136	42024	5,23	5	100,00	Cytoplasm	cellular component organization
1	Actin, cytoplasmic 2	P68138	41766	5,31	10	100,00	Cytoplasm	cellular component organization
1	Alpha-actinin-4	P57780	104911	5,25	10	99,77	Cytoplasm	cellular component organization
1	Alpha-enolase	P17182	47111	6,37	9	100,00	Secreted	metabolic process
3	Annexin A2	P07356	38652	7,55	15	100,00	Secreted	cellular process
2	Annexin A5	P48036	35730	4,83	16	100,00	Plasma membrane	cell comunnication
1	ATP synthase subunit beta, mitochondrial	P56480	56265	5,19	6	100,00	Mitochondrion	metabolic process
2	Beta-actin-like protein 2	Q8BFZ4	41977	5,3	5	100,00	Cytoplasm	cellular component organization
1	Calicin	Q8CDE2	66707	8,49	5	96,93	Cytoplasm	metabolic process
1	Cathepsin D	P18242	44925	6,71	4	99,91	Lysosome	metabolic process

Table 3 - Proteins expressed in Diffentiation identify by SDS-PAGE MALDI-TOF/TOF (Cont.)

Nº identifications	Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
1	Ceruloplasmin	Q61147	121074	5,53	16	100,00	Secreted	cell comunnication
2	Collagen alpha-1(II) chain	P28482	141886	6,58	32	100,00	Extracellular matrix	cell comunnication
2	Complement C3	P01027	186365	6,39	22	100,00	Secreted	cell comunnication
2	Connective tissue growth factor	P29268	37768	8,22	2	99,90	Extracellular matrix	cell comunnication
1	Elongation factor 1-alpha 1	P10126	50082	9,1	3	95,13	Cytoplasm	cellular process
2	Fibronectin	P11276	272319	5,39	33	100,00	Extracellular matrix	cell comunnication
1	Galectin-3	P16110	27498	8,46	6	99,82	Extracellular matrix	cellular process
2	Heat shock cognate 71 kDa protein	P63017	70827	5,37	16	100,00	Cytoplasm	response to stimulus
1	Heat shock-related 70 kDa protein 2	P17156	69698	5,58	9	99,99	Cytoplasm	response to stimulus
1	Hexokinase-1	P17710	108233	6,44	14	99,99	Mitochondrion	metabolic process
1	Histone H2A type 1-F	Q8CGP5	14153	11,05	5	100,00	nucleus	cellular process

Table 3 - Proteins expressed in Diffentiation identify by SDS-PAGE MALDI-TOF/TOF (Cont.)

Nº identifications	Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
1	Histone H2A type 1-H	Q8CGP6	13942	11,03	3	99,96	nucleus	cellular process
1	Histone H3.1	Q6LBF0	15394	11,13	9	100,00	nucleus	cellular process
1	Histone H4	P62804	11360	11,36	9	100,00	nucleus	cellular process
2	Insulin-like growth factor-binding protein 5	Q07079	30353	8,48	5	100,00	Secreted	cell comunnication
1	Keratin, type I cytoskeletal 14	Q61781	52834	5,1	9	100,00	Sarcolemma	cellular component organization
1	Keratin, type II cytoskeletal 73	Q6NXH9	58875	8,36	7	96,08	Cytoplasm	cellular component organization
1	Malate dehydrogenase, mitochondrial	P04636	35661	8,93	4	100,00	Mitochondrion	metabolic process
2	Neutrophil gelatinase-associated lipocalin	P11672	22861	8,96	6	100,00	Secreted	cell comunnication
3	Periostin	Q62009	93085	7,27	25	100,00	Secreted	cell comunnication
2	Protein disulfide-isomerase	P09103	57108	4,79	9	100,00	Cell membrane	metabolic process
1	Pyruvate kinase isozymes M1/M2	P52480	57808	7,18	12	100,00	Cytoplasm	cell comunnication

Table 3 - Proteins expressed in Diffentiation identify by SDS-PAGE MALDI-TOF/TOF (Cont.)

Nº identifications	Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
1	Rho GDP-dissociation inhibitor 1	Q5XI73	23393	5,12	7	100,00	Cytoplasm	cellular process
5	Serum albumin	P07724	68648	5,75	8	100,00	Secreted	transport
1	Sex comb on midleg-like protein 4	Q80VG1	44417	9,81	9	99,25	nucleus	cell cycle
2	SPARC	P07214	34428	4,77	11	100,00	Extracellular matrix	cell comunnication
1	TRIO and F-actin-binding protein	Q99KW3	223232	8,36	21	99,76	Cytoplasm	cell comunnication
1	Tubulin alpha-1C chain	P68373	49877	4,96	7	100,00	Cytoplasm	cellular component organization
1	Tubulin beta-5 chain	P99024	49639	4,78	11	98,29	Cytoplasm	cellular component organization

Table 4 - Proteins expressed in Pre-diffentiation 24h identify by SDS-PAGE MALDI-TOF/TOF

Nº identifications	Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
1	14-3-3 protein sigma	O70456	27696	4,75	8	99,99	Secreted	cell cycle
1	78 kDa glucose-regulated protein	P06761	72302	5,07	17	100,00	Cytoplasm	response to stimulus
1	Actin, cytoplasmic 2	P63260	41766	5,31	9	100,00	Cytoplasm	cellular component organization
1	Alpha-actinin-4	P57780	104911	5,25	12	100,00	Cytoplasm	cellular component organization
3	Alpha-enolase	P17182	47111	6,37	13	100,00	Secreted	metabolic process
1	Annexin A3	O35639	36349	5,33	8	100,00	Cell membrane	cell comunnication
2	Calsyntenin-1	Q9EPL3	108831	5	11	100,00	Extracellular matrix	cell comunnication
1	Cathepsin D	P18242	44925	6,71	7	99,91	Secreted	metabolic process
1	Ceruloplasmin	Q61147	121074	5,53	10	99,15	Extracellular matrix	cell comunnication
1	Collagen alpha-1(II) chain	P28481	141886	6,58	32	100,00	Extracellular matrix	cell comunnication
1	Collagen alpha-1(XI) chain	Q61245	180853	5,09	7	100,00	Extracellular matrix	cell comunnication

Table 4 - Proteins expressed in Pre-diffentiation 24h identify by SDS-PAGE MALDI-TOF/TOF (Cont.)

Nº identifications	Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
1	Cystatin-C	P21460	15521	9,18	9	100,00	Extracellular matrix	metabolic process
1	Double-stranded RNA-binding protein Staufen homolog 2	Q8CJ67	62496	9,61	10	99,19	endoplasmic reticulum	cell cycle
1	Elongation factor 1-alpha 1	P10126	50082	9,1	1	98,77	Cytoplasm	cellular process
1	Elongation factor 1-gamma	Q9D8N0	50029	6,31	7	99,60	Cytoplasm	cellular process
5	Elongation factor 2	P58252	95223	6,41	17	100,00	Cytoplasm	cellular process
1	Fibronectin	P11276	272319	5,39	33	100,00	Secreted	cell comunnication
1	Fructose-bisphosphate aldolase A	P05064	39331	8,31	8	100,00	Cytoplasm	metabolic process
1	Glyceraldehyde-3-phosphate dehydrogenase	P16858	35787	8,44	4	100,00	Cytoplasm	metabolic process
1	Heat shock cognate 71 kDa protein	P19378	70761	5,24	17	100,00	Cytoplasm	response to stimulus
4	Heat shock protein HSP 90-beta	P11499	83273	4,97	22	100,00	Cytoplasm	response to stimulus
1	Insulin-like growth factor-binding protein 5	Q07079	30352	8,48	6	99,80	Secreted	cell comunnication

Table 4 - Proteins expressed in Pre-diffentiation 24h identify by SDS-PAGE MALDI-TOF/TOF (Cont.)

Nº identifications	Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
1	Liprin-alpha-3	P58252	95223	6,41	17	100,0	Cytoplasm	cellular process
1	Malate dehydrogenase, cytoplasmic	P11276	272319	5,39	33	100,0	Secreted.	cell comunnication
1	MyoD family inhibitor	P16858	35787	8,44	4	100,0	Cytoplasm	metabolic process
1	Neutrophil gelatinase-associated lipocalin	P19378	70761	5,24	17	100,0	Cytoplasm	response to stimulus
1	Periostin	P11499	83273	4,97	22	100,0	Cytoplasm	response to stimulus
2	Phosphoglycerate kinase 1	P60469	116208	5,72	4	98,0	cell surface	cellular process
1	Protein disulfide-isomerase	P09411	44522	8	4	100,0	Cytoplasm	metabolic process
6	Serum albumin	P07724	68648	5,75	6	100,0	Secreted	transport
1	SPARC	P07214	34428	4,77	10	100,00	Extracellular matrix	cell comunnication

Table 5 - Proteins expressed by HC11 cells under estrogenic stimulation with or without EGFR activation identify by 2D-PAGE MALDI-TOF/TOF

Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
14-3-3 protein gamma	P61983	28285	4,8	2	99,81	Cytoplasm	cell cycle
14-3-3 protein zeta/delta	P63102	27754	4,73	8	100,00	Cytoplasm	cell cycle
40S ribosomal protein SA	P14206	32817	4,8	3	100,00	nucleus	metabolic process
60 kDa heat shock protein, mitochondrial	P63039	60917	5,91	1	99,99	Cytoplasm	metabolic process
6-phosphogluconate dehydrogenase, decarboxylating	Q9DCD0	53213	6,81	3	100,00	extracellular matrix	metabolic process
78 kDa glucose-regulated protein	P20029	72377	5,07	13	99,97	Cytoplasm	response to stimulus
Actin, cytoplasmic 1	P48975	41711	5,22	1	100,00	cytoskeleton	cellular component organization
Actin-related protein 2/3 complex subunit 2	Q9CVB6	34336	6,84	1	96,55	cytoskeleton	cellular component organization
Actin-related protein 3	Q4V7C7	47327	5,61	3	96,71	cytoskeleton	cellular component organization
Acyl-coenzyme A thioesterase 3	Q9QYR7	47460	6,98	2	100,00	Cytoplasm	metabolic process
Alpha-actinin-4	P57780	104911	5,25	4	99,34	cytoskeleton	cellular component organization
Alpha-centractin	P85515	42587	6,19	3	100,00	cytoskeleton	cellular component organization

Table 5 - Proteins expressed by HC11 cells under estrogenic stimulation with or without EGFR activation identify by 2D-PAGE MALDI-TOF/TOF (Cont.)

Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
Alpha-enolase	P17182	47111	6,37	16	100,00	Secreted	metabolic process
Annexin A3	O35639	36349	5,33	11	100,00	Cell membrane	cell comunnication
Annexin A8	O35640	36820	5,56	2	100,00		cell comunnication
Aspartate aminotransferase, cytoplasmic	P05201	46202	6,68	8	100,00	Cytoplasm	metabolic process
Aspartate aminotransferase, mitochondrial	P00507	47284	9,13	3	97,26	Mitochondrion	metabolic process
Basement membrane-specific heparan sulfate proteoglycan core protein	Q05793	398039	5,88	2	100,00	extracellular matrix	cell comunnication
Beta-actin-like protein 2	Q8BFZ3	41977	5,3	3	100,00	cytoskeleton	cellular component organization
Beta-centractin	Q8R5C5	42255	5,98	1	98,63	cytoskeleton	cellular component organization
Bifunctional purine biosynthesis protein PURH	Q9CWJ9	64177	6,3	10	100,00	Mitochondrion	metabolic process
Calreticulin	P14211	47965	4,33	1	99,99	extracellular matrix	metabolic process
Ceruloplasmin	Q61147	121074	5,53	7	100,00	Secreted	cell comunnication
Chloride intracellular channel protein 1	Q6MG61	26964	5,09	2	100,00	cytoskeleton	cell comunnication

Table 5 - Proteins expressed by HC11 cells under estrogenic stimulation with or without EGFR activation identify by 2D-PAGE MALDI-TOF/TOF (Cont.)

Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
Cofilin-1	P45592	18521	8,22	3	100,00	cytoskeleton	cellular component organization
Complement C3	P01027	186365	6,39	2	100,00	Secreted	cell comunnication
E3 ubiquitin-protein ligase NEURL3	Q8CJC5	28022	6,45	3	97,77		
Elongation factor 1-alpha 1	P10126	50082	9,1	7	99,98	Cytoplasm	cellular process
Elongation factor 1-beta	O70251	24678	4,53	1	100,00	Cytoplasm	cellular process
Elongation factor 1-gamma	Q9D8N0	50029	6,31	3	98,42	Cytoplasm	cellular process
Elongation factor 2	P58252	95253	6,41	16	100,00	Cytoplasm	cellular process
Endoplasmin	P08113	92418	4,74	10	100,00	endoplasmic reticulum	metabolic process
Eukaryotic initiation factor 4A-II	P10630	46373	5,33	1	99,97		metabolic process
Eukaryotic translation initiation factor 3 subunit I	B0BNA7	36438	5,38	2	99,92	Cytoplasm	metabolic process
Eukaryotic translation initiation factor 5A-1	Q3T1J1	16821	5,08	1	97,83	Cytoplasm	metabolic process
Eukaryotic translation initiation factor 6	Q3KRD8	26554	4,63	3	100,00	Cytoplasm	metabolic process

Table 5 - Proteins expressed by HC11 cells under estrogenic stimulation with or without EGFR activation identify by 2D-PAGE MALDI-TOF/TOF (Cont.)

Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
Extracellular superoxide dismutase [Cu-Zn]	O09164	27375	6,36	1	95,46	Secreted	
F-actin-capping protein subunit alpha-1	P47753	32919	5,34	3	99,99	cytoskeleton	cellular component organization
F-actin-capping protein subunit alpha-2	Q3T1K5	32947	5,57	6	100,00	cytoskeleton	cellular component organization
Fascin	P85845	54457	6,29	3	97,13	Cell projection	cellular component organization
Fructose-bisphosphate aldolase A	P05064	39331	8,31	8	100,00	flagellum	metabolic process
Galectin-3	P16110	27498	8,46	6	100,00	extracellular matrix	cellular process
Glucose-6-phosphate 1-dehydrogenase X	Q00612	59225	6,06	8	100,00		metabolic process
Glutaredoxin-3	Q9JLZ1	37825	5,51	1	99,99	Cytoplasm	metabolic process
Glutathione S-transferase P 1	P19157	23594	7,68	4	100,00	Cytoplasm	response to stimulus
Glyceraldehyde-3-phosphate dehydrogenase	P16858	35787	8,44	4	100,00	Cytoplasm	metabolic process
Heat shock 70 kDa protein 1B	P17879	70133	5,53	12	99,63	Cytoplasm	response to stimulus
Heat shock 70 kDa protein 4	Q61316	94073	5,15	8	100,00	Cytoplasm	response to stimulus

Table 5 - Proteins expressed by HC11 cells under estrogenic stimulation with or without EGFR activation identify by 2D-PAGE MALDI-TOF/TOF (Cont.)

Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
Heat shock protein beta-1	P14602	23000	6,12	3	100,00	Cytoplasm	response to stimulus
Heat shock protein HSP 90-alpha	P07901	84762	4,93	4	100,00	Cytoplasm	response to stimulus
Heat shock protein HSP 90-beta	P11499	83273	4,97	6	100,00	Cytoplasm	response to stimulus
Heat shock-related 70 kDa protein 2	P17156	69698	5,58	7	96,22	Cytoplasm	response to stimulus
Heterogeneous nuclear ribonucleoprotein A/B	Q99020	30812	7,68	11	100,00	ribonucleoprotein complex	cellular process
Heterogeneous nuclear ribonucleoprotein A1	P04256	34191	9,2	4	100,00	ribonucleoprotein complex	cellular process
Heterogeneous nuclear ribonucleoprotein F	Q794E4	45701	5,31	1	100,00	ribonucleoprotein complex	cellular process
Heterogeneous nuclear ribonucleoproteins A2/B1	A7VJC2	37455	8,97	3	100,00	ribonucleoprotein complex	cellular process
Inorganic pyrophosphatase	Q9D819	32646	5,37	7	100,00	Cytoplasm	metabolic process
Keratin, type II cytoskeletal 8	P11679	54531	5,7	2	100,00	cytoskeleton	cellular component organization
Lactoylglutathione lyase	Q9CPU0	20796	5,24	3	99,98		
Macrophage-capping protein	P24452	39216	6,73	4	100,00	cytoskeleton	cellular component organization

Table 5 - Proteins expressed by HC11 cells under estrogenic stimulation with or without EGFR activation identify by 2D-PAGE MALDI-TOF/TOF (Cont.)

Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
Malate dehydrogenase, cytoplasmic	P14152	36460	6,16	1	100,00	Cytoplasm	metabolic process
Malate dehydrogenase, mitochondrial	P04636	35661	8,93	4	100,00	Mitochondrion	metabolic process
Metalloproteinase inhibitor 2	P25785	24340	7,45	2	100,00	Extracellular matrix	metabolic process
Multifunctional protein ADE2	Q9DCL9	47040	6,94	2	100,00		metabolic process
Neutrophil gelatinase-associated lipocalin	P11672	22861	8,96	3	100,00	Secreted.	cell comunnication
Nucleolin	P08199	77082	4,7	4	96,31	Cytoplasm	
Nucleoside diphosphate kinase B	Q01768	17352	6,97	6	100,00	Mitochondrion	metabolic process
Obg-like ATPase 1	Q9CZ30	44701	7,64	3	100,00		
Peptidyl-prolyl cis-trans isomerase D	Q6DGG0	40740	6,73	6	100,00	Cytoplasm	metabolic process
Periostin	Q62009	93085	7,27	25	100,00	Secreted	cell comunnication
Peroxiredoxin-1	P35700	22162	8,26	5	100,00	Cytoplasm	metabolic process
Peroxiredoxin-6	O08709	24855	5,71	3	100,00	Cytoplasm	metabolic process

Table 5 - Proteins expressed by HC11 cells under estrogenic stimulation with or without EGFR activation identify by 2D-PAGE MALDI-TOF/TOF (Cont.)

Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
Phosphatidylethanolamine-binding protein 1	P70296	20817	5,19	4	100,00	Cytoplasm	cellular process
Phosphatidylinositol transfer protein alpha isoform	P16446	31887	5,97	1	99,99	Cytoplasm	metabolic process
Phosphoglycerate kinase 1	P09411	44522	8,02	16	100,00	Cytoplasm	metabolic process
Phosphoserine aminotransferase	Q99K85	40447	8,15	11	100,00		metabolic process
Pre-mRNA-processing factor 19	Q9JMJ4	55204	6,14	3	99,58	Cytoplasm	metabolic process
Proliferating cell nuclear antigen	P17918	28766	4,66	2	100,00	nucleus	cell cycle
Proteasome activator complex subunit 1	P97371	28655	5,73	7	100,00	Cytoplasm	metabolic process
Proteasome subunit alpha type-5	Q9Z2U1	26394	4,74	4	100,00	Cytoplasm	metabolic process
Proteasome subunit alpha type-7	Q9Z2U0	27838	8,59	7	100,00	Cytoplasm	metabolic process
Proteasome subunit beta type-3	P40112	22949	6,15	1	100,00	Cytoplasm	metabolic process
Proteasome subunit beta type-5	O55234	28514	6,52	3	99,96	Cytoplasm	metabolic process
Protein disulfide-isomerase	P09103	57108	4,79	7	100,00	cell membranne	metabolic process

Table 5 - Proteins expressed by HC11 cells under estrogenic stimulation with or without EGFR activation identify by 2D-PAGE MALDI-TOF/TOF (Cont.)

Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
Protein disulfide-isomerase A3	P27773	56643	5,88	5	100,00	endoplasmic reticulum	metabolic process
Protein disulfide-isomerase A6	Q922R8	48070	5	6	100,00	cell membranne	metabolic process
Protein SET	Q63945	33386	4,22	2	99,41	Cytoplasm	cell cycle
Pyruvate kinase isozymes M1/M2	P52480	57808	7,18	10	100,00	Cytoplasm	metabolic process
Renin receptor	Q9CYN9	39067	5,36	3	100,00	Cell membrane	cell comunnication
Rho GDP-dissociation inhibitor 1	Q99PT1	23393	5,12	8	100,00	cytoplasm	cellular process
Scavenger mRNA-decapping enzyme DcpS	Q9DAR7	38964	6,02	2	100,00	cytoplasm	metabolic process
Septin-2	Q91Y81	41566	6,15	3	100,00	cytoskeleton	cell cycle
Septin-7	Q9WVC0	50476	8,82	6	100,00	cytoskeleton	cell cycle
Serpin B5	P70124	42085	5,55	7	100,00	Secreted.	metabolic process
Serum albumin	P07724	68648	5,75	4	97,00	Secreted	transport
S-formylglutathione hydrolase	Q9R0P3	31299	6,7	1	99,96	Cytoplasm	

Table 5 - Proteins expressed by HC11 cells under estrogenic stimulation with or without EGFR activation identify by 2D-PAGE MALDI-TOF/TOF (Cont.)

Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
SPARC	P07214	34428	4,77	3	99,82	Extracellular matrix	cell comunnication
Stress-induced-phosphoprotein 1	Q60864	62542	6,4	10	100,00	Cytoplasm	response to stimulus
T-complex protein 1 subunit alpha (Fragments)	P86208	13403	5,13	3	96,47	Cytoplasm	metabolic process
T-complex protein 1 subunit beta	P80314	57441	5,97	10	100,00	Cytoplasm	metabolic process
Thioredoxin-like protein 1	Q920J4	32229	4,84	1	100,00	Cytoplasm	cell comunnication
Tight junction-associated protein 1	Q9DCD5	59383	5,74	7	95,46	tight junction	
Transketolase	P40142	67588	7,23	13	100,00		metabolic process
Tropomyosin alpha-3 chain	P21107	32843	4,68	2	95,66	cytoskeleton	cellular component organization
Tropomyosin alpha-4 chain	Q6IRU2	28450	4,65	2	100,00	cytoskeleton	cellular component organization
Tubulin alpha-1C chain	P68373	49877	4,96	2	100,00	cytoskeleton	cellular component organization
Tubulin alpha-4A chain	Q5XIF6	49892	4,95	3	100,00	cytoskeleton	cellular component organization
Ubiquitin carboxyl-terminal hydrolase isozyme L3	Q9JKB1	26135	4,96	5	99,93	Cytoplasm	metabolic process

Table 5 - Proteins expressed by HC11 cells under estrogenic stimulation with or without EGFR activation identify by 2D-PAGE MALDI-TOF/TOF (Cont.)

Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
Ubiquitin-conjugating enzyme E2 K	P61087	22393	5,33	7	100,00	Cytoplasm	metabolic process
UPF0760 protein C2orf29 homolog	Q9CWN7	54924	6,11	7	95,56		
UTP--glucose-1-phosphate uridylyltransferase	Q91ZJ5	56944	7,18	3	100,00	Cytoplasm	metabolic process
Zinc finger protein 664	Q4VA44	30305	9,03	5	97,20	nucleus	metabolic process

Table 7 - Proteins expressed in Diffentiation identify by 2D-PAGE MALDI-TOF/TOF

Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
14-3-3 protein epsilon	P62259	29155	4,63	5	100,00	mitochondrion	cell cycle
14-3-3 protein gamma	P61983	28285	4,8	4	99,99	Cytoplasm	cell cycle
14-3-3 protein theta	P68255	27761	4,69	10	100,00	Cytoplasm	cell cycle
14-3-3 protein zeta/delta	P63102	27754	4,73	11	100,00	Cytoplasm	cell cycle
78 kDa glucose-regulated protein	P20029	72377	5,07	27	100,00	Cytoplasm	response to stimulus
Actin, cytoplasmic 1	P48975	41711	5,22	10	100,00	cytoskeleton	cellular component organization
Annexin A5	P48036	35730	4,83	19	100,00	membrane	cell comunnication
Cadherin-1	P09803	98195	4,69	5	100,00	Cell membrane	cellular process
Calmodulin	P62161	16827	4,09	3	100,00	Cytoplasm	
Calreticulin	P14211	47965	4,33	16	100,00	extracellular matrix	metabolic process
Carboxypeptidase E	Q00493	53222	5,07	10	100,00	Secreted	metabolic process
Chloride intracellular channel protein 1	Q6MG61	26964	5,09	9	100,00	cytoskeleton	cell comunnication

Table 7 - Proteins expressed in Diffentiation identify by 2D-PAGE MALDI-TOF/TOF (Cont.)

Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
Coatomer subunit epsilon	O89079	34545	4,94	2	99,83	Cytoplasm	transport
Dynactin subunit 2	Q99KJ8	44090	5,14	6	100,00	Cytoplasm	
EF-hand domain-containing protein D2	Q9D8Y0	26775	5,01	9	100,00	cell membrane	cellular process
Elongation factor 1-delta	P57776	31274	4,91	5	100,00	Cytoplasm	cellular process
Endoplasmin	P08113	92418	4,74	9	99,98	endoplasmic reticulum	metabolic process
Eukaryotic translation initiation factor 2 subunit 1	P68101	36085	5,02	6	100,00	Cytoplasm	metabolic process
Eukaryotic translation initiation factor 5A-1	Q3T1J1	16821	5,08	5	99,98	Cytoplasm	metabolic process
F-actin-capping protein subunit alpha-1	P47753	32919	5,34	4	99,47	cytoskeleton	cellular component organization
Heat shock protein HSP 90-alpha	P07901	84735	4,93	8	100,00	Cytoplasm	response to stimulus
Heat shock protein HSP 90-beta	P11499	83273	4,97	18	100,00	Cytoplasm	response to stimulus
Heterogeneous nuclear ribonucleoprotein F	Q794E4	45701	5,31	5	100,00	ribonucleoprotein complex	cellular process
Inositol monophosphatase 1	O55023	30416	5,08	6	100,00	Cytoplasm	metabolic process

Table 7 - Proteins expressed in Diffentiation identify by 2D-PAGE MALDI-TOF/TOF (Cont.)

Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
Microtubule-associated protein RP/EB family member 1	Q61166	29997	5,12	9	100,00	cytoskeleton	cell cycle
Nascent polypeptide-associated complex subunit alpha	Q60817	23370	4,52	5	100,00	Cytoplasm	transport
Proliferating cell nuclear antigen	P17918	28766	4,66	1	96,71	nucleus	cell cycle
Prostate and testis expressed protein 3	B3GLJ3	11652	9,14	5	97,92	Secreted	
Proteasome subunit alpha type-3	P18422	28401	5,29	8	100,00	Cytoplasm	metabolic process
Protein disulfide-isomerase	P09103	57108	4,79	18	100,00	cell membranne	metabolic process
Protein disulfide-isomerase A6	Q922R8	48143	5	6	99,93	cell membranne	metabolic process
Protein SET	Q63945	33386	4,22	11	100,00	Cytoplasm	cell cycle
Ran-specific GTPase-activating protein	P34022	23582	5,15	9	100,00	nucleus	
Renin receptor	Q9CYN9	39067	5,36	6	100,00	Cell membrane	cell comunnication
Rho GDP-dissociation inhibitor 1	Q99PT1	23393	5,12	3	99,11	cytoplasm	cellular process
Serine-threonine kinase receptor-associated protein	Q5XIG8	38432	4,99	3	100,00	cytoplasm	cellular process

Table 7 - Proteins expressed in Diffentiation identify by 2D-PAGE MALDI-TOF/TOF (Cont.)

Protein name	Accession number	Protein MW (Da)	Protein PI	Peptide count	Protein Score C.I. (%)	Subcellular location	Biological process
SPARC	P07214	34428	4,77	5	100,00	Extracellular matrix	cell comunnication
S-phase kinase-associated protein 1	Q9WTX5	18660	4,4	3	100,00	cytoplasm	cellular process
SUMO-activating enzyme subunit 1	Q9R1T2	38596	5,24	8	100,00	nucleus	cellular process
Thioredoxin-like protein 1	Q920J4	32229	4,84	9	100,00	Cytoplasm	cell comunnication
Thrombospondin-1	P35441	129564	4,72	11	100,00	Extracellular matrix	response to stimulus
Translationally-controlled tumor protein	P63029	19450	4,76	4	100,00	Cytoplasm	cellular component organization
Tropomyosin alpha-1 chain	P58771	32661	4,69	2	99,71	cytoskeleton	cellular component organization
Tropomyosin alpha-3 chain	P21107	32843	4,68	11	100,00	cytoskeleton	cellular component organization
Tropomyosin alpha-4 chain	Q6IRU2	28450	4,65	12	100,00	cytoskeleton	cellular component organization
Tubulin alpha-1C chain	P68373	49877	4,96	15	100,00	Cytoplasm	cellular component organization
Tubulin beta-5 chain	P99024	49639	4,78	13	100,00	Cytoplasm	cellular component organization