Published in final edited form as:
*J Proteome Res*. 2011 February 4; 10(2): 447–458. doi:10.1021/pr100533k.

# 2D Difference Gel Electrophoresis Analysis of Different Time Points during the Course of Neoplastic Transformation of Human Mammary Epithelial Cells

**J. Tyson DeAngelis**[†], **Yuanyuan Li**[†], **Natalie Mitchell**[†], **Landon Wilson**[‡], **Helen Kim**[‡,§,||,⊥], and **Trygve O. Tollefsbol**[*,†,‡,||,#]
Department of Biology, UAB Comprehensive Cancer Center, UAB Center for Nutrient-Gene Interaction, UAB Nutrition Obesity Research Center, Department of Pharmacology and Toxicology, and Center for Aging, University of Alabama at Birmingham, Birmingham, Alabama 35294-1170, United States

## Abstract

Cell culture models of oncogenesis that use cellular reprogramming to generate a neoplastic cell from a normal cell provide one of the few opportunities to study the early stages of breast cancer development. Human mammary epithelial cells (HMECs) were induced to undergo a neoplastic transformation using defined genetic elements to generate transformed HMECs (THMECs). To identify proteins that displayed significantly different levels of abundance at three consecutive time points in oncogenesis over an 80 day period, protein extracts were analyzed by two-dimensional difference gel electrophoresis (2D-DIGE). Nine proteins were found to be significantly different in abundance: keratin 1, keratin 7, heat shock protein 4A-like, t-complex protein 1, stathmin, gelsolin, FK506 binding protein 5, ribosomal protein P0, and maspin. Keratin 7 and maspin displayed a linear down-regulation over 80 days. All of these proteins have been shown to be involved in the maintenance of a metastatic state including cytoskeletal modifications and motility. We conclude that, following neoplastic induction, THMECs display an early and progressive increase in metastatic potential. Further investigations into the function and regulatory mechanisms of these proteins will provide an unparalleled understanding of the initial states through which a breast cancer cell transitions following acquisition of the genetic abnormalities required for oncogenesis.

### Keywords

2D-DIGE; neoplastic transformation; human mammary epithelial cells; maspin; metastasis; breast cancer; Ingenuity Pathway Analysis



[*]To whom correspondence should be addressed: Trygve O. Tollefsbol, Ph.D., D.O.; tel, 205-934-4573; fax, 205-975-6097; trygve@uab.edu.
[†]Department of Biology.
[‡]UAB Comprehensive Cancer Center.
[§]UAB Center for Nutrient-Gene Interaction.
[||]UAB Nutrition Obesity Research Center.
[⊥]Department of Pharmacology and Toxicology.
[#]Center for Aging.

**Supporting Information Available:** A schematic of the experimental design. This material is available free of charge via the Internet at http://pubs.acs.org.

## Introduction

One of the most important yet least understood aspects of cancer biology is oncogenesis, or the process by which a normal, somatic cell transitions to an oncogenic state. There currently is no method to study this process *in vivo*; however, a procedure known as neoplastic reprogramming confers the ability to study this process *in vitro*. First used to generate cancer cells from normal fetal lung fibroblasts, this procedure has also been used to generate breast cancer cells from human mammary epithelial cells (HMECs).[1,2] HMECs are normal, somatic cells that display a finite lifespan and lack the ability to form tumors. Neoplastically transformed HMECs display most if not all of the characteristics of breast cancer cells, including the ability to form tumors in nude mice.

Two general types of cellular reprogramming exist, reprogramming to a pluripotent or stem state, and reprogramming to a neoplastic state.[1,3] These two groups are not mutually exclusive as they both use some of the same genomic regions to achieve the desired induced state and possess many of the newly acquired traits in common. The therapeutic importance of pluripotent reprogramming is very high, whereas neoplastic programming has little therapeutic value. The true value of neoplastic reprogramming lies in the ability to monitor the mechanisms driving the transition of a normal cell to a neoplastic cell in real-time. Neoplastic reprogramming in human cells using defined genetic elements was first discovered by Weinberg et al.[1] In 1999, Weinberg and colleagues published a landmark procedure illustrating that normal human cells can be induced to undergo a neoplastic transformation through retroviral-mediated addition of three defined genetic elements: Simian Virus 40 Early Region (*SV40 ER*), exogenous human telomerase reverse transcriptase (*hTERT*), and an activated, oncogenic form of hRAS (*hRAS-V12*).[1] Addition of each of the three genes either activates or inactivates specific pathways in a manner similar to how these pathways become dysregulated during oncogenesis *in vivo*. The addition of *SV40 ER* inactivates the defense mechanisms of the cell by sequestering p53 and pRB.[4] Stable expression of exogenous *hTERT*, the catalytic component of telomerase, promotes an increase in telomerase activity and provides the cell with the ability to replicate indefinitely.[5] Lastly, the introduction of activated hRAS mimics a sustained supply of growth signals and causes a consistent activation of one of the major pathways responsible for cellular division and the acquisition of a malignant phenotype.

The ability to neoplastically transform normal human cells provides the opportunity to monitor the very early events of tumorigenesis in real time. For example, we have previously shown that neoplastically transformed human fetal lung fibro-blasts display a progressive increase in both mRNA and protein expression of members of the DNA methyltransferase (DNMT) family of enzymes following antibiotic selection of the third construct, *hRas-V12*.[6] It is common knowledge that the DNMTs play a pivotal role in cancer biology by modifying the normal patterns of DNA methylation.[7] In fully developed tumors, the progressive nature of the change in expression would not be observed; only a difference in expression would be seen. It is believed that the key to understanding oncogenesis lies in the ability to identify which genes possess this progressive nature and then to understand what is driving the oncogenesis-related progressive alterations in gene expression.

Two-dimensional difference in gel electrophoresis (2D-DIGE) is one of the best available methods for discovery-based proteomic research; it is a relatively low cost “-omic” assay, that is highly quantitative and extremely accurate. The ability to analyze 24 individual samples on 12 gels in one experimental run virtually eliminates experimental variance due to minor differences in reagents, conditions, and so forth. Additionally, inclusion of the same

internal standard on each gel confers the ability to quantitatively compare all 24 samples across the 2D gel data set.

We hypothesized that additional proteins, as important to the maintenance of the oncogenic state as the DNMTs, would also display this same type of progressive alteration in protein expression in a breast cancer model of oncogenesis. To test this hypothesis, we applied 2D-DIGE to identify differentially expressed proteins in neoplastically transforming HMECs in the first 80 days (d) after selection for *hRas-V12*. Eighteen spots were identified as differentially expressed, with 11 spots each containing a single protein. Of these 11 spots, two spots were duplicates, leaving a total of 9 unique identifications: keratin 1, keratin 7, heat shock protein 4A-like, t-complex protein 1, stathmin, gelsolin, FK506 binding protein 5, ribosomal protein P0, and maspin. These proteins have been shown to be involved in the loss of cellular rigidity and the gain of cellular motility, alterations that are associated with a metastatic phenotype. Of these 9 proteins, three (ribosomal protein P0, maspin and keratin 7) displayed a near perfect linear down-regulation. *SERPINB5*, the gene that codes for the maspin protein, is considered to be a metastasis suppressor gene (MSG) and the observed linear down-regulation over an 80 day period is consistent with a progressive trend toward a metastatic phenotype. Identification of genes that display a progressive alteration in protein expression confirmed our hypothesis and will provide the basis for future studies involving the regulation and function of both oncogenesis-related proteins.

## Materials and Methods

### Cell Culture

HMECs, at early passage, were obtained from Lonza (Basel, Switzerland) and grown in serum-free Mammary Epithelial Growth Medium (MEGM) (Lonza; Basel, Switzerland) without sodium bicarbonate at 37 °C and 0.1% $CO_2$. The omission of sodium bicarbonate allowed HMECs to be grown at a lower $CO_2$ level, which helped to maintain HMECs in the best possible condition. MEGM was changed daily and HMECs were subcultured at 75% confluency. HMECs at all stages of transformation were cultured according to the cell culture procedures described by Hammond and Stampfer.[8,9] Transient retrovirus was produced using Phoenix gag-pol 293 Human Embryonic Kidney cells (Φnx) created by Dr. Gary Nolan (Stanford University) and obtained through ATCC (Manassas, VA).[10] Φnx, MCF-7, and MDA-MB-231 cells were grown in DMEM supplemented with 10% heat-inactivated fetal bovine serum and 1% APS (Amphoterocin B, Penicillin, Streptomycin). Photographs of HMEC-SV40-hTERT (SHMEC) and Transforming HMECs (THMECs) were taken with a CoolPix5100 (Nikon; Melville, NY) at 40× magnification on an Eclipse TS100 microscope (Nikon; Melville, NY). Four cell lines of the MCF10 Isogenic Breast Cancer Metastatic Progression Model, MCF10A, MCF10AT, MCF10CAa.1α, and MCF10CAd.1, were acquired as a kind gift from Dr. Danny Welch (University of Alabama at Birmingham). Cells were cultured as previously described.[11,12] MCF10A are the immortalized nontumorigenic cell line. MCF10AT cells are the tumorigenic, premetastatic cell line derived from MCF10A. MCF10CAa.1α and MCF10CAd.1 are the two metastatic clonal lines. Each of the metastatic variants were created from a separate carcinoma that was generated by tail vein injections of MCF10AT cells into severe combined immunocompromised mice.

### Plasmid Preparation

Each of the three retroviral plasmids, *pBabe-neo-SV40* (10891), *pBabe-hygro-hTERT* (1773), *pBabe-puro-RASV12* (1768) and the pseudotyping plasmid, *pCL-VSVG* (1733) were obtained from ADDGENE (Cambridge, MA) with each plasmid's respective ID number in parentheses.[1] Plasmids were prepared according to the following procedures:

Following bacterial culture and subsequent purification, purity of each plasmid was assessed by determination of the A260/A280 ratio. Once purity was verified, plasmid sequence integrity was verified using pBabe sequencing primer forward (5′-CTTTATC-CAGCCCTCAC-3′).

### Retroviral Production and HMEC Infection

Transient retrovirus was produced according to the following procedures: Φnx cells were plated on gelatin-coated 100 mm plates and grown to a confluency of 85%. For each retroviral infection, 15 $\mu$g of *pBabe* plasmid and 15 $\mu$g of VSV-G plasmid was transfected into Φnx cells using Virapack Transfection Kit (Stratagene; Cedar Creek, TX). At 3 h post-transfection, cells were washed with PBS and fresh media added. At 48 h post-transfection retroviral supernatant was filtered through a 0.45 $\mu$m filter and Polybrene added to a final concentration of 4 $\mu$g/mL. Retroviral supernatant (2 mL) was then applied to HMECs at 35% confluency. After 3 h, 3 mL of MEGM was added and infection was carried out for 48 h. Cells were washed with PBS at 48 h postinfection to remove all traces of retrovirus and fresh MEGM was added. Infected HMECs were then allowed to recover from infection for 2 d before beginning selection. For each infection, antibiotic selection was carried out according to Elenbaas et al.[2]

### Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR) and Western Blot

RT-PCR was used to confirm expression of *SV40-ER* and *exogenous hTERT* (*exo-hTERT*). Total RNA was harvested using an RNEASY kit (Qiagen; Valencia, CA). Purified RNA was then subjected to DNase1 treatment and RNA repurified. To prevent RNA degradation, immediately following DNase1 treatment and subsequent purification, 2 $\mu$g of RNA was reverse transcribed to cDNA using cDNA First Strand Synthesis kit (Invitrogen; Carlsbad, CA). PCR was performed using primers, specific to *SV40 Large T* (F, 5′-GCTTTGCAAA-GATGGATAAAG; R, 5′-ACTAAACACAGCATGACTC), *exo-hTERT* (F, 5′-GACACACATTCCACAGGTCG; R, 5′-GACTCGACACCGT-GTCACCTAC), and *Glyceraldehyde 3-phosphate dehydrogenase* (GAPDH) (F, 5′-GAGAGACCCTCACTGCTG; R, 5′-GATGGTA-CATGACAAGGTGC) cDNAs.[1] Electrophoresis was carried out at 100 V. Bands were visualized under UV light and analyzed with Kodak Digital Science Software (Kodak; Rochester, NY). Western blotting of protein extracts was used to determine stable expression of hRASV12. Nuclear protein extracts were isolated using NE-PER Nuclear Extraction Kit (Thermo Scientific; Rockford, IL). Protein concentration was determined with the Bradford method of protein quantification using the BioRad Protein Assay (BioRad; Hercules, CA). Nuclear protein extract (20 $\mu$g) was separated by electrophoresis at 100 V, transferred to a nitrocellulose membrane, and blocked in 5% dry milk in Tris buffered saline solution with 1% Tween (TBST) overnight. Primary antibody incubation was carried out overnight at 4 °C using mouse monoclonal antibodies specific to hRas (sc-53958). Membranes were then probed with goat anti-mouse IgG-HRP (sc-2005). Protein bands were visualized using enhanced chemiluminescence (ECL) reagents (Thermo Scientific; Rockford, IL). All antibodies used in immunoblotting were obtained from Santa Cruz Biotechnology (Santa Cruz, CA) with catalog numbers listed in parentheses.

### Telomeric Repeat Amplification Protocol (TRAP)

A TRAP assay using TRAPeze Kit (Millipore Corporation; Billerica, MA) was performed to confirm activation of telomerase activity as a result of stable expression of *exo-hTERT* as previously described.[13] TRAP assay was performed according to procedures described in the *TRAPeze* manual. Total protein extracts were harvested using CHAPs lysis buffer. Protein concentration was determined as described above and 250 ng of protein extract was used for each reaction. Heat-inactivated samples were incubated at 85 °C for 10 min. Protein

extract from MDA-MB-231 cells served as a positive control and one reaction lacking protein extracts served as a negative control. Samples were incubated at 30 °C for 30 min, followed by a 33 cycle PCR. PCR conditions were followed according to the kit manual. Reaction products (25 $\mu$L) were loaded onto a 12.5% nondenaturing PAGE and electrophoresed for 1.5 h at 300 V. Gels were stained with SYBR Green (Lonza; Basel, Switzerland) and visualized under UV radiation.

### THMEC Xenografts

Female Balb/c nude mice, 4–6 weeks of age, were obtained from Charles River Laboratories (Wilmington, MA). THMECs at 40 d post-hRas introduction were injected subcutaneously in the left flank of each mouse. For each injection, 0.1 mL of cell culture medium containing $1 \times 10^6$ cells was mixed with an equal volume of Matrigel and the entire 0.2 mL was injected. Mice were sacrificed at 6 weeks postinjection and average tumor volume ($mm^3$) was calculated as follows: tumor volume (TV) on day 42 = $(a \times b^2)/2$, where $a$ is the largest diameter (in mm) and $b$ is the diameter (in mm) perpendicular to $a$.

### Experimental Design

We chose three points in the neoplastic progression to assess proteomic differences. These three time points in the neoplastic progression were as follows: SHMEC, THMEC 40 d post-hRAS, and THMEC 80 d post-hRAS. SHMEC cells are HMECs stably expressing *SV40ER* and exo-*hTERT* and represent premalignant breast cells. THMECs are HMECs stably expressing all three genetic elements, *SV40ER*, exo-*hTERT*, and hRAS-V12. For each time point, 4 biological replicates were generated for a total of 12 samples. For THMECs, two independent samples were generated for each time point (40 and 80 days) for each of two separate THMEC clonal lines generated from the same SHMEC clonal line. A schematic of the experimental design is shown in Supplementary Figure 1. THMECs at 80 days were derived from the same cell populations as those at 40 days. Four biological replicates of SHMEC were generated from one SHMEC clonal line, which was the same SHMEC cell line used to generate both THMEC cell lines. Experiments were designed in this fashion to optimize a balance between a sufficient number of biological replicates and a minimalization of variation and sample number. Finally, each sample was dye-swapped for a final $n = 8$ for each time point.

### RNA and Protein Extraction for Real-Time PCR and 2D-DIGE

RNA, protein, and DNA were extracted using Tri Reagent (Ambion, Austin, TX) from HMEC-SV40-hTERT, THMEC 40 d post-hRas transfection, THMEC 80 d post-hRas, MCF10A, MCF10AT, MCF10CAa.1α, and MCF10CAd.1 following the manufacturer's protocol. For each sample, 4 biological replicates were generated as described above. Following the final ethanol wash, the RNA pellet was resuspended in 100% deionized formamide and stored at −80 °C.[14] Protein pellets were stored in 100% ethanol at −20 °C.

### Preparation of Protein Extracts for 2D-DIGE and CyDye Labeling

Preparation of protein extracts and 2D-DIGE experiments were carried out as previously described by Kim et al.[15] Precipitated protein from each cell sample was resuspended in isoelectric focusing (IEF) buffer and protein concentrations were determined using the 2D Quant Kit (GE Healthcare, Piscataway, NJ). All three CyDyes were obtained from GE Healthcare (Piscataway, NJ). To eliminate dye-specific bias, all samples were dye-swapped by labeling with both Cy3 and Cy5. An internal standard (IS) was generated by pooling 50 $\mu$g of each of the 12 samples and labeling all of this mixture with 200 pmol of Cy2. Protein (50 $\mu$g) for each sample was labeled according to the manufacturer's instructions with 200 pmol of either Cy3 or Cy5. The Cy3 and Cy5 sample for each gel can be viewed in Table 1.

### 2D Difference Gel Electrophoresis-First Dimension Separation

For each 2D-DIGE gel, 50 $\mu$g of a Cy3-labled sample, 50 $\mu$g of a Cy5-labled sample, and 50 $\mu$g of the IS were pooled and the final volume brought up to 110 $\mu$L with IEF buffer before diluting 1:1 with rehydration buffer (7 M urea/2 M thiourea/4% CHAPS/40 mM Tris-HCl, pH 8.8) containing 1% IPGphor ampholytes, p*I* range 4–7 (GE Healthcare; Piscataway, NJ) and 30 mM dithiothreitol (DTT). Twelve 11.0 cm immobilized pH gradient (IPG) DryStrips pH gradient 4–7 (GE Healthcare; Piscataway, NJ) were rehydrated overnight for 20 h at room temperature (RT) in individual troughs in the IPG strip rehydrating chamber (GE Healthcare; Piscataway, NJ). The choice was made to use a pH 4–7 strip instead of the broader pH 3–10 strip based on previous experience of the authors and preliminary testing prior to the start of data collection.[15] Each DryStrip was placed gel-side down on top of the respective 220 $\mu$L sample and overlaid with mineral oil. The following day, DryStrips were laid gel-side down and focused on an Ettan IPGphor II (GE Healthcare; Piscataway, NJ) at RT overnight using the following protocol: 6 h at 500 V, 1 h at 1000 V, 2.5 h at 6000 V, 2 h at 6000 V, 20 h at 300 V. Wicks were changed every 30 min for the first 1.5 h. The next day, DryStrips were placed at −80 °C.

### 2D Difference Gel Electrophoresis-Second Dimension Separation

DryStrips were thawed at RT for 15 min; and then equilibrated in SDS-sample buffer (6 M urea, 75 mM Tris-HCL, pH 8.8, 20% glycerol, 2% SDS, 65 mM DTT) twice for 15 min each and then once for 15 min in SDS-sample buffer containing freshly added iodoacetamide (135 mM) and a trace of bromophenol blue. Each strip was then placed on top of a 12.5% Criterion precast gel (Bio-Rad; Hercules, CA). Second-dimension separation for all 12 gels was carried out in a Dodeca Cell (Bio-Rad; Hercules, CA) at 150 V for 1.5 h at RT until bromophenol blue reached the bottom of the gel. Water chilled to 4 °C was circulated through the coils in the cell to maintain a low temperature.

### Gel Image Acquisition

Following second-dimension separation, each gel was scanned on a Typhoon Trio+ Variable Mode Imaging System (GE Healthcare; Piscataway, NJ), using the specific laser band-pass filters for each dye's excitation and emission wavelengths. The excitation/emission wavelength combinations were (480 ± 35 nm)/(530 ± 30 nm) for Cy2, (540 ± 25 nm)/(590 ± 35 nm) for Cy3, and (620 ± 30 nm)/(680 ± 30 nm) for Cy5. Each gel was scanned individually and Photo Multiplier Tube (PMT) voltages were adjusted for maximum image quality with minimal signal saturation and clipping. Images were cropped and exported as 16-bit GEL files using ImageQuant TL (GE Healthcare; Piscataway, NJ) and imported into DeCyder Image Analysis (GE Healthcare; Piscataway, NJ). During the scanning process, gel 3 tore (HMEC-SV40-hTERT replicate B-Cy3; THMEC 80 d replicate A-Cy5) and was excluded from image analysis.

### Image and Statistical Analysis

GEL files were analyzed using the Batch Processor, Difference In-gel Analysis (DIA), and Biological Variation Analysis (BVA) modules within DeCyder Image Analysis Software. Batch processing was accomplished by setting 1250 as the upper spot limit. DIA was performed within individual gels comparing the Cy3 or Cy5 spot pattern to the Cy2 spot pattern. Ratios of Cy3/Cy5 spot volumes on each gel in the 11 gel data set were normalized against the Cy2 spot pattern. Results were obtained as abundance ratios for each spot. The following ratios were used: SHMEC/THMEC 40 d and THMEC 40 d/THMEC 80 d. To determine statistical significance, a Student's *t* test was calculated for each ratio with Decyder's False Discovery Rate applied. Spots determined as significantly different (FDR corrected $p < 0.05$) were then processed for identification. Expression graphs were generated

as follows: the abundance for each spot was normalized to the internal standard (Cy2) on its gel to generate a standard abundance ratio for each spot. Abundance ratios for each spot at each time point were then averaged to generate a mean for SHMEC, THMEC 40 d, and THMEC 80 d.

### Protein Identification

Following scanning, each gel was poststained with Sypro Ruby (Invitrogen; Carlsbad, CA) for spot picking. Spot picking was carried out using the ProPic (Isogen Life Science; Netherlands) robotic spot picker. Gel plugs were rinsed three times with 1 mL of 50% aqueous acetonitrile in 10 mM ammonium bicarbonate, pH 8.0, to remove Sypro Ruby stain. After evaporation of solvent by SpeedVac, each gel plug was rehydrated in 25 $\mu$L of 10 mM ammonium bicarbonate, pH 8.0. Trypsin Gold (10 $\mu$L of 12.5 $\mu$g/$\mu$L) was added to each gel plug solution, and the mixture was agitated overnight for 16 h. The supernatant was removed and plugs were washed twice with 10 mM ammonium bicarbonate. These rinses were combined, evaporated to dryness on a SpeedVac, then reconstituted in 10 $\mu$L of 0.1% formic acid. Peptide sequences were determined using an ABI-Sciex 4000 Qtrap (Applied Biosystems: Carlsbad, CA). The list of peptide sequences was exported and searched for matches within the nonredundant NCBI database (9/27/07) using the MASCOT search engine at www.matrix-science.com. Mass accuracy was set at 100 ppm, missed trypsin cuts at 1, Carbamidomethyl (C) fixed modifications, and Oxidation (HW) and Oxidation (M) variable modifications. The “gi” accession numbers obtained from MASCOT were used to search the UniProt Knowledgebase (UniProtKB) to obtain Swiss Protein Database ID numbers for each individual protein.

### Analysis of Maspin mRNA and Protein Expression

Real-time PCR was used to assess mRNA levels of maspin (SERBINB5) in cell samples. RNA was prepared as described above using Trizol. RNA suspended in formamide was purified using RNEasy kit and resuspended in RNase free water. Purified RNA was reverse-transcribed as described above and cDNA was purified using a QiaQuick PCR Purification Kit (Qiagen; Valencia, CA). Purified cDNA (10 ng) was analyzed for expression of SERBINB5 (Hs00985283_m1) using Taqman Assays (Applied Biosystems; Carlsbad, CA) on a MiniOpticon Real-time Thermocycler (Bio-Rad; Hercules, CA). Specific Assay IDs are in parentheses. GAPDH served as the endogenous control. Expression values for THMEC 40 d and THMEC 80 d were calculated using the Ct method relative to HMEC-SV40-hTERT expression.[16] Expression values for MCF10CAa.1α and MCF10CAd.1 were calculated using the ΔΔCt method relative to MCF10AT expression.[16] Western blots were performed as described above to assess maspin protein levels in THMEC model and MCF10 model. Maspin (sc-22762) antibodies were obtained from Santa Cruz Biotechnology (Santa Cruz, CA), with catalog number in parentheses. Actin served as the internal control.

### Network Analysis Using Ingenuity Pathway Analysis Software

A data set, consisting of HUGO gene names as identifiers and the corresponding fold change (SHMEC/THMEC 40d) in abundance for each of the 9 proteins identified to have significantly different abundance ratios were uploaded into Ingenuity Pathway Analysis software (Application version 8.8, build: 105318, release date 11-13-2010; Content version 3204, build: ing_nashira, release date 10-27-2010). Each identifier was mapped to its corresponding object in Ingenuity’s Knowledge Base. These molecules, called Network Eligible molecules, were overlaid onto a global molecular network developed from information contained in Ingenuity’s Knowledge Base. A Network of Network Eligible Molecules was then algorithmically generated based on connectivity. Functional Analysis of a network identified the biological functions that were significant ($p< 0.05$) to the molecules in the network. The network molecules associated with biological functions in Ingenuity’s

Knowledge Base were considered for the analysis. Right-tailed Fisher's exact test was used to calculate a *p*-value determining the probability that each biological function and/or disease assigned to that network is due to chance alone.

## Results

### Generation of Induced Neoplastic Breast Cells

Using the protocols established by the Weinberg lab, we induced HMECs to undergo an oncogenic transformation to produce neoplastically transformed human mammary epithelial cells (THMECs).[2] We first sought to confirm stable expression of all three genetic elements (*SV40 ER*, *exo-hTERT*, and *hRAS-v12*) illustrate an upregulation in telomerase activity, and highlight the tumorigenicity of our THMECs *in vivo*. It should be noted that we have chosen to use a different terminology than Elenbaas et al. used when referring to neoplastically transformed HMECs.[2] We refer to HMECs expressing *SV40 ER* and *exo-hTERT* as SHMECs, whereas Elenbaas et al. referred to their cell line as HMLE.[2] We refer to HMECs expressing all three constructs as THMECs, whereas Elenbaas et al. referred to their cell line as HMLER. This was done in order to eliminate any confusion on the clonal origin of our cell lines.

Stable expression of *SV40 ER* and *exo-hTERT* was confirmed using reverse transcriptase PCR (RT-PCR). Figure 1A displays the results of RT-PCR. The presence of bands in all THMECs clearly indicates a stable expression of both *SV40 ER* and *exo-hTERT* in THMECs. Neither HMECs nor MCF-7 breast cancer cells expressed either *SV40 ER* or *exo-hTERT*. A lack of expression of *exo-hTERT* in MCF-7 cells is especially important as it indicates that the primers specific to *exo-hTERT* do not amplify *endo-hTERT* transcripts, as MCF-7 cells have been shown to both express *endo-hTERT* and display a high degree of telomerase activity.[1,17] Telomerase activity was analyzed using the telomerase repeat amplification protocol (TRAP). A TRAP assay was performed using protein extracts harvested from HMECs, HMEC-SV40-hTERT (SHMEC), T-HMECs 10d post-hRAS introduction, and MDA-MB-231 cells (Figure 1B). The presence of a intense laddering effect, as well as the similarity between SHMEC and T-HMECs 10d and the positive control, MDA-MB-231, a breast cancer cell line known to display a high level of telomerase activity, is clearly indicative of a positive result for telomerase activity in SHMEC and T-HMECs 10d. The black arrow indicates the 36 bp internal control (IC) band that is present in all samples. The presence of telomerase activity as confirmed by a TRAP assay, as well as the presence of *exo-hTERT* transcripts as confirmed by RT-PCR, is indicative of a successful stable addition of the *exo-hTERT* gene to HMECs-SV40. To confirm a stable introduction of hRASv12, protein extracts were analyzed using a Western blot. Figure 1C displays the results of immunoblotting and clearly indicates an increase in hRas expression in THMEC 10d, 20d, 30d compared to HMEC and SHMEC.

### Cell Culture Observations and Assessment of Tumorigenicity

Figure 1D,E displays photographs taken of SHMECs and THMECs 30 d. Figure 1D shows SHMECs displaying a cobblestone-like appearance that is typical of an epithelial morphology. Figure 1E shows THMECs 30 d post-hRas. At 30 d post-hRas, THMECs have lost both contact inhibition and the cobblestone-like appearance seen in Figure 1D and now display a much more cylindrical phenotype indicated by the arrow in Figure 1E. This observation is supported by the fact that neoplastically transformed HMECs have been shown to undergo an epithelial to mesenchymal transition (EMT). We next assessed THMEC tumorigenicity *in vivo* by injecting THMECs at 40 d post-hRas introduction into immune-compromised nude mice. We choose 40 d post-hRas because at this point in the neoplastic induction foci formation was a common occurrence. All subcutaneous injections

produced tumors (3 independent trials using 5 mice for a total $n = 15$). The mean tumor volume for all 15 mice was 126.22 $mm^3$. A photograph of 3 representative tumors and the corresponding mice they were derived from can be viewed in Figure 1F. Arrows indicate the mice and location from which the tumors were removed. Injection of normal HMECs as a negative control was deemed unnecessary as HMECs are normal cells that lack the ability to form tumors.

### Identification of Differentially Expressed Proteins during a Neoplastic Transformation

2D-DIGE was applied to assess differential expression at the proteomic level. A total of 4 protein samples for each of the following time points were analyzed: SHMEC, THMEC 40 d, and THMEC 80 d. We then dye-swapped each protein sample to remove any dye-specific bias, for a total $n = 8$ for each time point (4 bio reps, with 2 tech reps for each bio rep). All 24 labeled samples were separated two dimensionally on 12 gels, with each gel also containing an internal standard, generated by combining equal amounts of all proteins, for cross-gel comparison. A schematic of the images generated for each time point can be viewed in Figure 2A and the exact gel assignments can be viewed in Table 1. A Student's *t* test (performed on both pairs; SHMEC vs THMEC 40 d and THMEC 40 d vs THMEC 80 d) and a One-Way ANOVA were used to identify spots that were significantly (FDR corrected $p < 0.05$) different in abundance over the course of neoplastic induction, in at least one of the three statistical tests performed. Selection of spots for identification was done by selecting the most significant spots for different trend types. Of the spots that were significantly different, 18 spots were chosen for identification. The location of each spot on a representative 2D gel can be viewed in Figure 2B. Spots identified to be differentially expressed are circled in pink. Of these 18 spots, 11 spots were identified as containing a single protein. These 11 spots are circled and numbered in Figure 2B and shown with their respective topographic peaks outlined in yellow. Additionally, 7 spots contained a mixture of proteins and further analysis is needed to identify which proteins are associated with the observed changes in expression. The 7 spots identified to have a mixture of proteins are circled but not numbered in Figure 2B. Spot numbers correspond to the master spot numbers assigned by Decyder. Topographic maps were exported from Decyder. Of these 11 spots, 2 spots had duplicate identities. In total, our methods produced 9 unique protein identifications (gene names are listed in parentheses): Heat Shock Protein 4A like (*HSPA4L*), Gelsolin (*GSN*), Keratin 1 (*KRT1*), T-complex Protein 1 (*TCP1*), FK506 binding protein 5 (*FKBP5*), Keratin 7 (*KRT7*), Maspin (*SERPINB5*), ribosomal protein P0 (*RPLP0*), and Stathmin (*STMN1*). Gelsolin and maspin each had two spots with individual IDs, most likely corresponding to differing post-translational modifications given the similarity in the migration in the second dimension. Proteomic data on all 18 spots can be viewed in Table 2, with spots containing a single protein in bold. Relative abundance graphs for the 11 spots containing a single protein are shown in Figure 3. FK506 binding protein 5, gelsolin, and stathmin increased in abundance during neoplastic induction, whereas heat shock protein 4-like, keratin 1, t-complex protein 1, keratin 7, maspin, and ribosomal protein P0 were decreased. For all spots that illustrated statistically significant change, with the exception of T-complex protein, the fold change (FC) > 1.20. Four spots displayed a near perfect linear down-regulation in protein abundance over 80 days as indicated by their respective coefficients of determination: keratin 7 ($r^2 = 0.9913$); maspin ($r^2 = 0.9947$); maspin ($r^2 = 0.9997$), ribosomal protein P0 ($r^2 = 0.9971$).

### Analysis of SERPINB5 mRNA and Protein Expression in THMECs and Metastatic Progression Model

To validate our proteomic data, we used Western blotting to assess maspin (*SERPINB5*) protein expression and real-time PCR to assess *SERPINB5* mRNA levels at all three time points of the THMEC model that were analyzed in 2D-DIGE (Figure 4A). Consistent with

the proteomic data, *SERPINB5* expression was progressively down-regulated over the course of tumorigenesis. The down-regulation also displayed the same linear trend as indicated by the coefficient of determination $r^2 = 0.9849$. *SERPINB5* is known as a metastasis suppressor gene (MSG).[18] MSGs suppress characteristics of metastasis such as invasion and motility, but have no effect on cellular division, immortality, or other nonmetastatic characteristics of cancer. The progressive down-regulation of *SERPINB5* expression illustrates a trend toward a metastatic phenotype. To confirm that *SERPINB5* is indeed associated with metastasis, we analyzed mRNA levels of *SERPINB5* in three cell lines of the MCF10AT metastatic progression model (Figure 4B). MCF10AT is the premetastatic cell line, whereas MCF10CAa.1α. and MCF10CAd.1 are two metastatic clones each generated from separate mice by tail vein injections of MCF10AT.[11] The loss of *SERPINB5* expression observed in THMECs can be observed in both metastatic clones relative to their premalignant origin (MCF10AT). *SERPINB5* expression in both metastatic cell lines was roughly half of that observed in MCF10AT cells. Real-time data for SERPINB5 can be viewed in Figure 4A,B. We also assessed maspin protein levels in both models using Western blot analysis. Alterations in the protein expression of maspin in THMECs correlated with the observed changes in maspin mRNA expression. In the metastatic progression model, maspin expression was less in premalignant MCF10AT cells than in immortalized MCF10A cells. Maspin protein expression was undetectable in both metastatic carcinomas, MCF10CAd.1α and MCF10CAa.1. To summarize our findings on *SERPINB5* expression, a near identical linear down-regulation over the entire 80 d period was observed in protein spots (#919 and #920), real-time assays, Western blotting, and microarray analysis (microarray data not shown) in THMEC model.

## Network and Functional Analysis

To determine if these nine molecules displayed any commonalities with regard to function and association with other common molecules, we uploaded the 9 gene names of the proteins identified to be differentially expressed into Ingenuity Pathway Analysis (IPA) software. IPA uses a manually curated database of molecule interactions as well as biological functions and pathways to drive its network and functional analysis algorithms. A single network was generated from the gene list that contained all nine gene names (Figure 5). Molecules are represented by shapes based on their biological role, according to the legend in Figure 5. Direct interactions are represented by a solid line and indirect relationships are indicated by a dashed line. This network is supported by 131 publications, of which we have cited a selection of publications related to the nine genes identified using proteomic analysis.[19–25] The network has a single primary node: TGFB1. TGFB1 displayed the most molecular interactions within the network and was within a 2 molecule association of all but one molecule (RPS8) in the network. The molecule with the highest degree of connectivity that can be found in the initial data set was KRT1. The tumor suppressor TP53 and GLI2, a transcription factor involved in sonic hedgehog signaling, also displayed common molecular interactions with molecules identified to be significantly different in abundance. TP53, GLI2, and TGFB1 have previously been shown to play a pivotal role in the etiology of breast cancer and have been shown to interact with one another to modulate breast cancer metastasis.[26–28]

We next overlaid relevant biological functions associated with the network that involved the 9 differentially expressed proteins. All biological processes were determined to be statistically significant ($p < 0.05$) using a Fisher's Exact test. The biological functions and their associated $p$-values are as follows: cell movement of tumor cells ($p = 9.24 \times 10^{-4}$), cell movement of mammary tumor cells ($p = 2.49 \times 10^{-4}$), polymerization of filaments ($p = 3.97 \times 10^{-6}$), and a malignant tumor phenotype ($p = 1.93 \times 10^{-3}$).[20,29–40] Molecules associated with a particular biological function are connected with a pink line. Those molecules

associated with a malignant phenotype are indicated with a pink arrow. The association of *KRT7* and *SERPINB5*, both genes shown to be progressively down-regulated over the course of tumorigenesis, with malignancy and cell movement, respectively, exemplifies a trend toward a metastatic phenotype.

## Discussion

The inability to predict when a cell will begin a transformation from a normal state to a neoplastic state *in vivo* makes it impossible to study the initial states that a neoplastic cell transitions through as it acquires oncogenic features. This problem is circumvented through the use of *in vitro* cellular reprogramming. Neoplastic transformation is one method of cellular reprogramming that is capable of producing a neoplastic breast cell from a normal breast epithelial cell, thus, enabling the ability to analyze the early cellular events of oncogenesis in real time. Using previously published protocols to generate breast cancer cells from normal somatic breast cells, we analyzed three time points in the neoplastic progression to breast cancer to identify proteins that were progressively differentially expressed. We have identified nine differentially expressed proteins, three of which displayed a linear progressive loss of abundance over 80 d following selection for *hRAS*. Network analysis revealed a common association with three molecules well-known to play a role in breast cancer and metastasis: TP53, TGFB1, and GLI2. TGFB1 and GLI2 have also been shown to play a pivotal role in the generation of breast cancer stem cells.[41,42] Functional analysis of the IPA-generated network illustrates a strong correlation between the proteins significantly different in abundance and maintenance of a metastatic state. The progressive nature of the trend toward a metastatic state is exemplified by the linear down-regulation of the metastasis suppressor gene (MSG), *SERPINB5*, and the epithelial cytokeratin, KRT7. As important as the presence and nature of alterations in expression are the biological functions that are acquired due to differential expression of these nine proteins. Polymerization of filaments and cellular movement are two biological functions associated with our data set that are both instrumental in the acquisition of a metastatic state.

Stathmin and gelsolin are both implicated in the regulation of the cytoskeleton.[43,44] Stathmin affects microtubules by binding and sequestering tubulin subunits.[43] Gelsolin binds and severs actin filaments.[44] Thus, they both serve to promote cytoskeletal depolymerization. One characteristic of normal, epithelial cells is a static cellular structure inferred by a rigid cytoskeleton. During the acquisition of a metastatic phenotype, this static, rigid structure transitions to a dynamic, pliable structure, more suited for motility. An integral aspect in the gain of cellular motility is the ability to break down and rebuild the cytoskeleton in a highly dynamic fashion. The observed increases in stathmin and gelsolin abundance in THMECs promote the cytoskeletal disassembly, a process that is not required for the rigid state of an epithelial cell. Although stathmin and gelsolin are not solely responsible for the acquisition of metastatic characteristics, their increased abundance increases the metastatic potential of THMECs through the attenuation of the static, rigid cytoskeleton that is characteristic of nonmotile cells.

*SERPINB5*, the gene that codes for the maspin protein, is a MSG that has been shown to regulate cellular motility, although the exact mechanisms are poorly understood.[20,45] *SERPINB5* is related to the serpin family of extracellular protease inhibitors and, although structurally very similar to other members of the serpin family, maspin does not undergo the conformational change characteristic of inhibitory serpins. Originally thought to only function in the extracellular environment, maspin has recently been shown to localize intracellularly, possibly even to the nucleus.[45] Overexpression of maspin has also been demonstrated to result in a decrease in both cellular motility and metastasis; however, very little is known about the mechanisms by which maspin elicits its inhibitory effects on these

processes. We have clearly shown here that *SERPINB5* mRNA and protein level expression is progressively down-regulated in early tumorigenesis and is consistent with a metastatic phenotype. Additionally, with regard to maspin, we have established two cell culture-based progression models in which to begin elucidation of the currently unknown biological function of maspin. With regard to regulation of maspin expression, the association of maspin with TGF$\beta$1 signaling pathways may help to explain the cause for the observed progressive trend in maspin expression. TGF$\beta$1 has been shown to promote the conversion of neoplastically transformed HMECs to a stem-like state in a time-dependent manner, suggesting that maspin may play a role in not only the acquisition of a metastatic phenotype, but also in the maintenance of stemness. Overexpression of FKBP5 has previously been shown to lead to a nuclear factor kappa B-mediated increase in TGF$\beta$1 secretion, thus, suggesting that FKBP5 may be involved in the regulation of maspin expression as well. Future studies will be aimed at assessing the role of maspin in breast cancer stem cells and the extent to which FKBP5 modulates maspin expression.

Two previous studies have utilized 2D-DIGE to analyze alterations in the proteome in breast cancer models of tumorigenesis.[46,47] The first study utilized an erbB2 model of breast cancer.[46] In that study, they compared cells overex-pressing erbB2, a gene that has been previously shown to be associated with 25–30% of breast cancers, to normal human mammary luminal epithelial cells.[48] Similar to our study, they identified multiple cytokeratins that were differentially expressed, including keratins 13A, 17, and 19. Interestingly, they did not identify cytokeratin 1 or 7 in their study to be differentially expressed, suggesting that the regulation of these two genes in breast cancer occurs independent of the erbB2 pathway. Additionally, the erbB2 overexpressing cells were shown to overexpress maspin, a protein know to inhibit metastatic processes. An overexpression of maspin in these cells suggests that this model is indicative of a less metastatic variant of breast cancer. The second study compared proteomic differences between a cell line, EM-G3, derived from an infiltrating ductal breast carcinoma and normal mammary epithelial cells.[47] They also found differences in multiple cytokeratins, including cytokeratin 19. Interestingly, they identified cytokeratin 7 as showing no significant change in protein expression. The differences involving the differential expression of keratin 7 and cytokeratin 19 are most likely due to the different types of breast cancer model systems used. Additionally, an increase in cytokeratin 19 expression occurring with the addition of *SV40ER* prior to generation of SHMEC cells by addition of *exo-hTERT* would result in an inability to detect the differential expression of cytokeratin 19. The detection of cytokeratin 7 differential expression in our study suggests that the loss of cytokeratin 7 may be associated with breast cancer cells that possess an increased potential to metastasize. Future studies will be aimed at understanding the effect of the loss of cytokeratin 7 on metastatic processes.

The initial upregulation in *GSN* and *STMN1* illustrates an early increase in metastatic potential, whereas the linear nature of the observed trend in maspin and *KRT7* expression illustrates that the increase in metastatic potential is also a progressive event in oncogenesis. We intend to use these proteomic findings, in combination with data from microarray experiments, as the basis for future studies into understanding the loss/gain of the biological functions associated with a meta-static state. Future directions for this project also include methylomic analysis of DNA methylation patterns and further proteomic analysis using large format gels and a broader pH range for separation in the first dimension.

## Acknowledgments

The authors would like to acknowledge Lana Grinberg, Anna Pendleton, and Su Ngyuen for their help in generating THMECs. We would like to acknowledge the Targeted Metabolics and Proteomics Laboratory members Dr. Mark

Cope, Gloria Robinson, and Richie Herring for assistance with the 2D-DIGE protocols. We would like to acknowledge the Tollefsbol lab members who provided informal advice during all aspects of the projects. J.T.D. was supported by an NCI Cancer Prevention and Control Training Program (R25CA047888). This work was also supported in part with grants from Susan G. Komen for the Cure (T.O.T., PI) and the NIH (R01 CA129415, T.O.T., PI). The instrumentation for running, imaging, and processing 2D-gels was provided by a NCRR Shared Instrumentation grant (S10 RR16849, H.K., PI). The mass spectrometers were purchased from funds provided by NCRR grants (S10 RR11329, RR13795, RR19231, S. Barnes, PI). Operation of the Comprehensive Cancer Center Proteomics-Mass Spectrometry Shared Facility was partially supported by funds from a P30 Core support grant (CA13148-35, E. Partridge, PI) to the UAB Comprehensive Cancer Center. We would like to acknowledge our collaboration with the lab of Dr. Danny Welch (University of Alabama at Birmingham) for providing us with the metastatic variants of the MCF10AT metastatic progression model.

## References

1. Hahn W, Counter C, Lundberg A, Beijersbergen R, Brooks M, Weinberg R. Creation of human tumour cells with defined genetic elements. Nature. 1999; 400(6743):464–8. [PubMed: 10440377]
2. Elenbaas B, Spirio L, Koerner F, Fleming MD, Zimonjic DB, Donaher JL, Popescu NC, Hahn WC, Weinberg RA. Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells. Genes Dev. 2001; 15(1):50–65. [PubMed: 11156605]
3. Hochedlinger K, Plath K. Epigenetic reprogramming and induced pluripotency. Development. 2009; 136(4):509–23. [PubMed: 19168672]
4. Hahn W, Dessain S, Brooks M, King J, Elenbaas B, Sabatini D, DeCaprio J, Weinberg R. Enumeration of the simian virus 40 early region elements necessary for human cell transformation. Mol Cell Biol. 2002; 22(7):2111–23. [PubMed: 11884599]
5. Counter C, Hahn W, Wei W, Caddle S, Beijersbergen R, Lansdorp P, Sedivy J, Weinberg R. Dissociation among in vitro telomerase activity, telomere maintenance, and cellular immortalization. Proc Natl Acad Sci USA. 1998; 95(25):14723–8. [PubMed: 9843956]
6. Casillas MA Jr, Lopatina N, Andrews LG, Tollefsbol TO. Transcriptional control of the DNA methyltransferases is altered in aging and neoplastically-transformed human fibroblasts. Mol Cell Biochem. 2003; 252:33. [PubMed: 14577574]
7. Jones P, Baylin S. The fundamental role of epigenetic events in cancer. Nat Rev Genet. 2002; 3(6): 415–28. [PubMed: 12042769]
8. Hammond S, Ham R, Stampfer M. Serum-free growth of human mammary epithelial cells: rapid clonal growth in defined medium and extended serial passage with pituitary extract. Proc Natl Acad Sci USA. 1984; 81(17):5435–9. [PubMed: 6591199]
9. Stampfer M, Bartley J. Induction of transformation and continuous cell lines from normal human mammary epithelial cells after exposure to benzo[a]pyrene. Proc Natl Acad Sci USA. 1985; 82(8): 2394–8. [PubMed: 3857588]
10. Pear WS, Nolan GP, Scott ML, Baltimore D. Production of high-titer helper-free retroviruses by transient transfection. Proc Natl Acad Sci USA. 1993; 90(18):8392–6. [PubMed: 7690960]
11. Hurst D, Xie Y, Edmonds M, Welch D. Multiple forms of BRMS1 are differentially expressed in the MCF10 isogenic breast cancer progression model. Clin Exp Metastasis. 2009; 26(2):89–96. [PubMed: 18841483]
12. Santner S, Dawson P, Tait L, Soule H, Eliason J, Mohamed A, Wolman S, Heppner G, Miller F. Malignant MCF10CA1 cell lines derived from premalignant human breast epithelial MCF10AT cells. Breast Cancer Res Treat. 2001; 65(2):101–10. [PubMed: 11261825]
13. Saldanha S, Andrews L, Tollefsbol T. Analysis of telomerase activity and detection of its catalytic subunit, hTERT. Anal Biochem. 2003; 315(1):1–21. [PubMed: 12672407]
14. Chomczynski P. Solubilization in formamide protects RNA from degradation. Nucleic Acids Res. 1992; 20(14):3791–2. [PubMed: 1379361]
15. Kim H, Cope MB, Herring R, Robinson G, Wilson L, Page GP, Barnes S. 2D difference gel electrophoresis of prepubertal and pubertal rat mammary gland proteomes. J Proteome Res. 2008; 7(11):4638–50. [PubMed: 18767877]
16. Livak K, Schmittgen T. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods. 2001; 25(4):402–8. [PubMed: 11846609]

17. Ramachandran C, Fonseca H, Jhabvala P, Escalon E, Melnick S. Curcumin inhibits telomerase activity through human telomerase reverse transcritpase in MCF-7 breast cancer cell line. Cancer Lett. 2002; 184(1):1–6. [PubMed: 12104041]

18. Shi HY, Zhang W, Liang R, Kittrell F, Templeton NS, Medina D, Zhang M. Modeling human breast cancer metastasis in mice: maspin as a paradigm. Histol Histopathol. 2003; 18(1):201–6. [PubMed: 12507299]

19. Wang D, Park J, Chu J, Krakowski A, Luo K, Chen D, Li S. Proteomic profiling of bone marrow mesenchymal stem cells upon transforming growth factor beta1 stimulation. J Biol Chem. 2004; 279(42):43725–34. [PubMed: 15302865]

20. Wang SE, Narasanna A, Whitell CW, Wu FY, Friedman DB, Arteaga CL. Convergence of p53 and Transforming Growth Factor beta (TGFbeta) Signaling on Activating Expression of the Tumor Suppressor Gene maspin in Mammary Epithelial Cells. J Biol Chem. 2007; 282(8):5661–9. [PubMed: 17204482]

21. Koch H, Zhang R, Verdoodt B, Bailey A, Zhang C, Yates, Menssen A, Hermeking H. Large-scale identification of c-MYC-associated proteins using a combined TAP/MudPIT approach. Cell Cycle. 2007; 6(2):205–17. [PubMed: 17314511]

22. Wilson CA, Cajulis EE, Green JL, Olsen TM, Chung YA, Damore MA, Dering J, Calzone FJ, Slamon DJ. HER-2 overexpression differentially alters transforming growth factor-beta responses in luminal versus mesenchymal human breast cancer cells. Breast Cancer Res. 2005; 7(6):R1058–79. [PubMed: 16457687]

23. Stasyk T, Dubrovska A, Lomnytska M, Yakymovych I, Wernstedt C, Heldin CH, Hellman U, Souchelnytskyi S. Phosphoproteome Profiling of Transforming Growth Factor (TGF)-{beta} Signaling: Abrogation of TGF{beta}1-dependent Phosphorylation of Transcription Factor-II-I (TFII-I) Enhances Cooperation of TFII-I and Smad3 in Transcription. Mol Biol Cell. 2005; 16(10):4765–80. [PubMed: 16055503]

24. Ewing RM, Chu P, Elisma F, Li H, Taylor P, Climie S, McBroom-Cerajewski L, Robinson MD, O'Connor L, Li M, Taylor R, Dharsee M, Ho Y, Heilbut A, Moore L, Zhang S, Ornatsky O, Bukhman YV, Ethier M, Sheng Y, Vasilescu J, Abu-Farha M, Lambert JP, Duewel HS, Stewart II, Kuehl B, Hogue K, Colwill K, Gladwish K, Muskat B, Kinach R, Adams SL, Moran MF, Morin GB, Topaloglou T, Figeys D. Large-scale mapping of human protein-protein interactions by mass spectrometry. Mol Syst Biol. 2007; 3:89. [PubMed: 17353931]

25. Lohr K, Moritz C, Contente A, Dobbelstein M. p21/CDKN1A mediates negative regulation of transcription by p53. J Biol Chem. 2003; 278:32507–16. [PubMed: 12748190]

26. Abe Y, Oda-Sato E, Tobiume K, Kawauchi K, Taya Y, Okamoto K, Oren M, Tanaka N. Hedgehog signaling overrides p53-mediated tumor suppression by activating Mdm2. Proc Natl Acad Sci USA. 2008; 105(12):4838–43. [PubMed: 18359851]

27. Dennler S, Andre J, Verrecchia F, Mauviel A. Cloning of the human GLI2 Promoter: transcriptional activation by transforming growth factor-beta via SMAD3/beta-catenin cooperation. J Biol Chem. 2009; 284(46):31523–31. [PubMed: 19797115]

28. Araki S, Eitel JA, Batuello CN, Bijangi-Vishehsaraei K, Xie XJ, Danielpour D, Pollok KE, Boothman DA, Mayo LD. TGF-beta1-induced expression of human Mdm2 correlates with late-stage metastatic breast cancer. J Clin Invest. 2010; 120(1):290–302. [PubMed: 19955655]

29. Sheng S, Carey J, Seftor EA, Dias L, Hendrix MJ, Sager R. Maspin acts at the cell membrane to inhibit invasion and motility of mammary and prostatic cancer cells. Proc Natl Acad Sci USA. 1996; 93(21):11669–74. [PubMed: 8876194]

30. Cunningham CC, Vegners R, Bucki R, Funaki M, Korde N, Hartwig JH, Stossel TP, Janmey PA. Cell permeant polyphosphoinositide-binding peptides that block cell motility and actin assembly. J Biol Chem. 2001; 276(46):43390–9. [PubMed: 11533030]

31. Shehata M, Schwarzmeier JD, Hilgarth M, Hubmann R, Duechler M, Gisslinger H. TGF-beta1 induces bone marrow reticulin fibrosis in hairy cell leukemia. J Clin Invest. 2004; 113(5):676–85. [PubMed: 14991065]

32. Kwiatkowski DJ, Stossel TP, Orkin SH, Mole JE, Colten HR, Yin HL. Plasma and cytoplasmic gelsolins are encoded by a single gene and contain a duplicated actin-binding domain. Nature. 1986; 323(6087):455–8. [PubMed: 3020431]

33. Alli E, Bash-Babula J, Yang JM, Hait WN. Effect of stathmin on the sensitivity to antimicrotubule drugs in human breast cancer. Cancer Res. 2002; 62(23):6864–9. [PubMed: 12460900]
34. Baldassarre G, Belletti B, Nicoloso MS, Schiappacassi M, Vecchione A, Spessotto P, Morrione A, Canzonieri V, Colombatti A. p27(Kip1)-stathmin interaction influences sarcoma cell migration and invasion. Cancer Cell. 2005; 7(1):51–63. [PubMed: 15652749]
35. Vardouli L, Moustakas A, Stournaras C. LIM-kinase 2 and Cofilin Phosphorylation Mediate Actin Cytoskeleton Reorganization Induced by Transforming Growth Factor-{beta}. J Biol Chem. 2005; 280(12):11448–57. [PubMed: 15647284]
36. Steinmetz MO, Jahnke W, Towbin H, García-Echeverría C, Voshol H, Müller D, van Oostrum J. Phosphorylation disrupts the central helix in Op18/stathmin and suppresses binding to tubulin. EMBO Rep. 2001; 2(6):505–10. [PubMed: 11415983]
37. Holmfeldt P, Stenmark S, Gullberg M. Interphase-specific phosphorylation-mediated regulation of tubulin dimer partitioning in human cells. Mol Biol Cell. 2007; 18(5):1909–17. [PubMed: 17344472]
38. Gavert N, Sheffer M, Raveh S, Spaderna S, Shtutman M, Brabletz T, Barany F, Paty P, Notterman D, Domany E, Ben-Ze'ev A. Expression of L1-CAM and ADAM10 in human colon cancer cells induces metastasis. Cancer Res. 2007; 67(16):7703–12. [PubMed: 17699774]
39. Thor AD, Edgerton SM, Liu S, Moore DH, Kwiatkowski DJ. Gelsolin as a negative prognostic factor and effector of motility in erbB-2-positive epidermal growth factor receptor-positive breast cancers. Clin Cancer Res. 2001; 7(8):2415–24. [PubMed: 11489821]
40. Ashida S, Nakagawa H, Katagiri T, Furihata M, Iiizumi M, Anazawa Y, Tsunoda T, Takata R, Kasahara K, Miki T, Fujioka T, Shuin T, Nakamura Y. Molecular features of the transition from prostatic intraepithelial neoplasia (PIN) to prostate cancer: genome-wide gene-expression profiles of prostate cancers and PINs. Cancer Res. 2004; 64(17):5963–72. [PubMed: 15342375]
41. Mani S, Guo W, Liao M, Eaton E, Ayyanan A, Zhou A, Brooks M, Reinhard F, Zhang C, Shipitsin M, Campbell L, Polyak K, Brisken C, Yang J, Weinberg R. The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell. 2008; 133(4):704–15. [PubMed: 18485877]
42. Katoh Y, Katoh M. Hedgehog target genes: mechanisms of carcinogenesis induced by aberrant hedgehog signaling activation. Curr Mol Med. 2009; 9(7):873–86. [PubMed: 19860666]
43. Rana S, Maples P, Senzer N, Nemunaitis J. Stathmin 1: a novel therapeutic target for anticancer activity. Expert Rev Anticancer Ther. 2008; 8(9):1461–70. [PubMed: 18759697]
44. Renz M, Langowski J. Dynamics of the CapG actin-binding protein in the cell nucleus studied by FRAP and FCS. Chromosome Res. 2008; 16(3):427–37. [PubMed: 18461482]
45. Teoh SS, Whisstock JC, Bird PI. Maspin (serpinb5) is an obligate intracellular serpin. J Biol Chem. 2010; 285:10862–9. [PubMed: 20123984]
46. Chan H, Gharbi S, Gaffney P, Cramer R, Waterfield M, Timms J. Proteomic analysis of redox- and ErbB2-dependent changes in mammary luminal epithelial cells using cysteine- and lysine-labelling two-dimensional difference gel electrophoresis. Proteomics. 2005; 5(11):2908–26. [PubMed: 15954156]
47. Selicharová I, Smutná K, Sanda M, Ubik K, Matousková E, Bursíková E, Brozová M, Vydra J, Jirácek J. 2-DE analysis of a new human cell line EM-G3 derived from breast cancer progenitor cells and comparison with normal mammary epithelial cells. Proteomics. 2007; 7(9):1549–59. [PubMed: 17366476]
48. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science. 1987; 235(4785):177–82. [PubMed: 3798106]

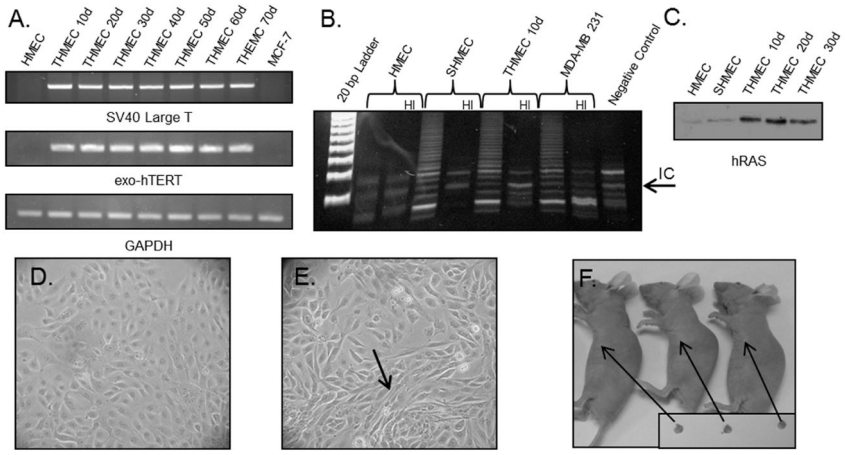


**Figure 1.**

Confirmation of gene transfer and neoplastic transformation. (A) Expression of SV40 ER, hTERT, and hRAS. RT-PCR of RNA extracted from HMEC, HMEC-SV40-hTERT-hRas (THMEC) 10d-70d, and MCF-7 cells using primers specific for *SV40 Large T* and exogenous *hTERT* was performed. *GAPDH* expression was analyzed to ensure equal loading. MCF-7 cells served as a negative PCR control. (B) Upregulation of telomerase activity. The Telomeric Repeat Amplification Protocol (TRAP) assay was performed using protein extracts from HMEC, SHMEC, THMEC 10d and positive control, MDA-MB-231. Heat-Inactivated (HI) samples were incubated at 85 °C for 10 min to inactivate telomerase. A reaction using 2 $\mu$L of DI water instead of protein extract served as a negative control testing for the presence of PCR artifacts. The laddering effect observed in SHMEC, T-HMEC 10d, and MDA-MB-231 is indicative of telomerase activity. The black arrow indicates the 36 bp internal control (IC) band. (C) Expression of hRAS. Western blot of total protein extracted from HMECs, HMEC-SV40-hTERT (SHMEC), and THMECS 10d, 20d, and 30d using antibodies specific for hRAS. (D and E) Images of SHMEC (D) and THMEC 30d (E). Arrow indicates cylindrical, mesenchymal like morphology. All pictures were taken with a CoolPix P5100 (Nikon; Melville, NY) at 40× magnification on an Eclipse TS100 microscope (Nikon). HMECs stably expressing SV40 and hTERT (SHMEC) display typical epithelial morphology. The cell population of THMEC (HMEC-SV40-hTERT-hRASV12) at 30 days post-hRAS transfection was composed of cells having a mesenchymal-like morphology. (F) Analysis of tumorigenicity. Female nude mice were injected subcutaneously in the left flank of each mouse with $1 \times 10^6$ THMEC (40 days post-hRAS) cells. Arrows indicate the location of the tumor on the corresponding mouse.

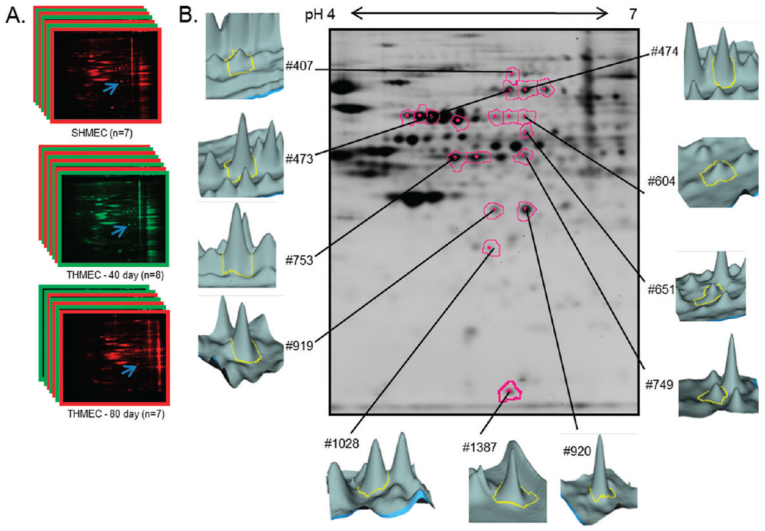


**Figure 2.**

Two-dimensional difference gel electrophoresis of protein extracts from SHMEC, THMEC 40 d, THMEC 80 d. (A) Schematic of images per time point. Green and red gels correspond to Cy3 and Cy5 labeled samples, respectively. Blue arrow indicates the location of maspin (Master Spot #919) on a 2D-Gel image from each time point. (B) A representative image of a 2D gel. This image is of Gel 10 Internal Standard labeled with Cy2. Circled spots are those identified as being significantly (FDR corrected $p < 0.05$) differentially expressed between HMEC-SV40-hTERT/THMEC 40 d or HMEC-SV40-hTERT/THMEC 80 d, or ANOVA. Spot numbers for unique IDs correspond to the Master Spot numbers assigned by DeCyder. For each numbered spot, a topographic map of signal intensity is displayed with the area used to calculate intensity shown in yellow.

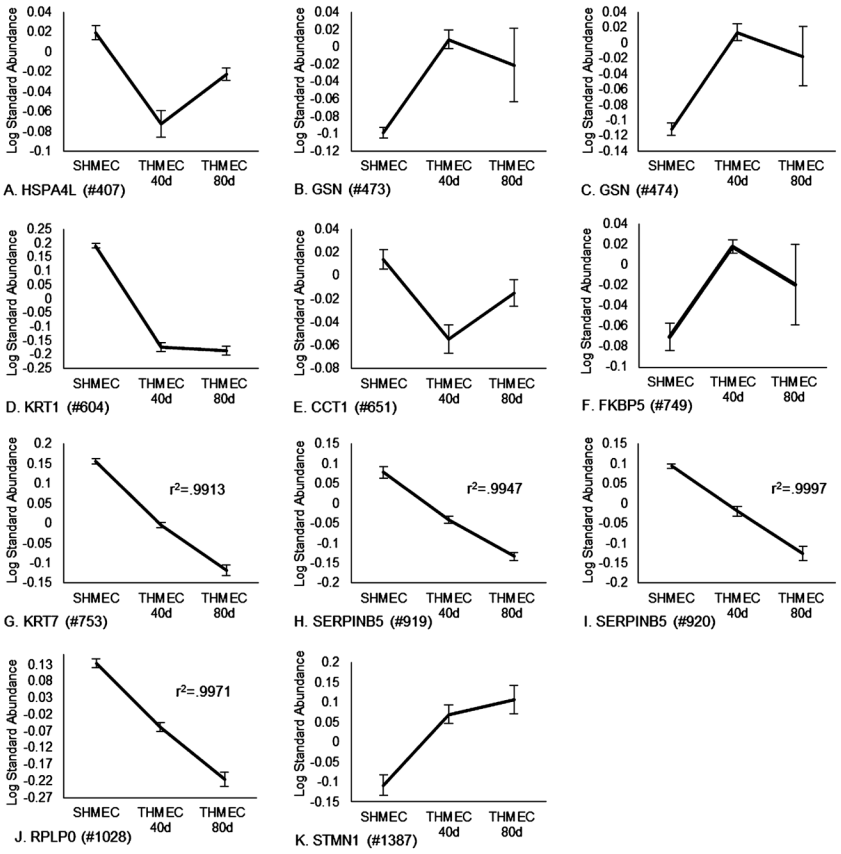


**Figure 3.**

Protein abundance graphs of significantly differentially expressed proteins identified by 2D-DIGE. To the left of each graph is the corresponding gene name and master spot# assigned by Decyder. Protein expression results are displayed as mean log standard abundance ratios. See Materials and Methods for derivation of protein standard abundance ratios. Error bars reflect SEM Coefficients of Determination ($r^2$) calculated using Microsoft Excel. (A) Heat Shock Protein 4A Like (*HSP4AL*); (B) Gelsolin (*GSN*); (C) Gelsolin (*GSN*); (D) Keratin 1 (*KRT1*); (E) T-Complex Protein 1A (*TCP1*); (F) FK506 binding protein 5 (*FKBP5*); (G) Keratin 7 (*KRT7*); (H) Maspin (*SERPINB5*); (I) Maspin (*SERPINB5*); (J) Ribosomal Protein P0 (*RPLP0*); (K) Stahmin (*STMN1*).

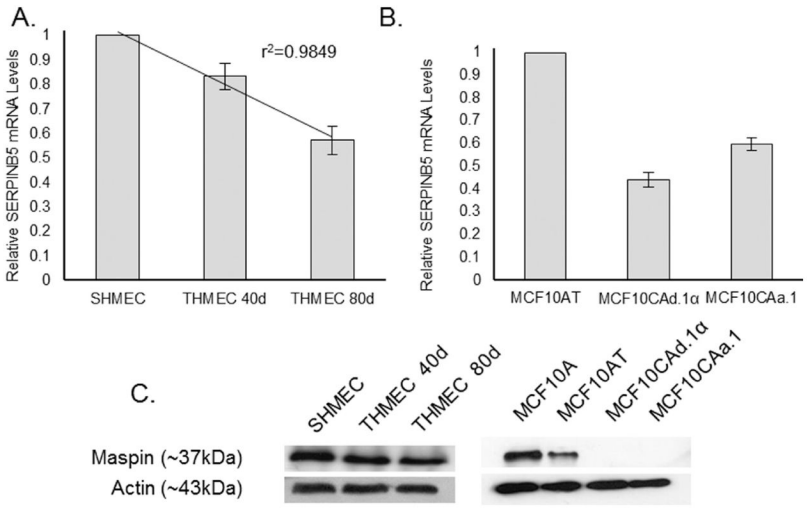


**Figure 4.**

Real-Time PCR analysis of mRNA expression of maspin (*SERBINB5*) in THMECs and the MCF10AT Metastatic Progression Model. (A) mRNA expression of *SERPINB5* in SHMEC, THMEC 40 d, 80 d. (B) mRNA expression of *SERPINB5* in metastatic progression model (MCF10AT, MCF10CAd.1α, and MCF10CAa.). Expression ratios for THMEC 40 d, THMEC 80 d, MCF10CAd.1α, and MCF10CAa.1 were calculated using the ΔΔCt method relative to the expression of HMEC-SV40-hTERT (for THMECs) and MCF10AT (for MCF10CAd.1α and MCF10CAa.1). SHMEC and MCF10AT expression was set to 1.0. *GAPDH* which served as endogenous control. Error bars represent SEM. The columns represent mean relative expression. Coefficient of determination ($r^2$) for maspin in THMEC model was calculated using Microsoft Excel. (C) Western blot of maspin protein expression in neoplastic progression model and metastatic progression model. Maspin protein expression was down-regulated over the course of tumorigenesis in THMECs. Maspin expression was less in premalignant MCF10AT cells than an in immortalized MCF10A cells. Maspin expression was undetectable in both metastatic carcinomas, MCF10CAd.1α and MCF10CAa.1. In both models, actin served as internal control.

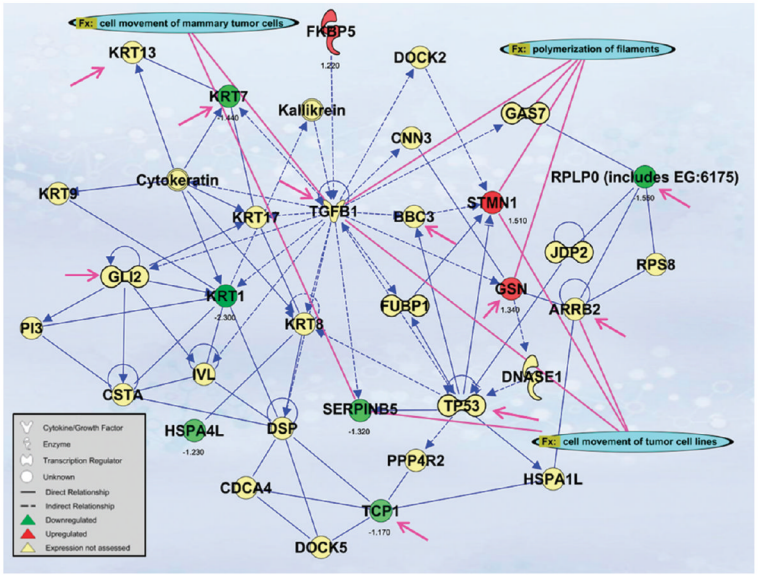


**Figure 5.**

Network analysis of differentially expressed proteins. A network is a graphical representation of the molecular relationships between molecules. Molecules are represented as nodes, and the biological relationship between two nodes is represented as an edge (line). All edges are supported by at least one reference from the literature, from a textbook, or from canonical information stored in the Ingenuity Pathways Knowledge Base. Human, mouse, and rat orthologs of a gene are stored as separate objects in the Ingenuity Pathways Knowledge Base, but are represented as a single node in the network. The node color indicates the directional change in protein expression of up- (red) or down- (green) and fold change (SHMEC → THMEC 40 d) in protein abundance is displayed below each molecule. Yellow nodes were added in by IPA to build the network and protein expression was not assessed. Nodes are displayed using various shapes that represent the functional class of the gene product. Edges are displayed as solid (direct) or dashed (indirect) to indicate the nature of the relationship. Functional analysis of a network identified the biological functions that were significant ($p < 0.05$) to the molecules in the network. The network molecules associated with biological functions in Ingenuity's Knowledge Base were considered for the analysis. Right-tailed Fisher's exact test was used to calculate a $p$-value determining the probability that each biological function and/or disease assigned to that network is due to chance alone. The biological functions cell movement of tumor cells ($p = 9.24 \times 10^{-4}$), cell movement of mammary tumor cells ($p = 2.49 \times 10^{-4}$), and polymerization of filaments ($p = 3.97 \times 10^{-6}$) are connected to the respective molecules with pink lines. Molecules associated with the biological process of a malignant tumor ($p = 1.93 \times 10^{-3}$) phenotype are indicated with pink arrows.

## Table 1

| gel no. | Cy 5 | Cy 3 | Cy 2 |
|---|---|---|---|
| 1 | HMEC-SV40-hTERT(A) | THMEC-40 d(A | Internal Standard |
| 2 | THMEC-40 d(A) | THMEC-80 d(A) | Internal Standard |
| 3 | THMEC-80 d(A) | HMEC-SV40-hTERT(B) | Internal Standard |
| 4 | HMEC-SV40-hTERT(B) | THMEC-40 d(B) | Internal Standard |
| 5 | THMEC-40 d(B) | THMEC-80 d(B) | Internal Standard |
| 6 | THMEC-80 d(B) | HMEC-SV40-hTERT(C) | Internal Standard |
| 7 | HMEC-SV40-hTERT(C) | THMEC-40 d(C) | Internal Standard |
| 8 | THMEC-40 d(C) | THMEC-80 d(C) | Internal Standard |
| 9 | THMEC-80 d(C) | HMEC-SV40-hTERT(D) | Internal Standard |
| 10 | HMEC-SV40-hTERT(D) | THMEC-40 d(D) | Internal Standard |
| 11 | THMEC-40 d(D) | THMEC-80 d(D) | Internal Standard |
| 12 | THMEC-80 d(D) | HMEC-SV40-hTERT(A) | Internal Standard |

## Table 2

Proteomic and Statistical Data of Spots That Were Significantly Different in Abundance over the Course of Breast Tumorigenesis[a]

| master spot number[b] | FC[c] | FC[d] | FDR[e] | protein[f] | ID[g] | MASCOT no.[h] | MOWSE[i] | % cov.[j] | MW[k] | pI[l] |
|---|---|---|---|---|---|---|---|---|---|---|
| **407** | **−1.23 (0.0024)** | | **0.0008** | **Heat shock 70 kDa protein 4-like** | **O95757** | **gi\|31541941** | **76** | **2%** | **95479** | **5.63** |
| 468 | 1.28 (0.00046) | | | Mitochondrial inner membrane protein | Q16891 | gi\|48145703 | 392 | 12% | 84027 | 6.08 |
| | | | | Gelsolin isoform b | P06396 | gi\|38044288 | 246 | 5% | 80876 | 5.58 |
| | | | | Procollagen-lysine, 2-oxoglutarate 5-dioxygenase 3 precursor | O60568 | gi\|4505891 | 42 | 1% | 85302 | 5.69 |
| **473** | **1.28 (0.00015)** | | | **Gelsolin isoform b** | **P06396** | **gi\|38044288** | **216** | **7%** | **80876** | **5.58** |
| **474** | **1.34 ($5.80 \times 10^{-5}$)** | | **0.047** | **Gelsolin isoform b** | **P06396** | **gi\|38044288** | **308** | **7%** | **80876** | **5.58** |
| 592 | −1.23 (0.018) | | 0.012 | Heat shock 70 kDa protein 8 isoform 1 | P11142 | gi\|5729877 | 493 | 20% | 71082 | 5.37 |
| | | | | Lamin B1 | P20700 | gi\|5031877 | 320 | 12% | 66653 | 5.11 |
| 598 | −1.32 ($3.4 \times 10^{-6}$) | | $4.00 \times 10^{-8}$ | Heat shock 70 kDa protein 8 isoform 1 | P11142 | gi\|5729877 | 539 | 24% | 71082 | 5.37 |
| | | | | V-type proton ATPase catalytic subunit A | P38606 | gi\|291868 | 177 | 7% | 68677 | 5.35 |
| **604** | **−2.30 ($4.30 \times 10^{-7}$)** | | **$1.10 \times 10^{-10}$** | **Keratin, type II cytoskeletal 1** | **P04264** | **gi\|11935049** | **116** | **4%** | **66198** | **8.16** |
| 606 | −1.29 ($5.0 \times 10^{-7}$) | | $9.00 \times 10^{-8}$ | Heat shock 70 kDa protein 8 isoform 1 | P11142 | gi\|5729877 | 587 | 20% | 71082 | 5.37 |
| | | | | Heat shock 70 kDa protein 1 | P08107 | gi\|4529893 | 275 | 7% | 70280 | 5.48 |
| | | | | Lamin B2 | Q03252 | gi\|27436951 | 247 | 7% | 67762 | 5.29 |
| | | | | V-type proton ATPase catalytic subunit A | P38606 | gi\|291868 | 101 | 3% | 68677 | 5.35 |
| 615 | −1.20 ($2.1 \times 10^{-6}$) | | 0.03 | Heat shock 70 kDa protein 1 | P08107 | gi\|4529892 | 628 | 23% | 70267 | 5.48 |
| | | | | Plastin 3 | P13797 | gi\|7549809 | 84 | 2% | 71279 | 5.41 |
| | | | | Lamin B2 | Q03252 | gi\|345758 | 66 | 1% | 59079 | 5.87 |
| **651** | **−1.17 (0.016)** | | **0.027** | **T-complex protein 1 isoform a** | **P17987** | **gi\|57863257** | **101** | **4%** | **60819** | **5.80** |
| **749** | **1.22 (0.0017)** | | | **FK506-binding protein 5** | **Q13451** | **gi\|1145816** | **38** | **2%** | **50859** | **5.80** |
| **753** | **−1.44 ($4.30 \times 10^{-7}$)** | **−1.30 (0.01)** | **$1.10 \times 10^{-10}$** | **Keratin, type II cytoskeletal 7** | **P08729** | **gi\|67782365** | **482** | **23%** | **51411** | **5.42** |
| 754 | −1.30 ($5.00 \times 10^{-7}$) | | $3.50 \times 10^{-6}$ | Keratin, type II cytoskeletal 8 | P05787 | gi\|181573 | 511 | 22% | 53529 | 5.52 |
| | | | | Keratin, type II cytoskeletal 7 | P08729 | gi\|67782365 | 455 | 20% | 51411 | 5.40 |
| **919** | **−1.32 (0.00047)** | **−1.24 (0.01)** | **$6.90 \times 10^{-8}$** | **Chain A, The 2.1 A Structure Of A Tumour Suppressing Serpin** | **P36952** | **gi\|62738363** | **113** | **5%** | **42259** | **5.72** |
| **920** | **−1.29 (0.00015)** | **−1.27 (0.041)** | **$2.10 \times 10^{-7}$** | **Chain A, The 2.8 A Structure Of A Tumour Suppressing Serpin** | **P36952** | **gi\|62738525** | **137** | **8%** | **43071** | **5.72** |
| **1028** | **−1.55 ($2.30 \times 10^{-5}$)** | **−1.43 (0.01)** | **$1.30 \times 10^{-8}$** | **60S acidic ribosomal protein P0** | **P05388** | **gi\|4506667** | **88** | **5%** | **34423** | **5.71** |

| master spot number[b] | FC[c] | FC[d] | FDR[e] | protein[f] | ID[g] | MASCOT no.[h] | MOWSE[i] | % cov.[j] | MW[k] | pI[l] |
|---|---|---|---|---|---|---|---|---|---|---|
| **1387** | **1.51 (0.0019)** | | **.015** | **Stathmin 1 isoform b** | **P16949** | **gi\|224451142** | **131** | **18%** | **19925** | **6.51** |

[a] Bold text indicates spots that contained only one protein.

[b] Master spot number assigned by Decyder.

[c] Fold change between SHMEC and THMEC 40 d. Numbers in parentheses correspond to false discovery rate corrected *p*-values from a Student's *t* test. Only statistically significant data is displayed.

[d] Fold change between THMEC 40 d and THMEC 80 d. Numbers in parentheses correspond to false discovery rate corrected *p*-values from a Student's *t* test. Only statistically significant data is displayed.

[e] False discovery rate corrected *p*-values from a One-way ANOVA (SHMEC/THMEC40 d/THMEC80 d). Only statistically significant data is displayed.

[f] Protein name.

[g] Swiss-Prot ID.

[h] The "gi" accession numbuer obtained from MASCOT database.

[i] MOWSE score obtained from MASCOT database.

[j] Percent coverage of fragmented peptides vs total peptide length.

[k] Predicted molecular weight in daltons.

[l] Predicted isoelectric point (p*I*).