Investigating CSMD1 signalling pathways and evaluating their prognostic values in breast cancer patients

May 30th 2011

Prepared for Bibliotheca Alexandrina
The centre for Special Studies and Programs

Table of contents

List of Figures

Abstract

Introduction

Carcinoma of the breast is the most prevalent cancer among Egyptian women and constitutes 29% of Cairo National Cancer Institute cases. The prognosis and clinical management of patients with breast cancer is commonly determined by traditional clinical and pathological factors. Nevertheless, patients may present with significantly different clinical outcomes despite similar clinic pathological features. This has prompted intense research to find biological markers that may closely reflect tumour biology and thereby clinical outcome.

Recently, we have identified CUB and Sushi multiple domain protein 1 (CSMD1), a tumour suppressor gene (TSG), as a novel independent prognostic marker for breast cancer patients with invasive ductal carcinoma (IDC) (Kamal *et al*., 2010). Our strong preliminary data showed that down regulation of CSMD1 expression, using shRNA, modulates cell behaviour in a way makes it more metastatic (Kamal *et al*, 2011, unpublished data). However, the mechanism/s of CSMD1 action is still unknown. Understanding how CSMD1 regulates cell behaviour will not only help identifying better measures for breast cancer prognosis, but also will provide very important information about how we stop cancer metastasis.

Our hypothesis is that using the expression level of CSMD1 along with the expression status of one or more of its interactants will give better prediction of patient's prognosis than using CSMD1 or any of its interactants alone.

Aim of the study

In this study, we aim to investigate CSMD1 signalling pathways, by determining it`s interactants. Moreover, we will investigate whether using the expression of CSMD1 combined with the expression of one or more of its partners will give better prognostic information than using the expression of each marker individually in breast cancer.

Design of the study and discussion

A- Investigating CSMD1 signalling pathways

Recently, we have down regulated CSMD1 expression in three cell lines; MCF10A, MDA-MB-435 and LNCaP, using shRNA. In this study, qPCR multiplex arrays (Lonza, Switzerland) for 48 genes will be designed to test their expression in the normal breast cell line MCF10A shCSMD1. The expression pattern of genes which show big differences in their expression between MCF10A CSMD1 positive cells and MCF10A CSMD1 negative cells (termed CSMD1 affected genes/proteins) will be further confirmed, using specific qPCR assays, in MDA-MB435 and LNCaP cell lines. Moreover, the protein expression of these genes will be analysed in the three cell lines using western blotting assays. Co Immunoprecipitation (CoIP) experiments will be conducted to examine physical interactions between CSMD1 and CSMD1 affected proteins. CSMD1`s interactants may represent a rich source of drug targets for cancer therapy. They may also act as novel prognostic signatures for breast cancer.

B- Evaluating the prognostic value of CSMD1 signalling pathway/s in breast cancer patients.

After determining CSMD1 affected genes and CSMD1`s interactants, an intensive literature search will be done to investigate the importance of the affected genes and components of the identified CSMD1 signalling pathway/s in the progression of breast cancer. Proteins showing highest relevance to breast cancer progression will be involved in the study. The expression pattern of these proteins, along with CSMD1, will be investigated, using Immunohistochemistry (IHC), in formaline fixed paraffin embedded (FFPE) tissues from a hundred breast cancer patients. Comprehensive statistical analyses will be performed to evaluate their prognostic values individually or in combinations. Results from this experiment may give a very useful and strong prognostic signature. This will help clinicians to better differentiate between patients with good or poor prognosis and therefore make the right treatment decision.

1. Introduction

Carcinoma of the breast is histologically heterogeneous disease with two main groups; ductal and lobular carcinomas. The reason for segregating breast tumours into subtypes is to determine whether this gives prognostic information concerning tumour behaviour (e.g. the susceptibility to metastasize) or predicts responsiveness to various therapies (Sims *et al*., 2007). Thus, identifying factors (prognostic factors) which able to predict patient's prognosis and segregate them into different risk groups is a priority in the field of cancer research.

The prognosis and clinical management of patients with breast cancer is commonly determined by traditional clinical and pathological factors such as tumour type, lymph node status, tumour grade, tumour size, mitotic index, Nottingham prognostic index (NPI) and the steroid hormonal status. Nevertheless these markers are relatively crude measures which results in many patients being over treated or undertreated. This has prompted intense research to find biological markers that may closely reflect tumour biology and thereby clinical outcome. Hence cancer progression and metastasis are tightly regulated by oncogenes and tumor suppressor genes, these genes provide rich sources of novel biomarkers and also novel targets for cancer therapy.

Recently, we have identified CSMD1, a TSG, as a novel independent prognostic marker for breast cancer patients with IDC (Kamal *et al*., 2010). In a functional study, we have found that down regulation of CSMD1 expression, using shRNA, affects cell proliferation, adhesion, migration, invasion and differentiation, in a way makes cells more metastatic (Kamal *et al.*, 2011, unpublished). For information about the biological consequences of reduced CSMD1 expression, please see appendix 2. Mechanism/s of CSMD1 action is still unknown. Understanding of how CSMD1 regulates cell behaviour will not only help identifying better measures for breast cancer prognosis, but it will also provide very important information of how we stop cancer metastasis. Our hypothesis is that using the expression level of CSMD1 combined with the expression pattern of one or more of its interactants will give better prediction of breast cancer patient's prognosis than using CSMD1 or any of its interactants alone.

Therefore, in this study, we are aiming to investigate mechanisms of CSMD1 functions by determining the signalling pathway/s, which CSMD1 adopts to perform its roles in regulating cell processes. Moreover, we will investigate whether using combinations of CSMD1 expression and components of its signalling pathway/s gives more prognostic information than using the expression of each marker individually.

Conducting this research project is very important for breast cancer patients because we are aiming to identify a prognostic tool which gives better separation between patients with different risks outcomes over the current prognostic markers. This helps clinicians to choose the appropriate treatment strategy and minimize number of patients having the inappropriate treatment. Moreover, this project will identify potential target for breast cancer therapy as components of the identified prognostic tool will be further investigated in the future for this purpose.

2. Literature review

In breast cancer, many traditional prognostic factors such as tumour type, lymph node status, etc have been used in the management of breast cancer. However, these factors are relatively crude measures which results in many patients being over treated or undertreated. Thus, many new markers have been identified, most of which are associated with mechanisms of proliferation (such as S-phase fraction, Thymidine labelling index, Ki67 and MIB1), cell growth (such as HER-2, Epidermal growth factor receptor and pS2) and others such as mutations in the p53 tumour suppressor gene, Bcl-2 and Cathepsin D. Nevertheless, except HER-2, none of these factors has been introduced to the routine management of breast cancer patients (Bundred, 2001; Colozza *et al*., 2005; Esserman *et al*., 2011; Schnitt, 2001).

After the development of the complementary DNA (cDNA) microarrays technology, researchers have identified prognostic signatures to stratify breast cancer patients. For example, one study defined a 70-gene prognostic signature for premenopausal breast cancer. Another study reported a different prognostic model based on 76 genes. To investigate whether gene based classification adds information to that provided by the classical factors, Eden *et al*. have found that models based on classical markers performed as well as gene based models (Edén *et al*., 2004; Paik, 2006; van 't Veer *et al*., 2002). Thus identifying novel prognostic factors is still a priority in the field of breast cancer research.

CSMD1 is a membrane protein with unknown function. The frequent deletion of 8p23 (the *CSMD1* locus), reported mutations in *CSMD1* and loss of *CSMD1* mRNA expression in cancers along with sharing similarities with proteins known to play a role in tumour progression, makes *CSMD1* an excellent candidate as a TSG (Farrell *et al*., 2008; Ma *et al*., 2009; Sun *et al*., 2001; Toomes *et al*., 2003).Thus, we have, for the first time, investigated the protein expression of CSMD1 in breast cancer and found that CSMD1 expression was reduced in 28.7% of cases. Moreover, CSMD1 expression was significantly associated with high tumour grade and overall survival (OS). Multivariate analysis showed that CSMD1 is an independent predictor of OS, which suggests its use as a new prognostic biomarker in breast cancer (Kamal *et al*., 2010).

Since CSMD1 is a membrane protein with a 63 amino acid cytoplasmic domain contains a tyrosine phosphorylation site, one could predict that CSMD1 serves as a receptor or a co-receptor for unknown ligands and is involved in signal transduction (Kraus *et al*., 2006; Sun *et al*., 2001). Thus, using shRNA, we have identified the functional roles of CSMD1 in cancer. Suppression of CSMD1 expression caused a significant increase in cell proliferation (Figure 1A), migration (Figure 1B) and invasion. It also resulted in loss of cell-matrix and cell-cell adhesion (Figure 1C). On the top of that, in a three dimensional (3D) breast duct matrigel model, reduced CSMD1 expression resulted in poorly differentiated acini which failed to form lumens (Figure 1D). Failure of lumen formation in mammary glands is a crucial structural alteration in ductal breast carcinomas (Pratap *et al*., 2009)(Kamal et., 2011, unpublished data). This suggests that CSMD1 is involved in a major signal

cascade/s which regulates cancer metastasis.

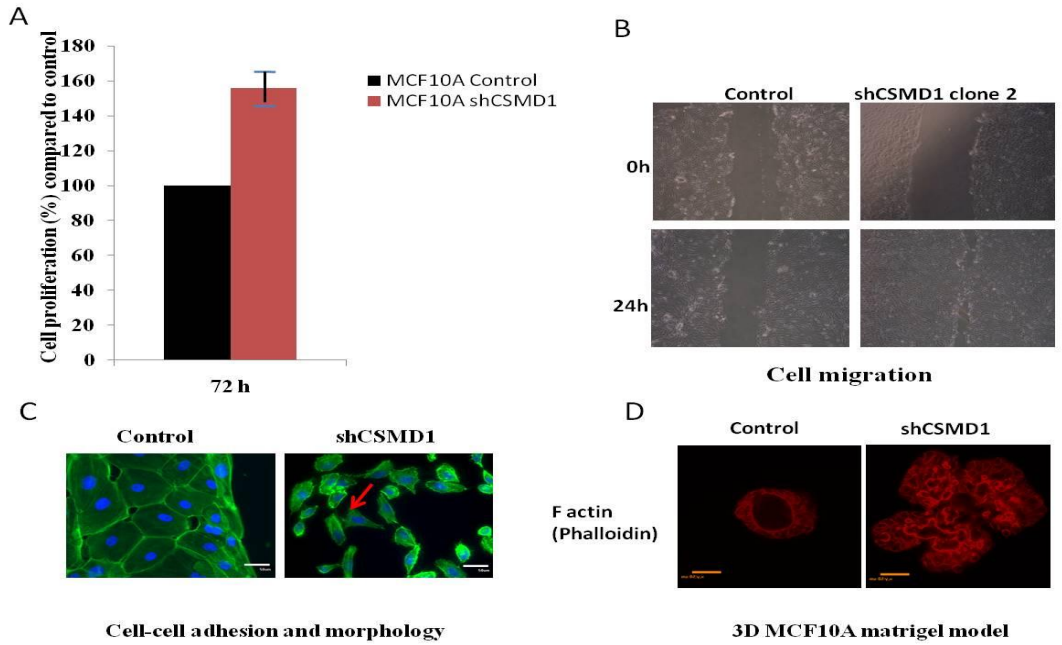


Figure 1: Reduced CSMD1 expression modulates MCF10A cells towards metastasis. A) Silencing of CSMD1 expression increased proliferation, after 72h, by 56% ($p<0.001$) compared to control cells. The percentages of cell proliferation, relative to shcontrols, were presented as the mean ±SD of at least three independent experiments. B) Wound healing assays revealed that loss of CSMD1 expression enhanced cell migration. Magnification, x4, scale bar 500µm. C) Phalloidin staining revealed that reduced CSMD1 expression resulted in misshapen cells lacking cell-cell contacts. Lots of cells are migrating with lamellipodia like protrusions (arrow). D) MCF10A 3D acini, at day 26, stained with phalloidin (red), Magnification, x40, scale bar 50µm. shCSMD1 acini are irregular in shape and heterogeneous in size with no lumen compared to control.

Because CSMD1 is a giant gene, spans over 2 Mb of genomic DNA, identifying its interactants using classical methodology (such as yeast two hybrid screens) is very difficult. So, in this project, CSMD1 partners will be investigated by determining the expression patterns of 48 genes in MCF10A shCSMD1 cell line, using qPCR arrays. Selected genes from this experiment will be further confirmed using western blotting and CoIP assays. The values of these interactants as drug targets to be further investigated in the future. However, in this project, the prognostic values of these interactants will be investigated in breast cancer patients. We expect that the use of CSMD1 expression combined with expression of one or more of its interactants (as a prognostic signature) will give a better separation between patients with different clinical outcomes, which will in turn help clinicians to provide better management for breast cancer progression.

3. Research Design and Methods

A- Investigating CSMD1 signalling pathways

Previously, we have down regulated CSMD1 expression in three cell lines; MCF10A, MDA-MB-435 and LNCaP, using shRNA. In this study, qPCR multiplex arrays (StellArray system, LONZA) for 48 genes will be designed to test their expression in MCF10A shCSMD1 cell line. These arrays are simple real-time PCR method provides the advantages of qRT-PCR plus high sensitivity and wider dynamic quantification range. GeneSieve System will be used for selection of gene candidates to be included in the array analysis. It extensively search tremendous number of abstracts for genes involved in cell processes, which are relevant to the gene under study (in our case; CSMD1). This will be followed by filtering thousands of genes using, for example, canonical signalling pathways, microarray data, downstream targets of transcription factors, etc. Finally, a list of the most pertinent genes will be created and included in the array.

Total RNA will be isolated from shCSMD1 cell lines using TriZol (Sigma) and first strand cDNA synthesis will be performed using Thermoscript and random hexamers (Promega). After confirmation of the quality of the cDNA, StellARray qPCR arrays will be loaded with sample-specific, SYBR Green master mix containing a chemically modified hot-start Taq polymerase (Takara, Japan). The array will be heat-sealed and run on a 7900HT Sequence Detection System (Applied Biosystems, Inc.) using default cycling parameters for 40 cycles.

Post-run data collection will involve the setting of a common threshold across all arrays within an experiment, exportation and collation of the Ct values. For comparison of the quantitative gene expression profiles of shControl cells and shCSMD1 cells, Ct values will be plotted. The linear regression and the resulting correlation coefficient will be calculated. For detailed expression analysis of the individual genes, Ct values will be analyzed by Global Pattern Recognition(GPR) data analysis tool, which utilizes a global normalization algorithm in which the data of each analyzed gene is normalized to that of every other gene without dependence on single gene normalization (Akilesh *et al*., 2003). Resulting fold changes of analyzed genes in shCSMD1 cells compared to shcontrol cells will be plotted and arranged in groups of fold changes. The expression level of the selected genes from the qPCR array experiments will be confirmed individually in the other two shCSMD1 cell lines using specific qPCR assays.

The protein expression of the selected genes will also be confirmed using western blotting analyses in all three cell lines. Cell lysates will be prepared and western blot analyses will be performed for the selected proteins along with beta-actin to normalize for loading and to allow comparisons of target protein expression to be made between all groups of cells.

Densitometric analysis will be performed using ImageJ software. For each western blot film, the densities of all bands will be measured and normalized to the background level and the change in the expression level of each protein will be calculated by normalization to the expression level of B actin of the correspondent cell group. One of the western blot limitations is that not all proteins have antibodies suitable for western. If we have this problem we will test the protein expression of these proteins using IHC or immunoflourescence.

To test physical interactions between CSMD1 and the selected proteins, CoIP assays will be performed. CSMD1 protein complexes will be precipitated by adding anti-CSMD1 antibody to a precleared cell lystaes. After eluting bound proteins in SDS sample buffer, western blotting analyses will be done using antibodies against proteins under study. In case any of the used antibodies does not work in CoIP, we plan to perform Co immunostaining.

B- Evaluating the prognostic value of CSMD1 signalling pathway/s in breast cancer patients.

The expression patterns of CSMD1 affected proteins will be investigated in a hundred breast cancer FFPE samples using IHC. IHC is an excellent technique that is able to show the exact site of a specific protein within the tissue. However, IHC technique may not be appropriate for all proteins in this study, in such a case we will use alternative methods such as ELIZA or western blotting (where possible).

Patient`s consent forms will be signed and paraffin blocks will be retrieved from the archives of the pathology department, school of medicine, Mansoura University, Egypt. IHC staining of proteins under study will be performed using the horse radish peroxidise enzyme (ChemMate DAKO Envision reagent, DakoCytomation) and its substrate diaminobenzine (DAB) chromogen. In each batch of IHC, some sections will be incubated with pre-immune serum instead of primary antibody and used as negative controls. Appropriate tissue sections will be used as positive controls.

Slides will be assessed depending on the proportion of the positively stained cells present on the section as a percentage. Expression status will be dichotomized and the cut-off point which shows the biggest difference between patients groups with regard to survival will be chosen. Cases will be scored blindly and independently by the principal investigator and a specialist consultant histopathologist. Discordant results will be re-evaluated jointly to reach consensus. Correlations between the expression patterns of all proteins under study will be performed. Correlations between expression pattern of each protein and patient`s clinicopathological data will be statistically analysed. Pearson correlation will be used to correlate IHC staining as a percentage of the positive cells versus tumour size, age of patient, number of positive nodes, and survival data as continuous variables. To perform survival curves, Kaplan–Meier method will be used, for each marker individually and for all possible combinations of markers. Differences in survival among groups will be analyzed by the log-rank test. Cox regression model will be established to confirm the Kaplan Meir results. To test whether any of the markers or combinations of them is an independent prognostic factor, a stepwise multivariable Cox regression model will be adjusted.

4. Project plan

4.1 Problem statement

4.1.1 Goal

The prognosis and clinical management of patients with breast cancer is commonly determined by traditional clinical and pathological factors such as tumour type, lymph node status, tumour grade, etc. Nevertheless these markers are relatively crude measures which results in many patients being over treated or undertreated. Although many new markers have been identified, including the recently identified gene expression signatures, none of these factors, except HER-2, has been introduced to the routine management of breast cancer patients. Thus a prognostic factor or signature which could provide accurate prediction of breast cancer prognosis is still missing. Therefore, our long-term goal is to identify novel biomarkers with superior prognostic values over the current markers being used in the clinic.

4.1.2 Rational and hypothesis

The rationale for the proposed research is that cancer progression and metastasis are tightly regulated by oncogenes and tumour suppressor genes, these genes provide rich sources of novel biomarkers and also novel targets for cancer therapy. Recently, we have identified a novel TSG (CSMD1) as an independent prognostic marker in breast cancer (Kamal *et al*., 2010). The actual function of CSMD1 in cancer was unknown until we have shown that loss of CSMD1 expression, using shRNA, seems to enhance epithelial mesenchymal transition (EMT) by adversely affecting major cell processes such as cell adhesion, proliferation, migration and invasion (Kamal *et al*., 2011, unpublished data). Since CSMD1 has not been reported as a member of any of the know signalling cascades, which regulate these processes, we hypothize that CSMD1 is implicated in novel metastatic pathways. Members of these pathways may represent rich sources of drug targets and biomarkers. Thus, the central hypothesis of this study is that expression pattern of CSMD1 combined with the expression status of one or more of its partners will represent a strong prognostic signature in breast cancer.

4.1.3. Objectives

Therefore, the specific objectives of this proposal are; first to determine CSMD1 signalling pathway/s. Second, is to examine the prognostic values of these pathways in breast cancer patients.

To achieve the first goal we will test the expression profile of 48 genes (relevant to cell processes, which are affected by CSMD1), using qPCR arrays in shRNA CSMD1 cells. The protein expression of the most affected genes will be confirmed using western blotting analyses. Possible interactions between these proteins with each other or with CSMD1 will be tested using CoIP assays.

To achieve the second goal, we will investigate the expression pattern of the selected proteins from the first goal, in breast cancer patients using IHC. Statistical analyses will be conducted to evaluate the prognostic values of CSMD1 expreesion combined with expression of its partners.

4.2 Project description

We will pursue these studies in two specific aims as shown in the diagram in appendix 1

Aim 1. Investigating CSMD1 signalling pathways

The expression level of at least 48 genes will be examined in MCF10A shCSMD1 stable cell line, which we have previously established, using qPCR multiplex arrays (StellArray system, LONZA). GeneSieve System will be used for selection of gene candidates to be included in the array analysis. Selection of these genes will be based on the functional relevance to CSMD1. The most affected genes, which showing the highest change in their expression between CSMD1 positive cells and CSMD1 negative cells (termed CSMD1 affected genes), will be identified. The expression level of these genes will be confirmed using specific qPCR assays in the other two shCSMD1 stable cell lines; MDA-MB-435 and LNCaP. Moreover, the protein expression of the selected genes will be tested in the three cell lines using western blotting assays. To examine physical interactions between selected proteins with each other or with CSMD1, CoIP experiments will be conducted.

Our working hypothesis for this aim is that CSMD1 is implicated in one of the major signalling pathways which regulate metastasis, because our strong preliminary results showed that loss of CSMD1 expression modulate cell behaviour towards metastasis.

Generally I have very strong experience working in CSMD1. I have investigated, for three years, the role of CSMD1 in cancer. This specific aim of the project will be achieved by performing lots of cDNA synthesis, qPCR, western blotting and immunoprecipitation techniques. During my phD study, I have mastered cDNA synthesis and qPCR experiments while I was establishing CSMD1 shRNA stable cell lines. In my postdoctoral post, I have gained intensive experience in western blotting and immunoprecitpitaion assays. Thus I am confident enough to accomplish this specific goal.

Aim 2. Evaluating the prognostic value of CSMD1 signalling pathway/s in breast cancer patients.

In this specific aim we will test the hypothesis that the expression pattern of CSMD1 combined with the expression pattern of one or more of its signalling cascades partners will give more prognostic information and better separation between breast cancer patients with good prognosis and those with poor prognosis.

After determining CSMD1 affected genes and (if possible) CSMD1 interactants (from specific aim 1), an intensive literature search will be done to investigate the importance of the affected genes in the progression of breast cancer. Proteins showing highest relevance to breast cancer progression will be involved in the study. After obtaining patient`s consent forms, paraffin blocks of a hundred breast cancer patient will be retrieved from the archives of the pathology department, school of medicine, Mansoura University, Egypt. The expression pattern of the selected proteins, along with CSMD1, will be investigated, using IHC. Comprehensive statistical analyses will be performed to test all possible correlations between the expression patterns of these proteins and patients clinicopathological data.

This specific aim of the project will be based on IHC technique and statistical analyses and it will mainly use breast cancer tissues. Previously, I have identified CSMD1 as an independent prognostic factor in breast cancer using IHC. This provided me with a solid experience in the field of prognostic factor studies and gave me intensive experiences in IHC and statistical analyses. Moreover I have gained strong background about the histology and pathology of the human breast. However, this phase of the project will be done in collaboration with a specialist histopathologist from the school of Medicine, Mansoura University, to guarantee high quality of the study. Thus, I am confident that this stage of the project will be achieved professionally and will reach highest standard.

4.3 Project significance

The proposed project is innovative because it investigates, for the first time, CSMD1 signalling pathways. At the completion of this project, we expect that the work proposed in Aim 1 will identify CSMD1 partners in the metastatic signalling pathways, which may be novel pathaways. Identification of these interactants will provide the research community with novel potential targets for cancer therapy which may lead to effective inhibition of cancer metastasis.

Based on our hypothesis that CSMD1 expression combined with the expression status of one or more of its interactants will give better prognostic information than using any of the markers alone, we expect work from aim 2 to identify a novel prognostic signature for breast cancer. This signature may give better separation between patients on the basis of their prognosis. Therefore, it will help clinicians to provide better management for breast cancer patients.

4.4 Risks

The first aim of this project includes qPCR and western blotting experiments using RNA and cell lystaes, which will be provided by our European collaborator. There is a risk that the RNA and / or cell lysates get degraded during the process of shipment. If this happens we will try every possible safe delivery method. In case we could not get high quality RNA and cell lysates, our collaborator will send us cells as frozen. Because there is no tissue culture facility in The faculty of Sciences, University of Benha, we are currently in the process of establishing collaborations with one of the well equipped tissue culture laboratories near Benha. For example: Cancer biology research Lab (University of Cairo), Oncology Diagnosis Unit (University of Ain Shams), Institute of Genetic Engineering (University of Menofya, Sadat city) or The city of Scientific Research and Technology applications (Borg Elarab, Alexandria). We will then grow up cells in one of these laboratories and isolate RNA and cell lysates as needed.

Also, as stated in the specific aims, the second aim of this project is to identify a prognostic signature for breast cancer using the protein expression of CSMD1 combined with the expression of its interactants. This study will be conducted by testing the expression pattern of these proteins in patients FFPE samples. FFPE samples will be provided by the department of pathology, Mansoura University. For any reason, we could not get enough samples or the quality of the samples are below the standards (in terms of sample preparation), or no enough clincopathological data, we plan to get samples from other institutions in Egypt. Thus√, we are currently establishing collaborations with other pathologists from different cancer centres in Egypt. If we could not get a high quality cohort of breast cancer samples from Egypt, we plan to establish collaborations with pathologists from around the globe.

7.5 Detailed Time Plan

Project Duration: (21 months) Number of Phases : 3 (including writing up)																								
Project Plan	Calendar																							
Project Phases	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20	M21	M22	M23	M24
Phase 1: Investigating CSMD1 signalling pathways	√	√	√	√	√	√	√	√	√	√	√	√	√											
Specific Objective 1: Identifying CSMD1 affected genes																								
Activity 1: Receiving RNA from our European counterpart, making up cDNA and testing its quality.	√	√																						
Activity 2: Selecting genes to be included in the qPCR arrays and placing the order (custom arrays take 6 weeks to arrive). After receiving them, performing qPCR arrays and analyzing their data to identify the most affected genes	√	√	√																					
Activity 3: Designing and optimizing specific qPCR assays for each one of the selected genes to be confirmed in all three cell lines.				√																				

Specific Objective 2: Testing protein expression of the selected genes in all three cell lines					√	√	√	√	√	√	√													
Activity 1: Optimizing western blotting assays for each protein in the three cell lines.					√	√	√	√	√	√														
Activity 2: Analayzing western blotting data and preparing a publication gel											√													
Specific objective 3: Testing physical interactions between the selected proteins.										√	√	√	√											
Activity 1: Optimizing CoIP experiments and testing all possible interactions between proteins under study. Analyzing data and preparing publication gels										√	√	√	√											
Phase 2: Evaluating prognostic values of CSMD1 signalling pathways	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√						
Specific Objective 1: Determining the expression pattern of all proteins under study in breast cancer tissues	√	√	√	√	√	√	√	√	√	√	√	√	√	√										

Activity 1: Collecting paraffin blocks from our pathologists collaborators, sectioning and stain them with haematoxylin and Eosin. Performing identification and histological characterization for all samples and preparing patients clinicopathological data.	√	√	√	√	√	√	√	√	√	√	√	√	√											
Activity 2: Optimizing IHC staining protocols for proteins under study and detect their expression pattern									√	√	√	√	√	√										
Specific objective 2: Evaluating the prognostic values of these proteins individually or in combinations.															√	√	√	√						
Activity 1: Scoring IHC staining and performing statistical analyses															√	√	√	√						
Phase 3: Data presentation																	√	√	√	√	√			

Activity 1: Travelling to the UK to meet with the European counterpart. The aim of the meeting is to discuss the progress of the project and its future directions and what the best way to publish the results. Also during the visit to the UK, data of phase 1 of the study will be presented in an international conference.																	√							
Activity 2: writing up results of the project for publication in peer reviewed international journals																	√	√	√	√				
Activity 3: writing up final report of the project																			√	√	√			

4.6 Resources

Phase 1 of this project involves mainly qPCR, western blotting and CoIP experiments. qPCR assays will use cDNA from shCSMD1 cell lines. Cells will be grown up in the tissue culture laboratory, in the Institute of Molecular Medicine (LIMM), University of Leeds, UK, by our partner. The RNA will be then isolated and sent to our laboratory in the faculty of Sciences, University of Benha, Egypt, where all the subsequent experiments will be set up. cDNA synthesis will need a water bath which is already available in our lab. All qPCR experiments will be set up in our laboratory and will be run in the (model of the machine to be determined) qPCR machine in the institute of genetic engineering, Sadat, Menofya. Consumable costs are requested to cover the costs of the qPCR arrays, primers, SYBRogreen master mix and the cDNA synthesis reagents. Western blotting assays and CoIP will be done in our lab. However a vertical electrophoresis and a western blotter are requested since our current ones are close to the end of their working life. Also costs are requested to cover the wide range of primary and secondary antibodies and general reagents for western blot and CoIP experiments.

Phase II of the project will involve mainly IHC staining. Although there is no established IHC facility, all these experiments will be done in our laboratory, department of zoology, faculty of sciences, Benha. To establish IHC facility in our lab, we already have a hot plate, microscope and oven. However we request a staining system, a pressure cooker for antigen retrieval and costs for the general reagents of the IHC staining. Establishing this facility in our department will not only allow this project to be achieved but also will allow other future projects to be done by us or by our colleagues in the department.

Costs for travelling to the UK are requested. The aim of this visit is to meet with the European partner and discuss the project progress. Also, during this visit, an international conference will be attended to present our results and get feed backs.

4.8 Expected outputs

The expected output of phase I is as following; the qPCR experiments will determine CSMD1 most affected genes, which will be confirmed by western blotting assays. The CoIP experiments will test physical interactions between the selected proteins. Thus, if these experiments showed positive results, it may unravel novel signalling cascades which regulate cancer metastasis. This will lead cancer researchers to further investigate how we could benefit from these pathways to either stop cancer metastasis or to provide better management of cancer progression.
The expected output of aim II is determination of a useful prognostic signature, which may be a combination of CSMD1 expression and the expression of one or more of its interactants. We expect this prognostic signature to provide better prediction of breast cancer prognosis over the current biomarkers. This because CSMD1 alone was shown to be an independent prognostic factor, and our preliminary data showed that it has functional roles in a wide range of cancer related cell processes. Together with the fact that CSMD1 has not been stated to be a member of the major signalling pathways, this suggests that CSMD1 is implicated in a novel multifunctional cancer regulating signalling cascade. So combining the expression status of CSMD1 with the expression status of other components of these pathways is expected to provide prognostic information superior to that provided by other traditional biomarkers.

4.9 Project evaluation

4.9.1. Process evaluation

The main measurable indicator for phase I of this project is identifying CSMD1 affected genes and CSMD1 interactants (if the CoIP experiments showed positive results) by month 13. Identification of these genes must be based on solid reproducible and confirmed experiments. Evaluating the quality of these experiments will be by the standard deviation which is an indicator for the assays reproducibility. Also all experiments must include the appropriate positive and negative controls and have been repeated at least three times. By the end of phase 1 of the project, I plan to visit Leeds Institute of Molecular Medicine, Leeds, UK. During this visit all the experiments which have been done to that date will be discussed and the quality of the data will be evaluated by our collaborator and other scientists in the institute, who are experts in the field. Also, while I am in the UK, I plan to attend and present these data in an international conference (to be determined) and feedbacks from world leads in this field will be considered.

The measurable indicators for phase II progress is as following; 1) getting each of the selected proteins stained, by IHC, in a hundred FFPE breast cancer tissues by month 14, 2) All IHC staining are scored and reviewed by our pathologists by month 16, 3) The prognostic values of these proteins in breast cancer patients are statistically evaluated, and reviewed by both our European counterpart and our pathologists, by month 18. Quality of the IHC staining will be evaluated by including positive and negative controls in each batch of staining.

4.9.2. Product evaluation

The end product of this project is supposed to be determining CSMD1 interactants and or CSMD1 affected genes and evaluating their prognostic importance in breast cancer patients. So by the end of this project we aim to have identified a group of proteins as the most affected proteins by CSMD1. Also we will have information about their roles in predicting breast cancer prognosis. These objectives will be evaluated by presenting their results in internal and external conferences along with publishing them in peer reviewed international journals, together with the final report of the project.

5. Dissemination of research findings

Results of this project will be disseminated internally and externally in many ways. First, oral presentations will be delivered at least twice in our departmental seminars; once after the completion of phase I of the study and the second will be after achieving the project. Second, data will be also present in our collaborators departments. We also intend to attend relevant internal conferences and present our results. Moreover, after the first phase of the study we are aiming to attend an international conference and present our data (as a poster or a talk, it depends on the conference). By the end of the study, data will be written up and sent for publication in peer reviewed international journals. In case this project results in findings in the interest of the public we will then disseminate our findings through media.

6. References

Akilesh, S., Shaffer, D. J., and Roopenian, D. (2003). Customized Molecular Phenotyping by Quantitative Gene Expression and Pattern Recognition Analysis. Genome Research *13*, 1719-1727.

Bundred, N. J. (2001). Prognostic and predictive factors in breast cancer. Cancer Treatment Reviews *27*, 137-142.

Colozza, M., Azambuja, E., Cardoso, F., Sotiriou, C., Larsimont, D., and Piccart, M. J. (2005). Proliferative markers as prognostic and predictive tools in early breast cancer: where are we now? Ann Oncol *16*, 1723-1739.

Edén, P., Ritz, C., Rose, C., Fernö, M., and Peterson, C. (2004). "Good Old" clinical markers have similar power in breast cancer prognosis as microarray gene expression profilers. European Journal of Cancer *40*, 1837-1841.

Esserman, L. J., Moore, D. H., Tsing, P. J., Chu, P. W., Yau, C., Ozanne, E., Chung, R. E., Tandon, V. J., Park, J. W., Baehner, F. L*., et al.* (2011). Biologic markers determine both the risk and the timing of recurrence in breast cancer. Breast Cancer Research and Treatment.

Farrell, L., Crimm, H., Meeh, P., Croshaw, R., Barbar, T., Vandersteenhoven, J., Butler, W., and Buckhaults, P. (2008). Somatic mutations to CSMD1 in colorectal adenocarcinomas. Cancer Biology & Therapy *7*, 609-613.

Kamal, M., Shaaban, A. M., Zhang, L., Walker, C., Gray, S., Thakker, N., Toomes, C., Speirs, V., and Bell, S. M. (2010). Loss of CSMD1 expression is associated with high tumour grade and poor survival in invasive ductal breast carcinoma. Breast Cancer Res Treat.

Kraus, D. M., Elliott, G. S., Chute, H., Horan, T., Pfenninger, K. H., Sanford, S. D., Foster, S., Scully, S., Welcher, A. A., and Holers, V. M. (2006). CSMD1 Is a Novel Multiple Domain Complement-Regulatory Protein Highly Expressed in the Central Nervous System and Epithelial Tissues. J Immunol *176*, 4419-4430.

Ma, C., Quesnelle, K. M., Sparano, A., Rao, S., Park, M. S., Cohen, M. A., Wang, Y., Samanta, M., Kumar, M. S., Aziz, M. U*., et al.* (2009). Characterization CSMD1 in a large set of primary lung, head and neck, breast and skin cancer tissues. Cancer Biol Ther *9;8*, 29-38.

Paik, S. (2006). Molecular profiling of breast cancer. Current Opinion in Obstetrics and Gynecology *18*, 59-63 10.1097/1001.gco.0000192970.0000152320.0000192929.

Pratap, J., Imbalzano, K. M., Underwood, J. M., Cohet, N., Gokul, K., Akech, J., van Wijnen, A. J., Stein, J. L., Imbalzano, A. N., Nickerson, J. A*., et al.* (2009). Ectopic Runx2 Expression in Mammary Epithelial Cells Disrupts Formation of Normal Acini Structure: Implications for Breast Cancer Progression. Cancer Res *69*, 6807-6814.

Schnitt, S. J. (2001). Traditional and newer pathologic factors. J Natl Cancer Inst Monogr, 22-26.

Sims, A. H., Howell, A., Howell, S. J., and Clarke, R. B. (2007). Origins of breast cancer subtypes and therapeutic implications. Nat Clin Prac Oncol *4*, 516-525.

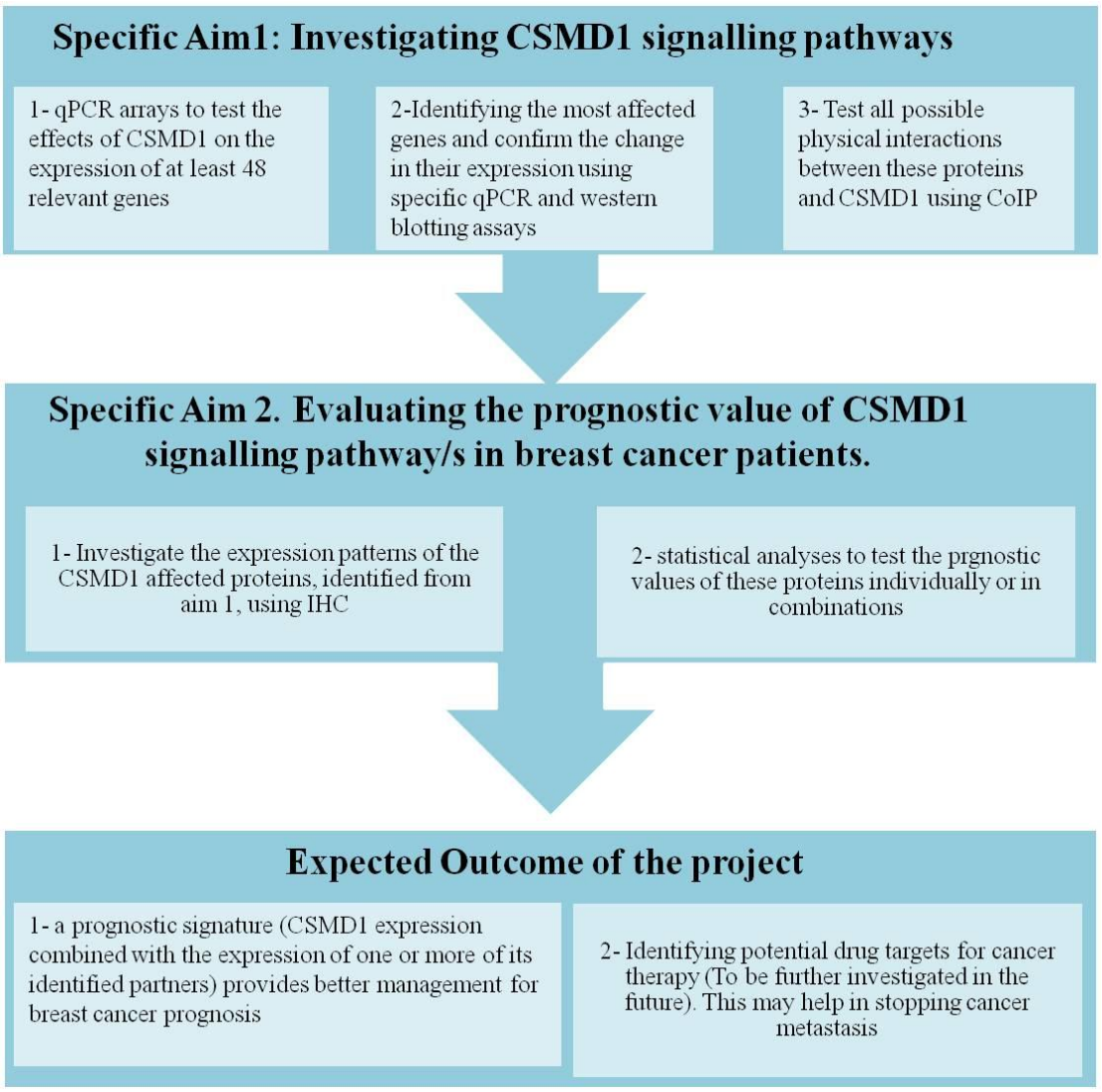
Sun, P. C., Uppaluri, R., Schmidt, A. P., Pashia, M. E., Quant, E. C., Sunwoo, J. B., Gollin, S. M., and Scholnick, S. B. (2001). Transcript Map of the 8p23 Putative Tumor Suppressor Region. Genomics *75*, 17-25.

Toomes, C., Jackson, A., Maguire, K., Wood, J., Gollin, S., Ishwad, C., Paterson, I., Prime, S., Parkinson, K., Bell, S*., et al.* (2003). The presence of multiple regions of homozygous deletion at the CSMD1 locus in oral squamous cell carcinoma question the role of CSMD1 in head and neck carcinogenesis. Genes, Chromosomes and Cancer *37*, 132-140.

van 't Veer, L. J., Dai, H., van de Vijver, M. J., He, Y. D., Hart, A. A. M., Mao, M., Peterse, H. L., van der Kooy, K., Marton, M. J., Witteveen, A. T*., et al.* (2002). Gene expression profiling predicts clinical outcome of breast cancer. Nature 415, 530-536.

7 appendices

Appendix 1: Diagram to illustrate the two main specific aims of the proposed project



Appendix 2: Preliminary data (unpublished)

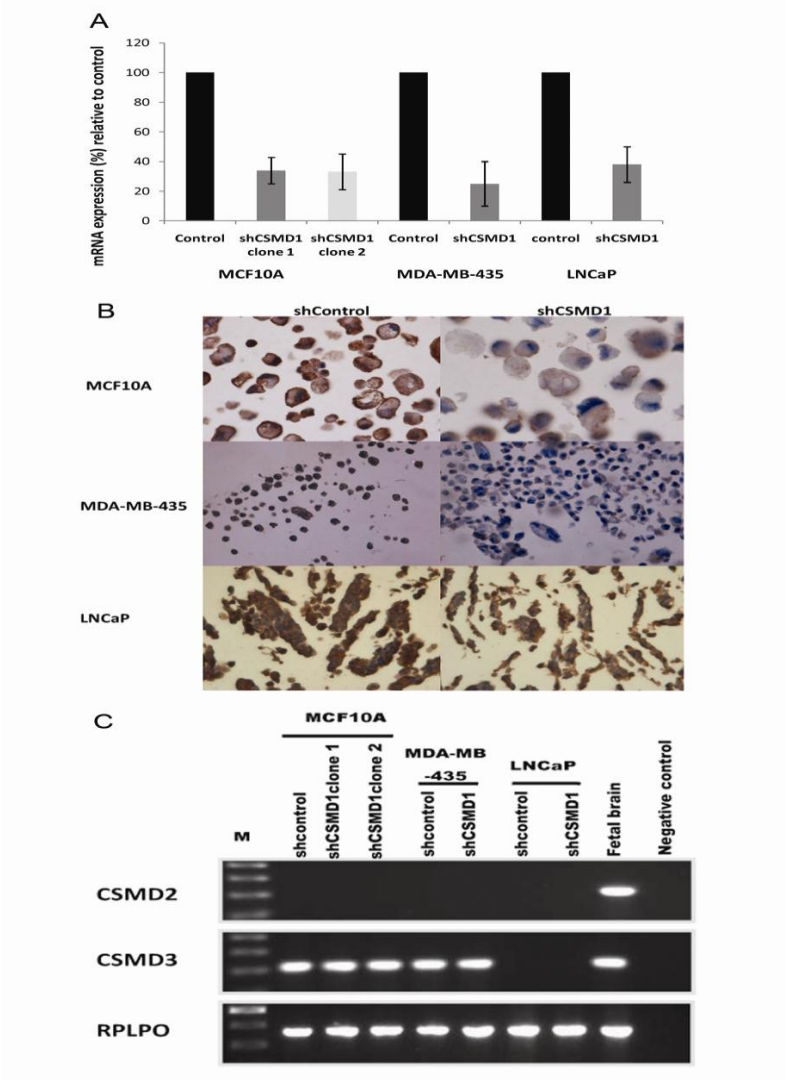


Figure Preliminary 1. Confirmation of CSMD1 expression knockdown A) qRT-PCR, using the relative standard curve method confirmed the shRNA vectors reduced *CSMD1* mRNA expression. MCF10A shCSMD1 clones 1 and 2, exhibited 66% and 67% reduction in CSMD1 expression, respectively. While, the knockdown efficiencies in MDA-MB-435 and LNCaP were 75% and 62%, respectively. The percentages of *CSMD1* expression (relative to shcontrols) were presented as means ±SD of at least 3 independent experiments. B) IHC using chicken anti-CSMD1 antibody on formalin fixed paraffin embedded cells revealed that shCSMD1cells exhibit lower CSMD1 protein expression levels compared to shcontrols. Magnification, x40. C) Confirmation of specificity of shRNA CSMD1 vectors. RT-PCR for *CSMD2* and *CSMD3* in shCSMD1 cells did not detect any changes in the

expression of these genes or the housekeeping gene *RPLPO*. Foetal brain cDNA was used as a positive control and water as a negative control. M = size standard marker.

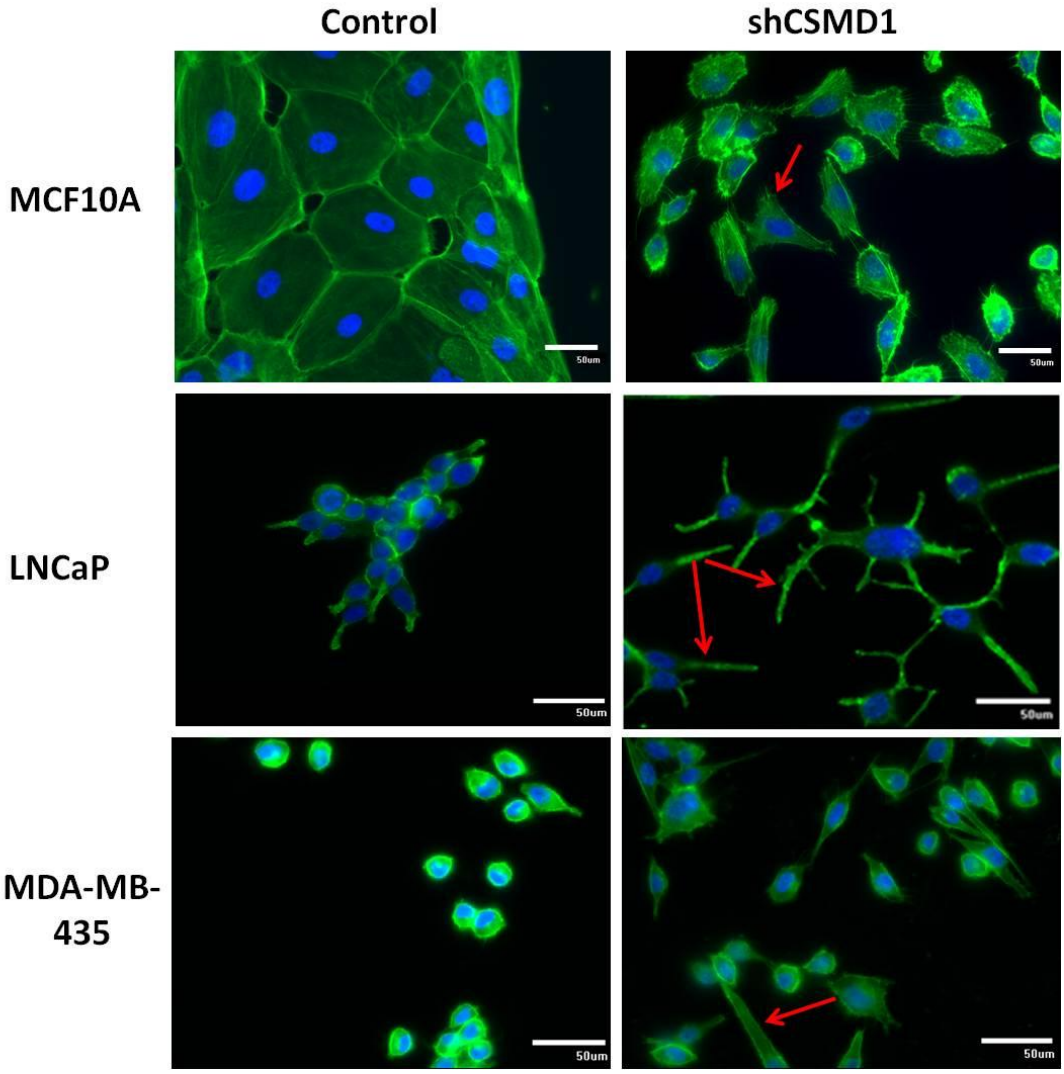


Figure Preliminary 2. Loss of CSMD1 expression disrupted cell morphology.
Phalloidin staining revealed that, in all cell lines, reduced CSMD1 expression resulted in misshapen cells lacking cell-cell contacts. Lots of MCF10A shCSMD1 cells are migrating with lamellipodia like protrusions (arrow head). MDA-MB-435 and LNCaP shCSMD1 cells exhibit long filopodia like protrusions (arrows). Nuclei are stained with DAPI. Magnification, x40, scale bar 50µm**.**

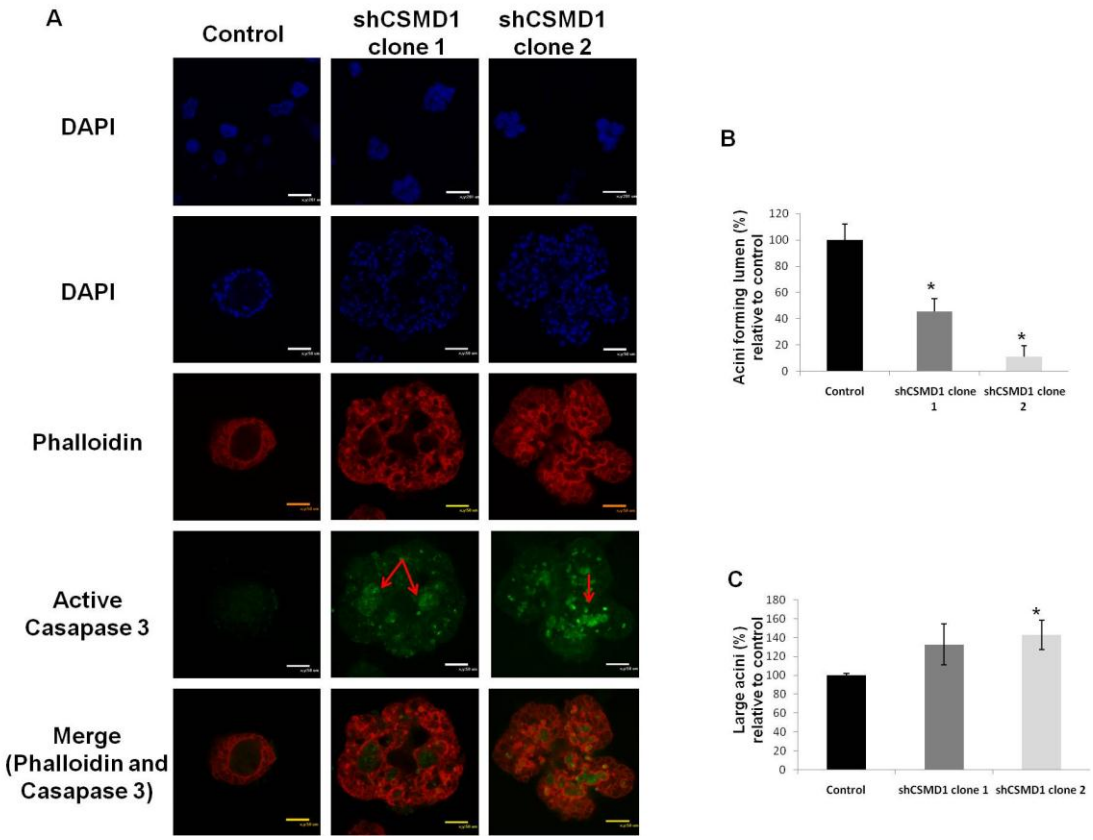


Figure Preliminary 3. Silencing CSMD1 expression disrupted mammary acini morphology and inhibited lumen formation in the MCF10A 3D model. A) Acini, at day 26, stained with DAPI (blue), phalloidin (red), and active caspase 3 antibody (green). Magnification, x40, scale bar 50µm. shCSMD1 acini are irregular in shape and heterogeneous in size with no lumen. Control acini showed weak staining for active caspase 3, while, lumens of shCSMD1 acini exhibit strong staining (arrows). B) The percentages of lumen forming acini, relative to shcontrol, were presented as the mean ±SD of at least three independent experiments. Reduced CSMD1 expression resulted in 55% (p=0.045)** and 90% (p=0.008)*** decrease in the percentage of lumen forming acini in MCF10A shCSMD1 clones 1 and clone 2, respectively. C) The percentages of large acini, relative to shcontrol, were presented as the mean ±SD of at least three independent experiments. Reduced CSMD1 expression resulted in 33% (p=0.1) and 43% (p=0.03)* increase in the percentage of large acini in MCF10A shCSMD1 clones 1 and 2, respectively.