THE ROLE OF SINGLE MINDED 2 SHORT

IN MAMMARY GLAND DEVELOPMENT AND BREAST CANCER

A Dissertation

by

HYEONG-IL KWAK

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2006

Major Subject: Toxicology

THE ROLE OF SINGLE MINDED 2 SHORT

IN MAMMARY GLAND DEVELOPMENT AND BREAST CANCER

A Dissertation

by

HYEONG-IL KWAK

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Weston W. Porter
Committee Members,	Robert C. Burghardt
	Stephen H. Safe
	Thomas E. Spencer
Chair of Toxicology Faculty,	Robert C. Burghardt

December 2006

Major Subject: Toxicology

ABSTRACT

The Role of Single Minded 2 Short

in Mammary Gland Development and Breast Cancer. (December 2006)

Hyeong-il Kwak, B.S., Chonnam National University;

M.S., Seoul National University

Chair of Advisory Committee: Dr. Weston Porter

Single minded 2 (Sim2) is a member of the basic helix-loop-helix Per-ARNT-Sim (Period-Arylhydrocarbon Nuclear Translocator-Single minded) family. Human SIM2 is involved in the etiology of the Down's phenotype. In addition to the physical and mental deficiencies associated with DS, it has become apparent that women with DS are 10-25 times less likely to develop breast cancer in comparison to age-matched normal populations. Such significant effects on breast cancer susceptibility are thought to result from gene dosage effects of one or more tumor suppressor genes on chromosome 21. Here we report the identification and transcriptional characterization of mouse Sim2s, a splice variant of Sim2, which is missing the carboxyl Pro/Ala-rich repressive domain. Similar to full-length Sim2, Sim2s interacts with ARNT and to a lesser extent, ARNT2. The effects of Sim2s on transcriptional regulation through hypoxia-, dioxin- and central midline response elements are different than that of full length Sim2. Specifically, Sim2s exerts a less repressive effect on hypoxia-induced gene expression than full length Sim2, but is just as effective as Sim2 at repressing TCDD-induced gene expression from a dioxin response element. Interestingly, Sim2s binds to and activates expression from a

central midline response element-controlled reporter through an ARNT transactivation domain-dependent mechanism.

Forced expression of SIM2s in MDA-MB-435 breast cancer cells significantly inhibited proliferation, reduced anchorage-independent growth, and decreased invasive potential. SIM2s directly decreased expression of matrix metalloprotease-3, a known mediator of breast cancer metastasis. In addition, loss of Sim2 in the mouse mammary gland increased ductal branching, accelerated lobuloalveolar-like precocious hyperplasia, and decreased cell apoptosis, suggesting that SIM2s is a mammary tumor suppressor. Sim2-/- mammary glands lose E-cadherin expression, suggesting that Sim2s plays a role in regulating E-cadherin/beta-catenin signaling. Loss of Sim2 in the mammary glands also resulted in dramatically increased MMP3 expression. The mechanism of SIM2s-mediated repression of MMP3 was found to be due to its ability to inhibit AP-1 binding to the MMP3 promoter. These results suggest that SIM2s contributes to the breast cancer protective effects observed in DS individuals.

DEDICATION

To my wife, Boyeon Lee,

my mother, Sangshim Kim,

parents-in-law, Jaewoo Lee and Sunghun Park,

and my pastors, Chong Kim and Youngsik Ahn.

ACKNOWLEDGEMENTS

I wish to thank my major advisor, Dr. Weston Porter, for his patience and support throughout my graduate career. I am also grateful to the members of my dissertation committee, Drs. Stephen Safe, Robert Burghardt, and Thomas Spencer.

I would also like to thank the following colleagues for their help and friendship: Dr. Richard Metz, Brian Laffin, Tanya Gustafson, Elizabeth Wellberg, Keelan Anderson, and Kristin Kucera. I would also like to acknowledge Kim Daniel for her administrative assistance.

TABLE OF CONTENTS

Page

LIST OF TABLES

TABLE Page

LIST OF FIGURES

FIGURE Page

FIGURE Page

FIGURE Page

I. INTRODUCTION*

1.1. Down Syndrome

Down Syndrome (DS), which results from trisomy 21, is the most common chromosomal abnormality and a leading cause of mental retardation, with a worldwide frequency of 1 in 800 births. DS was first described in 1866 by John Langdon Down. In 1959, Jerome Lejeune showed that DS is caused by trisomy of HSA21. In rapid succession, several studies reporting on DS owing to chromosomal translocations that involve HSA21 and mosaicism for Ts21 were published (Patterson & Costa, 2005). It has long been recognized that the risk of having a child with trisomy 21 increases with maternal age (Penrose, 1933). For example, the risk of having a liveborn with DS at maternal age 30 is 1 in 1,000 and at maternal age 40 is 9 in 1,000 (Hook et al., 1983). DS individuals have characteristic phenotypes such as brachycephaly, brachydactyly, broad hands, duodenal atresia, epicanthal folds, fifth finger clinodactyly, flat nasal bridge, hypotonia, lax ligaments, mental retardation, open mouth, short stature, and wide 1–2 toe gap (Roizen & Patterson, 2003).

The molecular and cellular events linking the presence of an extra chromosome to phenotypic features are yet unknown. It is the central hypothesis that characteristic features of DS individuals result from increased expression of genes on chromosome 21

The journal used as a model for this dissertation is Toxicology and Applied Pharmacology.

* Parts of this section are reprinted from Metz et al., 2006, J Biol Chem, 281(16), 10839-10848, Copyright © 2006 by the American Society for Biochemistry and Molecular Biology and Kwak et al., 2006, Carcinogenesis, in progress, Copyright © 2006 by Oxford University Press.

because of their higher gene dosage. Accordingly, an essential question in DS research is: are all trisomic genes overexpressed in all tissues and at all time points? If not, then which genes are overexpressed, and when and where? Answers to these questions are critical for determining which genes are relevant to phenotype development, for linking expression of specific genes to specific phenotypic features, and to account for phenotypic variability (Gardiner, 2004). The gene content of chromosome 21 is estimated to be 329, including 165 experimentally confirmed genes, 150 gene models based on expressed sequence tag databases, and 14 computer predictions (Roizen & Patterson, 2003). Given the large number of genes involved, determining which genes are overexpressed requires a large-scale approach, which is complicated by the small differences in expression level (50%) between normal and DS individuals predicted by gene dosage.

Interestingly, rare individuals with DS have partial trisomy 21. Nearly 92% of DS individuals have an extra chromosome 21 (Pueschel & Rynders, 1982) in all of their cells, resulting in a karyotype with 47 chromosomes, due to trisomy 21. About 4% of DS individuals have smaller portions triplicated because of unbalanced translocations (Pueschel & Rynders, 1982). Mosaicism of the chromosome 21 is present in 2 to 4% of the DS individuals. They present two kinds of cells, one with normal number of chromosome, and another with abnormal, i.e. 47. The main cause of mosaicism is the non-disjunction of chromosome 21 during mitosis in embryo. As lower the number of trisomic cells, lower is the phenotypic involvement (Mikkelsen, 1977). The study of individuals with partial trisomy 21 defined the genomic regions that harbor genes

associated with some DS phenotypes. A number of investigators have described a 'Down Syndrome critical region (DSCR)' that specifically contains genes that contribute to cognitive defects or other DS features. The definition of these regions has been controversial as there are individuals with partial triplications outside this region who, nevertheless, manifest some features of DS (Antonarakis, 2004). However, the idea of a DSCR implies that much of DS could be caused by extra copies of one or a small number of genes in this region (Korenberg et al., 1994).

It has not been simple to develop a mouse model for DS. Human chromosome 21 carries about 329 genes across the 33.5 million bases (Mb) of its long arm and the mouse orthologues are distributed across three chromosomes: 10, 16, and 17 (Fig. 1) (Davisson et al., 2001; Roizen & Patterson, 2003).

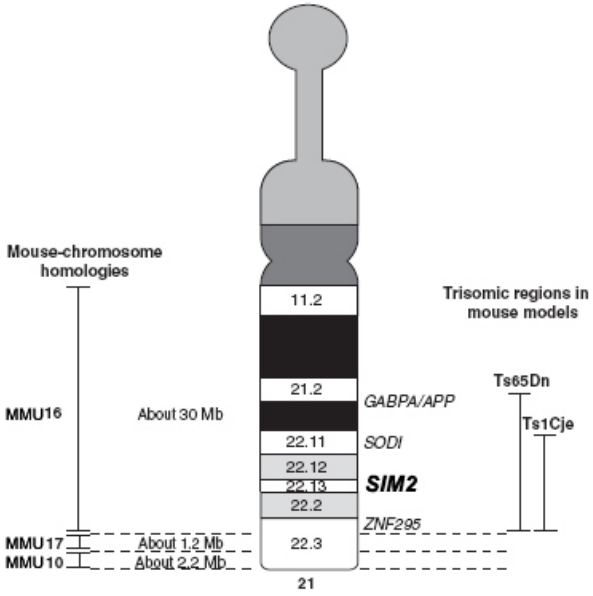


Figure 1. Human chromosome 21 and homologous regions in mouse models. Regions that are syntenic with mouse chromosomes are indicated on the left; those that are trisomic in the major mouse models are indicated on the right. Modified from Gardiner, 2004.

The centromere-proximal 30 Mb region of chromosome 21 up to and including the *ZNF295* gene is orthologous to the telomeric region of mouse chromosome 16 and the next two approximately 1-2 Mb segments of chromosome 21 are orthologous to regions of mouse chromosomes 17 and 10, respectively (Gardiner, 2004). Ts65Dn mice have three copies of 94 genes orthologous to human chromosome 21 genes, contained within chromosome 16 from the Gabpa/App gene cluster to the distal telomere (Reeves et al., 1995). Ts1Cje, a second partial trisomy mouse model, is trisomic for 71 orthologs of human chromosome 21 genes, within chromosome 16 distal to the superoxide dismutase 1 (Sod1) gene to the telomere (Fig. 1) (Sago et al., 1998). Although neither mouse was a perfect model for human trisomy 21, there were substantial similarities in phenotype, notably craniofacial changes that mimic the human condition, along with electrophysiological differences in brain activity and altered behavior.

Recently, several groups have examined gene dosage effects of trisomy 21, by screening microarrays with RNA from brains of human fetuses with DS (Mao et al., 2003) or a trisomic mouse models (Amano et al., 2004), by screening a cDNA array containing mouse orthologs of human chromosome 21 genes with RNA from several tissues of a mouse model (Kahlem et al., 2004), and by using quantitative reverse-transcriptase-coupled (RT) PCR analysis of RNA from different tissues of a mouse model at different ages (Lyle et al., 2004). The results of all the studies support the hypothesis that gene dosage effects exist in DS, but they also show that dosage effects may be specific to particular genes, alleles and/or tissues, and that background and stochastic or transient effects may be confounding factors (Table 1) (Gardiner, 2004).

Table 1. Overview of studies of gene-dosage effects in trisomies. Modified from Gardiner, 2004.

Species and strain	Humans with DS	Ts1Cje mice	Ts65Dn mice
Ages	17-20 weeks gestation	postnatal day 0	3-4 months
Tissues	Cerebrum and cortex-derived astrocyte cell lines	Whole brain	Cortex, mid brain, cerebellum, heart, testis, liver, kidney, lung, and muscle
Number of genes on trisomic segment		71	94 (77tested in 9 tissues)
Trisomic genes increased in expression	25 (Variation among individuals)	37 (Variation among individuals)	all genes (of 66 with detectable expression) except in muscle
Non-trisomic genes with altered expression			Not determined

Due to improvements in medical care, individuals with DS are living longer, healthier lives than was possible just a few decades ago. As a result, it has become apparent that the incidence of cancer in individuals with DS is unique. Most notably, people with DS are more susceptible to childhood leukemias and germ-line cancers, but are more than 50% less likely to develop solid tumors including cancers of the lung, colon, skin, head and neck (Boker et al., 2001; Hasle et al., 2000a; Hasle et al., 2000b; Hasle et al., 2000c; Yang et al., 2002a) (Table 2). Even more striking is the observation that women with DS are 10-25 times less likely to develop breast cancer in comparison to age-matched normal populations (Benard et al., 2005; Hasle et al., 2000a; Hasle et al., 2000b; Hasle et al., 2000c; Hill et al., 2003; Stage et al., 1998) (Table 3). Compared to an age-matched population, there were no cases of breast cancer in 1278 women with DS with 7.3 cases expected (Hasle et al., 2000a). Another study reported only one case of breast cancer in DS, compared with 11.7 expected in an age matched U.S. population (Scholl et al., 1982).

Table 2. Decrease in solid tumors in individuals with Down Syndrome. Adapted from Hasle et al., 2000.

Age group	Observed	Expected
0-9	1	1.42
10-19	1	1.50
20-29	1	3.89
30-39	4	6.47
40-49	5	11.52
50-59	8	14.28
60+	4	8.67
Total	**24**	**47.77**

Table 3. Distribution of solid tumors in individuals with Down Syndrome. Adapted from Hasle et al., 2000.

Site	Observed	Expected
Buccal Cavity	0	1.04
Digestive System	4	6.52
Respiratory System	1	4.96
Breast	**0**	**7.32**
Female Genital Organs	4	5.68
Male Genital Organs	4	2.82
Urinary Tract	4	2.97
Skin	2	8.14
Other	1	4.75
Secondary Sites	3	0.92
Non-Hodgkin's lymphoma	0	1.41
Hodgkin's Disease	0	0.92
All Solid Tumors	**24**	**47.77**

Women with DS rarely become pregnant, increasing their risk for breast cancer; however, they experience earlier menopause, leading to a decrease in risk. Hormonal factors alone are unlikely to explain the decreased rate of breast cancer, thus protective

factors are more likely to be genetically linked. These observations have led to the hypothesis that one or more tumor suppressor genes are present on chromosome 21 and prevent cancer development by gene dosage effects.

As introduced above, rare DS individuals carry chromosomal rearrangements resulting in triplication of only part of chromosome 21, although most cases of DS are attributable to the presence of three full copies of chromosome 21. Molecular characterization of these "partial trisomy" cases has allowed the delineation of a DSCR, located at the sub-band 21q22.2, which correlates with many DS abnormalities. Using the exon-trapping technique to isolate potential coding sequences within this region, two groups have identified exons that predict an open reading frame that is highly homologous to the *Drosophila* sim gene product. Based on its chromosomal location and sequence, the gene that maps to the DSCR is the human equivalent of the murine *Sim2* gene (Michaud & Fan, 1997). DS individuals are characterized by skeleton and cranio-facial defects, hypotonia, and heart defects. In addition, in DS individuals, several neural and neurochemical abnormalities were observed involving early Alzheimer's disease, mental retardation, hypoplasia of the hippocampus, the cortex and the, retarded cortical lamination and myelinization and neurotransmitter alterations. These neurological phenotypes may be a consequence of altered growth and/or differentiation of neuronal cells during Central Nervous System (CNS) development (Rachidi et al., 2005). Human *SIM2*, thus, has been thought as a candidate gene in the etiology of the Down's phenotype (Muenke et al., 1995). This is supported by the observation that transgenic mice trisomic for *SIM2*, with one BAC-containing copy of

hSIM2, and two endogenous copies of *mSim2*, have some aspects of Down's phenotype, exhibiting anxiety-related/reduced exploratory behaviour and reduced sensitivity to pain (Chrast et al., 2000). Neurological abnormalities have been reported in the Ts65Dn mouse which has the triplication of a segment of chromosome 16 syntenic to the DSCR and containing Sim2 (Reeves et al., 1995). It is also supported by a study on SIM2 spatial and temporal expression pattern during human central nervous system development, from embryonic to fetal stages. In embryonic stages, SIM2 was identified predominantly in restricted regions of CNS, in ventral part of D1/D2 diencephalic neuroepithelium, along the neural tube and in a few cell subsets of dorsal root ganglia. In fetal stages, SIM2 showed differential expression in pyramidal and granular cell layers of hippocampal formation, in cortical cells and in cerebellar external granular and Purkinje cell layers. SIM2 expression in embryonic and fetal brain could suggest a potential role in human CNS development, in agreement with *Drosophila* and mouse Sim mutant phenotypes and with the conservation of the Sim function in CNS development from *Drosophila* to Human. SIM2 expression in human fetal brain regions, which correspond to key structures for cognitive processes, correlates well with the behavioral phenotypes of *Drosophila* Sim mutants and transgenic mice overexpressing Sim2. In addition, SIM2-expressing brain regions correspond to the altered structures in DS individuals (Rachidi et al., 2005). Sim2−/− mice die within 3 days of birth due to breathing failure (Goshu et al., 2002), exhibiting reduced efficacy of lung inflation and numerous abnormalities within the structural components surrounding the pleural cavity.

1.2. Single minded 2

Sim is a member of basic helix-loop-helix Per-Arnt-Sim (bHLH-PAS) family that activates midline gene transcription and represses lateral CNS gene transcription (Crews, 1998). Sim gene was first identified in *Drosophila* melanogaster as a mutation affecting neurogenesis and causing early lethality. To the contrary, in transgenic flies, ubiquitously expressed sim caused many other cells of the ventral neurogenic region to differentiate into CNS midline cells (Nambu et al., 1991). Thus, *Drosophila* sim is proposed to act as a master regulator of midline cell development in the CNS, functioning as a transcriptional regulator in cell fate determination (Nambu et al., 1993). In *Drosophila*, expression of dsim is found in midline cells, anterior and posterior to the developing ventral cord throughout the germ band, and during post-embryonic development in the central complex of the brain (Pielage et al., 2002).

Two homologs of dsim, Sim1 and Sim2, have been identified (Ema et al., 1996; Moffett et al., 1996). Sim1 is localized on the proximal region of mouse chromosome 10, while Sim2 on the very distal end of chomosome 16. According to the mouse–human linkage homologies, it is likely that Sim1 maps to human 6q21 and Sim2 to 21q22 (Fan et al., 1996). A general feature of the expression of the mammalian Sim genes is that some of their sites of expression roughly parallel those of *Drosophila* sim. In *Drosophila*, sim is expressed in the brain, ventral nerve cord, gut, and muscles (Crews et al., 1988). Similarly, in mouse, Sim1 is expressed in the brain, ventral spinal cord, and foregut, and Sim2 is expressed in the brain and in muscles. In addition, both mammalian Sim genes are expressed in the developing mesonephros and, later, the

kidney. Importantly, whereas sim is a key regulator of CNS midline cell development in the fly, the mammalian Sim genes are not expressed in floor plate cells of the spinal cord. The only expression of the mouse Sim genes in the ventral midline of the CNS is in the mammillary area of the diencephalon. Sim2 is expressed in the forebrain in a restricted pattern in the caudal diencephalon at early stages of neural development (Fan et al., 1996). The importance of this expression pattern is in showing that the forebrain is already regionalized at the two-somite stage, prior to any overt morphological specialization of the diencephalons primordium (Fan et al., 1996).

1.3. Basic helix-loop-helix Per-ARNT-Sim

The bHLH-PAS proteins comprise a growing family of transcription factors that play key roles during development and in sensing and adapting to changes in the environment. Individual PAS proteins are known to control morphogenesis, circadian rhythmicity, responses to hypoxia and toxin metabolism. These proteins contain a bHLH motif, which mediates dimerization with other bHLH proteins and contributes to determining DNA binding specificity. The PAS domain, named after the founding members of this family (Period-Arylhydrocarbon nuclear translocator-Single minded), is a multifunctional protein surface responsible for such diverse activities as ligand binding, PAS protein dimerization and non-PAS protein interactions (Kewley et al., 2004). In addition to environmental adaptation, some members of the bHLH-PAS family regulate development.

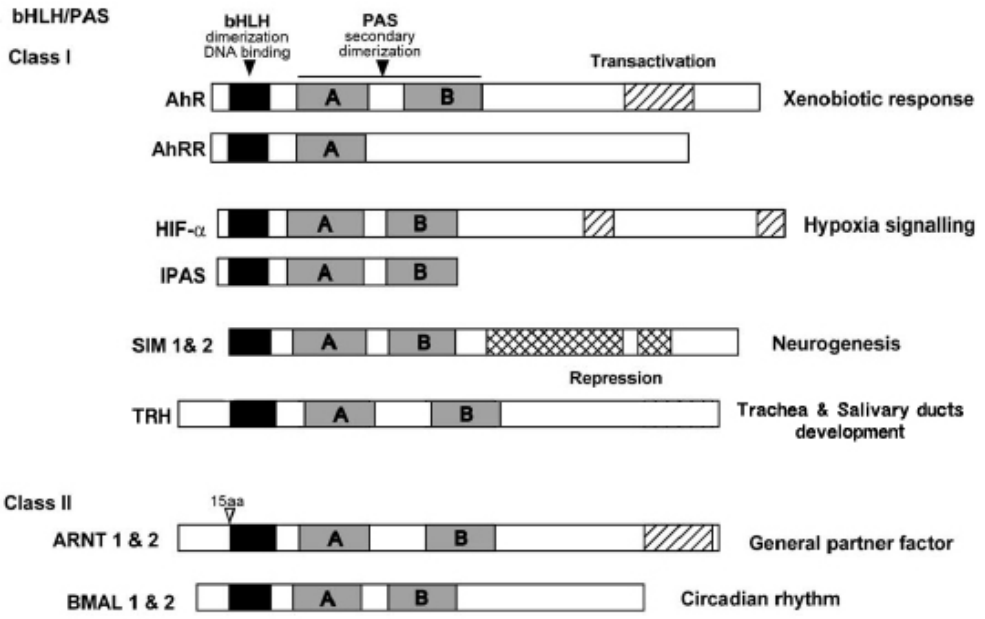


Figure 2. Schematic representation of the domain structure of bHLH-PAS family members. Modified from Kewley, 2004.

The bHLH-PAS proteins share a conserved sequence structure (Fig. 2). The bHLH domain is located near the amino terminus. The basic region binds DNA and the HLH domain promotes dimerization. These residues are followed closely by the PAS domain. The carboxy-terminal residues contain transcriptional activation domains or repression domains. The PAS domain found in bHLH-PAS proteins is ~260-310 amino acids long; it is subdivided into two well-conserved regions, PAS-A and PAS-B, separated by a poorly conserved spacer. Within both the A and B regions lie a copy of a 44-amino acid repeat referred to as the PAS repeat (Crews, 1998).

The bHLH-PAS family members can be grouped into two classes. Class I bHLH/PAS factors neither homodimerize nor heterodimerize with other Class I factors,

and include the aryl hydrocarbon receptor (AhR), the hypoxia inducible factors (HIF; HIF-1α, HIF-2α, and HIF-3α) and single minded proteins (SIM1 and SIM2). To form active transcription factor complexes they must dimerize with a Class II bHLH/PAS factor, which promiscuously homo- and heterodimerize. The best characterized Class II protein is the ubiquitous ARNT. Other members of this class include the tissue restricted ARNT2, and the circadian rhythm proteins BMAL1 and BMAL2 (Kewley et al., 2004).

The aryl hydrocarbon receptor (AhR), also known as the Dioxin receptor, is one of the best characterised bHLH/PAS proteins. In its inactive state, the AhR is found in the cytoplasm, stably associated with two molecules of heat shock protein 90 (Hsp90), p23 and hepatitis B virus X-associated protein (XAP2/AIP/Ara9). Following ligand binding, the AhR/Hsp90 complex translocates to the nucleus where Hsp90 is exchanged for partner protein ARNT. The ligand-bound AhR/ARNT heterodimer, then, binds xenobiotic response elements (XREs) that contain a –GCGTG- core binding sequence (Crews, 1998).

The bHLH–PAS proteins play an important role in responding to low oxygen levels in vertebrates and probably invertebrates. Responses to hypoxia are mediated by three bHLH/PAS proteins, HIF-1α, HIF-2α (Endothelial PAS domain protein 1) and HIF-3α. For example, HIF-1a protein is stabilized under hypoxic conditions by a poorly characterized O2 sensing pathway. It then dimerizes with Arnt and binds hypoxia response elements (HREs; ACGTG core sequence) on target genes (Crews & Fan, 1999). HIF-1a mediates the physiological response to both hypoglycemia and hypoxia by

upregulating genes that encode glycolytic enzymes, erythropoietin (EPO), and vascular endothelial growth factor (VEGF) (Semenza, 1998).

The trachealess (trh) gene encodes a bHLH-PAS protein that is specifically expressed in the developing trachea plus posterior spiracle, salivary gland placode, and salivary ducts. The functional similarity between Sim and Trh is remarkable; both are lineage-specific regulators, autoregulatory, and bHLH-PAS transcriptional activators. Work described below indicates that Sim and Trh control transcription in a similar fashion by binding the same DNA sequence element using Tgo as a dimerization partner (Crews, 1998).

Similar to other members of the bHLH/PAS family, dSim functions as a heterodimer with Tango, the *Drosophila* ortholog of Arnt (Crews et al., 1992). This complex binds central midline elements (CME: -TACGTG-) to regulate genes that direct midline cell development and axon growth, such as TGF-β, FGF-receptor and Slit (Lee et al., 1999; Estes et al., 2001; Crews et al., 1992) (Fig. 3A). For example, dSim activates Slit secretion from midline cells which blocks axon growth by binding the Slit receptor, Roundabout (Robo), to repel axons from crossing the midline or to prevent those that have crossed from migrating back (Brose et al., 1999) (Fig. 3B). Thus, *Drosophila* CNS midline cell regulation of axon branching and development is dependent upon dSim regulation of genes involved in cell fate determination and stromal-epithelial interactions.

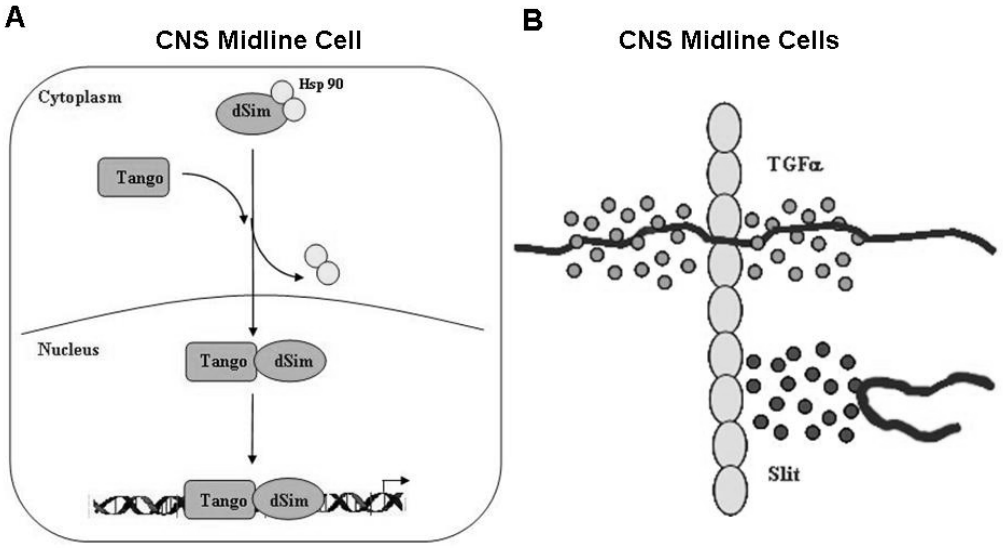


Figure 3. *Drosophila* Sim and development of CNS midline. dSim functions as a heterodimer with Tango and binds central midline elements (CME) to regulate genes that direct midline cell development and axon growth, such as TGF-β, FGF-receptor and Slit.

Two mammalian homologs of dsim, Sim1 and Sim2, share a high degree of similarity in their PAS domains, but little conservation is apparent in their carboxyl termini. Sim1 and Sim2 interact with ARNT and bind the CME (Fig. 4), but differ from *Drosophila* sim, the aryl hydrocarbon receptor (AHR) and hypoxia inducible factor (HIF) by functioning as transcriptional repressors (Moffett et al., 1997; Moffett & Pelletier, 2000). Mouse Sim1 and Sim2 show considerable sequence divergence from dsim over the C-terminal region and unlike the other members of the bHLH/PAS family, do not contain transactivation domains and instead act as repressor proteins when heterodimerised with ARNT (Ema et al., 1996; Moffett et al., 1997). Sim2-mediated repression occurs in two ways, through both sequestration of ARNT from other partner proteins and active transrepression, i.e. the direct repression of the transactivation domain of ARNT (Ema et al., 1996). A repression domain has been mapped to the C-terminus of Sim2, and

Sim2 repression is dependent on this domain (Moffett et al., 1997).

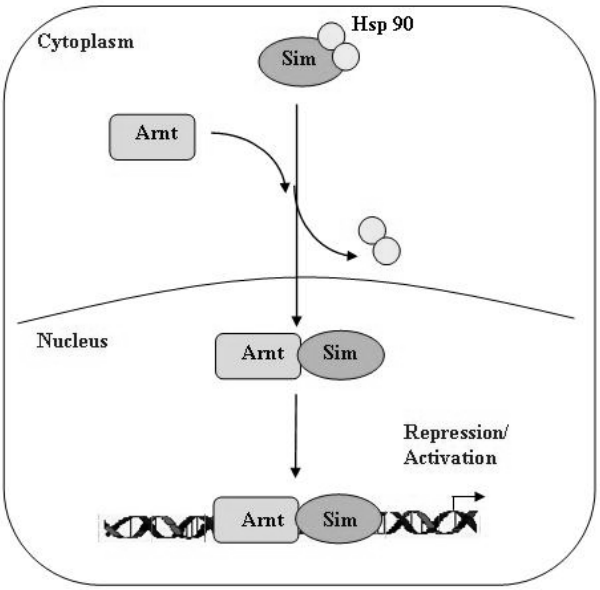


Figure 4. Transcriptional regulation by Sim. Sim interacts with ARNT and bind the CME to function as transcriptional repressors.

Sim1 and Sim2 are expressed in a variety of tissues including brain, kidney, lung and skeletal muscle where they play important developmental roles. After human SIM2 was first identified (Chrast et al., 1997), the mouse homolog was subsequently mapped to the syntenic region on chromosome 16. In addition, a splice variant of human SIM2, designated SIM2 short (SIM2s), has also been identified (Chrast et al., 1997) (Fig. 5). This splice variant, which is missing exon 11 and therefore lacks a portion of the region implicated in mediating the repressive effects of SIM2, is reported to be involved in cancer susceptibility (DeYoung et al., 2002; DeYoung et al., 2003); however, functional differences between these two isoforms have not been reported.

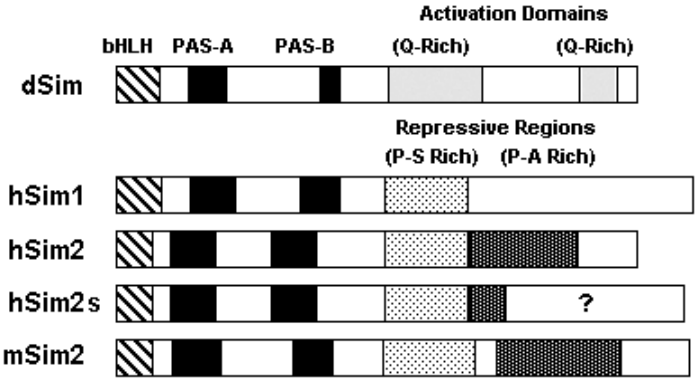


Figure 5. Comparison of *Drosophila*, human and mouse single minded genes.

Since the bHLH-PAS proteins share structural motifs and common binding partners, it is not surprising that cross-talk can occur between PAS-protein mediated pathways. The relationship among bHLH-PAS proteins is depicted in Figure 6.

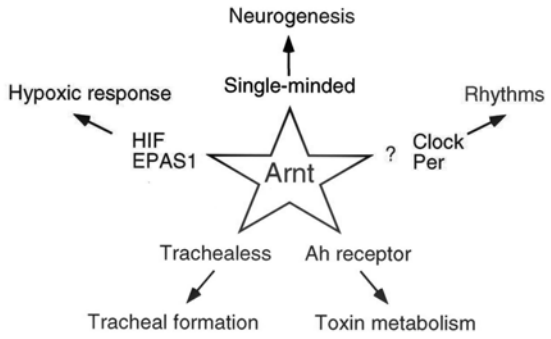


Figure 6. Array of Arnt interactions. The diagram shows the different proteins that Arnt/Tgo interacts with and the developmental and physiological processes these protein complexes control. The *in vivo* relationship between Arnt, Clock, and Per remains speculative.

In the case of the HIF proteins, both SIM1 and SIM2 can compete with HIFs for ARNT binding, and interact with a prototypical hypoxia response element (HRE) to affect gene expression (Moffett et al., 1997; Moffett & Pelletier, 2000; Woods and

Whitelaw, 2002). Interestingly, SIM1/ARNT, but not SIM2/ARNT, can induce transcription of an HRE-controlled reporter gene via the C-terminal transactivation domain of ARNT (Woods and Whitelaw, 2002). In contrast, ARNT-mediated transactivation of a CME-controlled reporter gene is severely impaired in the presence of SIM2 and is dependent upon the SIM2 dimerization domain and carboxyl terminus, which contains two separate repressive domains (Moffett et al., 1997; Moffett & Pelletier, 2000). Since repression by SIM2 is not specific for ARNT, as SIM2-Gal4 fusion constructs have repressive effects on a thymidine kinase promoter (Chrast et al., 1997), it is thought that SIM2-mediated repression can also occur through direct interactions with the basal transcription machinery. Similar to dSim, SIM1 and SIM2 also bind and regulate transcription through a consensus CME, which isn't surprising since the CME core sequence (5'-ACGTG-3') is identical to that of the HRE (Fig. 7). As was seen for the HRE, SIM1 strongly activates transcription of a CME-controlled gene through the transactivation domain of ARNT, whereas SIM2 is repressive (Moffett et al., 1997).

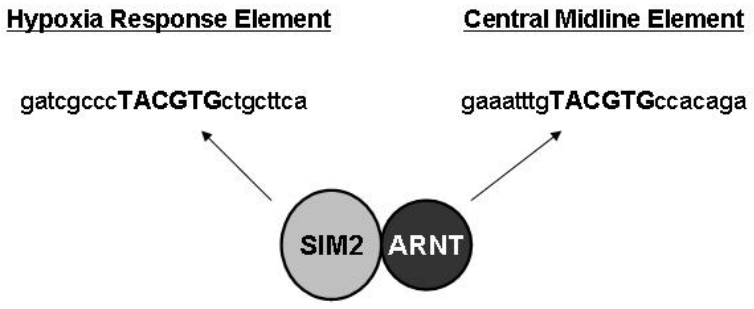


Figure 7. Comparison of HRE and CME.

1.4. Mammary development

Branching morphogenesis is regulated by positive and negative factors that direct the migration of cells during development in numerous branching tissues and organs including lung, kidney, central nervous system and mammary glands. Interactions between the epithelium and its surrounding stroma are the major driving force for induction of branching. A problem faced by all branching organs is that of confining growth to a select few cells, for example those at the tips of the terminal end bud (TEB) in the developing mammary gland (Fig. 8). During tumorigenesis this balance between the stroma and epithelium is disrupted leading to uncontrolled growth and metastasis.

The mammary gland is a dynamic organ the structure of which changes throughout the female reproductive cycle. Development of the gland occurs in defined stages that are connected to sexual development and reproduction, namely embryonic, prepubertal and pubertal stages, pregnancy, lactation and involution. Two cellular compartments constitute the gland: the epithelium and the surrounding stroma, which are derived embryologically from ectoderm and mesoderm respectively (Parmar & Cunha, 2004). Mammary gland development relies on interactions between the epithelium and stroma to drive cell migration and differentiation. Disruption of the interaction between the epithelium and stroma can both induce and promote breast cancer. Crosstalk between the mammary epithelium and stroma is also crucial for the proper patterning and function of the normal mammary gland. During its developmental cycle the mammary gland displays many properties associated with breast cancer. Moreover, many of the factors implicated in breast cancer are also vital for mammary development.

Understanding how these factors function in normal development may provide a better understanding of how tumors begin and thrive (Wiseman and Werb, 2002).

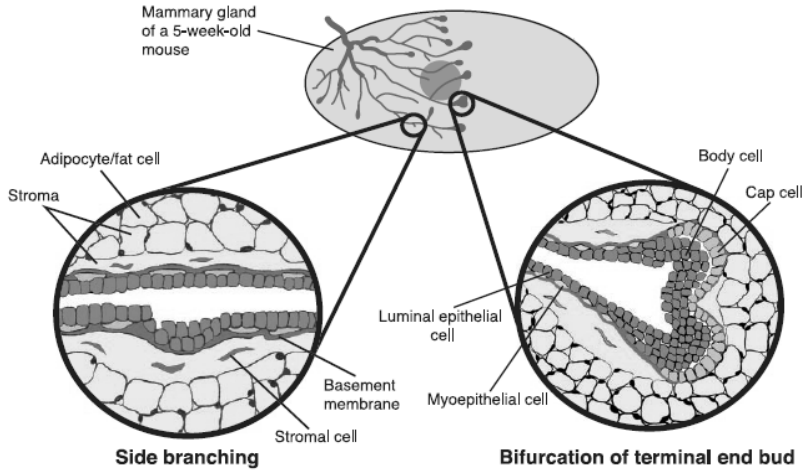


Figure 8. The two distinct mechanisms of branching morphogenesis in the pubertal mouse mammary gland. The mouse mammary gland branches through two distinct mechanisms: bifurcation of TEBs and side branching. Bifurcation of TEBs to form primary and secondary branches occurs only from immature ducts. The branch point is formed through deposition of stroma at the cleft site, and the ducts extend directly into adipose tissue, without myoepithelial cells or stroma and with only a minimal basement membrane at their invasive front. In contrast, in side branching, a new branch forms from a mature duct. First, the region where the bud is to form must be defined. Then the emerging bud extrudes through and remodels a region containing layers of myoepithelial cells, basement membrane, and periductal stroma. Distinct molecules have been implicated in each type of branching. Factors involved in side branching include the progesterone receptor, Wnts, HSPGs, nuclear factor kB (NFkB), MMPs, TIMP-1, TGFβ and its receptor (TGFβIIR), gelsolin, P-cadherin, CCAAT/enhancer binding protein b (C/EBPβ), CSF-1, Stat5a, and Stat5b. Factors involved in TEB formation include b1 integrin, laminin-1, MMPs, discoidin domain receptor-1 (DDR-1), GH, IGF-I and its receptor IGF-IR, Ptc-1, inhibins and activins, and p27Kip-1. Adapted from Wiseman & Werb, 2002.

Mammary gland development is unique because it mainly occurs post-embryonically during adolescence while most vertebrate organs are patterned during embryogenesis and then maintain their basic structure throughout adult life. In the mouse, mammary gland development begins at approximately embryonic day 10, when the epithelium

grows from the nipple region into the surrounding stroma establishing a rudimentary mammary tree. At puberty, ovarian hormones initiate a proliferative phase of ductal elongation, which is driven by the TEB (Chepko and Smith, 1999; Daniel et al., 1989; Muller and Neville, 2001).

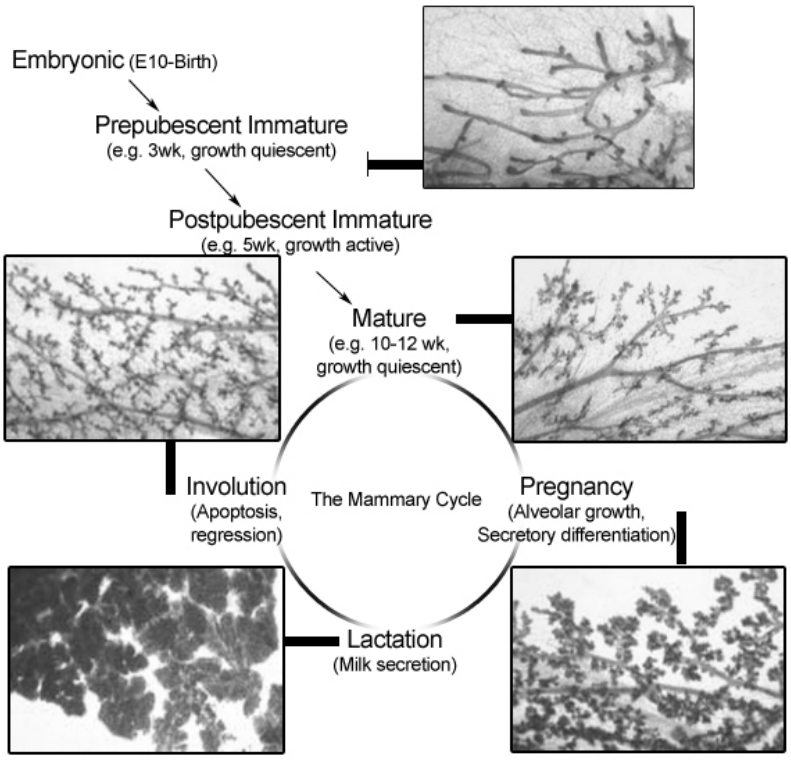


Figure 9. Mammary gland development in mice

The TEB is a bulb-like structure that is highly proliferative within a single "cap cell" layer that invades the fat pad leaving differentiated ducts behind (Silberstein, 2001). During pregnancy, the mammary gland undergoes expansion and terminal differentiation of the mammary epithelium into milk-producing lobuloalveolar structures and

differentiation of the large fat cells into tiny preadipocytes (Hennighausen & Robinson, 2001). Terminal differentiation of the alveolar epithelium is completed at the end of gestation with the onset of lactation. After weaning, the entire alveolar epithelium of the mammary gland dies by apoptosis, the fat cells redifferentiate, and the gland is remodeled back to a state resembling that of the adult nulliparous mouse (Alexander et al., 2001). The mammary gland development cycle is depicted in Figure 9.

1.4.1. Stroma plays an important role in mammary gland development

During the mammary development, the mammary epithelium proliferates and invades into the stromal tissue. However, the mammary stroma also plays an important role by providing the epithelial cells with both instructive and permissive signals (Wiseman & Werb, 2002). The stroma consists of many cell types, including fibroblasts, adipocytes, endothelial cells, extracellular matrix and inflammatory cells, each subject to regulation throughout the entire developmental cycle. In the embryonic stage, mouse mammary epithelial buds emerge from the embryonic epidermis as a result of stromal induction. This initial stage of mammary gland development depends on reciprocal signaling between epithelium and the stroma instead of systemic cues (Parmar and Cunha, 2004). When embryonic mammary epithelium was co-cultured with embryonic salivary mesenchyme, the mammary tissue developed salivary gland-like lobules (Sakakura, 1991). Sakakura also showed that adult mammary tissue also responded in this way to salivary stromal signals. At puberty, the TEB develop on the end of the ducts, the site of intense DNA synthesis that elicits ductal elongation. In addition to the terminal end buds, lateral buds develop along mature ducts, grow a short distance, and then stop

because of the presence of other competing ductal elements. Alveolar buds develop in most mouse strains only in response to the hormones of pregnancy. Postnatal mammary gland development and function are highly dependent upon the actions of pituitary, adrenal and ovarian hormones such as estrogen, progesterone, prolactin and corticosteroids (Parmar & Cunha, 2004).

Table 4. ***In vivo*** **murine mammary morphogens.** Adapted from Fata et al., 2004.

Morphogen	Branching*	Mode of application
Extracellular factors		
Estrogen	↑	Administered
Estrogen	↑	Slow-release pellet
Anti-estrogen	↓	Slow-release pellet
Progesterone	↓	Slow-release pellet
Parathyroid-hormone-related peptide	↓	Homozygous deletion
Parathyroid-hormone-related peptide	↓	Transgenic overexpression
Parathyroid hormone	↓	Transgenic overexpression
Corticosterone	↓	Administered
Epidermal growth factor (EGF)	↑	Slow-release pellet
TGF-β1	↓	Slow-release pellet
TGF-β2	↓	Slow-release pellet
TGF-β3	↓	Slow-release pellet
Insulin-like growth factor 1	↓	Homozygous deletion
Growth hormone	↑	Slow-release pellet
Hepatocyte growth factor	↑	Transgenic overexpression
Relaxin	↑	Systemically administered
Amphiregulin	↑	Overexpressing transplants
Amphiregulin	↑	Slow-release pellet
Hereguln-α	↑	Slow-release pellet
Heregulin-β	↑	Slow-release pellet
Wnt-4	↓	Homozygous deletion
Wnt-4	↑	Overexpressing transplants
C-neu	↓	Transgenic overexpression
Colony-stimulating factor	↓	Homozygous deletion
Eotaxin	↓	Homozygous deletion
MMP-3 (stromelysin-1)	↑	Transgenic overexpression
TIMP-1	↑	Transgenic antisense TIMP-1
TIMP-1	↓	Slow-release pellet
Receptors		
EGF receptor (EGFR)	↓	Homozygous deletion
Epidermal-growth-factor receptor	↑	Transgenic/dominant negative
Estrogen receptor α	↓	Homozygous deletion
Progesterone receptor A (PRA)	↑	Transgenic overexpression
Progresterone receptor A/B	↓	Homozygous deletion
Prolactin receptor	↓	Homozygous deletion
TGF-β2 receptor	↑	Transgenic/dominant negative
Intracellular factors		
CCAAT/enhancer binding protein β	↓	Homozygous deletion
Gelsolin	↓	Homozygous deletion
Hox6C	↓	Homozygous deletion

* ↑, augmentation of branching; ↓, inhibition of branching. MMP, matrix metalloproteinase; TGF, transforming growth factor; TIMP, tissue inhibitor of metalloproteinase.

Two mechanistically distinct processes, TEB bifurcation and sprouting of side branches from mature ducts, in mammary gland branching are depicted in Figure 8. During TEB bifurcation, the distal epithelial cells abut the fat cells through a sparse basement membrane, and stromal matrix is deposited to form a cleft at the site of bifurcation. In contrast, side branches must extend through the layer of myoepithelial cells, degrade the basement membrane that surrounds the mature epithelial ducts, and invade a periductal layer of fibrous stromal tissue that separates the epithelium from the fat cells of the mammary fat pad (Wiseman & Werb, 2002) (Fig. 8). A number of paracrine, juxtacrine, and autocrine factors are known to affect mammary gland branching morphogenesis and summarized in Table 4 (Fata et al., 2004).

1.4.2. The ECM and stromal factors in mammary branching and tumorigenesis

Interaction between the epithelium and the extracellular matrix (ECM) plays a major role in mammary gland branching morphogenesis. Mammary gland branching morphogenesis is dependent, in part, on the ECM, ECM receptors, such as integrins, and ECM-degrading enzymes, including matrix metalloproteinases (MMPs) and their inhibitors, tissue inhibitors of metalloproteinases (TIMPs). There is some evidence that these ECM cues affect one or more of the following processes: cell survival, polarity, proliferation, differentiation, adhesion, and migration (Fata et al., 2004). Direct attachment of epithelial cells to ECM occurs through basally located integrins (Table 5) and nonintegrin ECM receptors (Table 6).

Table 5. Integrin heterodimers and their corresponding ECM ligands. Adapted from Fata et al., 2004.

Integrin	ECM ligands
$\alpha_1\beta_1$	Laminins, collagens
$\alpha_2\beta_1$	Laminins, collagens, tenascin
$\alpha_3\beta_1$	Laminins, collagens, fibronectin, entactin
$\alpha_4\beta_1$	Fibronectin, osteopontin
$\alpha_5\beta_1$	Fibronectin, tenascin
$\alpha_6\beta_1$	Laminins
$\alpha_7\beta_1$	Laminins
$\alpha_8\beta_1$	Fibronectin, tenascin, vitronectin
$\alpha_9\beta_1$	Collagen type I, laminin, tenascin, osteopontin
$\alpha_V\beta_1$	Vitronectin, fibronectin, osteopontin
$\alpha_2\beta_1$	Collagen type I
$\alpha_{IIb}\beta_3$	Fibronectin, vitronectin, thrombospondin
$\alpha_V\beta_3$	Vitronectin, fibronectin, osteopontin, entactin, thrombospondin, denatured collagens, tenascin, laminin
$\alpha_6\beta_4$	Laminins
$\alpha_V\beta_5$	Vitronectin, osteopontin, fibronectin
$\alpha_V\beta_6$	Fibronectin
$\alpha_V\beta_7$	Fibronectin
$\alpha_V\beta_8$	Fibronectin, vitronectin

Table 6. Nonintegrins implicated in branching morphogenesis. Adapted from Fata et al., 2004.

Nonintegrin	ECM ligands
β-1,4-glactosyltransferase	*N*-acetylglucosamine residues
Dystroglycan	Laminin 1
Discoidin domain receptor 1 (DDR1)	Collagens
Galectin	Laminin, fibronectin, vitronectin, *N*-acetylglucosamine residues

TEB formation and ductal invasion requires ECM factors, such as discoidin domain receptor–1, which can serve as a collagen receptor, β1 integrin, which recognizes many ECM proteins, the ECM protein laminin-1, and several matrix metalloproteinases (MMPs), which cleave ECM and other proteins in the cellular microenvironment. Specific MMPs refine the mammary branching pattern by distinct mechanisms during

mammary branching morphogenesis. Wiseman et al. (2003) proposed that MMP2 and 3 regulates mammary gland branching in different phases (Fig. 10).

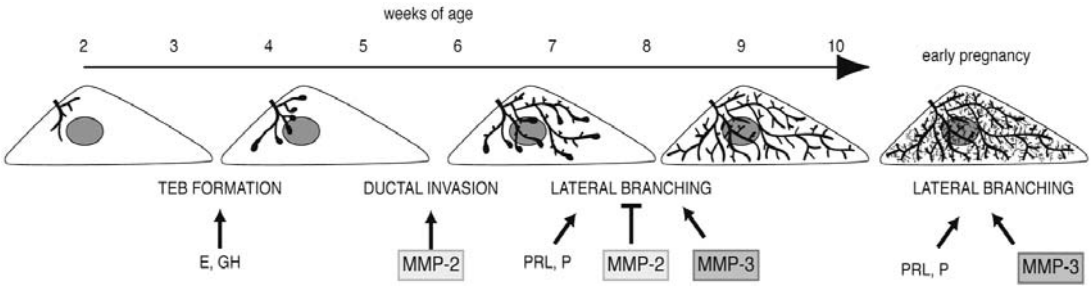


Figure 10. Model for different phases of mammary gland branching morphogenesis. Before puberty, the mammary epithelial is small and simply branched. In response to the release of estrogen (E) and growth hormone (GH), at 3 wk old TEBs form. MMP-2 is then involved in inducing TEBs to invade and the ducts begin to fill the fat pad by branching dichotomously through bifurcation. At 6–8 wk old, the mammary ducts branch laterally. This process is suppressed by MMP-2 and induced by MMP-3 and may be related to changes in the response of the gland to progesterone (P) and prolactin (PRL). The fat pad is filled with ducts at 10 wk old and is relatively quiescent until pregnancy, when there is another wave of lateral branching that is regulated by MMP-3, P, and PRL before the formation of lobular alveoli. Adapted from Wiseman et al., 2003.

DDR1-deficient mice have excessive mammary collagen deposition, delayed ductal development, enlarged terminal end buds, hyperproliferative ducts, and incomplete lactational differentiation (Vogel et al., 2001). Blocking β1 integrin or the γ1 chain of laminin induced terminal end bud regression, fewer terminal end buds, and decreased ductal elongation in developing mammary glands (Klinowska et al., 1999). MT1-MMP cleaved laminin-5 to release bioactive laminin fragments that induce the migration of breast epithelial cells (Koshikawa et al., 2000). Proper side branching also requires that the ECM and the cellular microenvironment surrounding the ductal epithelium be maintained. Unrestrained side branching often results in tumorigenesis. Indeed,

overexpression of MMP-3, MMP-14, Wnt-1 induced excessive side branching and eventual tumorigenesis in the mouse mammary gland. In contrast, a reduction in side branching occurs in mice deficient in MMP-3 and Wnt-4 (Wiseman & Werb, 2002). Liu and colleagues (2003) proposed that TCF-dependent transactivational activity is suppressed in 50% of cells in Sydecan-1 knockout glands, and conclude that the major effect of Sdc1 does not map to the activity of the Wnt signaling complex, but to another pathway to create or stabilize the b-catenin/TCF-responsive tumor precursor cells in mouse mammary gland. Transforming growth factor-β1 (TGF-β1) influences mammary gland development by altering the composition of the ECM. Localized release of TGF-β1 from a pellet implanted in the stromal compartment induced intense collagen I gene expression (Silberstein et al., 1990). Transforming growth factor-ß1 also positively regulates expression of ECM proteins such as fibronectin, collagen IV, and laminin in human mammary epithelial cells (Stampfer et al., 1993). The virgin P-cadherin–null female mice displayed precocious differentiation of the mammary gland and the P-cadherin mutant mice develop hyperplasia and dysplasia of the mammary epithelium with age (Radice et al., 1997).

1.4.3. Stromal regulation of involution

After weaning, the gland goes through a process of death and remodeling termed involution, which is initiated by milk stasis once milk removal has ceased (Quarrie et al., 1996). During involution, 90% of the epithelium dies by apoptosis and fat cells replace that tissue. There are three stages of involution. The first stage of involution is triggered by local factors related to milk accumulation rather than changes in systemic hormone

levels. When suckling is interrupted, the first need of the mother is to, at least temporarily, reduce milk production. However, these changes are reversible if suckling is resumed within 48 hrs. This provides a mechanism for both the mother and her offspring to survive short separations. In this stage, individual mammary epithelial cells die by apoptosis, but the general structure of the mammary gland is maintained (Furth et al., 1997). This stage is regulated locally by milk stasis, is dependent on the tumor suppressor gene p53 (Wiseman & Werb, 2002). If suckling is not resumed, then the irreversible second stage of involution begins. This stage is characterized by decreased systemic hormone levels and proteinase activation. Irreversible tissue remodelling follows with near complete loss of the epithelial cell compartment (Furth et al., 1997). The second stage is also characterized by apoptosis, but this is mediated by lactogenic hormones and is independent of p53. This stage is irreversible and is dependent on proteinases. The third stage is a biosynthetic phase in which the mammary stroma is remodeled and repopulated with adipogenic cells (Wiseman & Werb, 2002).

In the first stage of involution, apoptosis is inhibited by the transcription factors Stat5a and interferon regulatory factor–1 (IRF-1) and promoted by the transcription factor Stat3 and the growth factor TGF-β3. Stat3 may induce apoptosis by inducing a known promoter of apoptosis, IGF binding protein–5 (IGFBP-5), and by down-regulating Stat5a (Wiseman & Werb, 2002).

During the second stage of involution, apoptosis likely occurs because the epithelial cells lose their adhesion to a basement membrane, which is destroyed by the increased proteinase activity. As a result, the cells lose survival signals from the ECM (Wiseman

& Werb, 2002). The expression of TIMPs is downregulated around day 3 of involution with the concomitant upregulation of MMPs and serine proteinases such as stromelysin-1, stromelysin-3, gelatinase A and urokinase-type plasminogen activator (Marti et al., 1999).

Involution of the mammary gland can be achieved more quickly in a situation of reduced proteinase activity. TIMP-1 overexpressing mice or MMP-3 deficient mice showed an accelerated involution which was due to an increase in adipogenesis rather than an alteration in apoptosis (Alexander et al., 2001).

1.4.4. Development of the human mammary gland

It will be meaningful to take a closer look at human mammary development. Parmar and Cunha (2004) presented a profound description on the development of the human mammary gland in their recent review. The first visible indication of human mammary gland development can be found during day 35 (4th week), with the proliferation of paired areas of epithelial cells in the epidermis of the thoracic region. Subsequently, these areas of proliferation extend in a line between the fetal axilla and inguinal region and form two indistinct ridges called the mammary ridges or milk streaks. By the end of week 6, the mammary ridges have regressed back to two areas in the thoracic region, where two solid epithelial masses (the mammary buds) begin to grow downwards into the underlying mesenchyme (Fig. 11). This solid core of cells continues to evaginate into the underlying mesenchyme and becomes surrounded by a more cellular zone of fibroblast-like cells within a dense collagenous stroma. The 9th week is the cone stage, at which there is further inward growth of the mammary placode. Between the 10th and

12th weeks, called the budding stage, epithelial buds sprout from the invading placode, and the buds become lobular in shape, with notching (indentation) of the epithelial–stromal border. In the epidermis overlying the developing gland, the nipple begins to form. During the same period, the mesenchymal cells differentiate into fibroblasts, smooth muscle cells, capillary endothelial cells and adipocytes. Further branching into 15–25 solid epithelial cords marks the branching stage at 15 weeks.

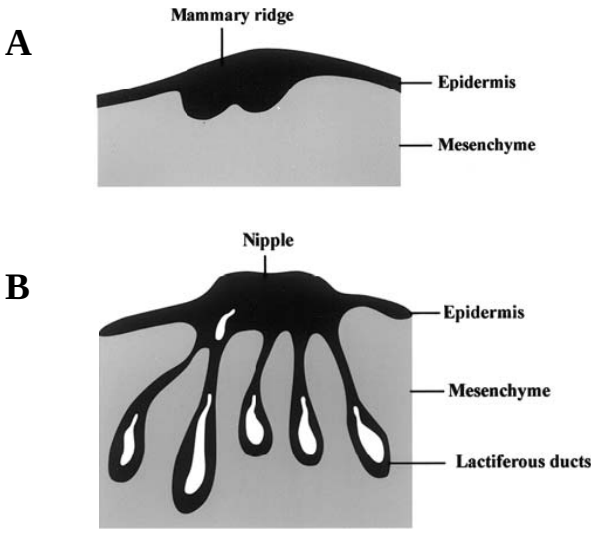


Figure 11. Schematic drawing of embryonic human breast development. A. Early epidermal mammary ridge. B. From the mammary ridge, solid epithelial sprouts grow into underlying mesenchyme to form lactiferous ducts. Adapted from Parmar and Cunha, 2004.

Around week 20, the solid mammary cords canalize, and the epidermis in the region of the nipple becomes depressed, forming the mammary pit. The epithelial cells lining the ducts first appear as a bilayer of cuboidal cells. The luminal layer rapidly acquires the characteristics of secretory cells, whereas the basal layer becomes myoepithelial. By

6 months of gestation, the basic tubular architecture of the fetal gland has become established. In contrast to rodent development, glands develop similarly in both the female and male human fetuses (Howard & Gusterson 2000).

After birth, the mammary gland becomes quiescent until the onset of puberty in girls, when it resumes growth of both the glandular tissue and the surrounding stroma. The ducts elongate, branch and form club-shaped terminal end buds, as seen in the mouse. The terminal end buds give rise to new branches and small ducts or alveolar buds (Parmar & Cunha, 2004). As in the fetus, from birth to puberty there are no identifiable morphological differences in the development of the breasts in boys and girls (Howard & Gusterson 2000).

At puberty, changes occur in both the epithelium and stroma. In the stroma, there is an increase in the amount of fibrous and fatty tissue, with the adult non-lactating breast consisting of 80% or more of stroma. Indeed, the extension of ducts in the pubertal human breast is preceded by proliferation of connective tissue; fatty tissue is believed to inhibit growth of human breast epithelium (Howard & Gusterson 2000). Small bundles of primary and secondary ducts grow, divide, and form club-shaped terminal end buds. Terminal end buds give origin to new branches, twigs, and small ductules or alveolar buds. Alveolar buds cluster around a terminal duct, forming the lobule type 1 or virginal lobule (Fig. 12) and each cluster is composed of approximately 11 alveolar buds. Lobule formation in the female breast occurs within 1–2 years after onset of the first menstrual period. Full differentiation of the mammary gland is a gradual process taking many years, and in some cases, if pregnancy does not supervene, is never attained

(Russo & Russo, 2004).

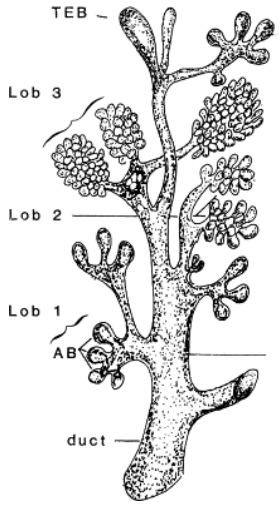


Figure 12. Diagrammatic representation of the lobular structures of the human breast. Adapted from Russo and Russo, 2004.

The lobules have been classified as types 1, 2, 3 and 4 lobules (Russo & Russo 1987). In type 1 (virginal lobule), alveolar buds cluster around a terminal duct. Terminal ducts or alveolar buds are lined by a bi-layered epithelium, whereas four layers of epithelial cells line terminal end buds. The transition from lobule type 1 to type 2 to type 3 is a gradual process of sprouting of new alveolar buds. In type 2 and type 3 lobules, ductules increase in number from about 11 per type 1 lobule to 47 and 80 ductules per type 2 and type 3 lobules respectively. Type 1 lobules are mainly found in the breast of nulliparous young women, whereas lobules type 2 and type 3 are more frequent in the gland of parous women. Type 4 lobules are the maximal expression of development and differentiation in the adult gland, seen in pregnancy when glands are secreting milk and

have undergone complete functional differentiation (Figs. 12 & 13).

During pregnancy, the breast attains its maximum development; it occurs in two distinctly dominant phases characteristic of the early and late states of pregnancy. The early stage is characterized by growth consisting of proliferation of the distal elements of the ductal tree, resulting in the formation of ductules that at this stage, can be called acini, thus developing a lobule type 3 into a lobule type 4. The intensity of budding and degree of lobule formation goes beyond what has been observed in the virginal breast (Russo & Russo, 2004).

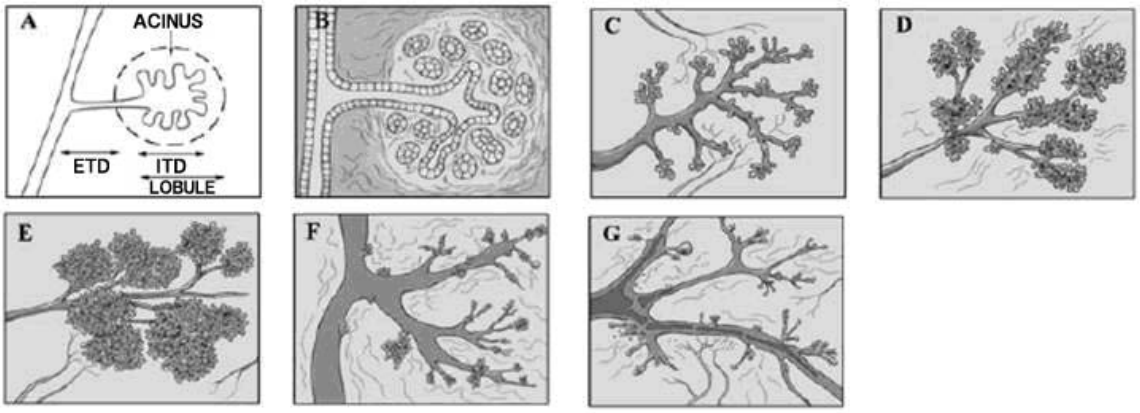


Figure 13. Drawings depicting the functional development of the human breast, the basic functional unit of which is the terminal ductal–lobular unit (TDLU). A. Diagram of a TDLU emerging from the extralobular terminal duct (ETD). B. Diagram depicting the histology of a normal TDLU in the resting nulliparous state, showing the terminal duct and associated acini. C. Drawing of terminal branching structures from a 15-year-old human breast. An extralobular duct is shown with rudimentary terminal branching structures. True lobules and acini are not yet present. D. Drawing of TDLUs from a 22-yearold nulliparous human mammary gland having well developed lobules and acini. E. Drawing of breast structures of a 30-year old pregnant woman. Note well developed TDLUs. F. Drawing of breast tissue of a 55-year-old parous menopausal woman. Ducts are dilated, and the TDLUs are atrophied. G. Drawing of breast tissue of a 80-year-old woman. Marked atrophy of ducts and TDLUs can be seen. Adapted from Parmar and Cunha, 2004.

1.5. Breast cancer

In the United States, breast cancer is the most frequently diagnosed cancer among women with an estimated 211,240 invasive breast cancer diagnoses in 2005. It is second only to lung cancer as a cause of cancer death among women and over 40,000 women will die from breast cancer each year in the United States. It was estimated that about 58,490 carcinomas *in situ* of the breast and 46,170 melanoma *in situ* would be newly diagnosed in 2005 (Jemal et al., 2005).

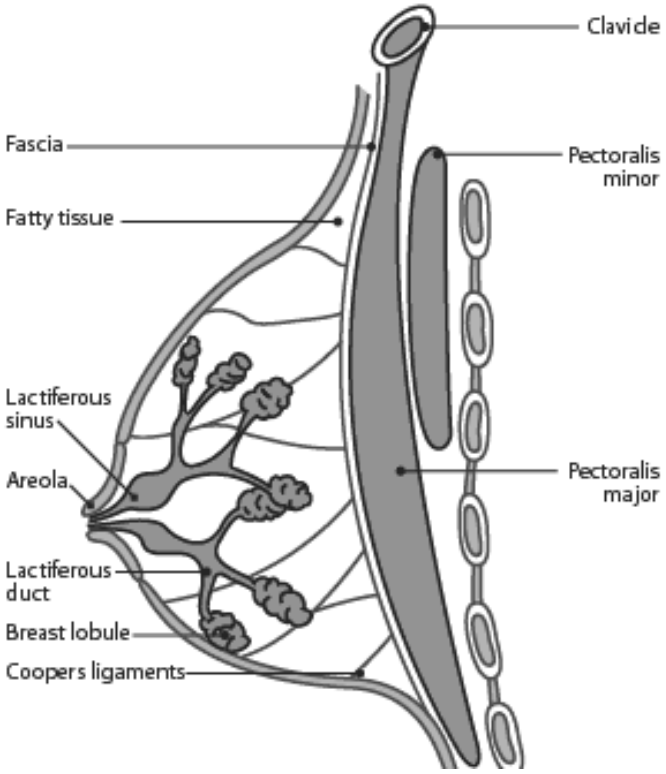


Figure 14. Anatomy of female breast. Adapted from Osborne, 2000.

The breasts lie over the pectoral muscles which attach them by a layer of connective tissue known as the fascia. The mammary gland is supported by two layers of Cooper's ligaments. These arise from stromal elements in the gland and insert into the skin and pectoral fascia, holding the breast against the chest. The breast is made up of the secretory glandular tissue and surrounding adipose tissue (Fig. 14).

The glandular tissue comprises between 15 and 20 lobes separated by septa of connective tissue. Each lobe has its own duct system and connects to a lactiferous duct, several of which converge to form a lactiferous sinus. These sinuses empty into the nipple where there are a number of duct openings. Each lobe has a lactiferous duct so there are a series of openings in the nipple. Each lobe consists of many lobules; hence the main lactiferous ducts give rise to many branches, the intralobular ducts. The nipple is surrounded by a pigmented area, the areola, which is lubricated by secretions from the sebaceous glands (Osborne, 2000).

Breast cancer refers to the uncontrolled growth and proliferation of cells that originate in the breast tissue and there are several types of tumors that may develop within different areas of the breast. Some cancers are called *in situ* because cancer cells are confined to the ducts and do not invade surrounding fatty and connective tissues of the breast. They are also called 'noninvasive'. Ductal carcinoma *in situ* (DCIS) is the most common form of non-invasive breast cancer (90%). Lobular carcinoma *in situ* (LCIS) is less common and many physicians do not classify LCIS as breast cancer. Other cancerous breast tumors are invasive or infiltrating. In this case, cancer cells

break through the duct and lobular wall and invade the surrounding fatty and connective tissues of the breast. Cancer can be invasive without being metastatic to the lymph nodes or other organs. Infiltrating lobular carcinoma (ILC) or invasive lobular carcinoma begins in the lobules of the breast, but often spreads to other regions of the body. ILC accounts for 10% to 15% of breast cancers. Infiltrating ductal carcinoma (IDC) or invasive ductal carcinoma begins in the ducts of the breast and penetrates the wall of the duct, invading the fatty tissue of the breast and possibly other regions of the body. IDC is the most common type of breast cancer, accounting for 80% of breast cancer diagnoses (Hoehne & Taylor E, 2005). Although less common, there are several other types of breast cancers: medullary carcinoma, mucinous carcinoma, tubular carcinoma, inflammatory breast cancer, Paget's disease of the nipple, and phylloides tumor (Hoehne & Taylor E, 2005).

1.5.1. Risk factors for breast cancer

Although the exact cause of breast cancer is unknown, it has been found that women with certain risk factors are more likely to develop breast cancer. A risk factor is anything that increases the chances that a particular disease will develop.

Genetics: Alteration of certain genes may cause some cells to become cancerous. It has been shown that 5-10% of breast cancer cases are hereditary (Lux et al., 2006). Genetic diseases predisposing an individual to cancer have led to the identification of numerous tumor suppressor genes including BRCA-1 and -2 (familial breast cancer), PTEN (Cowden's Disease), APC (familial polyposis), LKB1 (Peutz-Jeghers syndrome), WT1 (Wilms Tumor) and P53 (Li-Fraumeni syndrome) (Groden et al.,1991; Haber et al.,1990;

Hemminki et al., 1998; Liaw et al., 1997; Malkin,1994; Moynahan et al., 2002). Women who carry mutations of BRCA1 (breast cancer gene 1) or BRCA2 (breast cancer gene 2) are at higher risk of developing both breast and ovarian cancer than women who do not have these genetic mutations (Moynahan, 2002). Currently, women with BRCA1 mutations account for 5% of all breast cancer cases. Jewish women of Northeastern European decent (Ashkenazie women) have a higher prevalence of BRCA1. Human epidermal growth factor receptor 2(HER2) also plays a key role in regulating cell growth. The over-expression of HER2 causes increased cell growth and reproduction, often resulting in more aggressive tumor cells. HER2 protein over-expression affects over 20% of breast cancer patients (Martin, 2006). Germline p53 mutations have been identified in patients with Li-Fraumeni cancer susceptibility syndrome, an autosomal dominant disorder. It is characterized by a markedly increased risk of breast cancer with early-onset (Malkin, 1994).

Gender: Breast cancer is the most common form of cancer among women in the United States, while it is uncommon in males. One third of cancer diagnosis and 15% of cancer deaths are attributed to breast cancer in women (Lacey et al., 2002). The risk of development of breast cancer is 100 times higher in females compared to males (Andersen et al., 1992). Female breast cells are constantly exposed to the growth-promoting effects of estrogen and progesterone, thus making breast cancer much more common in women than men.

Age: The risk of developing breast cancer increases with age, with most breast cancers occurring after age 50 (Callebaut and Mornon, 1997). About 18% of breast cancer

diagnoses are among women in their 40s, while about 77% of women with breast cancer are older than 50 when they are diagnosed. In contrast, breast cancer incidence is less than 10 new cases per 100,000 women before age 25 and increases up to 100-fold by age 45 (Hulka & Moorman, 2001).

Race: White women are slightly more likely to develop breast cancer than are African-American women. But African-American women are more likely to die of this cancer. Asian, Hispanic, and Native American women have a lower risk of developing and dying from breast cancer (Dumitrescu & Cotarla, 2005).

Geographical location: Breast cancer incidence and mortality vary considerably by world region. In general, the incidence is high (greater than 80 per 100,000) in developed regions of the world and low (less than 30 per 100,000), though increasing, in developing regions; the range of mortality rates is much less (approximately 6-23 per 100,000) because of the more favorable survival of breast cancer in (high-incidence) developed regions (Parkin & Fernandez, 2006). Incidence rates are high in most of the developed areas (except for Japan, where it is third after colorectal and stomach cancers), with the highest age-standardized incidence in North America (99.4 per 100,000). The incidence is more modest in Eastern Europe, South America, Southern Africa, and Western Asia, but it is still the most common cancer of women in these geographic regions. The rates are low (<30 per 100,000) in most of Africa (with the exception of South Africa) and in most of Asia. The lowest incidence is in Central Africa (ASR, 16.5 per 100,000) (Parkin et al., 2005). The large variation of breast cancer incidence among or within different regions of the world may be attributed to genetic differences among

populations and/or differences in lifestyle, including diet and environmental exposures. Interestingly, breast cancer incidence increases in people who move from a region with low breast cancer incidence to other locations with higher breast cancer incidence. John et al. (2005) reported that breast cancer risk was 50% lower in foreign-born Hispanics than U.S.-born Hispanics. They also found that risk increased with increasing duration of residence in the United States, decreasing age at migration, and increasing acculturation. Migrants from urban areas had a risk 30% higher than migrants from rural areas. Migrants who had lived in the West for a decade or longer had a risk 80% higher than more recent migrants (Ziegler et al., 1993). These imply the crucial contribution of the environmental and sociocultural factors to breast cancer risk.

Reproductive factors: Reproductive hormones have been shown to play a critical role in breast cancer etiology in many epidemiologic data, animal models, and *in vitro* studies. Lifetime exposure to endogenous sex hormones is determined by several variables including age at menarche, age at first full-term pregnancy, number of pregnancies and age at menopause, which have been studied all in relation to breast cancer risk (Dumitrescu & Cotarla, 2005). Early menarche, late menopause, older age at first full-term pregnancy, lower parity, and fewer numbers of pregnancies may increase risk of breast cancer by affecting the endogenous reproductive hormones (Hulka & Mooreman, 2001).

Early age at menarche (less than 12 years of age) has been associated with an increase in breast cancer risk. Breast cancer risk declines 10–20% with each year that menarche is delayed (Titus-Ernstoff et al., 1998). Late age at menopause is also associated with

increased breast cancer risk. Longer menopause represents a longer time period of exposure to ovulatory menstrual cycles. Breast cancer risk is about two times greater for those with a last menstrual period at age 55 years or later than for those whose last period occurred at age 45 or younger (Bernstein, 2002). Early pregnancy has a protective effect against breast cancer. Both early age (less than 20 years versus more than 30 years) at first full-term pregnancy and higher parity decrease breast cancer risk to half of the risk of nulliparous women. Early age at second pregnancy further reduces the risk of breast cancer. In contrast, nulliparity and late age at first birth contribute towards an increased risk of developing breast cancer. Interestingly, women with their first birth after age of 35 are even at higher risk than nulliparous women (Dumitrescu & Cotarla, 2005). It also has been reported that prolonged lactation reduces the risk of breast cancer (Enger et al. 1998). Beral (2002) reported a 4.3% decrease in the relative risk of breast cancer for every 12 months of breastfeeding, in addition to a decrease of 7.0% for each birth. The decrease of breast cancer risk due to prolonged lactation may be explained by a markedly reduced susceptibility of the fully differentiated mammary gland to carcinogens, by the altered hormonal environment during pregnancy, and by the reduction of total number of ovulatory menstrual cycles and consequently cumulative ovarian hormone exposure (Dumitrescu & Cotarla, 2005).

Family history of breast cancer: Family history of breast cancer increases a woman's risk of developing the disease. Collaborative Group on Hormonal Factors in Breast Cancer (2001) reported that 87% of breast cancers occurred in women with no affected first-degree relatives, 12% in women with one affected relative, and 1% in women with

two or more affected relatives. About 20% to 30% of women with breast cancer have a family member with this disease.

Personal history of breast cancer: Women with a history of prior invasive breast cancer or history of noninvasive breast lesions have an increased risk for developing invasive breast cancer (Vokes and Golomb, 2003).

Hormone replacement therapy: It has become clear that long-term use of hormone replacement therapy (HRT) after menopause, particularly estrogens and progesterone combined increase the risk of breast cancer (Ross et al., 2000).

Alcohol: Many studies showed a slight increase in breast cancer risk with alcohol consumption. Jain et al. (2001) reported a slightly elevated risk when 10 g of ethanol were consumed. Breast cancer risk is higher for women consuming moderate to high levels of alcohol (3 drinks/day) compared with abstainers, and that there is a significant dose-response relationship beginning with intakes as low as 1 to 2 drinks per day. The proportion of breast cancer cases among US women that can be attributed to alcohol intake has been estimated to be as low as 2%, but the proportion is as high as 15% in Italy, where the average alcohol intake is higher (Singletary & Gapstur, 2001). A great number of epidemiological studies have clearly identified chronic alcohol consumption even in moderate amounts as a risk factor for breast cancer. 84% of the 69 case–control and 76% of the 21 cohort studies published so far show a positive association between ethanol intake and breast cancer. So have all six meta-analyses to date. It was calculated that 4% of all newly diagnosed breast cancer cases in the US are primarily due to alcohol (Poschl & Seitz, 2004). It has been demonstrated that postmenopausal

women who have a higher-alcohol and low-folate intake have an increased risk for developing estrogen receptor-negative tumors (Sellers et al., 2002).

Obesity: Obesity has been found to be a breast cancer risk in all studies, especially for women after menopause (Adderley-Kelly & Williams-Stephens, 2003; McTiernan et al., 2003). Weight gain during adult life has been associated with increased risk of postmenopausal, but not premenopausal, breast cancer (Huang et al., 1997; Magnusson et al., 1998). Huang et al. (1997) have also showed a stronger association between obesity and postmenopausal breast cancer risk in women who have never used hormone replacement therapy. Postmenopausal obesity increases the risk for developing breast cancer due to higher levels of endogenous estrogen, because adipose tissue is a primary source of circulating estrogens in postmenopausal women (McTiernan et al., 2003).

Other risk factors: There are many other breast cancer risk factors such as history of breast benign diseases, mammographic density, bone density, prolactin, diet, previous radiation, smoke, physical activity and so on (Dumitrescu & Cotarla, 2005).

1.6. Connection between mammary development and breast cancer

The developing mammary gland displays many of the properties associated with tumor progression, such as invasion, reinitiation of cell proliferation, resistance to apoptosis, and angiogenesis. The TEB invades into stromal tissue as a solid tumor does. Furthermore, the epithelium must retain the ability to initiate proliferation throughout its lifetime. The lactating mammary gland protects it from premature involution, actively resisting apoptotic signals. In addition, as the mammary gland needs the blood supply to

be adjusted, the mammary gland induces angiogenic remodeling as tumors do. As the mammary gland retains many of these properties throughout its lifetime, many of the factors essential for mammary gland development are also associated with cancer. In the mammary gland, a variety of genes have been implicated in the developmental processes, and many of these genes have also been linked to tumorigenesis (Wiseman and Werb, 2002).

1.7. Tumor suppressor genes

Tumor suppressor genes refer to those genes whose loss of function results in the promotion of malignancy. Tumor suppressor genes are usually negative regulators of growth or other functions that may affect invasive and metastatic potential, such as cell adhesion and regulation of protease activity (Osborne et al., 2004). The inactivation of various tumor suppressor genes is important in the development and progression of breast cancer. Tumor suppressor genes are considered to act mostly in a recessive fashion, i.e. some abnormality must affect both gene alleles. The classical inactivation of tumor suppressor genes, i.e. the Knudson 'Two-Hit' hypothesis, is caused by tumor suppressor gene loss due to chromosomal loss of one allele, and mutation of the other remaining allele. The two-hit hypothesis first proposed by Knudson predicts that loss of one copy of a tumor suppressor gene renders an organism more susceptible to tumorigenesis (Knudson et al., 2000). Consistent with this hypothesis, haplosufficiency for several genes including p27, p53, Ink4a, PTEN, BRCA1 and BRCA2 has been shown to enhance the rate of mammary cancer (D'Amico et al., 2003; Medina et al.,

2002; Berton et al., 2003; Muraoka et al., 2002; Kuperwasser et al., 2000; Li et al., 2001).

Besides genetic alteration, the inactivation of tumor suppressor genes can be result from nongenetic (or epigenetic) mechanisms, including hypermethylation (Foster et al. 1998), increased degradation (Storey et al. 1998), mislocalization (Chen et al. 1995), or abnormalities in other proteins that interact with the gene product.

Tumor suppressor genes have various mechanism of action. Some tumor suppressor genes, such as Rb or INK family of cdk inhibitors, are directly involved in controlling proliferation by regulating cell cycle checkpoints. The retinoblastoma (Rb) gene was the first tumor suppressor to be discovered. In breast cancer, mutation or loss of Rb is present in up to 30% of cases, and it has been associated with a greater likelihood of progression (Andersen et al., 1992).

In contrast, certain tumor suppressor genes have an indirect effect on growth. They are responsible for genome integrity, and changes in such genes lead to genome instability. For example, The ATM gene senses DNA damage and activates checkpoints and DNA repair pathways. Loss of both alleles of the ATM gene causes ataxia-telangiectasia, a syndrome of progressive cerebellar degeneration, blood vessel fragility, immunological defects, and predisposition to lymphoid malignancies (Shiloh, 2003).

Tumor suppressor genes are also are involved in epithelial-stromal and epithelial-epithelial interactions. E-cadherin is a major mammary epithelial cell-cell adhesion molecule and loss of E-cadherin function contributes to breast cancer progression by promoting cell proliferation, invasion and metastasis (Hu et al., 2003).

1.8. Epithelial-mesenchymal transitions

Polarized epithelial cells are converted into motile mesenchymal cells during embryonic development and in tumor progression. Epithelial-mesenchymal transitions (EMTs) occur as key steps during embryonic morphogenesis in many animal species. This process is very important of early development. In vertebrates it facilitates the formation of a three-layered embryo by gastrulation. EMT has also been implicated in the ontogeny of other structures including heart endocardium, musculoskeletal system, most craniofacial structures and the peripheral nervous system (Thiery, 2002). The development of the mammary gland probably employs these mechanisms. EMT is one of the most drastic aspects of epithelial cell plasticity. Some of the molecular programs of EMT might be involved in the development of the mammary gland, particularly at terminal end buds or possibly in lateral branching. However, EMT is also likely to be important in tumor progression. Increasing evidence suggests that EMT plays a specific role in the migration of cells from a primary tumor into the circulation.

One of the best markers of EMT associated with breast cancer is the loss of E-cadherin, and this is controlled in part by Snail family members, as is EMT associated with development (Vincent-Salomon and Thiery, 2003). Cadherins are essential for the development and maintenance of mammary luminal epithelial cells. Adhesion between vertebrate cells is generally mediated by three types of adhesion junction: ‘tight junctions (TJs)’, ‘adherens junctions (AJs)’, and desmosomes. Cadherins are the principal components of AJs and desmosomes, and cluster at sites of cell–cell contact in

most solid tissues (Cavallaro & Christofori, 2004) (Fig. 15).

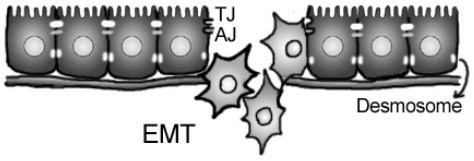


Figure 15. Three types of adhesion junctions and EMT.

Epithelial(E)-cadherin is a major mammary epithelial cell-cell adhesion molecule and has been implicated as an important signaling component in maintaining normal growth and differentiation (Delmas et al., 1999; Boussadia et al., 2002; Stockinger et al., 2001). Cellular polarity and inhibition of migration are promoted and controlled by E-cadherin-containing adherens junctions, and loss of E-cadherin is an early step in EMT (Berx et al., 1996; Nathke et al., 1994; Berx et al., 1995; Savagner et al., 1997). E-cadherin mediated adhesion is regulated, in part, by β-catenin, which couples it to the actin cytoskeleton (Nathke et al., 1994). In addition to determining cell fate, cadherin-catenin complexes stabilize mammary gland architecture, dictate epithelial cell polarity, promote survival and inhibit motility. E-cadherin is the prototypic type I cadherin. Type I cadherins mediate homophilic interactions by forming adhesive bonds between one or several immunoglobulin (Ig) domains in their extracellular region and connecting to actin microfilaments indirectly via α-catenin and β-catenin in the cytoplasm. The *de novo* production of E-cadherin in normal and transformed mesenchymal cells can induce the formation of stable cell–cell contacts and the development of AJ. In epithelial cells,

early contacts are also mediated by E-cadherin molecules that cluster into small junctional complexes, which then expand to establish stable AJ and promote the formation of desmosomes. E-cadherin is required for the maintenance of stable junctions: anti-E-cadherin antibodies can disrupt these contacts and induce a mesenchymal phenotype (Thiery, 2002). Inactivation of E-cadherin is associated with invasive breast cancer, premature apoptosis and involution in mice. Consequently, expression of a stabilized β-catenin in the mouse mammary gland causes precocious alveolar development, hyperplasia, delayed involution and mammary tumors (Imbert et al., 2001). Wnt/β-catenin singaling pathway is also involved in EMT (Fig. 16). Non-sequestered, free α-catenin and β-catenin are rapidly phosphorylated by glycogen synthase kinase 3 (GSK-3) in the adenomatous polyposis coli (APC)–axin–GSK-3 complex and are subsequently degraded by the ubiquitin–proteasome pathway. If the tumour suppressor APC is non-functional, as in many colon cancer cells, or if GSK-3 activity is blocked by the activated WNT-signalling pathway, β-catenin accumulates at high levels in the cytoplasm. Subsequently, it translocates to the nucleus, where it binds to members of the TCF/LEF1 family of transcription factors and modulates the expression of their target genes, including c-MYC, cyclin D1, fibronectin, MMP7, ID2, CD44, NrCAM, axin-2, TCF1 and others, which are mostly genes implicated in cell proliferation and tumour progression. The loss of E-cadherin function would subsequently lead to the activation of the WNT signalling pathway. In various cellular systems, sequestration of β-catenin by E-cadherin can compete with the β-catenin/TCF-mediated transcriptional activity of the canonical WNT signalling pathway (Cavallaro &

Christofori, 2004).

EMT can be induced in several epithelial cell lines by growth factors activating tyrosine kinase surface receptors, which include hepatocyte growth factor, fibroblast growth factors, epithelial growth factor (EGF) family members, and insulin-like growth factors 1 and 2 (Thiery, 2002). Although the transient activation of Src, phosphoinositide 3-kinase (PI3K) and Rac has an effect on particular aspects of EMT in some cell lines, the Ras/mitogen-activated protein kinase (MAPK) pathway has a crucial role in inducing EMT in the most cases. TGF-β is a potent inducer of EMT in co-operation with the Ras pathway (Janda et al., 2002). MAPK and PI3K have also been implicated in TGF-β signaling and more direct signalling, both through the conventional Smad pathways and other as yet unknown pathways. The classical TGF-β signaling pathway leading to inhibition of cell growth or even the induction of apoptosis is abrogated by activated Ras in part through Raf/MAPK or PI3K activation. In addition, it has been clearly shown in the skin carcinogenesis model that nuclear accumulation of Smad2 by oncogenic Ras is required for progression towards the spindle cell carcinoma stage (Vincent-Salomon and Thiery, 2003). Snail or Slug, a closely related member of the Snail superfamily, control gastrulation and neural-crest EMT. They control EMT by repressing E-cadherin expression in breast cancer (Cano et al., 2000; Hajra et al., 2002). Similarly, Twist, a bHLH transcription factor, represses E-cadherin to cause EMT (Thiery & Morgan). EMT can also be induced by ECM components, including collagens and laminin 5 (Grassi et al., 1999). Stromelysin 1 (matrix metalloprotease [MMP]-3) also induces EMT *in vitro* and tumorigenesis *in vivo* (Sternlicht et al., 1999).

Figure 17 shows the possible signaling pathway involving E-cadherin and other factors.

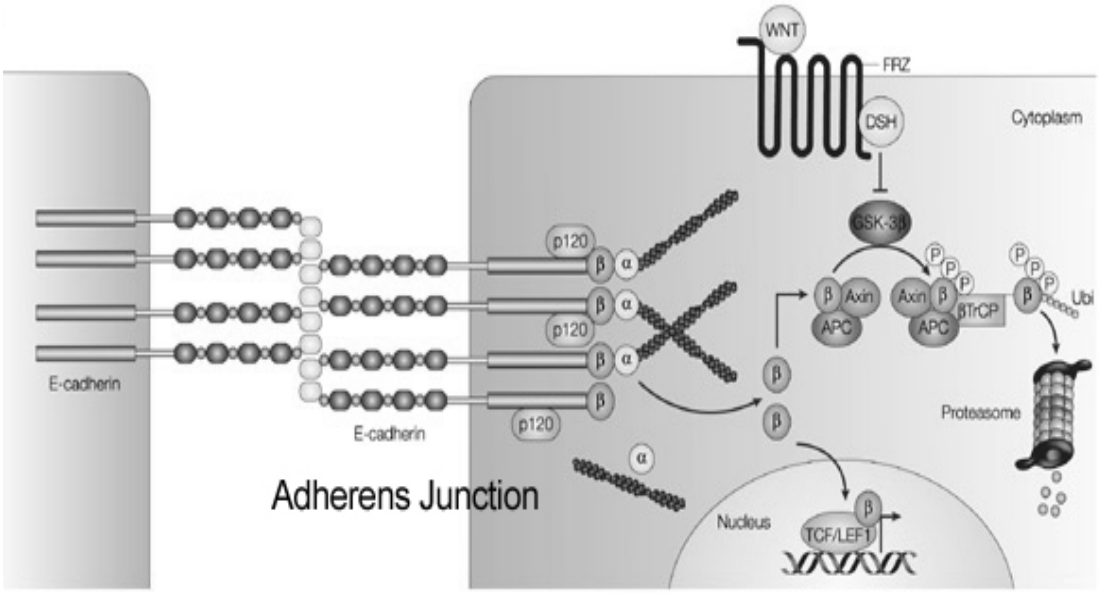


Figure 16. Potential signalling pathways affected by loss of E-cadherin function. After loss of epithelial (E)-cadherin function and disassembly of the cytoplasmic cell-adhesion complex (CCC), catenins are released and accumulate in the cytoplasm. β-Catenin (β) is then sequestered by the adenomatous polyposis coli (APC)–axin–glycogen synthase kinase 3 (GSK-3) complex and phosphorylated by GSK-3. Phosphorylated β-catenin is specifically bound by TrCP, a subunit of the E3 ubiquitin-ligase complex. However, on activation of the WNT signalling pathway, GSK-3 is repressed and β-catenin is no longer phosphorylated. It translocates to the nucleus where, together with the TCF/LEF1 transcription factors, it modulates the expression of several target genes that are known to be involved in cell proliferation and tumour progression. Adapted from Cavallaro & Christofori, 2004

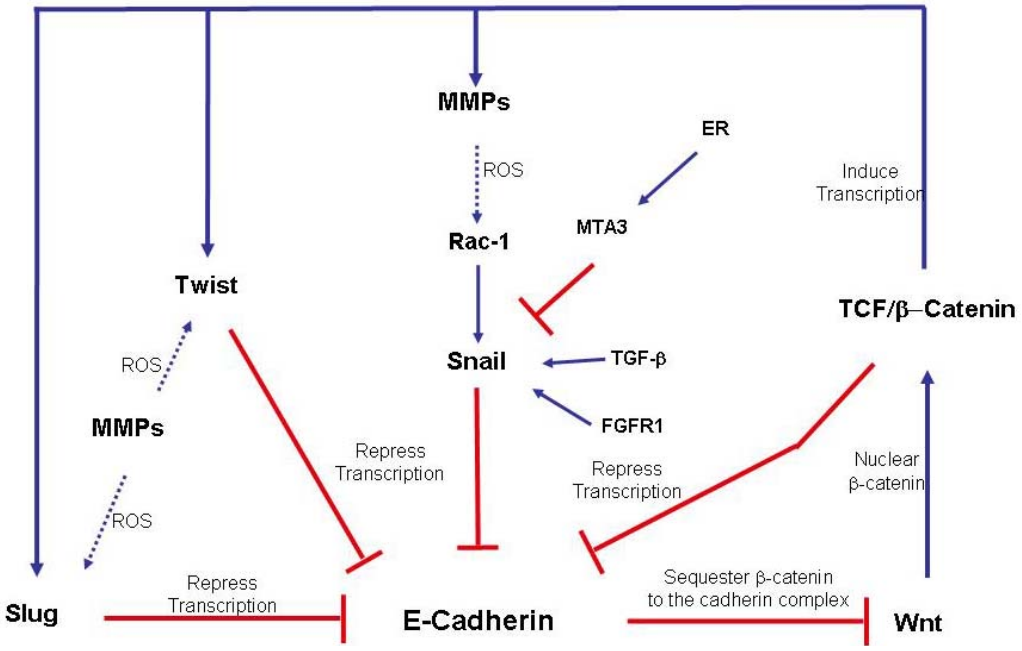


Figure 17. Signaling pathways involving E-cadherin.

1.9. Matrix metalloproteases

The mammary gland is composed of two compartments: epithelium and stroma which are separated by the basement membrane. Therefore, tissue remodeling requires degradation of the basement membrane and surrounding extracellular matrix. This occurs during normal mammary gland development, and under pathological conditions such as tumor progression and metastasis (Andarawewa et al., 2003; Alexander et al., 2001; Ha et al., 2001; Lee et al., 2000). To establish metastatic growth, cancer cells must pass through the basement membranes at least three times (Fig. 18). Breast cancer cells initially cross these membranes when an *in situ* carcinoma becomes an invasive lesion. Later, malignant cells transverse these structures during both entry into and exit from the blood stream (Duffy et al., 2000). Figure 18 depicts the process in detail.

Normal epithelia lined by a basement membrane can proliferate locally to give rise to an adenoma. Further transformation by epigenetic changes and genetic alterations leads to a carcinoma *in situ*, still outlined by an intact basement membrane. Further alterations can induce local dissemination of carcinoma cells, possibly through an EMT, and the basement membrane becomes fragmented. The cells can intravasate into lymph or blood vessels, allowing their passive transport to distant organs. At secondary sites, solitary carcinoma cells can extravasate and either remain solitary or they can form a new carcinoma through a mesenchymal-epithelial transition (Thiery, 2002).

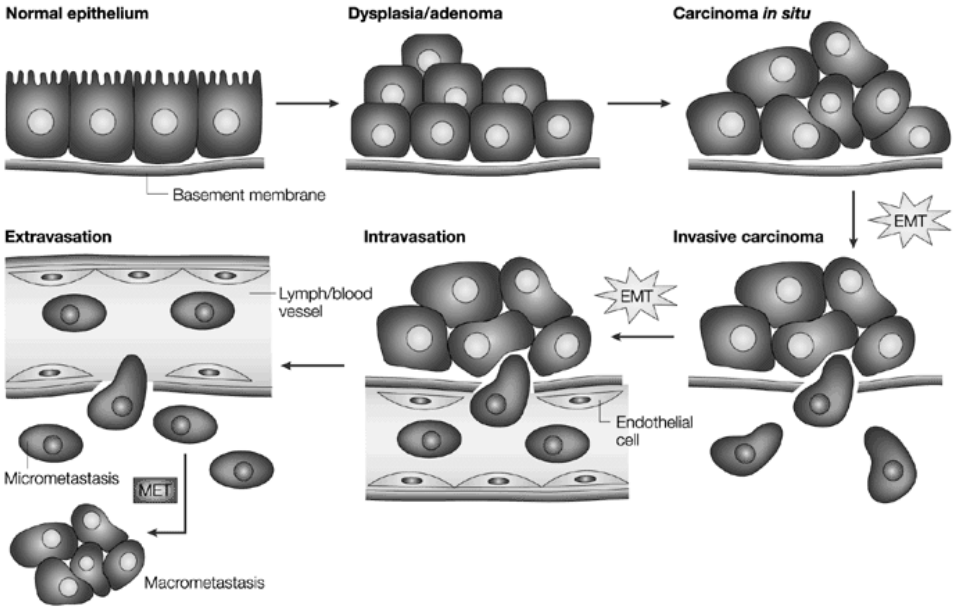


Figure 18. Process of EMT and MET in the emergence and progression of carcinoma. Adapted from Thiery, 2002.

MMPs play a critical role in this process by actively degrading ECM components. The MMPs are endopeptidases that can cleave virtually any component of the ECM.

There are eight distinct structural classes of MMPs: five are secreted and three are membrane-type MMPs (MT-MMPs) (Fig. 19).

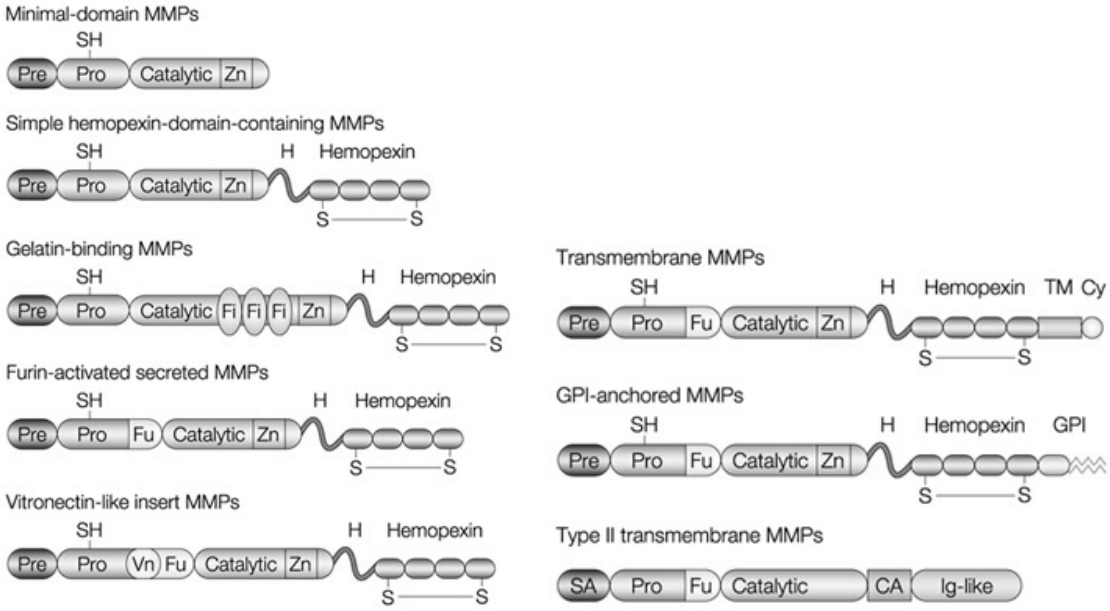


Figure 19. The protein structure of the MMPs. Secreted MMPs: The minimal-domain MMPs contain an amino-terminal signal sequence (Pre) that directs them to the endoplasmic reticulum, a propeptide (Pro) with a zinc-interacting thiol (SH) group that maintains them as inactive zymogens and a catalytic domain with a zinc-binding site (Zn). In addition to the domains that are found in the minimal domain MMPs, the simple hemopexin-domain-containing MMPs have a hemopexin-like domain - that is connected to the catalytic domain by a hinge (H) - which mediates interactions with tissue inhibitors of metalloproteinases, cell-surface molecules and proteolytic substrates. The first and the last of the four repeats in the hemopexin-like domain are linked by a disulphide bond (S–S). The gelatin-binding MMPs contain inserts that resemble collagen-binding type II repeats of fibronectin (Fi). The furin-activated secreted MMPs contain a recognition motif for intracellular furin-like serine proteinases (Fu) between their propeptide and catalytic domains that allows intracellular activation by these proteinases. This motif is also found in the vitronectin-like insert (Vn) MMPs and the membrane-type MMPs (MT-MMPs). MT-MMPs: MT-MMPs include transmembrane MMPs that have a carboxy-terminal, single-span transmembrane domain (TM) and a very short cytoplasmic domain (Cy), and the glycosylphosphatidylinositol (GPI)-anchored MMPs. MMP-23 represents a third type of membrane-linked MMP. It has an N-terminal signal anchor (SA) that targets it to the cell membrane, and so is a type II transmembrane MMP. MMP-23 is also characterized by its unique cysteine array (CA) and immunoglobulin (Ig)-like domains. Adapted from Egeblad and Werb, 2002.

Secreted MMPs, the minimal-domain MMPs, contain an amino-terminal signal sequence that directs them to the endoplasmic reticulum, a propeptide with a zinc-interacting thiol group that maintains them as inactive zymogens and a catalytic domain with a zinc-binding site. In addition to the domains that are found in the minimal domain MMPs, the simple hemopexin-domain-containing MMPs have a hemopexin-like domain - that is connected to the catalytic domain by a hinge - which mediates interactions with tissue inhibitors of metalloproteinases, cell-surface molecules and proteolytic substrates. The first and the last of the four repeats in the hemopexin-like domain are linked by a disulphide bond. The gelatin-binding MMPs contain inserts that resemble collagen-binding type II repeats of fibronectin. The furin-activated secreted MMPs contain a recognition motif for intracellular furin-like serine proteinases between their propeptide and catalytic domains that allows intracellular activation by these proteinases. This motif is also found in the vitronectin-like insert MMPs and the membrane-type MMPs (MT-MMPs). MT-MMPs include transmembrane MMPs that have a carboxy-terminal, single-span transmembrane domain and a very short cytoplasmic domain, and the glycosylphosphatidyl-inositol-anchored MMPs. MMP-23 represents a third type of membrane-linked MMP. It has an N-terminal signal anchor that targets it to the cell membrane, and so is a type II transmembrane MMP. MMP-23 is also characterized by its unique cysteine array and immunoglobulin-like domains (Egeblad and Werb, 2002).

1.9.1. MMPs are involved in breast cancer initiation and growth

Batimastat, a synthetic inhibitor of MMPs, inhibited tumor formation and metastasis

in a nude mouse injected with MDA-MB-435 breast cancer cells (Sledge et al., 1995). Overexpression of MMP3 in transgenic mice gave rise to preneoplastic and malignant mammary gland tumors (Wiesen and Werb, 1996). Transgenic mice expressing matrilysin under the control of mouse mammary tumour virus (MMTV-long terminal repeat promoter/enhancer) developed premalignant hyperplastic nodules. Mating MMTV-matrilysin mice with MMTV-*neu* transgenic mice resulted in offspring that developed primary mammary tumors about 13 weeks earlier than MMTV-*neu* controls **(**Rudolph-Owen et al., 1998**).**

Cancer initiation or to tumor cell growth by MMPs may be involved with increased angiogenesis, activation of stimulating growth factors or their receptors, and inactivation of inhibiting growth factors. Angiogenesis, the recruitment of new blood vessels, is an essential component of the metastatic pathway. These vessels provide the principal route by which tumor cells exit the primary tumor site and enter the circulation (Zetter, 1998). The process begins with local degradation of the basement membranes that surround capillaries, followed by invasion of the surrounding stroma by the underlying endothelial cells in the direction of the angiogenic signal. Endothelial cell migration is accomplished by cell growth at the leading edge of the migrating column. The endothelial cells then organize themselves into three-dimensional structures to form new capillary tubes (Duffy et al., 2000). MMPs degrade the extracellular matrix and thereby allowing endothelial cell invasion; and liberate factors that promote or maintain the angiogenic phenotype. For example, MMP2 enhances mammary epithelial growth by the degradation of the ECM protein laminin-5. Both MMP-1 and MMP-3 release basic

FGF to breakdown endothelial-derived perlecan (Stetler-Stevenenson, 1999).

Cancer initiation or to tumor cell growth by MMPs is involved with increased activation of stimulating growth factors or their receptors. MMP inhibitors were shown to reduce cell proliferation in direct proportion to their effect on transforming growth factor-α release (Dong et al., 1999). MMP-1 was shown to specifically degrade insulin-like growth factor (IGF) binding proteins 2, 3 and 5, fibroblast growth factor (FGF) binding protein and transforming growth factor (TGF)-β binding protein and release IGF, FGF and TGF-β (Vu & Web, 2000). Matrix metalloproteinase 2 releases active soluble ectodomain of fibroblast growth factor receptor 1. Because the hydrolyzed ectodomain retains its ability to bind FGF, it has the potential to modulate the mitogenic and angiogenic activities of FGF (Levi et al., 1996). IGFs are released when IGF-BPs are cleaved by MMP9 (Manes et al, 1999).

1.9.2. MMPs are also strongly related with invasion and metastasis

MMPs are strongly associated with cancer-cell invasion and metastasis. Reduction of MT1-MMP from MDA-MB-231 breast cancer cell resulted in significant reduction of *in vitro* invasiveness and loss of response to an invasion stimulus, HGF (Jiang et al., 2006). Batimistat reduced spontaneous metastasis of a highly malignant rat mammary cancer (Eccles et al., 1996). When MMP3 is overexpressed in transgenic mice the mammary gland undergoes precocious involution and is predisposed to forming a reactive stroma resembling that of a wound site or a tumor (Wiesen et al. 1996). E-cadherin is cleaved by MMP3 or MMP7 and the released fragment of E-cadherin promotes tumour-cell invasion (Noe et al., 2001)

The MMPs are synthesized as inactive zymogens (pro-MMPs). They are kept inactive by an interaction between a cysteine-sulphydryl group in the propeptide domain and the zinc ion bound to the catalytic domain: activation requires proteolytic removal of the propeptide prodomain. Most of the MMPs are activated outside the cell by other activated MMPs or serine proteinases (Egeblad & Werb, 2002). The ratio of active to latent form of MMP-2 increased with tumour progression in invasive breast cancers (Davies et al., 1997)

1.10. Research objectives

Single minded-2, a member of the bHLH-PAS family of proteins, plays a key role in the developing central nervous system (CNS) by regulating a number of genes that guide neuron growth. SIM2 was initially identified by positional cloning on chromosome 21 and contributes to many of the physiological abnormalities associated with DS as a result of trisomy-21. Sim2 is also expressed in a number of branching organs including the lung, kidney and mammary gland; however, its function in these tissues has not been determined. The pattern of occurrence of malignant disorders in people with DS is unique with a very high risk of leukemia in childhood extending into young adulthood and a decrease in the risk of solid tumors at all ages. Women experience a lower risk than men, mainly due to the absence of breast cancer and this may be a result of a gene dosage effect of tumor suppressor genes on chromosome 21. Although SIM2 is associated with critical biological functions, very little is known regarding its role in development outside of the CNS, and even less is known about its role in cancer

susceptibility.

1.10.1. Objective 1

In *Drosophila*, Single minded acts as the master regulator of central nervous system midline development by controlling expression of many genes required for differentiation. *Drosophila* sim forms a heterodimer with ARNT and bind CME in the regulatory regions of target genes to activate expression of proteins required for proper central midline establishment. In mammalian, SIM interact with ARNT, but differ from *Drosophila* sim, the AHR and HIF by functioning as transcriptional repressors. In mouse, Sim1 and Sim2 are expressed in a variety of tissues including brain, kidney, lung and skeletal muscle where they play important developmental roles. After hSIM2 was first identified, the mouse homolog was subsequently mapped to the syntenic region on chromosome 16. In addition, SIM2s has also been identified. This splice variant lacks a portion of the region implicated in mediating the repressive effects of SIM2. However, functional differences between these two isoforms have not been reported. Thus, differential transcriptional regulation of Sim2s will be intensively studied by isolating mouse Sim2s homolog and characterizing its expression profile and transcriptional properties.

1.10.2. Objective 2

In *Drosophila*, the central midline of the developing CNS serves a number of critical functions by providing signals to guide the path of growing axons and regulate epidermis development. dsim serves as a master regulator of *Drosophila* midline differentiation by controlling all steps of midline development. *Drosophila* CNS midline cell regulation

of axon branching and development is dependent upon dsim regulation of genes involved in cell fate determination and stromal-epithelial interactions. Transgenic mice over-expressing Sim2 display many of the mental characteristics seen in DS individuals including learning impairment and reduced fear response, implying that Sim2 plays a key role in many of the physiological phenotypes seen in DS. Thus, the working hypothesis is that Sim2s plays a crucial role in determining and/or maintaining mammary epithelial cell fate and function. We will characterize the overall role of Sim2s, using C57Bl/6J mice and Sim2 knock-out mice.

1.10.3. Objective 3

Individuals with DS are more than 50% less likely to develop solid tumors including cancers of the lung, colon, skin, head and neck. Especially, women with DS are 10-25 times less likely to develop breast cancer in comparison to age-matched normal populations. Women with DS rarely become pregnant, increasing their risk for breast cancer; however, they experience earlier menopause, leading to a decrease in risk. Thus, hormonal factors alone are unlikely to explain the significantly decreased rate of breast cancer and protective factors are more likely to be genetically linked. Based on the facts, we hypothesized that SIM2 is a breast cancer suppressor gene. This will be tested by reestablishing SIM2s into MDA-MB-435, a highly invasive breast cancer cell line. This study will also provide valuable insights into possible mechanisms of SIM2s tumor suppressor activity.

II. MATERIALS AND METHODS*

2.1. Animals

C57Bl/6J, *Sim2-/-*, and athymic nude mice were housed under standard 12 hr lights on/off conditions in a temperature and humidity controlled facility with food and water provided ad libitum. All animal housing and treatments were approved by and conformed to the Animal Use Protocols at Texas A&M University.

2.2. Cell lines

The cell lines used in these studies were obtained from American Type Culture Collection (ATCC). The 16N normal breast epithelial cells and 21T breast cancer cells were a gift from Dr. Heide Ford at the University of Colorado Health Science Center and were grown as previously described (Coletta et al., 2004). HEK-293T and HepG2 cells were cultured in DMEM (Life Technologies, Gaithersburg, MD) containing 10% FBS and 1% Penicillin-streptomycin. HEK-293t Ampho-Phoenix packaging cells were obtained from ATCC with permission from Dr. Gary Nolan at Stanford University and maintained as recommended. All cells were grown in 5% CO_2 at 37°C.

2.3. RNA isolation and reverse transcription

Total RNA was initially isolated from mouse tissues using Trizol reagent (Invitrogen

* Parts of this section are reprinted from Metz et al., 2006, J Biol Chem, 281(16), 10839-10848, Copyright © 2006 by the American Society for Biochemistry and Molecular Biology and Kwak et al., 2006, Carcinogenesis, in progress, Copyright © 2006 by Oxford University Press.

Carlsbad, CA) and was further purified using RNeasy kits (Qiagen, Valencia, CA) with on-column DNase digestion (Qiagen). One microgram of total RNA was reverse transcribed using oligo dT and Superscript II reverse transcriptase (Invitrogen).

2.4. 3' Rapid Amplification of cDNA Ends (RACE)

Total RNA (2 μg) was reverse transcribed using an oligo-dT-based adapter primer (AP-dT17) (5'-GAT CAG GAC GTT CGT TTG AGT TTT TTT TTT TTT TTT T-3'). An initial round of PCR was performed using 1ul of a 1:100 dilution of the RT reactions with the adapter primer (AP) (5'-GAT CAG GAC GTT CGT TTG AGT-3') and a Sim2-specific primer located in exon 9 (P1; 5'-AGT CCC AGG AGT CCT GGA-3'). Subsequent rounds of PCR were performed on 1ul of a 1:100 dilution of the previous reaction using the adapter primer and nested Sim2 primers (P2; 5'-AAA TCA GCT AAA CCC AAA AAC ACA AA-3', P3; 5'-AG AAC CAA CCC ATA TCC CC-3' P4; 5'-CTT CTC CTG TGA ATG CTG C-3'). The PCR products were cloned into pCR II-TOPO (Invitrogen), and screened for inserts by restriction enzyme analyses. Clones positive for the insert were sequenced (DNA Technologies Lab, Texas A&M University) and compared to the reported mouse Sim2 and chromosome 16 sequences. The sequence of the Sim2s 3-RACE product was submitted to GenBank and was assigned the accession number AY963781.

2.5. RT and real-time PCR

Total RNA was isolated using RNeasy Mini kits (Qiagen, Valencia, CA) along with

the RNase-free DNase Set (Qiagen, Valencia, CA) to remove genomic DNA. One µg total RNA was reverse transcribed into cDNA using oligo (dT) and Superscript II reverse transcriptase (Invitrogen, Carlsbad, CA). PCR reactions contained 1µl cDNA template with 1µM each primer, 0.4mM dNTPs and 1 unit Taq DNA Polymerase (Promega, Madison, WI) per 25µl reaction. Primers used were: Sim1 (FP; 5'-GGT CAG CCA GTG AAT CTG GT-3', RP; 5'- TGG TCT CCT GCT GTC TGA TG -3'), Sim2s (FP; 5'- AAA TCA GCT AAA CCC AAA AAC ACA AAG ATG-3', RP; 5'- TCC CTG TGA GGT CAC GAA TA -3'), Sim2 (FP; 5'- AAA TCA GCT AAA CCC AAA AAC ACA AAG ATG -3', RP; 5'- TAG TGG CCG CAG CTC GGG AA -3', Total (pan) SIM2 (FP: 5'-TGA AAT GTG TCT TGG CGA AAA-3', RP: 5'-GCG TAG CGG AGG TGG AAC-3'), SIM2s (FP: 5'-GCT GAG AAC AAA CCC TTA CC-3', RP: 5'-GAA GCA GAA AGA GGG CAA GTT-3'), SIM2L (FP: 5'-CAG GTT CGG CGA GGA CAC-3', RP: 5'-ACC TCC CGT TGG TGA TGA T-3') and GAPDH (FP: 5'-AAT CCC ATC ACC ATC TTC CA-3', RP: 5'-GTC ATC ATA TTT GGC AGG TT-3'). Real time PCR was performed using SYBR Green Master Mix (Applied Biosystems, Foster City, CA). Primers for analyzing total SIM2 by real time PCR were FP: 5'- AGA CAA AGC TGA GAA CAA ACC CTT A -3' and RP: 5'- CCG CAT TCC AGT TTG TCC AT -3'. MMP3 was analyzed using the primers FP: 5'- TTC CTG ATG TTG GTC ACT TCA GA -3' and RP: 5'- TCC TGT ATG TAA GGT GGG TTT TCC -3'. TBP was used as a normalizing gene and was assayed using the primers FP: 5'- TGC ACA GGA GCC AAG AGT GAA -3' and RP: 5'- CAC ATC ACA GCT CCC CAC CA -3'. Product specificity was determined by dissociation curve analyses after each run and

product identities were confirmed by sequencing. Real time PCR data were analyzed by the $\Delta\Delta C_T$ method (Hettinger et al., 2001). The relative positions of the SIM2-specific primers used are indicated in Fig. 20A.

2.6. Plasmid construction

The full-length mouse Sim2s plasmid (pcDNA-mSim2s) was made by amplifying the 5' most 1548-bp of Sim2 from pmSim2-GAL4-HA (kindly provided by Dr. Jerry Pelletier, McGill University, Montreal, Quebec, Canada) using HiFi Taq DNA polymerase (Roche, Indianapolis, IN) and the primers 5'-ATG AAG GAG AAG TCC AAA AAT GC-3' and 5'-AGC ATT CAC AGG AGA AGG CTC AGA A-3'. This fragment was cloned into pCR II-TOPO, and the 5' Spe I fragment was cut out and ligated to Spe I/Xba I cut pCRmSim2s-8 (original Sim2s 3'RACE clone) to create pCRFLmSim2s. After the sequence was confirmed, the entire insert was removed by EcoR I digestion, and cloned into EcoR I-cut pcDNA3. pmSim1-HA and pmSim2-HA were a kind gift of Dr. Jerry Pelletier. Full-length Sim2L was removed from pmSim2-GAL4-HA by EcoR I digestion and cloned into Eco RI-cut vectors. Full-length ARNT and ARNT2 were amplified from mouse kidney and cloned into pCR II-TOPO. Once the sequences were verified, the inserts were removed and cloned into pGBK-T7 and pGADT7-Rec.

2.7. Reporter plasmids

pDRE-TATA-luc, pGL2-TATA and pß-gal were kindly provided by Dr. Stephen

Safe at Texas A&M University (College Station, TX, USA) (Morrow et al., 2004). pHRE-TATA-luc was constructed by cloning annealed oligonucleotides corresponding to the top and bottom strands of the hypoxia-responsive region of the EPO promoter into pGL2-TATA (Chun et al., 2001). pCME-luc was a kind gift of Dr. Jerry Pelletier. pLNCX2-SIM2s was made by subcloning SIM2s cDNA into the EcoRI site of the pLNCX2 retroviral expression vector (Clontech). pCDNA3-SIM2s was made by cloning the entire SIM2s insert into the EcoRI site of pcDNA3.1. The insert from the largest human MMP3-luc construct (-2264 to +37) was amplified from human genomic DNA using the primers 5'-CCT GTT TGA CAT TTG CTA TG-3' and 5'-TTG TCT CTA TGC CTT GCT G-3' and cloned into pCR-2.1 TOPO (Invitrogen). The entire insert was removed and cloned into pGL3-Basic. Deletions were made by removing portions of -2264 MMP3-luc by restriction enzyme digest (Spe I for -1575 to +37 and EcoRI for -275 to +37) and re-ligating the plasmids. Oligonucleotides corresponding to three adjacent AP-1 consensus sites (top strand 5'-CGC TTG ATG AGT CAG CCG GAA CGC TTG ATG AGT CAG CCG GAA CGC TTG ATG AGT CAG CCG GAA -3') were synthesized, annealed and ligated into pGL2-TATA, which is described above.

2.8. Co-immunoprecipitation

Full-length ARNT, ARNT2, Sim2 and Sim2s were made from pcDNA3-based expression plasmids using an *in vitro* translation kit (Promega). Sim2 and Sim2s were made in the presence of [35S] Methionine (Amersham). Five µl of ARNT or ARNT2 product were mixed with 15µl radiolabeled Sim2 or Sim2s and allowed to incubate at

RT for 2 hr. Five µg anti-ARNT (Upstate), anti-ARNT2 (Santa Cruz Biotechnology, Inc.), or rabbit IgG (Upstate) were added, and volumes increased to 150µl with water and 2× Co-IP buffer to make a 1Xsolution [10mM Tris-HCl (pH 7.4), 150mM NaCl, 1mM EDTA, 1mM EGTA, 1% Triton X-100 and 0.5% Igepal]. Following 2 hr incubation at RT, 20µl of agarose A bead slurry was added and samples were incubated at RT for 1.5 hr with gently mixing. Beads were recovered at 5000 rpm for 1 min, and washed 3 times in 500µl 1× Co-IP buffer. Final pellets were resuspended in loading buffer, boiled and separated on polyacrylamide gels. Gels were vacuum-dried and exposed to film for one week.

2.9. Transient transfection

Sim2 and Sim2s expression was confirmed by Western analyses before large-scale experiments were conducted. For each transfection, cells were seeded at 4×10^4 cells per well in 24-well plates the day before transfection. The following morning, 200 ng appropriate reporter plasmid was cotransfected with 100 ng internal control (pß-Gal) and various amounts of test plasmids using Lipofecamine and Plus reagent (Invitrogen). Twenty-four hrs later, cells were incubated under hypoxic (1% O2) or normoxic (21% O2) conditions (pHRE-TATA-luc transfected HEK-293T cells) or in the presence of vehicle (DMSO) or 10 nM 2,3,7,8-tetrachlorodibenzo-pdioxin (TCDD) (pDRE-TATA-luc transfected HepG2 cells) for 40 hr. For MMP3 reporter assays, MDA-MB-435 cells were added to 12-well plates (1×10^5 cells per well) 24 hr before transfection. Plasmids (0.45 µg hMMP3-luc, 0.15 µg pcDNA3 or pcDNA3-hSim2sHA and 0.2µg ß-Gal

expression vector) were transiently transfected using a combination of Lipofectamine (2µl) and Lipofectamine Plus (5µl) reagents (Invitrogen). Following a 24 hr recovery period, fresh media containing DMSO or PMA (50 ng/ml) was added. Cells were harvested, and the luciferase and ß-galactosidease activities of the cell lysates determined by dual luciferase assay using Luciferin (Molecular probes) and Galacto-Light (Tropix, Applied Biosystems). Luciferase activities were normalized to the internal control values and are represented as the means ± S.E. for three wells per condition. Significant differences were determined using Student's t test.

2.10. Retroviral transduction

For retroviral transduction, SIM2s cloned into the pLNCX2 vector (Clontech), was transfected into Phoenix packaging cells for 6-8 hrs with 25 µl of lipofectamine and Plus reagents (Invitrogen). The medium was replaced 24 hrs later and virus-containing supernatants were collected after 48 hrs. Supernatants were filtered through a 0.45 µm filter syringe and added to MDA-MB-435 cells in 6 well plates in the presence of 8 µg/ml polybrene (Sigma). The plates were then centrifuged at 1000 × g at room temperature for 45 min. Following a 24 hr incubation, the medium was replaced and stable transduced cells were selected in the presence of puromycin. The cells were used within the first 10-12 passages and critical experiments were performed within the first few passages. At least three independent infections were utilized and obtained similar results. SIM2s expression was confirmed by RT-PCR, real time RT-PCR and immunofluorescence.

2.11. Chromatin immunoprecipitation assay

Twenty-four hr after transfection, formalin (270 µl per 10 ml medium) was added to cells and allowed to incubate at 37°C for 10 min. Crosslinking was stopped by addition of glycine to a final concentration of 125 mM. Following a series of PBS washes, cells were scraped and recovered by centrifugation. Cells (500,00 cells per ChIP assay) were resuspended in SDS lysis buffer (50 mM Tris-HCl, pH 8.1, 1% SDS, 10 mM EDTA) and sonicated on ice. Chromatin was recovered, and mixed with ChIP dilution buffer (16.7 mM Tris-HCl, pH 8.1, 0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 167 mM NaCl) containing 1× Complete Protease Inhibitor Cocktail (Roche). Chromatin was cleared twice by incubation with agarose bead slurry at 4°C for 30 min. followed by centrifugation. Antibody was added (10µg for Sim2, or C-Jun, 1µg for IgG) and incubated at 4°C overnight with mixing. Agarose bead slurry was added and reactions were incubated at 4°C for 1 hr with mixing. The agarose beads were recovered and washed successively for 5 min at 4°C in low salt (20 mM Tris-HCl, pH 8.1, 150 mM NaCl, 0.1% SDS, 1% Triton X-100, 2 mM EDTA), high salt (20 mM Tris-HCl, pH 8.1, 500 mM NaCl, 0.1% SDS, 1% Triton X-100, 2 mM EDTA). LiCl (10 mM Tris-HCl, pH 8.1, 250 mM LiCl, 1% IGEPAL-CA630, 1% deoxycholic acid, 1 mM EDTA) and TE (10 mM Tris-HCl, pH 8.0, 1 mM EDTA) buffers. Chromatin was eluted at RT by incubation in freshly prepared elution buffer (1% SDS, 0.1 M NaHCO3) for 15 min. Samples were then treated with RNase A in the presence of 200 mM NaCl and crosslinking reversed by incubation at 65°C for 4 hrs. Chromatin was ethanol

precipitated and resuspended in buffer containing 40 mM Tris-HCl, pH 8.1, 10 mM EDTA and 200μg/ml proteinase K. DNA was purified using Qiaquick PCR purification columns (Qiagen, La Jolla, CA) and eluted in 50 μl sterile water. PCR detection of pGL2-CME immunoprecipitation targets was performed in 25 μl reactions containing 2 μl chromatin and the primers pGL2-ChIP-F1 (5'-CCC CCT GAA CCT GAA ACA TA-3') and pGL2-ChIP-R1 (5'-GCC TTA TGC AGT TGC TCT CC-3'). Thermocycling was done with an initial denaturation at 95°C for 5 min followed by 27 cycles of 95°C for 30 sec, 50°C for 30 sec, and 72°C for 30 sec followed by a final elongation step at 72°C for 5 min. The JUN antibody was obtained from Santa Cruz (catalog number sc-44)

2.12. Immunostaining, immunofluorescence and western blot

The following primary antibodies were used: anti-SIM2 (Chemicon, catalog number AB4145) and anti-SIM2s (Santa Cruz, catalog number sc-8715). Secondary antibodies were: FITC anti–rabbit; and Texas red anti–rabbit (Molecular Probes). Human breast and breast tumor tissue arrays were purchased from Chemicon (Temecula, CA). The sections were baked at 55^{o}C for 15 min then dehydrated by sequential washes in xylene and a series of graded ethanol washes. Antigen retrieval was performed for 20 min at 98°C in 0.01 mol/L sodium citrate buffer, pH 6.4, in a microwave oven. For immunostaining, sections were incubated in 0.3% hydrogen peroxide for 30 min to block endogenous peroxidase activity. After 30 min block in 5% milk, the sections were incubated with 1:100 dilution of primary antibody 1 hr at room temperature. The

sections were washed in phosphate-buffered saline containing 0.1 % tween-20 and then incubated with either biotinylated donkey anti-goat (Vector Laboratories, Burlingame, CA) followed by avidin peroxidase using the Vectastain ABC elite kit (Vector Laboratories, Burlingame, CA) or Alexa-conjugated secondary (Molecular Probes). The chromogenic reaction was carried out with 3–3' diaminobenzidine (Sigma). The slides were mounted using Permount (Sigma) or Vectashield containing DAPI (Vector Laboratories) and evaluated under a microscope.

2.13. Zymography

To determine MMP2 and MMP3 activity, conditioned medium from treated cells was concentrated ~20-fold using Centricon 10 spin concentrators (Amicon). Samples were quantitated by Bradford analysis and equal amounts of protein were mixed with Laemmli sample buffer without reducing agents, incubated for 15 min at 37°C, and separated on precast gradient SDS-polyacrylamide slab gels (Invitrogen) containing 1 mg/ml gelatin for MMP2 or casein for MMP3. Following electrophoresis, gels were placed in 2.5% Triton X-100 for 30 min then incubated at 37°C in 50 mM Tris-HCl, pH 7.4, containing 5 mM $CaCl_2$ for 18 hrs. MMP2 or MMP3 activity was visualized by Coomassie blue staining. A non-specific band at 77 Kd was used as a loading control.

2.14. Flow cytometry analysis

Control or SIM2s-transduced MDA-MB-435 cells were collected for cell cycle analysis, washed once with PBS, suspended in 500 μl of cold (–20°C) ethanol for at least

1 hr at 4°C, washed twice with PBS, and treated for 1 hr at room temperature with ribonuclease A (RNase A, 100 μg/ml) and propidium iodide (PI, 50 μg/ml). RNase A and PI were from Sigma-Aldrich. Cell cycle analysis was carried out with an EPICS XL 2 (Beckman-Coulter). PI binding was measured by using the FL2 photomultiplier tube (bandpass 575 nm) with events discriminated on forward and side scatter to exclude debris. The flow rate was set at about 300 cells per second, and at least 2×10^4 cells were analyzed from each sample.

2.15. Soft agar and invasion

Control or SIM2s-transduced MDA-MB-435 cells (5×10^3 cells per plate) in 2 ml of medium (DMEM + 10% FBS) supplemented with 0.35% agarose were layered onto a 1.5-ml base of medium containing 0.5% agarose in 35-mm dishes. Media was changed daily and after 14 days of growth, cells were stained with 0.005% crystal violet, photographed and colonies >200μm were counted. Data presented are the summary of three separate experiments. For cell invasion assays, 1×10^5 cells in 0.5 ml of serum-free DMEM were plated in Boyden chambers containing 8 μm pore membranes coated with Matrigel (Becton Dickinson). The lower chamber contained 10% FBS in DMEM to serve as a chemoattractant. Following a 24 hr incubation, cells that had not left the top chamber were removed from the membrane by scraping. Cells that had migrated to the bottom surface of the membrane were stained with Diff-Quik reagent (American Scientific Products, McGaw Park, IL) and examined under a bright-field microscope. Values are the average number of cells per 5 fields per membrane and analyzed by

Students t test.

2.16. Transplantation

Mammary tissues from 1 day-old pups were dissected out under a microscope and transferred into the number four cleared fat pad of nude mice. Pups for this experiment were obtained by mating *Sim2*+/- male and female mice. During the surgery, the tails of pups were collected and used later for genotyping. The transplanted glands were examined at 8 weeks after the transplantation.

2.17. Staining

For whole mount staining, mammary glands were fixed in 10% neutral-buffered formalin overnight. Tissues were washed three times for 1 hr in acetone, then for 1 hr each in 100% ethanol and 95% ethanol. Glands were then stained for 3 hr in a solution of 95% ethanol, pH 1.25, containing 2.5mM $FeCl_3$ and 0.09% hematoxylin. Whole mounts were then washed twice for 1 hr in tap water, then in increasing concentrations of ethanol for 1 hr each (50, 70, 95 and 100%) and a final 1 hr rinse was performed in xylene. H&E and trichrome staining was performed by the Veterinary Anatomy and Public Health Histology Core Facility at Texas A&M University.

III. RESULTS*

3.1. Differential transcriptional regulation by mSim2s

3.1.1. Cloning of mSim2s

Previous studies have identified a splice variant of human SIM2; however, a similar isoform has not been reported in mice, nor have functional differences between SIM2 and SIM2s been reported. In order to better characterize biochemical and physiological differences between Sim2 and Sim2s in mice, we set out to identify the mouse Sim2s transcript. Total RNA from mouse kidney and liver tissues were used for 3'RACE experiments since Sim2 is expressed at high levels in the kidney, but not in liver. A unique adapter primer (AP-dT17) was used to reverse transcribe total RNA from the tissues. This reaction was then used as a template for PCR using a primer directed to the AP region of the 3' adapter and a Sim2-specific forward primer (P1, Fig. 20A). Subsequent PCR reactions were performed on diluted PCR products using the adapter primer (AP) and nested Sim2-specific primers (P2-P4, Fig. 20A). The resulting PCR products from each reaction were submitted to restriction enzyme analyses for predicted sites to support the identity of the PCR products as Sim2 (Fig. 20B). Potential Sim2 clones with predicted restriction digest products were only obtained in kidney, and were too small to be full length Sim2. In humans, the SIM2s splice variant is missing exon 11, and contains an extension of exon 10 (Chrast et al., 1997). Based on the sequence of

---

* Parts of this section are reprinted from Metz et al., 2006, J Biol Chem, 281(16), 10839-10848, Copyright © 2006 by the American Society for Biochemistry and Molecular Biology and Kwak et al., 2006, Carcinogenesis, in progress, Copyright © 2006 by Oxford University Press.

mouse chromosome 16 near the region of mouse Sim2 exon 10, a single Hae III site is predicted to be present in the Sim2s transcript. This was confirmed in the fourth round 3' RACE PCR product (Fig. 20B, right panel) suggesting that this clone is the mouse homolog of human SIM2s. 3' RACE clones were sequenced and compared to reported mouse Sim2 sequences and chromosome 16.

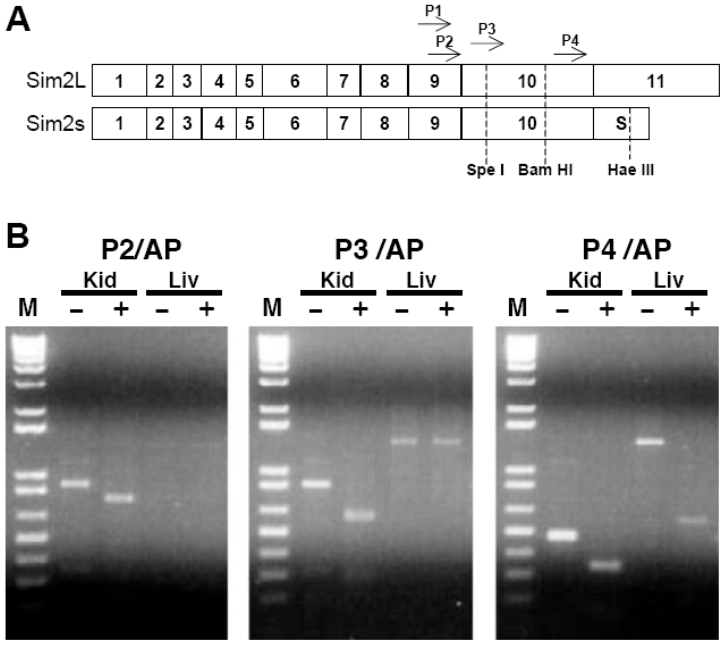
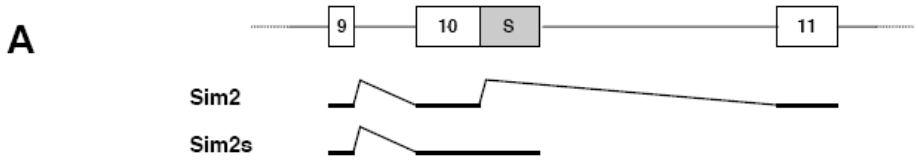


Figure 20. Isolation of mouse Sim2s by 3' rapid amplification of cDNA ends (3' RACE). A. Structure of mouse Sim2 and Sim2s cDNAs. Numbered squares indicate exons. Numbered arrows indicate relative positions of mouse Sim2-specific primers used in 3' RACE PCR reactions. Dotted lines indicate relative positions of restriction sites used to verify identity of 3'RACE products. B. Restriction enzyme analyses of mouse kidney and liver 3'RACE products. Two micrograms of total RNA from mouse kidney or liver was reverse transcribed using an oligo dT-based adapter primer (AP-dT17). An initial round of PCR was performed on 1µl of a 1:100 dilution of the RT reaction using the adapter primer and a mouse Sim2-specific forward primer (P1). The products of this initial PCR reaction were not visible (data not shown). Subsequent PCR reactions were performed on 1µl of a 1:100 dilution of the previous reaction using the adapter primer and nested Sim2-specific primers (P2, P3 and P4 in Fig. 20A). For visualization and confirmation of identity, 10µl of PCR product plus (+) or minus (-) restriction enzyme was analyzed on a 1.2% agarose gel. Mouse Sim2s 3'RACE products are predicted to have Spe I (P2/AP), Bam HI (P3/AP) or Hae III (P4/AP) sites. Sim2-specific products were only detected in kidney RNA samples, and only Sim2s was detected.



B

```
                                  ← Exon 10 |
Sim2s             CTTGTGCCAAACTATGAAGGTAGGTCC-CATTTACAGGAGGGGAAGGGCAGGGGTTCAGC
Chromosome 16     CTTGTGCCAAACTATGAAGGTAGGTCC-CATTTACAGGAGGGGAAGGGCAGGGGTTCAGC
Sim2              CTTGTGCCAAACTATGAAGCGCCCTCCGCAGCCGCGCGCAGGTTCTGCGAGGACCCCGCG 1620

Sim2s             CTCAGTAGCAGT--GGGTGGTGGATGGAAAAAGAATCCTCATGCTTTGGGCAAACTGGCC
Chromosome 16     CTCAGTAGCAGT--GGGTGGTGGATGGAAAAAGAATCCTCATGCTTTGGGCAAACTGGCC
Sim2              CCCCCAAGCTTCCCGAGCTGCGGCCACTACCGCGAGGAGCCGGCGCTGGGC-------CC 1673

Sim2s             TCTTCTGCTTCTAAGCTGGGATTTCTGTGCTTCCTACTGCCCATCATGGTTTCATGGGCG
Chromosome 16     TCTTCTGCTTCTAAGCTGGGATTTCTGTGCTTCCTACTGCCCATCATGGTTTCATGGGCG
Sim2              TGCCAAGGCGCCTCGCCAGGCGTCCCGGGACGCAGCCCGGCTGGCGCTGGCCCGAGCGCC 1733

Sim2s             T---GAGTTACATATAACAAGAACACCGGAGATGGCATAAACAGCAGAGAGAAGCTGAGT
Chromosome 16     T---GAGTTACATATAACAAGAACACCGGAGATGGCATATACAGCAGAGAGAAGCTGAGT
SIM2              CCCCGAGTGCTGCGCGCCGCCCGCCCCCGAGCCGCAGGCGCCGGCGCAGTTGCCCTTCGT 1793

Sim2s             GCC---CAGTCCCCATATAAGTTCTGACTTTATTCGTGACCTCACAGGGAGGCAAAAGTG
Chromosome 16     GCC---CAGTCCCCATATAAGTTCTGGCTTTATTCGTGACCTCACAGGGAGGCAAAAGTG
SimM2             GCTGCTCAACTACCAC-CGCGTGCTGGCGCGCCGCGGGCCTCTGGGGAGCGCTGCGCCCG 1852

Sim2s             ATTTTCCAAACCAGCTCCCGTGTTTGACCACGCTCCAGCGGCATCCTTGAGGTGCCTTTT
Chromosome 16     ATTTTCCAAACCAGCTCCCATGTTTGACCACGCTCCAGCGGCAACCTTGAGGTGCCTTTT
Sim2              GAGCCCCCGA--GGCGGCCG-GCTCGCTGAGGCCCCGACACCCAGGCCCCGTCGCCGCCT 1909

Sim2s             CGCCGTTCCTCAGAGTGAATGGCAGATGTTAGTTCTCTTCCCAGACCGGT-TCTCAAACT
Chromosome 16     CGCCGTTCCTCAGAGTGAATGGCAGATGTTAGTTCTCTTCCCAGACCGGT-TCTCAAACT
Sim2              CTGCG---CCCGGCGCGCCCCGACCGCATTACCTGGGCGCCTCGGTCATCATCACCAACG 1966

Sim2s             ACAGAAAAAAAAAAAAAAA
Chromosome 16     ACAG----------------
Sim2              GCAGGTGA----------- 1974
```

Figure 21. Structure and comparison of mouse and human SIM2s coding sequences. A. Genomic structure of the mouse Sim2 locus surrounding exons 9-11. Exons are indicated by numbered boxes and the "s" exon is represented by a shaded box. B. Comparison of Sim2s with Sim2 and mouse chromosome 16. The 3' boundary of exon 10 is indicated by the arrow and vertical line. Stop codons are bold and underlined. The mouse Sim2 sequence used (accession number U42554) is that of Moffett et al. (1996). C. Comparison of mouse and human SIM2s coding and predicted amino acid sequences. The 3' boundary of exon 10 is indicated by the arrow and vertical line. Stop codons are bold and underlined and represented as asterisks in the amino acid sequences. Regions of amino acid identity are boxed.

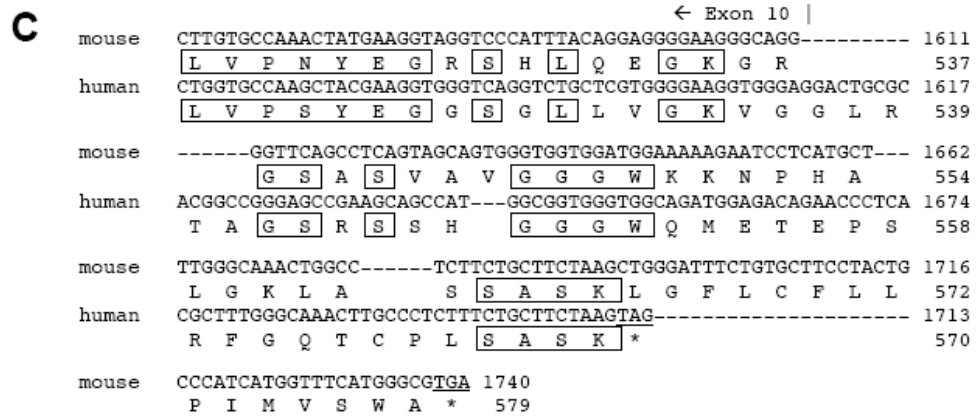


Figure 21. Continued.

Like the human form, mouse Sim2s is missing exon 11 and contains an extension of exon 10 (Fig. 21A). The mRNA (Fig. 21B) is predicted to encode a protein of 579 amino acids that is missing the last 131 amino acids and contains an additional 53 amino acids encoded by the "s" exon (Fig. 21C). Mouse Sim2s still contains the Ser/Thr- and Pro/Ser-rich regions shown previously to harbor repressive activities, but is missing the Pro/Ala-rich repressor region present in full-length Sim2 (Moffett et al., 1997). Mouse and human SIM2s share 86% identity at the RNA level and are 87% homologous at the protein level. Interestingly, the "s" portions of mouse and human SIM2s are not as homologous as the rest of the gene (Fig. 21C). The mRNA and protein sequences of mouse and human SIM2 from start to the end of exon 10 are 88% and 91% homologous, respectively. However, the DNA and protein sequences of the extended exon 10 regions of mouse and human SIM2s are only 45% and 20% homologous, respectively. Despite these differences, a few amino acid residues remain conserved, including 545-GGGW-548 and 561-SASK-564.

3.1.2. Sim2s preferentially interacts with ARNT over ARNT2

Interactions between Sim2 and ARNT proteins were confirmed biochemically by coimmunoprecipitation. 35S-Methionine-labeled Sim2 and Sim2s proteins were mixed with cold ARNT or ARNT2 and complexes precipitated with antibodies directed against ARNT or ARNT2. Anti-ARNT antibody was able to pull down both ARNT/Sim2 and ARNT/Sim2s complexes at relatively comparable levels (Fig. 22, lanes 1 and 2). In contrast, anti-ARNT2 antibody was able to pull down Sim2 slightly more efficiently than Sim2s (Fig. 22, lanes 3 and 4). Although IgG control samples show some ARNT binding (Fig. 22, lane 5), there was very little Sim2 or Sim2s pulled down by nonspecific binding (Fig. 22, lane 6).

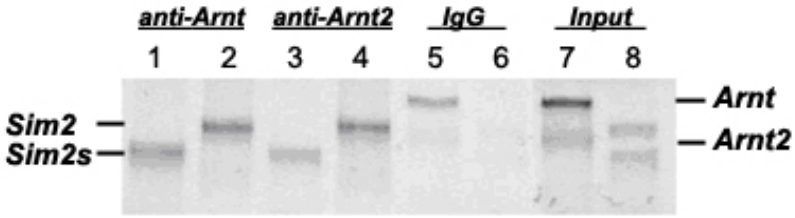


Figure 22. Sim2s preferentially interacts with ARNT over ARNT2. *In vitro* translated ARNT and ARNT2 Sim2s preferentially interacts with ARNT over ARNT2. *In vitro* translated ARNT and ARNT2 were mixed with radiolabeled Sim2 or Sim2s and co-immunoprecipitated using antibodies directed at either ARNT or ARNT2. Anti-ARNT antibody was highly efficient at pulling down Sim2s (lane 1) and Sim2 (lane 2) in the presence of *in vitro* translated ARNT. Similarly, anti-ARNT2 was able to pull down Sim2s (lane 3) and Sim2 (lane 4) in the presence of *in vitro* translated ARNT2, although Sim2s was not pulled down as well as Sim2. IgG pulled down ARNT to a greater extent than ARNT2 (lane 5), but was unable to pull down either Sim2 or Sim2s (lane 6). A diluted mixture of radiolabeled ARNT and ARNT2 (lane 7) as well as Sim2 and Sim2s (lane 8) were ran on the gel to verify translation efficiency.

3.1.3. Transcriptional activity of mouse Sim2 and Sim2s on a HRE- and DRE-controlled reporter genes

Direct targets of mammalian SIM proteins have not been determined; however,

indirect transcriptional activities of SIM1 and SIM2 have been reported. In those studies, it was found that SIM2 can inhibit the actions of other bHLH-PAS proteins through active repression or competition for ARNT binding (Moffett et al., 1997; Moffett & Pelletier, 2000; Woods & Whitelaw, 2002). Over-expression of Sim2 blocked hypoxia-induced expression from HRE-controlled genes by directly binding to the HRE and actively repressing expression of adjacent genes. This effect was dependent upon the Sim2 dimerization domain and carboxyl terminus. Sim2 can also inhibit TCDD-induced expression from XRE-controlled genes, but does so by competing with AHR for ARNT binding. In contrast, Sim2 represses expression of a CME driven reporter, while SIM1 can activate CME-mediated gene expression via the ARNT transactivation domain (Moffett & Pelletier, 2000). To investigate if Sim2s can affect the function of other bHLHPAS protein-mediated pathways, we performed co-transfection assays using various mouse Sim and Arnt expression plasmids with HRE-, XRE- and CME-controlled luciferase reporter constructs.

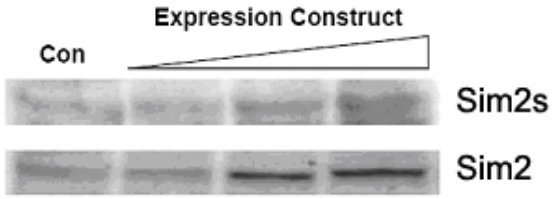


Figure 23. Confirmation of Sim2 expression in HEK293T cells by Western blot. HEK293T cells transfected with pcDNA (Con) or increasing (3, 9 and 15 ug plasmid in 100 mm plate format) amounts of Sim2 or Sim2sexpression plasmids were harvested and analyzed by Western blot to confirm Sim2 gene expression.

Expression of Sim2 and Sim2s in transfected HEK293T cells was confirmed by Western blotting (Fig. 23). HEK293T cells transfected with an HRE-controlled luciferase reporter and the pcDNA empty vector showed a significant increase in luciferase activity following a 40-hr incubation under hypoxic conditions (Fig. 24). Introduction of increasing amounts of mouse Sim2 or Sim2s significantly repressed hypoxia-induced reporter gene expression. Inclusion of 100 ng mouse ARNT expression vector increased the response to hypoxia and slightly overcame both Sim2 and Sim2s-mediated repression. Addition of mouse Arnt2 expression vector did not alter the response to hypoxia and both Sim2 and Sim2s repressed hypoxia-induced expression of the HRE-controlled reporter in the presence of ARNT2. Increasing amounts of Sim2 or Sim2s did not repress in a concentration-dependent manner suggesting that both Sim2 and Sim2s repress hypoxia-induced gene expression through direct interaction with the HRE and not by competing with HIF factors for ARNT binding.

Previous studies have suggested that Sim2 can interfere with AHR-mediated gene expression through competition with AHR for ARNT binding (Moffett et al., 1997). To determine if Sim2s exerts a similar effect on AHR-mediated signaling, HepG2 cells were co-transfected with a DRE-controlled reported gene with various combinations of Sim2 or Sim2s and Arnt expression plasmids. AHR activation was accomplished by addition of 10nM TCDD togrowth medium 24-hr before harvest. A robust TCDD response seen in pcDNA transfected cells was significantly repressed ($p<0.0001$) by inclusion of either Sim2 or Sim2s expression plasmids (Fig. 25).

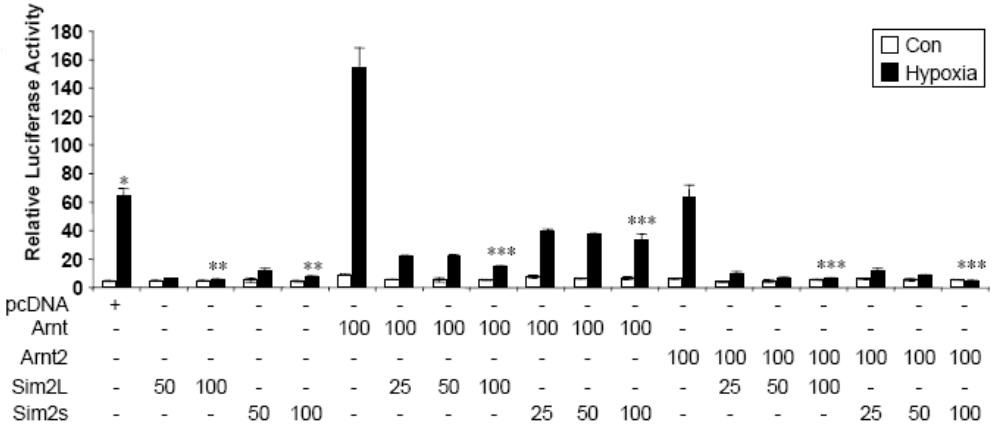


Figure 24. Transcriptional activity of mouse Sim2 and Sim2s on HRE-controlled reporter genes. HEK-293T cells were transfected with ARNT or ARNT2 expression plasmids in the presence of increasing amounts of Sim2 or Sim2s expression plasmids plus an HRE-controlled luciferase reporter vector and a ß-gal expression plasmid for normalization. The amount of each expression vector (in ng) is indicated under the figure. Following 40 hr incubation under normoxic (white bars) or hypoxic (black bars) conditions, cell lysates were collected and analyzed for ß-gal and luciferase activities. Data are expressed as the means ± SE for three plates per condition. *Luciferase expression under hypoxic conditions is higher than control ($p<0.002$). **Luciferase expression under hypoxic conditions is lower in Sim and Sim2s-transfected compared to pcDNA-transfected cells ($p<0.002$). ***Luciferase expression in the presence of ARNT during hypoxia is less in Sim2-transfected cells than in Sim2s-transfected cells ($p<0.01$).

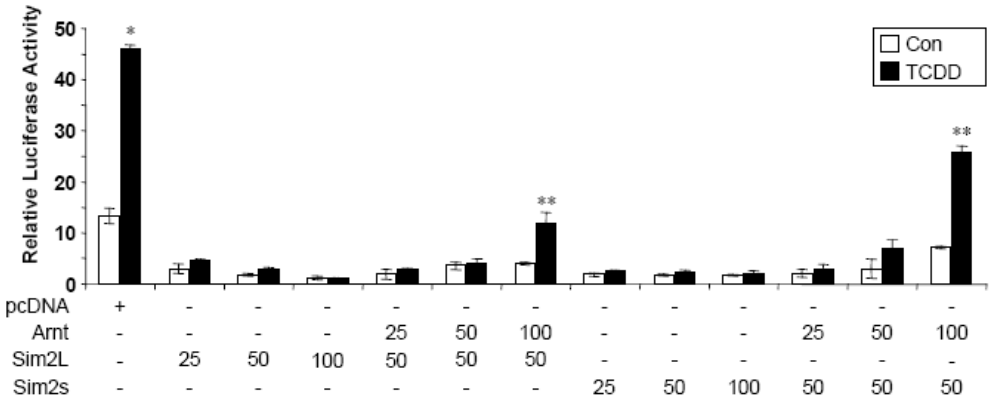


Figure 25. Transcriptional activity of mouse Sim2 and Sim2s on DRE-controlled reporter genes. HepG2 cells were transfected with varying amounts of either ARNT or ARNT2 expression plasmids in the presence of increasing amounts of Sim2 or Sim2s expression plasmid plus a DRE-controlled luciferase reporter vector and a ß-gal expression plasmid for normalization. The amount of each expression vector (in ng) is indicated under the figure. Following 40 hr incubation with vehicle (DMSO, white bars) or 10nM TCDD (black bars) cell lysates were collected and analyzed for ß-gal and luciferase activities. Data are expressed as the means ± SE for three plates per condition. * Higher due to TCDD treatment ($p<0.0001$). **TCDD-induced luciferase expression in the presence of ARNT is lower in Sim2-transfected cells than in Sim2s-transfected cells ($P<0.01$).

The degree of this repression was not significantly different between Sim2 and Sim2s; however, increasing amounts of ARNT attenuated the repression ($p<0.01$) in both Sim2 and Sim2s-transfected cells indicating that the mechanism of this repression involves competition for ARNT.

These results are consistent with previous studies (Moffett et al., 1997) and suggest that Sim2 and Sim2s do not differ in their ability to repress AHR-mediated gene expression, which occurs through direct competition for ARNT binding.

3.1.4. Transcriptional activity of mouse Sim2 and Sim2s on a CME-controlled reporter gene

In *Drosophila*, SIM regulates transcription through the CME, which contains the core sequence 5'-ACGTG-3' also found in an HRE. Previous studies have shown that murine Sim1 and Sim2, in concert with ARNT, can bind and regulate expression of a CME reporter construct (Moffett & Pelletier, 2000). Sim1 activated expression of a CME-controlled gene through the ARNT transactivation domain, whereas full-length Sim2 was repressive. Activation of a CME-controlled gene by Sim2 was accomplished when portions of the C-terminal repression domain were deleted. Since Sim2s is missing part of this repressive region, we anticipated that Sim2s would be less repressive than Sim2 on CME-mediated gene expression. Cotransfection of HEK293 cells with a CME-controlled reporter gene and increasing amounts of ARNT or ARNT2 had no effect on reporter expression (Fig. 26). Contrary to previous reports, we found that introduction of Sim2 with ARNT, but not ARNT2, resulted in a slight increase in luciferase activity (Fig. 26).

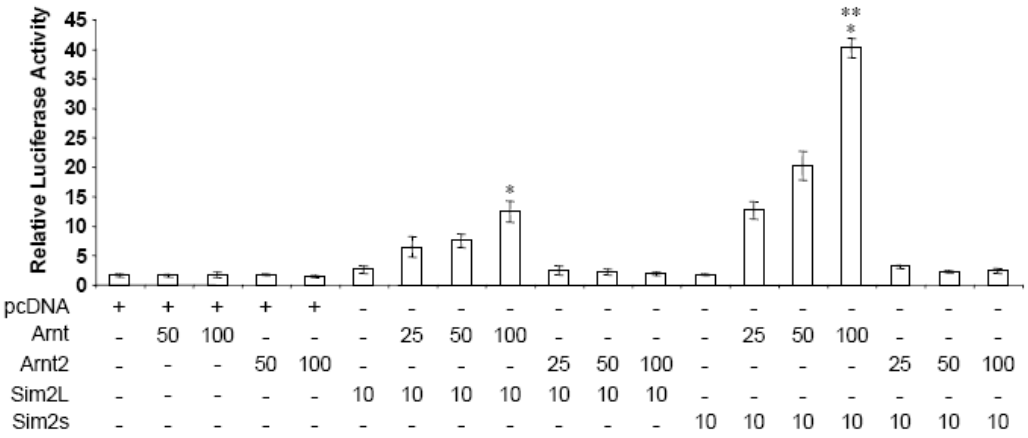


Figure 26. Transcriptional activity of mouse Sim2 and Sim2s on a CME-controlled reporter gene: the effects of increasing ARNT and ARNT2 on mouse Sim2- and Sim2s-mediated expression of a CME-controlled reporter. HEK-293T cells were transfected with increasing amounts of ARNT or ARNT2 expression vector in the presence of absence of Sim2 or Sim2s expression vectors plus a CME-controlled luciferase reporter and ßgal expression plasmid. The amount of each plasmid per transfection (in ng) is indicated under the figure. Luciferase activity was increased in both Sim2 and Sim2s-transfected cells, but only in the presence of ARNT, and not ARNT2. The degree of induction was higher in Sim2s-transfected cells in comparison to Sim2 cells. *Luciferase expression is higher in the presence of Sim and Sim2s in comparison to no ARNT controls ($p<0.005$). **Luciferase expression is higher in Sim2s cells in comparison to Sim2 cells ($p<0.01$).

Co-expression of Sim2s with ARNT, but not ARNT2, resulted in significantly increased reporter gene expression. This effect appears to be ARNT-dependent as luciferase activity increased with increasing amounts of ARNT. This conclusion was further supported by experiments utilizing constant ARNT and increasing amounts of Sim2 expression plasmids (Fig. 27). Although reporter gene expression increased with increasing amounts of Sim2 and Sim2s expression plasmids, this effect was only significant in cells receiving the highest amount of Sim2 and Sim2s expression plasmid.

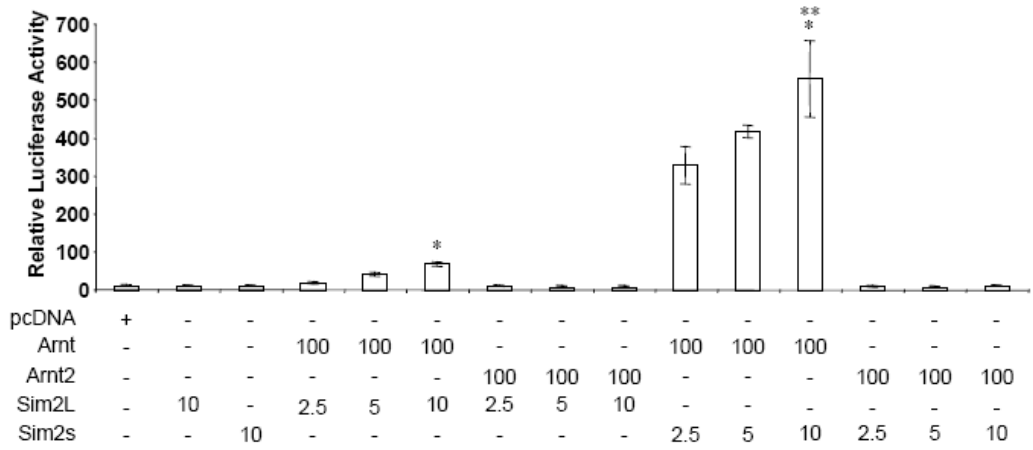


Figure 27. Transcriptional activity of mouse Sim2 and Sim2s on a CME-controlled reporter gene: the effects of increasing amounts of mouse Sim2 or Sim2s on CME-mediated gene expression. HEK-293T cells were transfected with or without 100ng ARNT or ARNT2 expression vector with increasing amounts of Sim2 or Sim2s expression vector plus a CME-controlled luciferase reporter and ß-gal expression plasmids. The amount of each plasmid per transfection (in ng) is indicated under the figure. Luciferase activity was increased in both Sim2 and Sim2s-transfected cells, but only when ARNT was present. The degree of induction was Sim2-dependent as higher amounts of expression vector increased luciferase activity. Again, Sim2s was better than Sim2 at inducing CME-controlled gene expression. *Luciferase expression is higher in the presence of SIM and Sim2s in comparison to no ARNT controls ($p<0.005$). **Luciferase expression is higher in Sim2s cells in comparison to Sim2 cells ($p<0.01$).

3.1.5. SIM-mediated regulation of a CME-controlled gene requires the transcriptional activation domain of ARNT

Confirmation of ARNT-dependent Sim2s transcriptional activation from a CME-controlled gene is presented in Figure 28. In these experiments, contransfection of a mutant ARNT (ArntΔTAD), which is missing the transactivation domain, abolished the ability of Sim1, Sim2 and Sim2s to increase expression from a CME. Expression of ArntΔTAD repressed basal CME-mediated gene expression in the absence of external SIM proteins (Fig. 28). Co-expression of Sim1 and ARNT resulted in significant reporter gene expression that was abolished when ARNT was replaced with the

ArntΔTAD expression vector. A similar effect was seen with Sim2, but the degree of gene activation was significantly lower than was seen with Sim1. Surprisingly, Sim2s was almost as potent as Sim1 in activating CME-controlled gene expression (Fig. 28). As observed in Sim1, this effect was completely abolished when ARNTΔTAD was substituted for ARNT, implying that this effect is entirely mediated by the activation domain of ARNT.

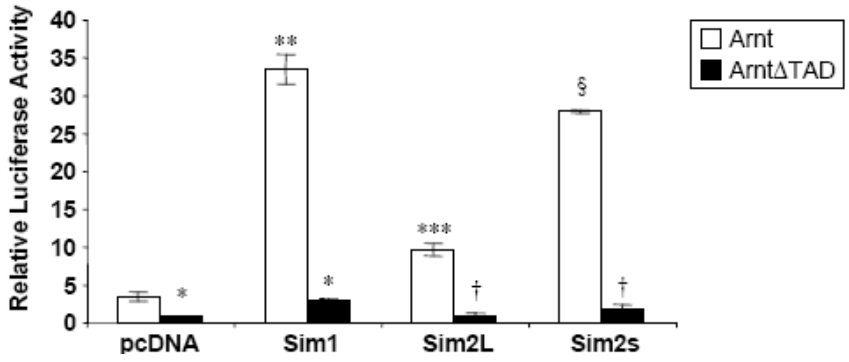


Figure 28. Sim-mediated regulation of a CME-controlled gene requires the transcriptional activation domain of ARNT. Control (pcDNA), Sim1, Sim2 or Sim2s expression plasmids were co-transfected with pCME-luc and a ß-gal expression vector along with full length Arnt (white bars) or a mutant Arnt (black bars) expression plasmid. In the presence of full length ARNT, robust luciferase expression was detected in Sim1 and Sim2s-transfected cells and to a lesser extent, in Sim2 cells. This effect was abolished when ARNT was replaced with the transactivation mutant ARNT (ARNTΔTAD).

*Luciferase expression is repressed in the presence of ARNTΔTAD ($p<0.03$). **Luciferase expression is induced in the presence of Sim1 ($p<0.02$). ***Luciferase expression is increased in the presence of Sim2 in comparison to pcDNA control cells ($p<0.01$). §Luciferase expression is higher in the presence of Sim2s in comparison to pcDNA-transfected cells ($p<0.0003$). †Luciferase expression is inhibited in the presence of ARNTΔTAD ($p<0.001$).

To confirm the interactions between Sim2 and Sim2s on a CME, ChIPs were performed on CME-luc transfected control- and Sim2/ARNT or Sim2s/ARNT cells. Chromatin was immunoprecipitated with an anti-SIM2 antibody that recognizes both

SIM2 and SIM2s, and was analyzed for CME binding using a set of PCR primers specific for the CME reporter plasmid. The presence of Sim2 on the CME was detectable in control cells, most likely due to endogenous Sim2 (data not shown). More importantly, the presence of both Sim2 and Sim2s on the CME was elevated in Sim2-transfected cells (Fig. 29). These data suggest that the differential outcomes of Sim2 isoform binding to a CME are not due to changes in DNA binding.

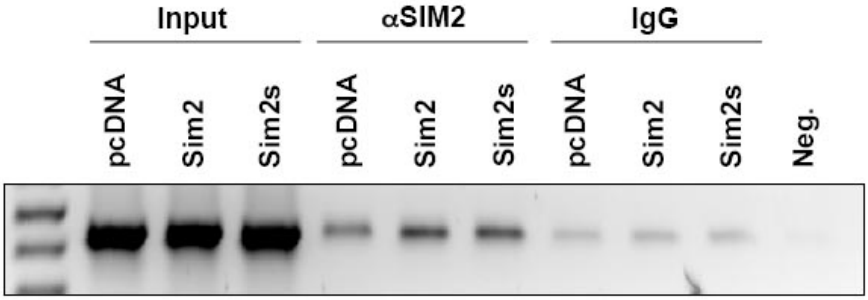


Figure 29. Sim2 and Sim2s bind directly to a CME. HEK-293T cells were co-transfected with CME-luc and either a control (pcDNA), Sim2 or Sim2s expression vector. Following formalin cross-linking, chromatin was sonicated and immunoprecipitated with an anti-Sim2 antibody and analyzed by PCR using primers that recognize the CME binding site in the reporter vector. Background was determined using IgG instead of anti-SIM2. The relatively small amount of product obtained in control cells is comparable to the IgG control, and may reflect binding of endogenous Sim2 to the CME. In contrast, both Sim2 and Sim2s binding to the CME was readily apparent in Sim2-transfected cells. Neg; PCR negative control containing no template.

3.2. Sim2s is an important regulator of mammary gland development

3.2.1. Sim2s is the predominant Sim2 isoform expressed in many mouse tissues

Sim expression was analyzed in mouse tissues by RT-PCR and real time RT-PCR (Fig. 30). Sim1 was expressed at high levels in kidney, brain, lung and skeletal muscle (Fig. 30A). In contrast, Sim2 and Sim2s were expressed at high levels in kidney and

skeletal muscle. The ratio of Sim2s and Sim2 mRNA expressed in kidney and skeletal muscle appeared to differ, suggesting that their expression is controlled in a tissue-specific manner. Therefore, expression of both isoforms of Sim2 were determined in kidney, liver and skeletal muscle by RT-PCR using a common forward primer located in exon 9, and both Sim2- and Sim2s-specific reverse primers located in exons 11 and "s", respectively. As expected, neither Sim2 isoform was detected in liver, but kidney and skeletal muscle expressed both Sim2 and Sim2s in different relative amounts (Fig. 30B). In agreement with our initial analyses (Fig. 30A), kidney expressed slightly more Sim2s than Sim2 while skeletal muscle expressed significantly more Sim2 than Sim2s. Quantitative real time RT-PCR analyses of kidney, liver and skeletal muscle RNA for total Sim2 (Fig. 30C), Sim2s (Fig. 30D) and full length Sim2 (Fig. 30E) corroborated our PCR analyses and found that, overall, skeletal muscle contains significantly higher levels of full length Sim2 mRNA than kidney.

3.2.2. Expression patterns of mSim2 in HC11 cells

Sim2s was also the predominant form of Sim2 in HC11 mouse mammary epithelial cell line without any Sim2L expression. Interestingly, we found there was more protein expression in differentiated than undifferentiated HC11 cells (Fig. 31B), while there was no difference in the mRNA expression level (Fig. 31A).

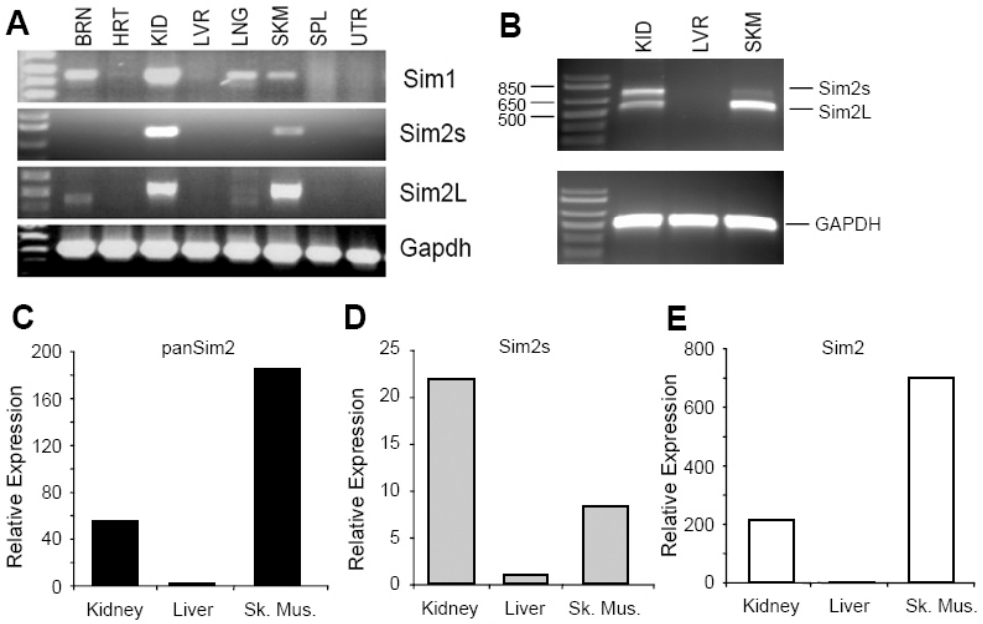


Figure 30. Tissue-specific expression of mouse Sim genes. A. RT-PCR analysis of Sim1, Sim2 and Sim2s expression in various mouse tissues. Total RNA from the indicated tissues (BRN; brain, HRT; heart, KID; kidney, LVR; liver, LNG; lung, SKM; skeletal muscle, SPL; spleen, UTR; uterus) was subjected to RT-PCR as described under Materials and Methods. Sim1 was expressed at relatively high levels in mouse kidney>brain>lung>skeletal muscle. Sim2s was expressed at high levels in kidney and was detectable in skeletal muscle. Full-length Sim2 was detected in skeletal muscle>kidney. B. Co-amplification of Sim2 and Sim2s by RTPCR in mouse kidney, liver and skeletal muscle. RT-PCR reactions were performed on total RNA from mouse tissues using a single forward primer and both Sim2s-, and Sim2 long form-specific reverse primers. Sizes of DNA ladder are indicated to the left. Neither isoform of Sim2 was detectable in liver, but both Sim2 and Sim2s were detected in kidney and skeletal muscle. The relative amounts of Sim2 and Sim2s expressed in these tissues differed with kidney expressing more Sim2s than Sim2 and vice versa in skeletal muscle. C-E. Real time PCR analyses of mouse kidney, liver and skeletal muscle mRNA for total Sim2 (C), Sim2s (D) and full-length Sim2 (E).

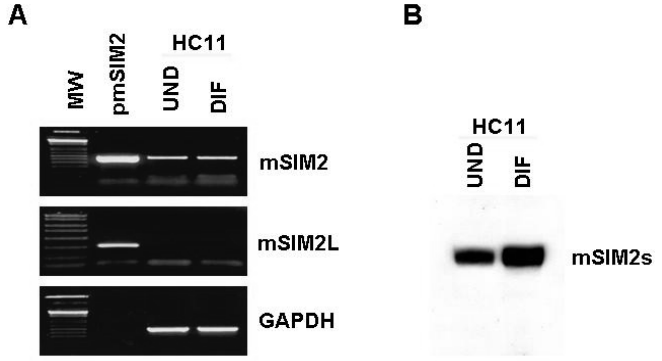


Figure 31. Expression pattern of mSim2 in HC11 cells. A. Expression of mSim2 mRNA in undifferentiated (UND) and differentiated (DIF) HC11 cells detectecd by RT-PCR. B. Expression of mSim2s protein in undifferentiated and differentiated HC11 cells detected by Western blot. mSim2-pcDNA (pmSim2) was used as positive control and GAPDH was used as internal control.

3.2.3. Sim2s is the predominant Sim2 isoform in the mouse mammary glands

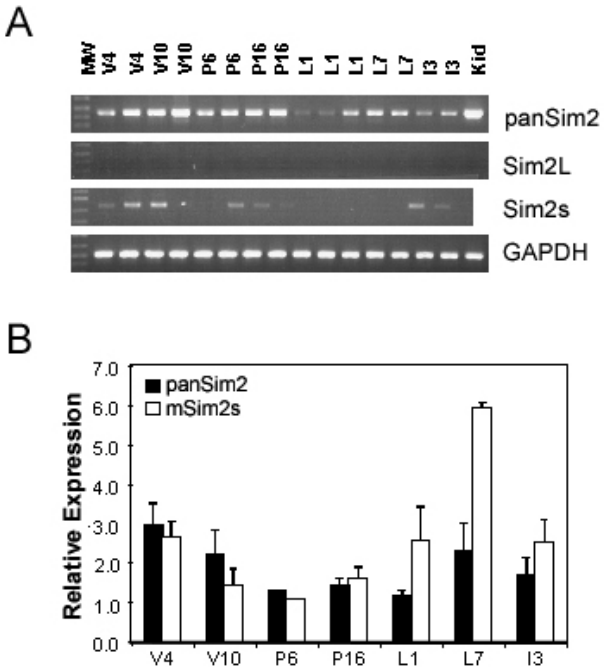


Figure 32. Expression of Sim genes in the mouse mammary gland. A. Expression of Sim genes in the developing mouse mammary gland. RT-PCR was performed on mammary gland RNA isolated from 4- and 10-week old virgin mice (V4 and V10), mice at days 6 and 16 of pregnancy (P6 and P16), mice at days 1 and 7 of lactation (L1 and L7) and mice at day 3 of involution (I3). Kidney RNA (Kid) was included as a positive control. B. Quantification of Sim2 gene expression in the developing mouse mammary gland. Real-time RT-PCR was used to quantify Sim2 mRNA levels in the developing mammary gland.

In the mammary gland, neither Sim1 nor the long form of Sim2 was detectable at any developmental time point (Fig. 32A). However, a form of Sim2 was expressed and developmentally regulated throughout pregnancy, lactation and involution (panSim2, Fig. 32A). Sim2s expression was similar to total Sim2 suggesting that Sim2s is the predominant Sim2 isoform expressed in the mammary gland. Quantitative real time PCR supported this observation and indicated that Sim2s expression is high in young

virgin mice and decreases in early adulthood (Fig. 32B). During pregnancy, Sim2s expression remains low, but begins increasing near parturition and through lactation. By involution day 3, Sim2s levels decrease back to adult nulliparous baseline levels (Fig. 32B).

3.2.4. Mammary gland Sim2s expression is developmentally regulated

Mammary gland development is a complicated process involving positive and negative interactions within and between the epithelial and stromal compartments. Since our data showed that Sim2 message levels are developmentally regulated in the mouse mammary gland, we performed experiments to better characterize the spatial and temporal expression patterns of mammary Sim2s. To do this, sections of mouse mammary glands from different stages of development were analyzed by immunostaining using a human SIM2s-specific antibody. Our results indicate that Sim2s is present in the condensed stroma surrounding the terminal end buds in the actively developing 4-week old mammary gland (Fig. 33A-B). Sim2s remains stromal through late virgin development, but is detectable in both the epithelial and stromal compartments by 10-weeks of age (Fig. 33C). Interestingly, Sim2s is only present in luminal epithelial cells during pregnancy and lactation (Fig. 33D) and is significantly decreased throughout involution (data not shown). Co-immunostaining of 4-week old virgin mammary glands for Sim2s and Keratin 14, a myoepithelial cell marker, indicated that Sim2s is entirely stromal at this stage, and is not expressed in myoepithelial cells (Fig. 33E-F). Negative controls for immunostain secondary antibody and immunofluoresence secondary antibodies are also presented in Figure 33G and 33H,

respectively.

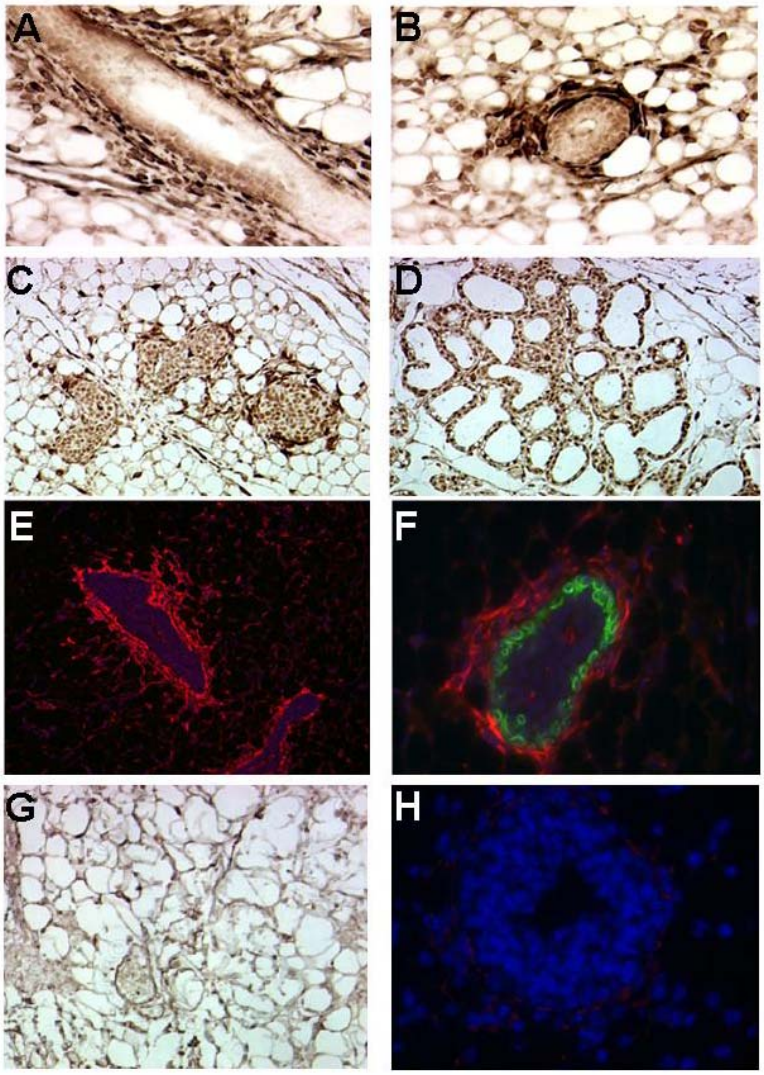


Figure 33. Immunolocalization of Sim2 expression in mouse mammary gland. A-B. Immunolocalization of mSim2s protein in a 4-week old virgin mouse mammary gland. Lateral (A) and cross-sectional (B) view of a 4 week-old mouse mammary duct. C. Immunolocalization of Sim2s protein in a 10-week old virgin mouse mammary gland. D. Immunolocalization of Sim2s protein in a day 7 lactating mouse mammary gland. E-F. Co-localization of Sim2s (red) and Keratin-14 (green) in 4 week-old mouse mammary gland. G. Negative control for immunostain secondary antibody. Donkey anti-goat. H. Negative control for immunofluoresence secondary antibodies. Co-stain with goat anti-rabbit and donkey anti-goat.

3.3. Sim2s is mammary tumor suppressor gene

3.3.1. SIM2s is down-regulated in breast cancer cells and primary breast cancer samples

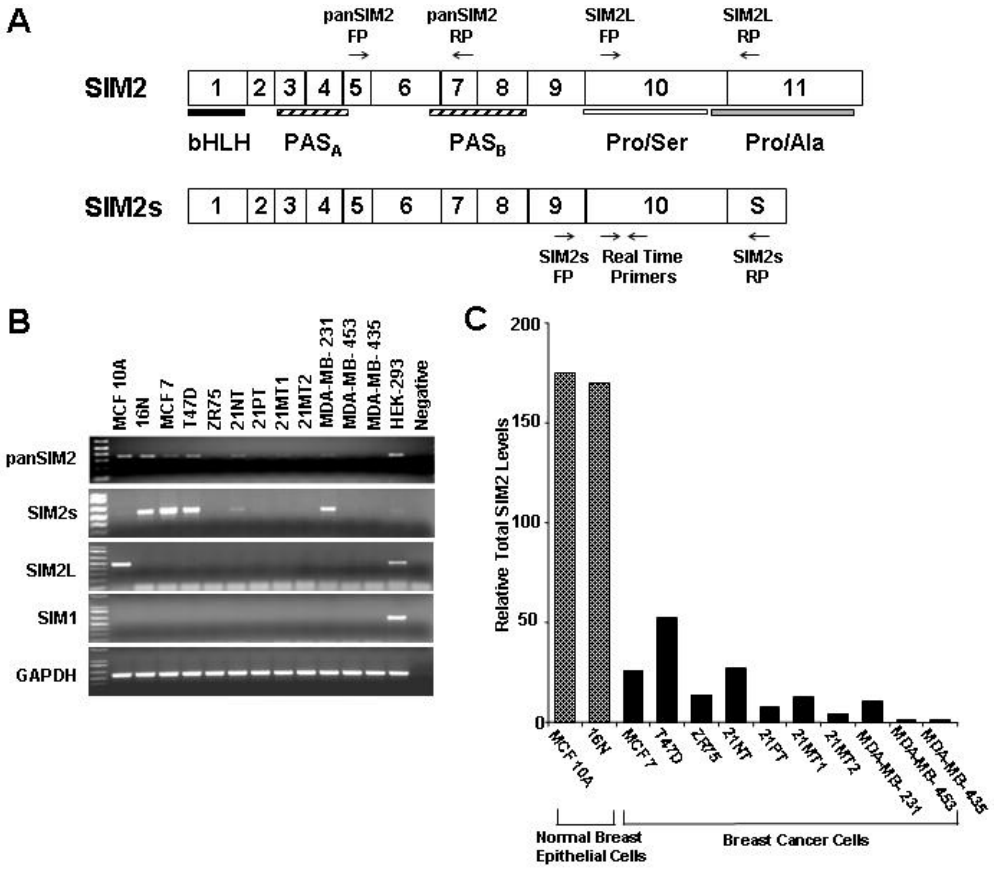


Figure 34. Single minded mRNA expression in breast epithelial- and cancer-derived cell lines. A. Schematic representation of SIM2 and SIM2s mRNAs with relative positions of PCR primers used in these studies indicated. Exons are indicated by numbered boxes. Structural motifs are represented by bars: bHLH (basic helix-loop-helix domain, black bar), PAS (Per-Arnt-Sim domains, stripped bars), Pro/Ser (proline/serine-rich, white bar) and Pro/Ala (proline/alanine-rich, grey bar) regions implicated in harboring SIM2 repression domains. PCR primer locations are indicated by name and arrows. B. Analyses of total SIM2 (pan SIM2), SIM2s, SIM2 long form (SIM2L) and SIM1 message levels in normal human breast and breast cancer-derived cell lines by RT-PCR. Cell lines are indicated at the top of the figure. C. Quantification of total SIM2 expression in human breast and breast cancer cells by real time RT-PCR. Data was quantified by the ΔΔCT method using TATA binding protein mRNA as the normalizer.

To determine if SIM2 plays a role in breast cancer, we examined a panel of human breast- and breast cancer-derived cell lines for SIM gene expression by standard and real time RT-PCR. The relative positions of the SIM2-specific primers used in these analyses are indicated in Figure 34A. The 21T cell line series was established from normal breast (16N) and primary (21NT and 21PT) and metastatic (21MT1 and 21MT2) breast tumor specimens from a single patient with infiltrating ductal and intraductal carcinoma (Band et al., 1990 & 1991; Coletta et al., 2004). Together with other human breast- and breast cancer-derived cell lines (i.e. MCF10A, MCF7, ZR75, T47D, MDA-MB-231, MDA-MB-453 and MDA-MB-435 cells), these cells provide a useful model to evaluate gene expression in a simulated tumor progression series. Total SIM2 mRNA was detectable at the highest levels in MCF10A and 16N normal breast epithelial cells (Fig. 34B). SIM2 expression was markedly decreased in the 21T breast cancer progression series and in the other breast cancer cell lines examined.

RT-PCR analyses using primers to discern between the short and long forms of SIM2 indicated that SIM2-positive cell lines expressed the SIM2s isoform, with the exception of MCF10A cells, which expressed only the long form of SIM2 (Fig. 34B). In contrast, SIM1 mRNA was not detected in any of the breast cells, but was abundant in kidney-derived HEK 293 cells.

Quantification of relative SIM2 expression in these cells by real time RT-PCR corroborated the RT-PCR results and confirmed that SIM2 expression is highest in non-transformed MCF10A and 16N cells, is much lower in the non-invasive, estrogen receptor positive MCF-7, T47D and ZR75 breast cancer cells and is dramatically

decreased in the highly invasive estrogen receptor negative MDA-MB-453 and MDA-MB-435 cell lines (Fig. 34C).

A subset of these cell lines was analyzed for SIM2s protein levels by Western blot (Fig. 35). In agreement with the PCR data, SIM2s protein was only detectable in MCF-7 and T47D cells. Interestingly, SIM2s protein was not detected in MDA-MB-231 cells despite having observable levels of SIM2s message. These results suggest that SIM2s is the most prevalent isoform of SIM2 expressed in human breast-derived cell lines.

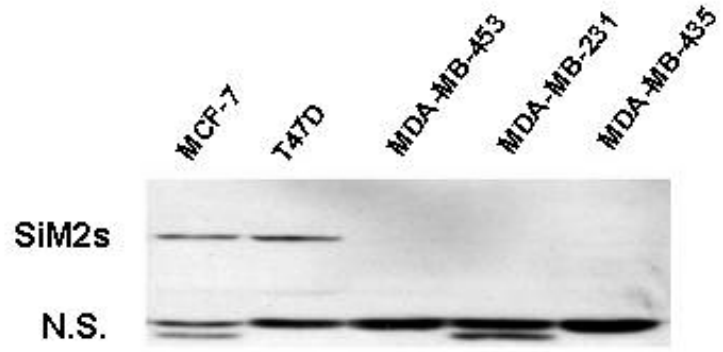


Figure 35. Western analysis of SIM2s expression in representative breast cancer cells using a SIM2s specific antibody.

This was confirmed by immunohistochemical staining of normal breast and primary breast tumor tissues using a SIM2s-specific antibody. SIM2s protein was readily detectable in the ductal epithelium of normal breast tissue (Fig. 36A), whereas SIM2s was not present in 18 out of 25 (72%) breast tumor samples (Fig. 36B-D).

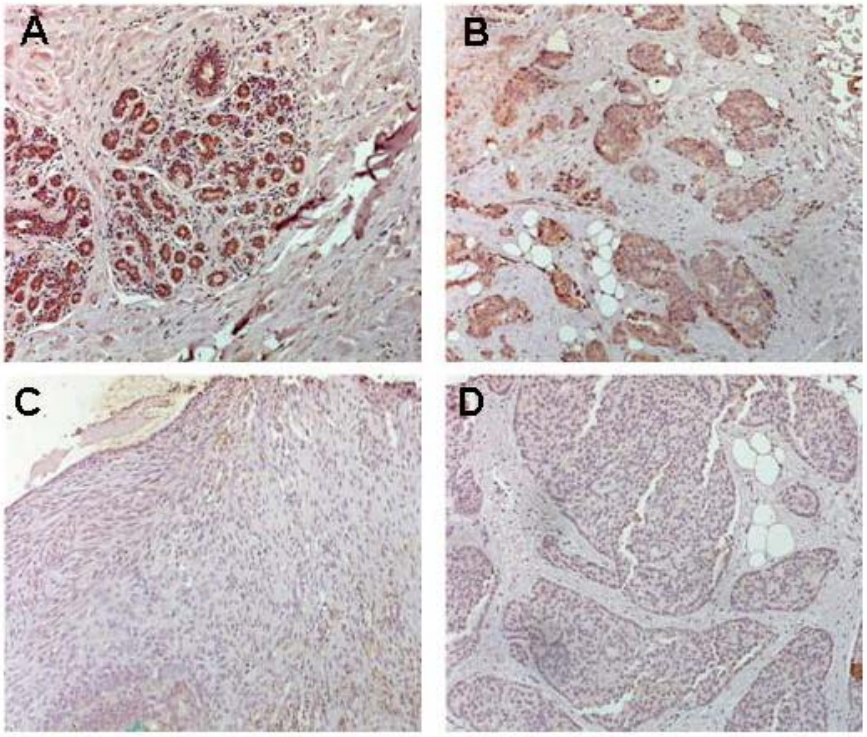


Figure 36. Immunohistochemical analysis of SIM2s protein levels in normal human breast and breast tumor tissues. A human breast cancer tissue array (Chemicon) containing 5 normal and 25 breast tumor tissue sections (in duplicate) was analyzed for SIM2s protein levels by immunohistochemistry as described in Materials and Methods. A. Representative image of a normal human breast tissue sample showing ductal epithelial cells with high levels of SIM2s staining. B. Representative image of a SIM2s-positive breast tumor sample. Note that the majority of breast tumor samples (18 out of 25, or 72%) were SIM2s negative. When present in tumor samples, SIM2s staining was considerably less intense than that observed in the normal breast tissues. C-D. Representative images of SIM2s-negative human breast tumor sections.

3.3.2. Inhibition of breast cancer growth by SIM2s

Our data indicates that SIM2s expression is lost during breast cancer progression. To determine if SIM2s has tumor suppressor activity, MDA-MB-435 breast cancer cells were transduced with a recombinant retrovirus expressing SIM2s. Re-establishment of SIM2s expression was confirmed by RT-PCR (Fig. 37A) and immunofluorescence (Fig. 37B).

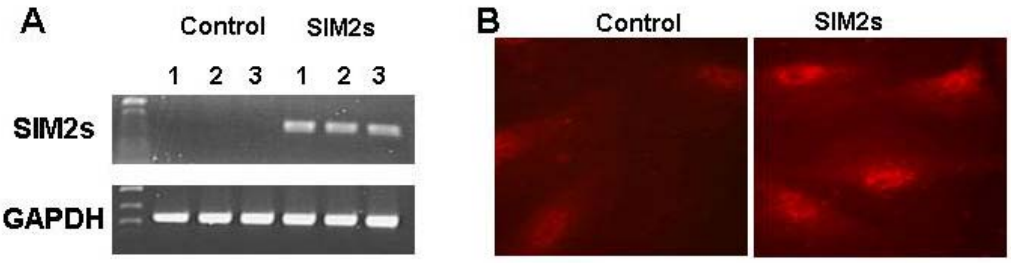


Figure 37. Reintroduction of SIM2s into MDA-MB-435 cells. MDA-MB-435 cells were infected with recombinant retroviruses expressing SIM2s or Puro only (Control vector). Two days after infection, puromycin was added to the cell medium and selection continued for 5 days. A. Confirmation of SIM2s expression in SIM2s-infected MDA-MB-435 cells by RT-PCR. B. Confirmation of SIM2s protein expression by immunofluorescence using a SIM2s polyclonal antibody.

Quantification of colonies after two weeks of selection showed that SIM2s significantly inhibited colony growth in MDA-MB-435 cells (Fig. 38A). Subsequently, a 7-day proliferation assay confirmed that stable expression of SIM2s blocked MDA-MB-435 cell growth (Fig. 38B). In addition to pooled samples, individual colonies of control- and SIM2s-transduced cells were isolated by serial dilution; however, these clones lost SIM2s expression after a few weeks of selection (data not shown), further supporting a role for SIM2s in inhibiting tumor cell growth.

The ability to grow in soft agar is a hallmark of the transformed phenotype. To assess if re-introduction of SIM2s reduces anchorage-independent growth of breast cancer cells, vector control- and SIM2s-infected MDA-MB-435 cells were plated in soft agar and assayed for colony formation over two weeks. Gross observation revealed that SIM2s-infected cells do not grow as well in soft agar as the control cells (Fig. 39). After 14 days, the number of colonies greater than 200μm was significantly lower in the SIM2s-infected cells (Fig. 39).

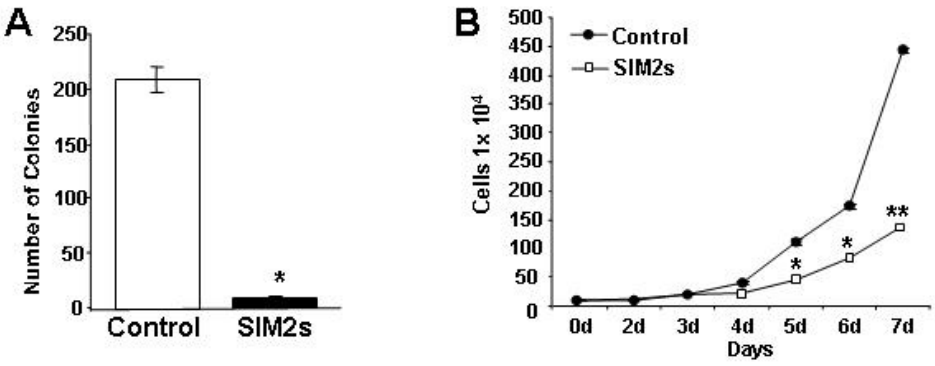


Figure 38. SIM2s suppresses growth of MDA-MB-435 breast cancer cells. A. Re-introduction of SIM2s inhibits expansion of MDA-MB-435 colonies under puromycin selection in comparison to vector-only cells. The number of colonies formed after two weeks of puromycin selection was lower (*P<0.0001) in SIM2s-transduced cells than in control cells as determined by Gimsa staining. The values shown are the mean ± SE of three separate samples. B. SIM2s inhibits breast cancer cell proliferation. Equal numbers of control and SIM2s-expressing cells were plated in six-well plates and counted daily over a 7-day period. Significant differences in cell proliferation due to reintroduction of SIM2s were apparent by day 5 (*P<0.001) and continued through day 7 (**P<0.0001). Data represent results from at least two independent experiments. The values shown are the mean ± SE of triplicates.

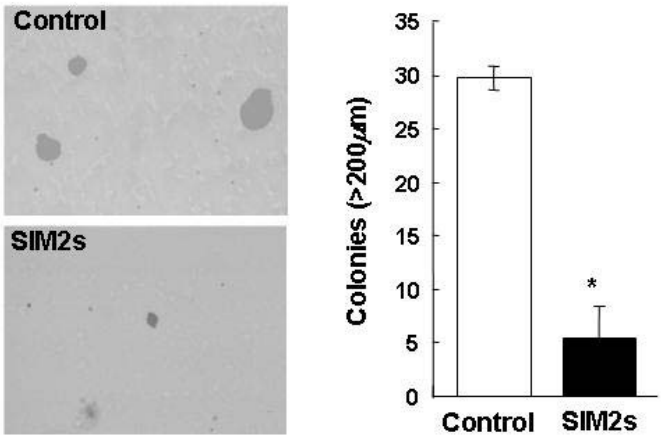
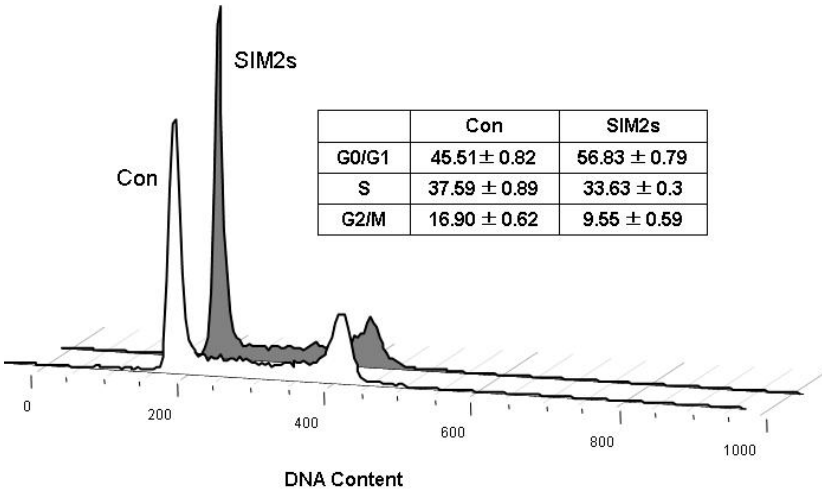


Figure 39. Reintroduction of SIM2s reduces anchorage-independent growth of breast cancer cells. SIM2s inhibits MDA-MB-435 cell growth on soft agar. The left portion of the figures shows representative images of colonies that grew on soft agar in control (top) and SIM2s (bottom) MDA-MB-435 cells. Following 14 days of growth, the number of colonies greater than 200 μm in size were counted (right portion of figure). Re-introduction of SIM2s into MDA-MB-435 cells reduced colony formation on soft agar (* P<0.005). Values are the mean number of colonies ± SE of three plates/group.

3.3.3. SIM2s induces G0/G1 cell cycle arrest

Reintroduction of SIM2s led to G0/G1 cell cycle arrest in MDA-MB-435 breast cancer cells. Compared with Control, SIM2-transduced MDA-MB-435 cells showed over 10% increase in G0/G1 phase and about 7% decrease in G2/M phase (Fig. 40).



	Con	SIM2s
G0/G1	45.51± 0.82	56.83 ± 0.79
S	37.59 ± 0.89	33.63 ± 0.3
G2/M	16.90 ± 0.62	9.55 ± 0.59

Figure 40. Reintroduction of SIM2s induced G0/G1 cell cycle arrest in MDA-MB-435 cells.

3.3.4. SIM2s regulates genes involved in cancer initiation and progression

Identification of SIM2s target genes by microarray analysis. We found that SIM2s regulates many genes involved in cancer initiation and progression including Wnt5b, Interleukin 21, Melanin A and MMP3 (Fig. 41A). This result was confirmed by Real Time PCR for Interleukin 21 and WNT5B (41B).

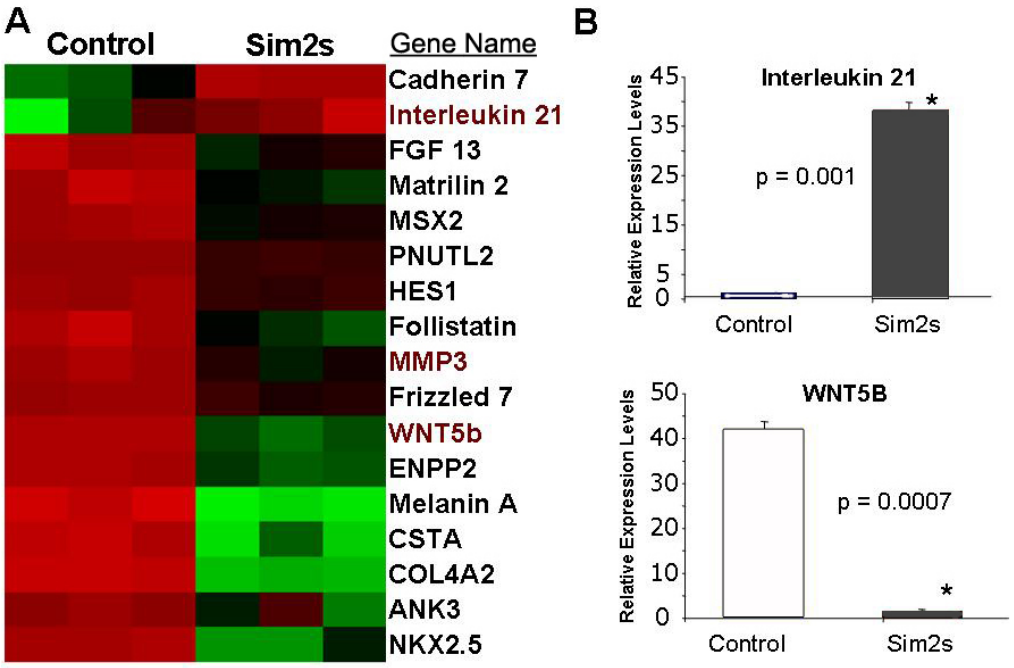


Figure 41. Identification and confirmation of SIM2s target genes by microarray analysis. A. Microarray analysis of MDA435 and MDA435-SIM2s cells (Codelink Bioarrays). Genes shown were regulated by at least 4-fold and p<0.001. B. Confirmation of SIM2s-dependentt regulation of Wnt5b and Interleukin-21 by real time RT-PCR.

3.3.5. Loss of mammary Sim2 expression results in loss of epithelial characteristics and hyperplasia

Since Sim2-/- mice die shortly after birth, we dissected mammary glands from newborn Sim2-/- pups and placed them into the cleared fat pads of nude mice. After 8 weeks of outgrowth, the transplanted glands were removed for histological analyses. Whole mounts of Sim2-/- mammary glands initially suggested that loss of Sim2 stimulates alveolar development since Sim2-/- glands appear to have a significant increase in the number of alveolar buds on the ductal tree (compare Figures 42A and 42B). However, H&E staining of sections through these glands indicate that the ducts of Sim2-/-

mammary glands failed to hollow (compare Figures 42C and 42D) and that the alveolar-like structures are actually solid masses of cells that appear to be spreading into the surrounding stroma (compare Figures 42E and 42F).

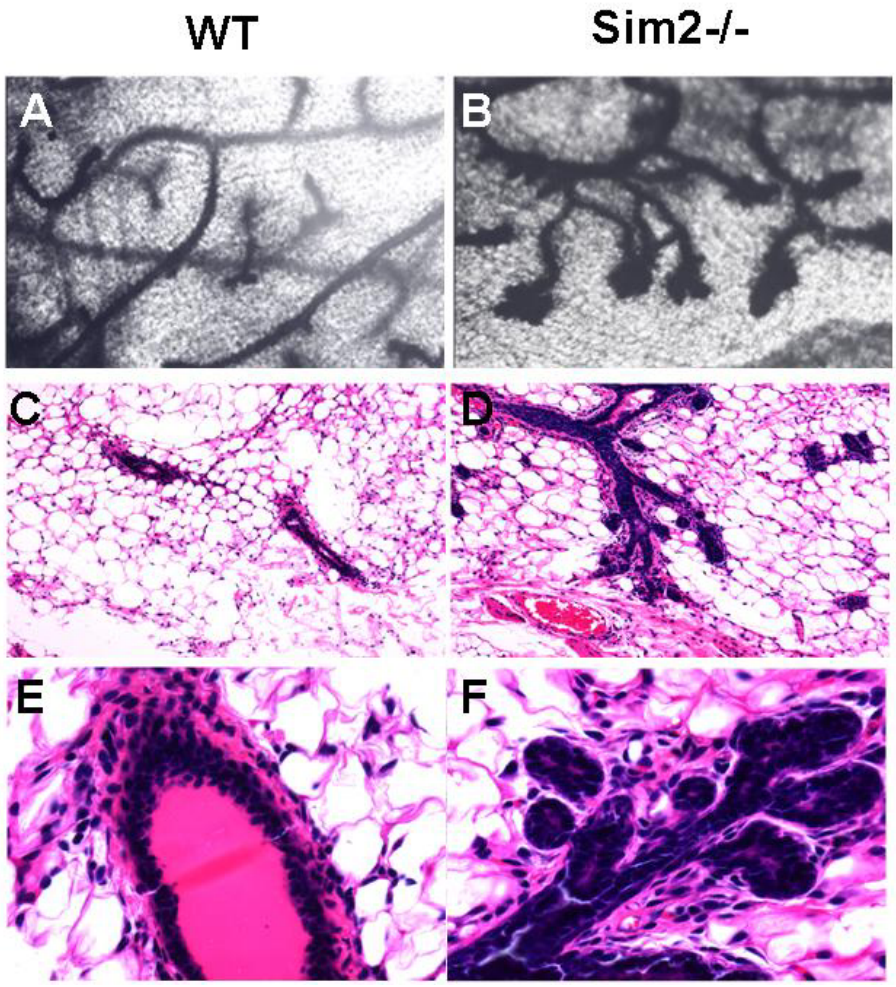


Figure 42. Mammary gland morphology in Sim2-/- mice. A & B. Whole mount analyses of wild type and Sim2s-/- mammary gland outgrowths 8 weeks post-transplantation. C-F. H&E staining of wild type and Sim2-/- mammary gland outgrowths.

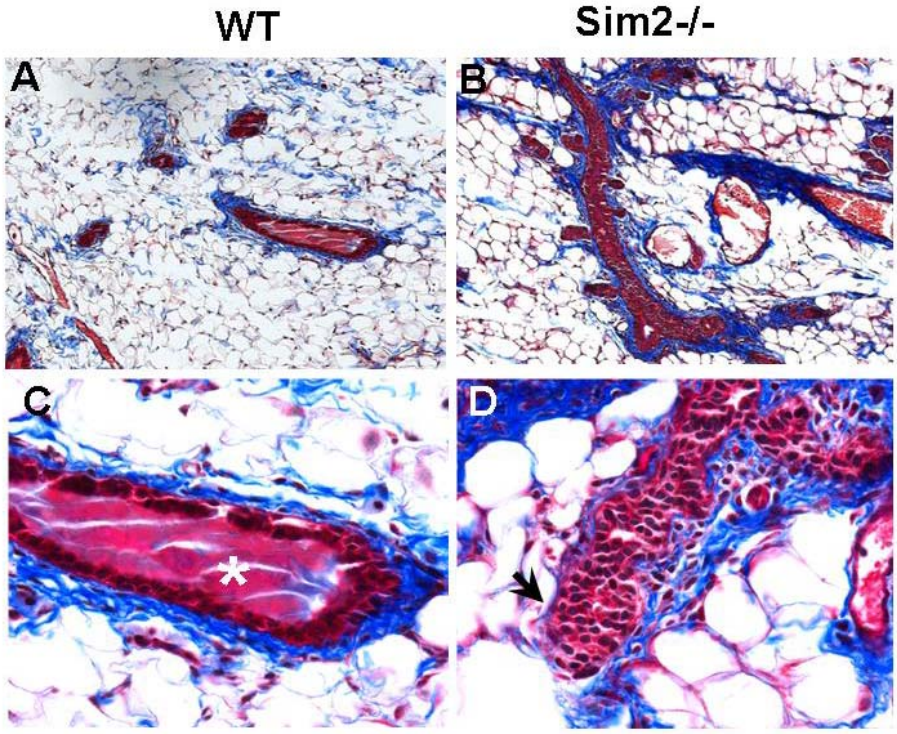


Figure 43. Trichrome staining of mouse mammary gland sections. Asterisk indicates hollow duct found in normal, but not Sim2-/- mammary glands. Arrow indicates ruffling edge of epithelial cells and thinning basement membrane.

Masson's trichrome staining of sections showed that the cells present in these masses are non-polarized and the collagen-rich basement membrane surrounding them is disorganized and reduced in thickness (compare Figure 43A to 43B and Figures 43C to 43D). In addition, the cells adjacent to the basement membrane in $Sim2^{-/-}$ mammary glands have ruffled edges suggesting that they are actively degrading the basement membrane and migrating into the surrounding stroma (arrow in Figure 43D). Furthermore, Sim2s heterozygote mice have an elevated frequency and severity of focal ductal dysplasia, demonstrating that loss of one copy of Sim2s also affects mammary gland development (data not shown). These results suggest that the cells in $Sim2^{-/-}$

mammary glands are undergoing an epithelial to mesenchymal transition and provide strong support for our hypothesis that SIM2s is an important mammary tumor suppressor gene.

3.4. SIM2s is a metastasis inhibitor

3.4.1. SIM2s inhibits metastasis and invasion

Since we found that SIM2s expression is inversely related to cellular invasiveness (Fig. 34), we next wanted to determine if SIM2s affected the invasive potential of MDA-MB-435 cells. Control- and SIM2s-transduced cells were grown in a Matrigel-coated modified Boyden chamber, which mimics the three-step process of invasion: adhesion, proteolytic dissolution of the extracellular matrix (ECM) and migration. Before the invasion assay, we also tested the migration rate of Control- and SIM2s-transduced cells. Consistent with our hypothesis that SIM2s is a tumor suppressor, forced expression of SIM2s significantly blocked both migration (Fig 44A) and invasion (Fig 44B) over 80% in comparison to the control cell line.

3.4.2. Loss of Sim2 leads to an epithelial to mesenchymal transition (EMT) in the mouse mammary gland

Since Sim2-/- mammary glands appear to be undergoing an EMT, we examined transplanted wild type and Sim2-/- mammary glands for E-cadherin levels by immunofluorescence. E-cadherin expression was dramatically decreased in Sim2-/- glands (compare Figures 45A and 45B). Confirmation of Sim2 expression in WT and Sim2-/- glands is shown in Figures 45E and 45F.

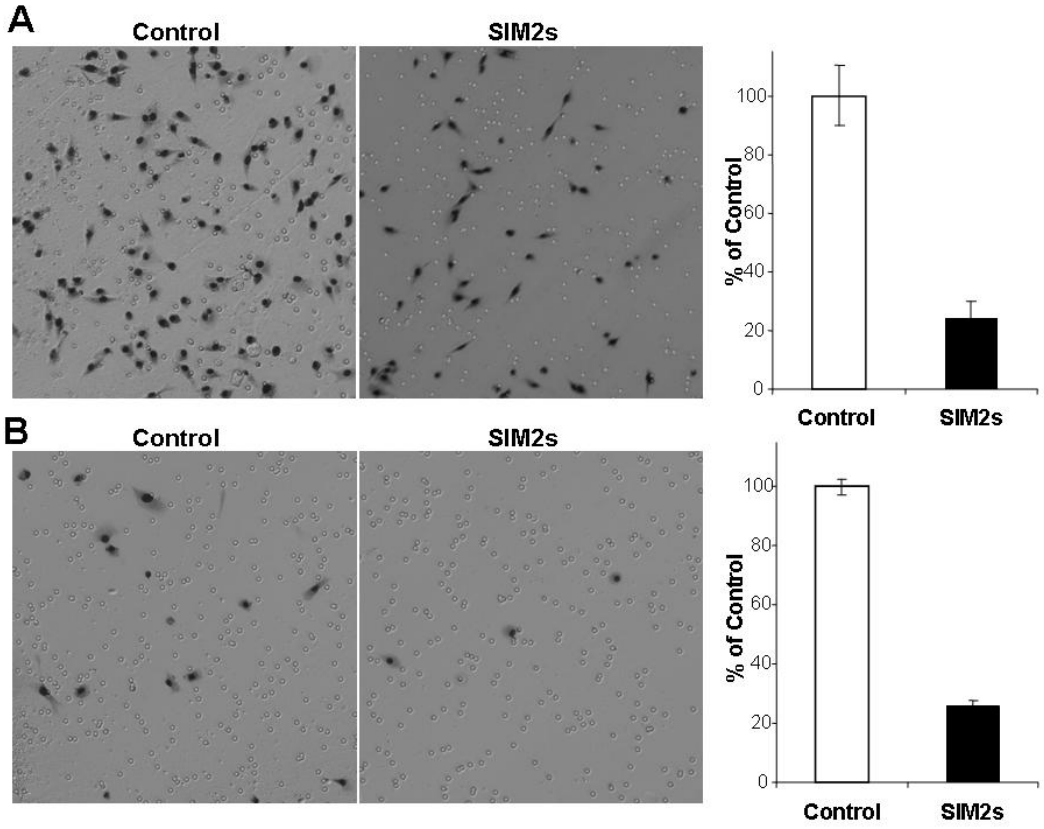


Figure 44. SIM2s inhibits both migration and invasion. SIM2s-dependent inhibition of migration and invasion potential in MDA-MB-435 cells. A. Migration assay. B. Invasion assay. MDA-MB-435 cells (1 × 105 cells per sample) were plated in serum free media in the upper compartment of a modified Boyden chamber coated with Matrigel. The lower compartment contained 10% FBS to serve as the chemo-attractant. After 24 hr, the number of cells that had migrated to the bottom surface of the membrane were stained with Diff-Quick and counted. Less SIM2s-expressing cells were able to invade and migrate to the lower chamber ($P<0.005$). Values are the average number of cells per five fields per membrane of three separate plates, and are expressed as percent of control (mean ± SE).

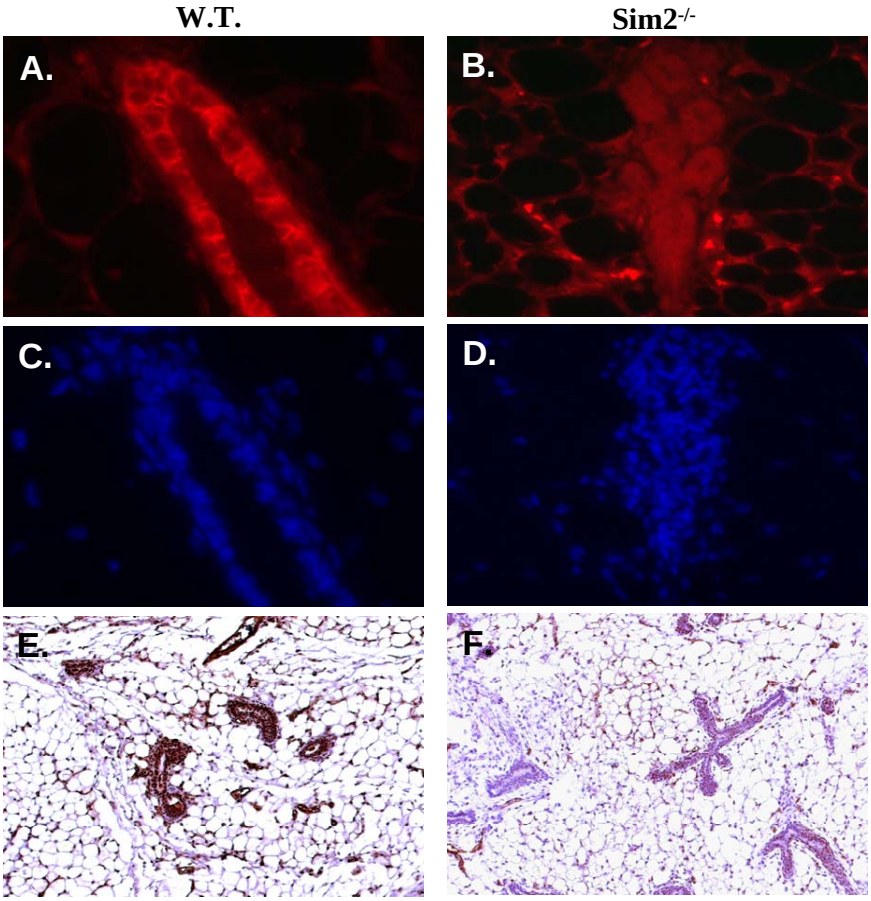


Figure 45. E-Cadherin and Sim2s expression in wild type and Sim2-/- mouse mammary glands. A-D. Loss of E-cadherin expression in Sim2-/- mammary glands. Indirect immunofluorescent staining for E-cadherin (A and B) and DAPI (C and D) in WT (A and C) and Sim2-/- (B and D) mammary glands. E-F Confirmation of Sim2 loss in Sim2-/- mammary glands Immunolocalization of Sim2s expression in wild type (E) and Sim2-/- (F) mammary glands by immunostaining.

3.4.3. SIM2s inhibits MMP gene expression and activity

The ability of SIM2s to decrease invasiveness of MDA-MB-435 cells suggested that SIM2s affects one or more genes responsible for ECM remodeling. Tissue remodeling requires degradation of the basement membrane and surrounding extracellular matrix. This occurs during normal mammary gland development, and under pathological conditions such as tumor progression and metastasis. Because MMPs are involved in

breast cancer initiation, proliferation and invasion, RT-PCR analyses of control- and SIM2s-transduced cells for expression of several MMP genes was performed. MMP 2, 3, 9, and 10 was decreased by SIM2s (Fig 46A). The effects of SIM2s on MMP2 expression correlated with MMP2 enzymatic activity as determined by gelatin gel zymography (Fig. 46B). Real time analysis for the MMP2 again confirmed that SIM2s decreases MMP2 mRNA expression level (Fig. 46C).

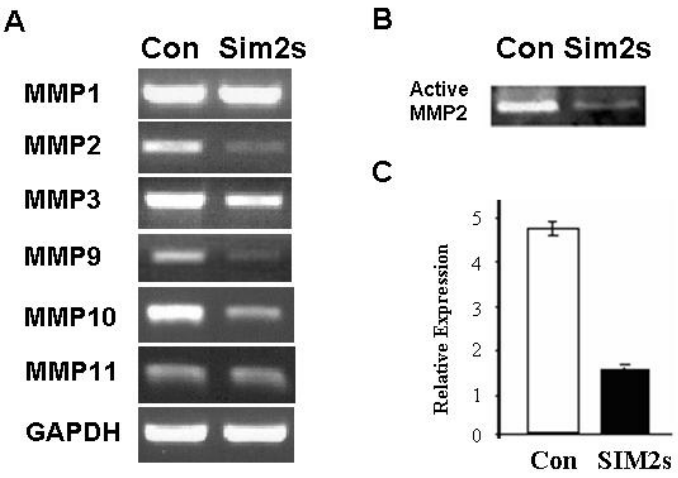


Figure 46. Down regulation of MMP expression and activity by SIM2s. A. Analysis of MMP gene expression in MDA-MB-435 and MDA-MB-435-SIM2s cells by RT-PCR. B. Analysis of MMP2 activity in MDA-MB-435 and MDA-MB-435-SIM2s cells using gelatin gel zymography. C. Quantification of MMP2 and MMP3 expression in MDA-MB-435 and MDA-MB-435-SIM2s cells by real time RT-PCR.

Of the MMP genes analyzed in these experiments, MMP3 was most significantly repressed by SIM2s. This finding was noteworthy as MMP3 is known to play important roles in cell migration, breast cancer progression and EMT (Ioachim et al., 1998; La Rocca et al., 2004). To better characterize the effects of SIM2s on MMP3 expression,

control- and SIM2s-transduced MDA-MB-435 cells were analyzed for MMP3 mRNA by real time RT-PCR. SIM2s significantly repressed both basal and PMA-induced MMP3 expression (Fig. 47A). The effects of SIM2s on MMP3 expression correlated with MMP3 protein and enzymatic activity as determined by Western blot and casein gel zymography (Fig. 47B).

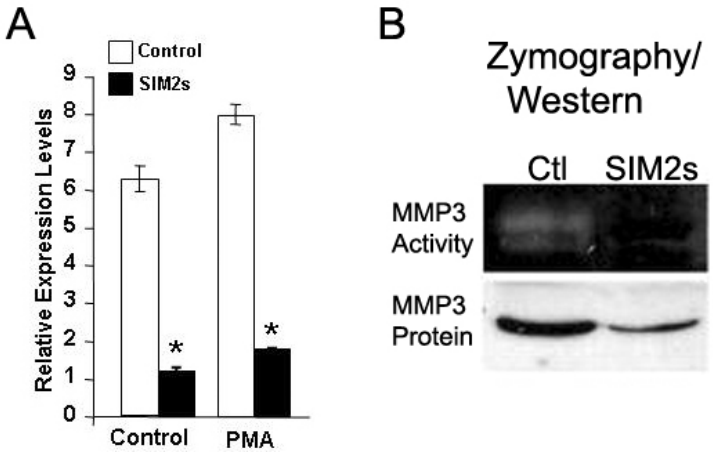


Figure 47. SIM2s-dependent inhibition of MMP3 expression and enzymatic activity in MDA-MB-435 cells. A. SIM2s inhibits basal and PMA-induced MMP3 expression in MDA-MB-435 cells. Quantification of MMP3 mRNA levels in control and SIM2s- transfected MDA-MB-435 cells by real time RT-PCR. The values shown are the mean ± SE of three separate samples, *P<0.005. B. Casein zymography and Western Blot analysis of MMP3 activity and protein, respectively, in control and SIM2s-infected cells.

3.4.4. Loss of Sim2 leads to results in increase in MMP3

The effect of SIM2s on MMP3 expression was also confirmed in the mouse mammary glands. We examined transplanted wild type and Sim2-/- mouse mammary glands for MMP3 levels using immunohistochemistry. MMP3 expression was dramatically increased in Sim2-/- glands (Fig. 48A & B). Confirmation of Sim2

expression in WT and Sim2-/- glands is shown in Figures 45E and 45F.

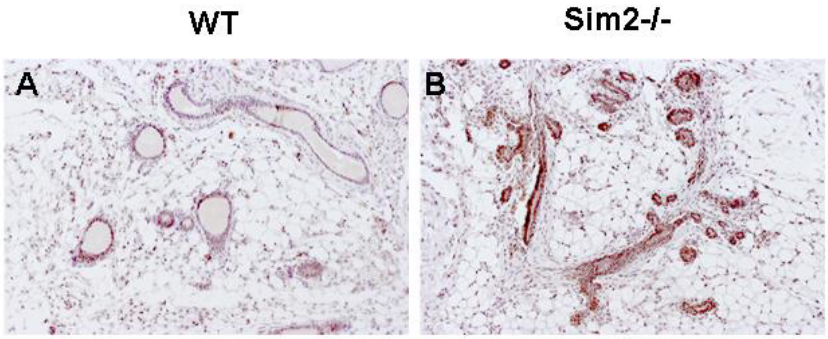


Figure 48. MMP3 expression in wild type and Sim2-/- mouse mammary glands. Loss of Sim2 results in increase in MMP3.

We also found that Sim2 -/- mammary gland is highly proliferating by immunohistochemistry for a proliferation marker, Ki-67. The expression level of Ki-67 was 4 times higher in the Sim2-/- mice, comparing with wild type mice (Fig. 49).

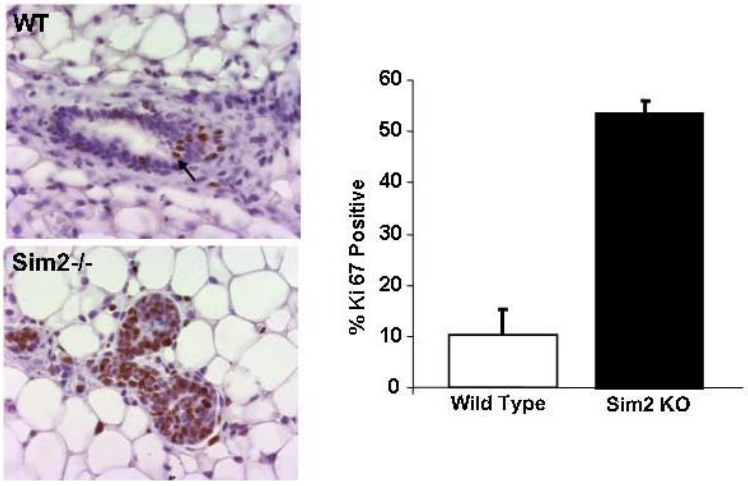


Figure 49. Proliferation in wild type and Sim2-/- mice.

3.4.5. SIM2s directly binds the MMP3 promoter to repress MMP3 expression

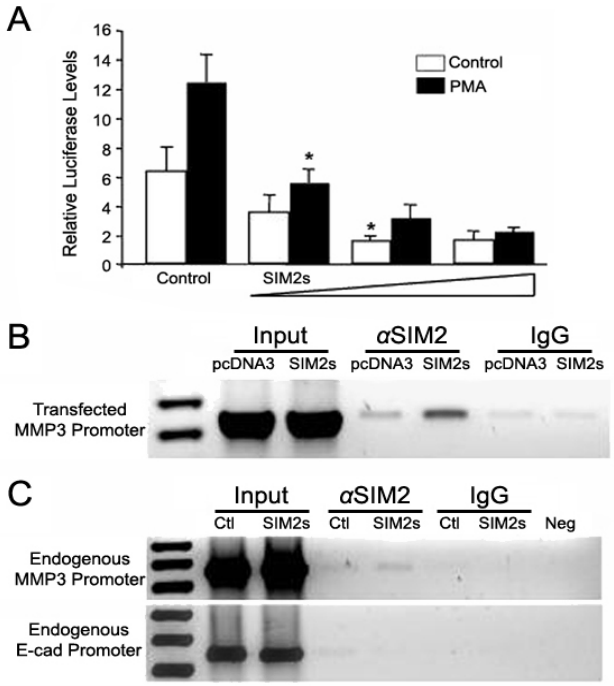


Figure 50. SIM2s represses MMP3 transcription by binding the MMP3 promoter. A. Effects of SIM2s on basal and PMA-induced expression from human MMP3 reporter constructs. MDAMB-435 cells containing -2264 to +37 of the human MMP3 promoter were tested for basal (open bars) and PMA-induced (black bars) luciferase expression in the presence of increasing amounts of SIM2s expression construct. Plasmids (0.45 µg MMP3-luc construct, various amounts of pcDNA3-SIM2sHA [0.0, 0.4, 1.0 or 2.0 µg] plus pcDNA3 [to equal 2.0 µg total expression construct] and 0.2µg ß-Gal expression vector) were introduced into MDA-MB-435 cells by transfection. Following a 24 hr recovery period, fresh medium containing DMSO or PMA (50 ng/ml) was added and luciferase and ß-galactosidase activities were measured 24 hr later. The values shown are the means ± SE for three separate plates. (B) Transfected SIM2s binds the MMP3-Luc reporter plasmid. 293t cells were co-transfected with the -2264 to +37 MMP3-Luc reporter and either a control (pcDNA3) or SIM2s expression vector. Following formalin crosslinking, chromatin was sonicated and immunoprecipitated with anti-SIM2 (Chemicon) or IgG (negative control) antibodies and analyzed by PCR using primers that target the reporter vector. (C) SIM2s binds the endogenous MMP3 promoter in SIM2s-transduced MDA-MB-435 cells. Chromatin immunoprecipitation analysis of the MMP3 and E-Cadherin (negative control) promoters in control and MDA-MB-435-SIM2s cells.

To determine if the effects of SIM2s on MMP3 expression are mediated at the

transcriptional level, a human MMP3 promoter-controlled luciferase reporter gene was analyzed for SIM2s responsiveness. MDA-MB-435 cells were transfected with a luciferase reporter gene under the control of the human MMP3 promoter in the presence or absence of SIM2s. SIM2s significantly decreased basal and PMA-induced luciferase activity in a dose-dependent manner (Fig. 50A). Consistent with SIM2s functioning as a transcriptional repressor, we determined by ChIP analyses that SIM2s bound the MMP3-Luc reporter in the presence of transfected SIM2s (Fig. 50B) and also interacted with the endogenous MMP3 promoter in SIM2s-transduced MDA-MB-435 cells (Fig. 50C). No SIM2s binding was observed on the E-Cadherin promoter, which was used as a negative control (Fig. 50C).

3.4.6. SIM2s inhibits MMP3 expression by inhibiting AP-1 binding to the MMP3 promoter

To determine if the effects of SIM2s on MMP3 expression are mediated at the transcriptional level, a human MMP3 promoter-controlled luciferase reporter gene was analyzed for SIM2s responsiveness. Introduction of a SIM2s expression construct significantly decreased basal, as well as PMA-induced, luciferase expression from an MMP3-luc reporter (Fig. 51A). Various 5' deletions of the MMP3 regulatory region (up to -288) were also significantly inhibited by SIM2s, implying that a sequence, or sequences, between -288 and +20 of the human MMP3 promoter are responsible for mediating the repressive effects of SIM2s on MMP3 expression (Fig. 51A).

MMP3 expression is low in normal tissues but can be up regulated following growth factor and/or cytokine stimulation. Many of the pathways regulating MMP3 expression

converge at an AP-1 site located between -75 and -65 of the MMP3 promoter, which binds JUN and FOS heterodimers. Increased AP-1 activation has been shown to play a role in transformation and invasion, most notably through up-regulation of MMPs.

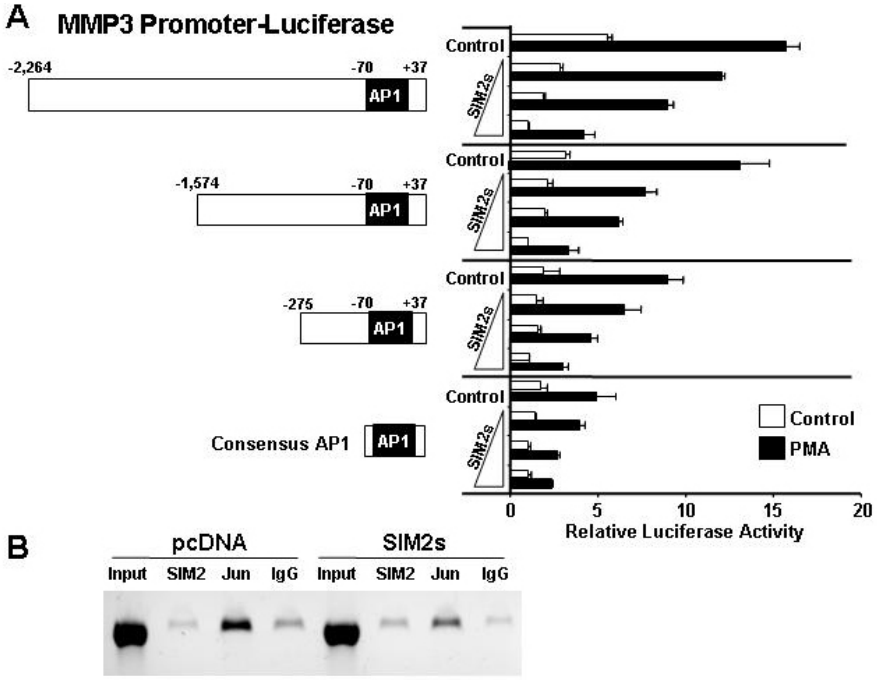


Figure 51. SIM2s represses the human MMP3 promoter by inhibiting AP-1 binding. A. Effects of SIM2s on basal and PMA-induced expression from human MMP3 reporter constructs. MDA-MB-435 cells containing various 5' deletions of the human MMP3 promoter, or a consensus AP-1 element, were tested for basal (open bars) and PMA-induced (black bars) luciferase expression in the presence of increasing amounts of SIM2s expression construct. Plasmids (0.45 µg MMP3-luc construct, various amounts of pcDNA3-SIM2sHA [0.0, 0.4, 1.0 or 2.0 µg] plus pcDNA3 [to equal 2.0 µg total expression construct] and 0.2µg ß-Gal expression vector) were introduced into MDA-MB-435 cells by transfection. Following a 24 hr recovery period, fresh medium containing DMSO or PMA (50 ng/ml) was added and luciferase and ß-galactosidase activities were measured 24 hr later. The values shown are the means ± SE for three separate plates. B. SIM2s blocks AP-1 binding to the MMP3 promoter. MDA-MB-435 cells were co-transfected with the -288 to +20 MMP3-Luc reporter and either a control (pcDNA3) or SIM2s expression vector. Following formalin cross-linking, chromatin was sonicated and immunoprecipitated with anti-JUN, anti-SIM2 or IgG (negative control) antibodies and analyzed by PCR using primers that target the reporter vector.

To determine if SIM2s mediates repression of MMP3 through this AP-1 site, MDA-MB-435 cells were co-transfected with a SIM2s expression construct and a consensus AP-1 luciferase reporter. Significant repression of basal and PMA-induced luciferase expression from the AP-1 reporter was observed in the presence of SIM2s (Fig. 51A). This response correlated with a decrease in JUN binding to the MMP3 AP-1 element as determined by ChIP analysis (Fig. 51B). These findings are similar to previous studies demonstrating that the Retinoic Acid and Glucocorticoid Receptors can repress MMP3 expression by inhibiting AP-1 binding (Dedieu & Lefebvre, 2006; Nicholson et al., 1990).

IV. DISCUSSION*

4.1. Differential transcriptional regulation by mSim2s

Transcriptional regulation occurs through multiple distinct mechanisms involving negative and positive interactions between regulatory factors. The mammalian SIM proteins are unique members of the bHLH-PAS family since they can exert negative effects on transcription. It has been determined that the repressive effects of SIM2 are mediated by two domains in its carboxyl terminus. One of these domains is rich in Pro and Ser residues while the other is Pro and Ala rich. These repressive domains appear to be non-specific as SIM2 can suppress activation of a Gal4 activation domain on a thymidine kinase promoter (Moffett & Pelletier, 2000; Probst et al., 1997) as well as ARNT-mediated transactivation. Similar hydrophobic domains are present in other transcriptional repressors including the *Drosophila* Kruppel transcription factors and Even-skipped, which inhibit transactivation by competing with TBP for TATA box binding thus preventing assembly of the preinitiation complex (McKay et al., 1999).

In this study, we have shown that mouse Sim2s, a splice variant of Sim2, has differential effects on CME- and HRE-mediated gene expression. Sim2s is less potent than full length Sim2 at repressing HIF1-α (Fig. 24), and can activate expression of a gene controlled by the *Drosophila* toll gene CME via ARNT's transactivation domain

---

* Parts of this section are reprinted from Metz et al., 2006, J Biol Chem, 281(16), 10839-10848, Copyright © 2006 by the American Society for Biochemistry and Molecular Biology and Kwak et al., 2006, Carcinogenesis, in progress, Copyright © 2006 by Oxford University Press.

(Fig. 26). The hypo-suppressive effects of Sim2s are not surprising given that Sim2s is missing the Pro/Ala-rich repression domain. What is surprising is that Sim2s is just as repressive as Sim2 on TCDD-mediated gene expression through a DRE (Fig. 25) and is able to increase expression of a CME-controlled gene (Fig. 26). These data suggest a model in which the response element dictates transcription factor domain-dependent suppression or activation.

On an HRE, both the Pro/Ser-rich and Pro/Ala-rich domains of Sim2 appear to exert repressive effects. This is based on the observation that Sim2s, which is missing the Pro/Ala-rich region, still exerts a repressive effect although it is not as strong as that observed with Sim2 (Fig. 24). In the case of the DRE, both Sim2 and Sim2s can repress gene expression to an equal extent suggesting that only the Pro/Ser-rich domain mediates Sim2-mediated repression from a DRE. In contrast, Sim2s can activate expression from a CME apparently by acting as a docking protein for ARNT (Figs. 28 and 29). This implies that the Pro/Ala-rich sequence present in Sim2, but not Sim2s, exerts a negative effect on CME-mediated gene expression. As Sim2s lacks this domain, interactions between Sim2s and ARNT on a CME result in ARNT-mediated activation of Sim2s targets.

Interactions between transcription factors and their cognate response elements are influenced by sequence flanking the core binding motif. Whitelaw, et al. (1994) reported that a region of the AHR ligand-binding domain exerted different degrees of repression on different DNA targets thus, providing an example of such promoter-specific influence on transcription factor function in the bHLH-PAS family (Whitelaw et

al., 1994). These response element-specific effects are most likely due to DNA-dependent conformational changes in the interacting factor, which may influence the ability of the transcription factor to recruit co-regulatory proteins to the promoter. For example, the POU domain-containing transcription factor POU1F1 (a.k.a. PIT1) represses transcription when bound to its response element in the growth hormone gene, but induces expression through a similar element in the prolactin promoter. This was shown by crystallography to be due to DNA-induced allosteric changes in PIT1 confirmation resulting in differential coregulator recruitment (Scully et al., 2000). In the case of SIM2s, such dynamic, DNA-mediated changes in transcriptional outcome may reflect the ability of SIM2s to exert differential effects on similar response elements. Further studies are necessary to elucidate the mechanisms governing SIM2s-mediated gene repression and activation.

In *Drosophila*, several SIM targets have been identified including slit, engrailed, breathless and spitz (Crews et al., 1992; Crews & Fan, 1999). No surprisingly, HIF1-α regulates many of the mammalian homologs of these genes since HIF1-α can bind the same core response element. Although definitive targets of SIM2 have not been identified in mammals, this study and others have shown that SIM2 can affect the actions of other bHLH-PAS proteins by active repression and interference (Moffett et al., 1997; Woods & Whitelaw, 2002).

4.2. Sim2s is an important regulator of normal mammary gland

Human SIM2s is expressed in normal kidney and tonsil as well as lung, and testes

(DeYoung et al., 2003a & b). We have found that Sim2s expression in mice is comparable, with high levels of Sim2s mRNA detected in kidney and skeletal muscle (Fig. 30A). Interestingly, the ratio of Sim2 to Sim2s differs between these tissues with Sim2s being expressed at higher levels than Sim2 in kidney and vice versa in skeletal muscle (Fig 30B). The significance of Sim2 isoform predominance in these tissues is unknown, but presumably could have substantial effects as we have demonstrated differences between binding partner specificity (Fig. 22) and transcriptional activities between Sim2 and Sim2s. Due to the complex interaction potential and overlapping expression patterns of Sim proteins and HIF1-α, a hypoxic switch has been proposed to operate in cells expressing both genes (Woods & Whitelaw, 2002). Such a switch could have profound implications for environmental regulation of developmental signaling pathways as developmental stage and organ-specific differences in response to systemic hypoxia have been reported (Minchenko et al., 2003; Stroka et al., 2001).

We found that Sim2s is a predominant Sim2 isoform in mammary gland and developmentally regulated. Sim2s expression is high in young virgin mice, decreases in early adulthood and remains low during pregnancy. Sim2s expression begins increasing near parturition and through lactation, and then decreasing back to adult nulliparous baseline levels (Fig. 32). As shown by immunostaining (Fig. 33), Sim2s was regulated not only temporally but also spatially in mouse mammary gland. Sim2s is present in the condensed stroma surrounding the terminal end buds in 4-week and in both the epithelial and stromal compartments by 10-weeks of age. Interestingly, Sim2s is only present in luminal epithelial cells during pregnancy and lactation, and then significantly decreased

throughout involution. Similar to the role of dSim in regulating axon growth and development in the *Drosophila* CNS (Crews et al., 1992), these results support a role for Sim2s in modulating stromal-epithelial interactions during the early stages of mammary gland development. Interestingly, expression of Sim2s was shown to be affected by cell differentiation as Sim2s expression levels increased upon differentiation of the HC11 cells (Fig. 31B). This is in agreement with the result showing that Sim2s expression increases near parturition and through lactation, and confined only in the luminal epithelial cells during pregnancy and lactation (Fig. 33D). This finding strongly implies that Sim2s plays a role at some stage of the differentiation program of mouse mammary epithelial cells.

Taken together, these results indicate that Sim2s is the predominant form of Sim expressed in the mouse mammary gland, and its expression is tightly regulated during adolescent development and throughout pregnancy. These data support our hypothesis that Sim2s plays an important role during mammary gland development and differentiation.

4.3. Sim2s is a mammary tumor suppressor gene

SIM2 was initially identified by positional cloning around the DS critical region of chromosome 21 and is believed to contribute to many of the physiological abnormalities associated with DS (Chrast et al., 2000; Ema et al., 1999). In these studies, we present evidence that a splice variant of SIM2, SIM2s, has tumor suppressor activity when re-expressed in invasive breast cancer cells. We found that forced expression of SIM2s in

MDA-MB-435 cells significantly inhibited proliferation (Fig. 38), anchorage-independent growth (Fig 39.) and invasion potential (Fig 44). Expression of SIM2s in MDA-MB-435 cells also caused G1/S cell cycle arrest (Fig. 40). Meng et al. (2006) found G1/S cell cycle arrest in the SIM2-transfected PN12 cell line, derived from a rat adrenal medullary pheochromocytoma tumor. The mRNA and protein expressions of cyclin E decreased in the SIM2-transfected cells, while p27 expression significantly increased.

The hypothesis that SIM2s has tumor suppressor activity is strongly supported by Sim2 knock out study. In this study, we showed that loss of Sim2s in the mammary gland resulted in dramatically increased ductal branching, accelerated lobuloalveolar-like precocious hyperplasia, and decreased cell apoptosis. These results suggest that the cells in Sim2-/- mammary glands are undergoing an epithelial to mesenchymal transition and provide strong support for our hypothesis that SIM2s is an important mammary tumor suppressor gene.

4.4. SIM2s is a metastasis inhibitor

E-cadherin is essential for the development and maintenance of mammary luminal epithelial cells (Boussadia et al., 2002). E-cadherin mediated adhesion is regulated, in part, by β-catenin, which couples it to the actin cytoskeleton (Barth et al., 1997). In addition to determining cell fate, cadherin-catenin complexes stabilize mammary gland architecture, dictate epithelial cell polarity, promote survival and inhibit motility. Inactivation of E-cadherin is associated with invasive breast cancer, premature apoptosis

and involution in mice. Consequently, expression of a stabilized β-catenin in the mouse mammary gland causes precocious alveolar development, hyperplasia, delayed involution and mammary tumors (Imbert et al., 2001). Sim2s-/- mammary glands lose E-cadherin expression and contain hyperplastic nodules (Figure 45), suggesting that Sim2s plays a role in regulating E-cadherin/ß-catenin signaling.

In this study, overexpression of SIM2s in MDA-MB-435 cells significantly inhibited invasion potential (Fig. 44). The decrease in invasion appears to be due, in part, to SIM2s' ability to repress MMP3 expression through direct inhibition of AP-1 binding to the MMP3 promoter (Fig 50). These results are consistent with the observation that SIM2 is a transcriptional repressor and imply that SIM2s is a mammary-specific tumor suppressor.

Matrix metalloproteases play a major role in a number of biological processes, including morphogenesis, polarity determination, apoptosis and angiogenesis (Alexander et al., 2001; Andarawewa et al., 2003; Ha et al., 2001; Lee et al., 2000). These processes occur at some stages in normal development, and during pathological conditions such as tumor progression and metastasis. Invasion requires active degradation of the extracellular matrix and basement membrane, which requires the actions of MMPs. Not surprisingly, over-expression of MMP3 in the mouse mammary gland causes increased lateral branching, precocious alveolar development, hyperplasia, EMT and cancer (Sympson et al., 1994 & 1995). MMP3 gene expression is primarily regulated at the promoter level and is induced by growth factors, cytokines, ECM components and EMT-dependent transcription factors like SLUG and SNAIL

(Chakraborti et al., 2003; Miyoshi et al., 2005, Miyoshi et al., 2004). We have found that re-expression of SIM2s decreases MMP3 mRNA levels by transcriptional repression of the MMP3 promoter through inhibition of AP-1 binding. This is consistent with other tumor suppressor genes such as the tumor metastasis suppressor nm23-H1, which inhibits invasion and metastasis by competing for and blocking activator binding to the YB1 response element on the MMP2 promoter (Cheng et. al., 2002). Similarly, expression of the ETS transcription factor TEL, abrogates RAS transformed NIH 3T3 cell growth and tumorigenicity by suppressing MMP3 mRNA and promoter activity (Fenrick et. al., 2000).

Our results are surprising since others have proposed that SIM2s is a tumor marker gene found to be up-regulated in prostate and pancreatic tumors (Benard et al., 2005; DeYoung et al., 2002; 2003a & 2003b). The observation that SIM2s may have contradictory roles in cancer progression is not unique. For example, in breast, ovarian and lung cancers, reduced caveolin-1 (CAV1) expression has been found to be associated with increased invasiveness and histological grade (Williams & Lisanti, 2005). This is supported by *in vivo* studies, which showed that mammary glands from $Cav1^{-/-}$ mice undergo precocious alveolar development, hyperplasia and increased incidence of tumor formation (Williams et al., 2004). In contrast, CAV1 is upregulated in metastatic bladder, thyroid and prostate carcinomas (Pflug et al., 1999; Williams et al., 2005). One possibility is that CAV1 is transiently down-regulated in migrating cells and re-established after intravasation. Perhaps there are tissue-specific differences in ancillary factors modifying CAV1 effects. Alternatively, it is possible that deletion of, or

mutations in, down-stream targets of CAV1 signaling could account for these differences. At present, very little is known about the function of SIM2, and even less is known about SIM2s. Thus, it is possible that the discrepancies between our results and those of DeYoung et al. may be due to similar, as yet uncharacterized phenomena.

The unique profile of cancer risk in DS individuals suggests that more than one gene may be responsible for the observed differences in cancer susceptibilities. Most likely, these effects are the result of complex interactions between gene products encoded by chromosome 21 and downstream targets of these genes on other chromosomes. Chromosome 21 contains over 300 genes, of which many have been implicated as playing a role in the DS-associated cancer promotion or protection effects. A recent report linked the increased incidence of DS acute megakaryoblastic leukemia to the transcription factor GATA1 (Greene et al., 2003; Wechsler et al., 2002). Another putative oncogene located on chromosome 21, *ETS2*, has been shown to be up-regulated in testicular germ cell tumors from DS subjects (Stage et al., 1997). Interestingly, despite evidence suggesting that ETS2 plays a role in breast cancer progression, over-expression of *ETS2* in DS individuals is not associated with increased breast cancer incidence.

The tumor suppressive activity of trisomy 21 may be the result of changes in extracellular matrix (ECM) and angiogenic regulators. Several genes regulating ECM composition have been linked to chromosome 21 including collagen 6a (*COL6A)*, superoxide dismutase 2 (*SOD2)*, ß2 integrin (*ITGB2*) and endostatin (*COL18A1*) (Benard et al., 2005; Zorick et al., 2001). In addition, it has recently been shown that

the DS critical region-1 protein (DSCR-1) is induced by VEGF and inhibits thrombin and VEGF signaling ultimately blocking endothelial proliferation and angiogenesis (Minami et al., 2004). Our results suggest that SIM2s contributes to the breast cancer-protective effects of DS by inhibiting breast cancer growth. In addition, we have found that SIM2s abrogates breast cancer cell invasion and MMP3 gene expression, implying that SIM2s may also function as a metastasis inhibitor.

V. FURTHER STUDIES

In this study, we found that SIM2s is decreased in breast cancer patients and differentially expressed in human breast epithelial and breast cancer cell lines with low levels observed in highly invasive MDA-MB-435 and MDA-MB-453 cell lines. However, the mechanism by which SIM2s is silenced in highly invasive breast cancer cells is unknown. Further studies are needed to decipher the mechanism of SIM2 suppression in breast cancer. We currently are taking a closer look at epigenetic mechanisms, such as hypermethylation, deacetylation, increased degradation, and mislocalization.

The effects of Sim2s overexpression in the mammary gland also need to be determined. Dominant gain of function experiments utilizing transgenic mice have proven useful in elucidating the roles of mammalian gene products in development and cancer. Thus, Developing MMTV-mSim2s transgenic mice will be beneficial to examine the role of Sim2s in mammary development and breast cancer. This can be accomplished by placing mSim2s under the control of the long terminal repeat of the mouse mammary tumor virus (MMTV), which directs expression to the mammary gland.

The role of Sim2s in mammary tumor initiation and progression can be determined by crossing Sim2s heterozygous and MMTV-mSim2s overexpressing mice with mammary cancer prone MMTV-Neu and MMTV-Myc transgenic animals. MMTV-Neu mice develop multi-focal metastatic mammary tumors after a relatively short latency period, whereas MMTV-Myc mice exhibit an extended latency period (Muller et al.,

1988; Stewart et al., 1984). MMTV-Myc tumors arise as stochastic foci within normal breast tissue; however, other factors are required to convert the tumor foci to a metastatic state. By using these two models of tumorigenesis, it will be possible dissect pathways involved in Sim2s-mediated tumor suppression.

VI. SUMMARY

SIM is a member of basic helix-loop-helix Per-ARNT-Sim (bHLH-PAS) family that actives midline gene transcription and represses lateral CNS gene transcription. In humans, SIM2 is involved in the etiology of the Down's phenotype. Interestingly, women with DS are 10-25 times less likely to develop breast cancer. Women with DS rarely become pregnant, increasing their risk for breast cancer; however, they experience earlier menopause, leading to a decrease in risk. Thus, hormonal factors alone are unlikely to explain the significantly decreased rate of breast cancer and protective factors are more likely to be genetically linked. These observations have led to the hypothesis that one or more tumor suppressor genes are present on chromosome 21 and prevent cancer development by gene dosage effect. Based on its role in the CNS midline development and its relation with DS, SIM2 was considered as a potential candidate.

In this study, we have shown that mouse Sim2s is less potent than full length Sim2 at repressing HIF1-α, and can activate expression of a gene controlled by the *Drosophila* toll gene CME via ARNT's transactivation domain. Sim2s is just as repressive as Sim2 on TCDD-mediated gene expression through a DRE and is able to increase expression of a CME-controlled gene. In contrast, Sim2s can activate expression from a CME apparently by acting as a docking protein for ARNT. This implies that the Pro/Ala-rich sequence present in Sim2, but not Sim2s, exerts a negative effect on CME-mediated gene expression.

We found that Sim2s is the predominant form of Sim expressed in the mouse

mammary gland, and its expression is tightly regulated during adolescent development and throughout pregnancy. These data support our hypothesis that Sim2s plays an important role during mammary gland development and differentiation.

We found that forced expression of SIM2s in MDA-MB-435 cells significantly inhibited proliferation, anchorage-independent growth and invasion potential. In addition, loss of Sim2s in the mammary gland resulted in dramatically increased ductal branching, accelerated lobuloalveolar-like precocious hyperplasia, and decreased cell apoptosis suggesting that SIM2s is a mammary tumor suppressor.

Sim2-/- mammary glands lose E-cadherin expression and contain hyperplastic nodules, suggesting that Sim2s plays a role in regulating E-cadherin/ß-catenin signaling. In SIM2s infected MDA-MB-435 cells, decrease in invasion appears to be due, in part, to SIM2s' ability to repress MMP3 expression through direct inhibition of AP-1 binding to the MMP3 promoter. Our results suggest that SIM2s contributes to the breast cancer-protective effects of DS by inhibiting breast cancer growth. In addition, we have found that SIM2s abrogates breast cancer cell invasion and MMP3 gene expression, implying that SIM2s may also function as a metastasis inhibitor.

REFERENCES

Adderley-Kelly, B., Williams-Stephens E., 2003. The relationship between obesity and breast cancer. ABNF J. 14(3), 61-5.

Alexander, C.M., Selvarajan, S., Mudgett, J., Werb, Z., 2001. Stromelysin-1 regulates adipogenesis during mammary gland involution. J. Cell Biol. 152, 693.

Amano, K., Sago, H., Uchikawa, C., Suzuki, T., Kotliarova, S.E., Nukina, N., Epstein, C.J., Yamakawa, K., 2004. Dosage-dependent over-expression of genes in the trisomic region of Ts1Cje mouse model for Down syndrome. Hum. Mol. Genet., 13, 1333-1340.

Andarawewa K.L., Boulay A., Masson R., Mathelin C., Stoll I., Tomasetto C., Chenard M.P., Gintz M., Bellocq J.P., Rio M.C., 2003. Dual stromelysin-3 function during natural mouse mammary tumor virus-ras tumor progression. Cancer Res. 63, 5844-9.

Andersen T.I., Gaustad A., Ottestad L., 1992. Genetic alterations of the tumour suppressor gene regions 3p, 11p, 13q, 17p, and 17q in human breast carcinomas. Genes Chromosomes Cancer 4, 113–121.

Antonarakis S.E., Lyle R., Dermitzakis E.T., Reymond A., Deutsch S., 2004. Chromosome 21 and Down Syndrome: from Genomics to Pathophysiology. Nat. Rev. Genet. 5, 725-738

Band V., Zajchowski D., Swisshelm K., Trask D., Kulesa V., Cohen C., Connolly J., Sager R., 1990. Tumor progression in four mammary epithelial cell lines derived from the same patient. Cancer Res. 50, 7351-7.

Band V., De Caprio J.A., Delmolino L., Kulesa V., Sager, R., 1991. Loss of p53 protein in human papillomavirus type 16 E6-immortalized human mammary epithelial cells. J. Virol. 65, 6671-6.

Barth A.I., Nathke I.S., Nelson W.J., 1997. Cadherins, catenins and APC protein: interplay between cytoskeletal complexes and signaling pathways. Curr. Opin. Cell. Biol. 9(5), 683-90.

Benard J., Beron-Gaillard N., Satge D., 2005. Down's syndrome protects against breast cancer: is a constitutional cell microenvironment the key? Int. J. Cancer 113, 168-70.

Beral V., 2002. Breast cancer and breastfeeding: collaborative reanalysis of individual

data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease, Lancet 360, 187-195.

Bernstein L., 2002. Epidemiology of endocrine-related risk factors for breast cancer. J. Mamm. Gland Biol. Neoplasia 7(1), 4-15

Boker L.K., Blumstein T., Sadetzki S., Luxenburg O., Litvak I., Akstein E., Modan B., 2001. Incidence of leukemia and other cancers in Down syndrome subjects in Israel. Int. J. Cancer 93, 741-4.

Boussadia O., Kutsch S., Hierholzer A., Delmas V., Kemler R., 2002. E-cadherin is a survival factor for the lactating mouse mammary gland. Mech. Dev. 115(1-2), 53-62.

Cano A., Perez-Moreno M.A., Rodrigo I., Locascio A., Blanco M.J., del Barrio M.G., Portillo F., Nieto M.A., 2000. The transcription factor snail controls epithelial–mesenchymal transitions by repressing E-cadherin expression. Nat. Cell. Biol. 2, 76-83

Chakraborti S., Mandal M., Das S., Mandal A., Chakraborti T., 2003. Regulation of matrix metalloproteinases: an overview. Mol. Cell. Biochem. 253, 269-85.

Chen Y., Chen C.F., Riley D.J., Allred D.C., Chen P.L., Von Hoff D., Osborne C.K., Lee W., 1995. Aberrant subcellular localization of BRCA1 in breast cancer. Science 270, 789-791.

Cheng S., Alfonso-Jaume M.A., Mertens P.R., Lovett D.H., 2002. Tumour metastasis suppressor, nm23-beta, inhibits gelatinase A transcription by interference with transactivator Y-box protein-1 (YB-1). Biochem. J. 366, 807-16.

Chrast R., Scott H.S., Chen H., Kudoh J., Rossier C., Minoshima S., Wang Y., Shimizu N., Antonarakis S.E., 1997. Cloning of two human homologs of the *Drosophila* single minded gene SIM1 on chromosome 6q and SIM2 on 21q within the Down syndrome chromosomal region. Genome Res. 7, 615-24.

Chrast R., Scott H.S., Madani R., Huber L., Wolfer D.P., Prinz M., Aguzzi A., Lipp H.P., Antonarakis S.E., 2000. Mice trisomic for a bacterial artificial chromosome with the single minded 2 gene (Sim2) show phenotypes similar to some of those present in the partial trisomy 16 mouse models of Down syndrome. Hum. Mol. Genet. 9, 1853-64.

Chun Y.S., Yeo E.J., Choi E., Teng C.M., Bae J.M., Kim M.S., Park J.W., 2001. Inhibitory effect of YC-1 on the hypoxic induction of erythropoietin and vascular endothelial growth factor in Hep3B cells. Biochem. Pharmacol. 61(8), 947-954

Coletta R.D., Christensen K., Reichenberger K.J., Lamb J., Micomonaco D., Huang L., Wolf D.M., Muller-Tidow C., Golub T.R., Kawakami K., Ford H.L., 2004. The Six1 homeoprotein stimulates tumorigenesis by reactivation of cyclin A1. Proc. Natl. Acad. Sci. USA 101, 6478-83.

Collaborative Group on Hormonal Factors in Breast Cancer. 2001. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. Lancet 358(9291).

Crews S., Franks R., Hu S., Matthews B., Nambu J., 1992. *Drosophila* single minded gene and the molecular genetics of CNS midline development. J. Exp. Zool. 261(3), 234-244

Crews S.T., 1998. Control of cell lineage-specific development and transcription by bHLH–PAS proteins, Genes Dev. 12, 607–620.

Crews S.T., Fan C.M., 1999. Remembrance of things PAS: regulation of development by bHLH-PAS proteins. Curr. Opin. Genet. Dev. 9(5), 580-587.

Darribere T., Skalski M., Cousin H.L., Gaultier A., Montmory C., Alfandari D., 2000. Integrins: regulators of embryogenesis. Biol. Cell. 92, 5–25.

Davisson M.T., Bechtel L.J., Akeson E.C., Fortna A., Slavov D., Gardiner K., 2001. Evolutionary breakpoints on human chromosome 21. Genomics 78, 99-106.

Dedieu S., Lefebvre P., 2006. Retinoids interfere with the AP1 signalling pathway in human breast cancer cells. Cell Signal. 18, 889-98.

Deyoung M.P., Scheurle D., Damania H., Zylberberg C., Narayanan R., 2002. Down's syndrome-associated single minded gene as a novel tumor marker. Anticancer Res. 22, 3149-57.

DeYoung M.P., Tress M., Narayanan R., 2003a. Down's syndrome-associated single minded 2 gene as a pancreatic cancer drug therapy target. Cancer Lett. 200; 25–31.

DeYoung M.P., Tress M., Narayanan R., 2003b. Identification of Down's syndrome critical locus gene SIM2-s as a drug therapy target for solid tumors. Proc. Natl. Acad. Sci. USA 100, 4760-5.

Dong J., Opresko L.K., Dempsey P.J., Lauffenburger D.A., Coffey R.J., Wiley H.S., 1999. Metalloprotease-mediated ligand release regulates autocrine signaling through the epidermal growth factor receptor. Proc. Natl. Acad. Sci. USA 96(11), 6235-40.

Duffy M.J., Maguire T.M., Hill A., McDermott E., O'Higgins A., 2000. Metalloproteinases: role in breast carcinogenesis, invasion and metastasis. Breast Cancer Res. 2, 252-257.

Dumitrescu R.G., Cotarla I., 2005. Understanding breast cancer risk -where do we stand in 2005? J. Cell. Mol. Med. 9(1), 208-221.

Eccles S.A., Box G.M., Court W.J., Bone E.A., Thomas W., Brown P.D., 1996. Control of lymphatic and hematogenous metastasis of a rat mammary carcinoma by the matrix metalloproteinase inhibitor batimistat (BB-94). Cancer Res. 56, 2815-2822.

Egeblad M., Werb Z., 2002. New functions for the matrix metalloproteases in cancer progression. Nature Reviews Cancer 2, 161-174.

Ema M., Ikegami S., Hosoya T., Mimura J., Ohtani H., Nakao K., Inokuchi K., Katsuki M., Fujii-Kuriyama Y., 1999. Mild impairment of learning and memory in mice overexpressing the mSim2 gene located on chromosome 16, an animal model of Down's syndrome. Hum. Mol. Genet. 8, 1409-15.

Ema M., Suzuki M., Morita M., Hirose K., Sogawa K., Matsuda Y., Gotoh O., Saijoh Y., Fujii H., Hamada H., Fujii-Kuriyama Y., 1996. cDNA cloning of a murine homologue of *Drosophila* single minded, its mRNA expression in mouse development, and chromosome localization. Biochem. Biophys. Res. Commun. 218, 588-94.

Enger S.M., Ross R.K., Paganini-Hill A., Bernstein L., 1998. Breastfeeding experience and breast cancer risk among postmenopausal women. Cancer Epidemiol. Biomark. Prev. 7, 365–369.

Estes P., Mosher J., Crews S.T., 2001. *Drosophila* single minded represses gene transcription by activating the expression of repressive factors. Dev. Biol. 232(1), 157-175

Fan C., Kuwana E., Bulfone A., Fletcher C.F., Copeland N.G., Jenkins N.A., Crews S., Martinez S., Puelles L., Rubenstein J.L., Tessier-Lavigne M., 1996. Expression patterns of two murine homologs of *Drosophila* single minded suggest possible roles in embryonic patterning and in the pathogenesis of Down Syndrome. Mol. Cell. Neuro. 7, 1–16.

Fata J.E., Werb Z., Bissell M.J., 2004. Regulation of mammary gland branching morphogenesis by the extracellular matrix and its remodeling enzymes. Breast Cancer Res. 6(1), 1–11.

Fenrick R., Wang L., Nip J., Amann J.M., Rooney R.J., Walker-Daniels J., Crawford

H.C., Hulboy D.L., Kinch M.S., Matrisian L.M., Hiebert S.W., 2000. TEL, a putative tumor suppressor, modulates cell growth and cell morphology of ras-transformed cells while repressing the transcription of stromelysin-1. Mol. Cell. Biol. 20, 5828-39.

Foster S.A., Wong D.J., Barrett M.T., Galloway D.A., 1998. Inactivation of p16 in human mammary epithelial cells by CpG island methylation. Mol. Cell. Biol. 18 1793-1801.

Gardiner K., 2004. Gene-dosage effects in Down syndrome and trisomic mouse models. Genome Biology 5, 244.

Grassi M., Moens G., Rousselle P., Thiery J.P., Jouanneau J.,1999. The SFL activity secreted by metastatic carcinoma cells is related to laminin 5 and mediates cell scattering in an integrin-independent manner. J. Cell Sci. 112, 2511-2520.

Greene M.E., Mundschau G., Wechsler J., McDevitt M., Gamis A., Karp J., Gurbuxani S., Arceci R., Crispino J.D., 2003. Mutations in GATA1 in both transient myeloproliferative disorder and acute megakaryoblastic leukemia of Down syndrome. Blood Cells Mol. Dis. 31, 351-6.

Groden J., Thliveris A., Samowitz W., Carlson M., Gelbert L., Albertsen H., Joslyn G., Stevens J., Spirio L., Robertson M., et al. 1991. Identification and characterization of the familial adenomatous polyposis coli gene. Cell 66, 589-600.

Ha H.Y., Moon H.B., Nam M.S., Lee J.W., Ryoo Z.Y., Lee T.H., Lee K.K., So B.J., Sato H., Seiki M., Yu D.Y., 2001. Overexpression of membrane-type matrix metalloproteinase-1 gene induces mammary gland abnormalities and adenocarcinoma in transgenic mice. Cancer Res. 61, 984-90.

Haber D.A., Buckler A.J., Glaser T., Call K.M., Pelletier J., Sohn R.L., Douglass E.C., Housman D.E., 1990. An internal deletion within an 11p13 zinc finger gene contributes to the development of Wilms' tumor. Cell 61, 1257-69.

Hajra K.M., Chen D.Y., Fearon E.R., 2002. The SLUG zinc-finger protein represses E-cadherin in breast cancer. Cancer Res. 62,1613-1618.

Hasle H., Clemmensen I.H., Mikkelsen M., 2000a, Incidence of cancer in individuals with Down syndrome. Tidsskr Nor Laegeforen 120, 2878-81.

Hasle H., Clemmensen I.H., Mikkelsen M., 2000b. Occurrence of cancer in individuals with Down syndrome. Ugeskr Laeger 162, 4535-9.

Hasle H., Clemmensen I.H., Mikkelsen M., 2000c. Risks of leukaemia and solid

tumours in individuals with Down's syndrome. Lancet 355, 165-9.

Hemminki A., Avizienyte E., Roth S., Loukola A., Aaltonen L.A., Jarvinen H., de la Chapelle A., 1998. A serine/threonine kinase gene defective in Peutz-Jeghers syndrome. Nature 391, 184-7.

Hennighausen L., Robinson G.W., 2001. Signaling pathways in mammary gland development. Dev. Cell 467-75

Hettinger A.M., Allen M.R., Zhang B.R., Goad D.W., Malayer J.R., Geisert R.D., 2001. Presence of the acute phase protein, bikunin, in the endometrium of gilts during estrous cycle and early pregnancy. Biol. Reprod. 65, 507-13.

Hill D.A., Gridley G., Cnattingius S., Mellemkjaer L., Linet M., Adami H.O., Olsen J.H., Nyren O., Fraumeni J.F. Jr., 2003. Mortality and cancer incidence among individuals with Down syndrome. Arch. Intern. Med. 163, 705-11.

Hoehne F.M., Taylor E., 2005. Trends in breast cancer at a county hospital, Am. Surg. 71(2), 159-63.

Hook E.B., Cross P.K., Schreinemachers D.M., 1983. Chromosomal abnormality rates at amniocentesis and in live-born infants. JAMA 249, 2034-2038.

Howard B.A., Gusterson B.A., 2000. Human breast development. J. Mam. Gland Biol. Neoplasia 5(2), 119-137

Hu X.C., Wong I.H., Chow L.W., 2003, Tumor-derived aberrant methylation in plasma of invasive ductal breast cancer patients: clinical implications. Oncol. Rep. 10, 1811-1815

Huang Z., Hankinson S.E., Colditz G.A., Stampfer M.J., Hunter D.J., Manson J.E., Hennekens C.H., Rosner B., Speizer F.E., Willett W.C., 1997. Dual effects of weight and weight gain on breast cancer risk. JAMA. 278(17), 1407-11.

Hulka B.S., Moorman P.G., 2001. Breast cancer: hormones and other risk factors, Maturitas 38, 103-113.

Ioachim E.E., Athanassiadou S.E., Kamina S., Carassavoglou K., Agnantis N.J., 1998. Matrix metalloproteinase expression in human breast cancer: an immunohistochemical study including correlation with cathepsin D, type IV collagen, laminin, fibronectin, EGFR, c-erbB-2 oncoprotein, p53, steroid receptors status and proliferative indices. Anticancer Res. 18, 1665-70.

Jain M.G., Ferrence R.G., Rehm J.T., Bondy S.J., Rohan T.E., Ashley M.J., Cohen J.E.,

Miller A.B., 2000. Alcohol and breast cancer mortality in a cohort study Breast Cancer Res. Treat. 64, 201–209.

Janda E., Lehmann K., Killisch I., Jechlinger M., Herzig M., Downward J., Beug H., Grunert S., 2002. Ras and TGF-β cooperatively regulate epithelial cell plasticity and metastasis: dissection of Ras signaling pathways. J. Cell. Biol. 156, 299-313.

Jemal A., Murray T., Ward E., Samuels A., Tiwari R., Ghafoor A., Feuer E., Thun M., 2005. Cancer statistics. CA Cancer. J. Clin. 55:10-30.

Jiang W.G., Davies G., Martin T.A., Parr C., Watkins G., Mason M.D., Mansel R.E., 2006. Expression of membrane type-1 matrix metalloproteinase, MT1-MMP in human breast cancer and its impact on invasiveness of breast cancer cells. Int. J. Mol. Med. 17(4), 583-90.

John E.M., Phipps A.I., Davis A., Koo J., 2005. Migration history, acculturation, and breast cancer risk in Hispanic women. Cancer Epidemiol. Biomarkers Prev. 14(12), 2905-13.

Kahlem P., Sultan M., Herwig R., Steinfath M., Balzereit D., Eppens B., Saran N.G., Pletcher M.T., South S.T., Stetten G., et al. 2004. Transcript level alterations reflect gene dosage effects across multiple tissues in a mouse model of Down syndrome. Genome Res. 14, 1258-1267.

Kewley R.J., Whitelawa M.L., Chapman-Smith A., 2004. The mammalian basic helix–loop–helix/PAS family of transcriptional regulators. Int. J. Bio. Cell Biol. 36(2), 189–204.

Klinowska T.C., Soriano J.V., Edwards G.M., Oliver J.M., Valentijn A.J., Montesano R., Streuli C.H., 1999. Laminin and beta1 integrins are crucial for normal mammary gland development in the mouse. Dev. Biol. 215, 13–32.

Korenberg J.R., Chen X.N., Schipper R., Sun Z., Gonsky R., Gerwehr S., Carpenter N., Daumer C., Dignan P., Disteche C., et al. 1994. Down syndrome phenotypes: the consequences of chromosomal imbalance. Proc. Natl. Acad. Sci. USA 91, 4997.

Koshikawa N., Giannelli G., Cirulli V., Miyazaki K., Quaranta V., 2000. Role of cell surface metalloprotease MT1-MMP in epithelial cell migration over laminin-5. J. Cell Biol. 2000, 148, 615–624.

La Rocca G., Pucci-Minafra I., Marrazzo A., Taormina P., Minafra S., 2004. Zymographic detection and clinical correlations of MMP-2 and MMP-9 in breast cancer sera. Br. J. Cancer 90, 1414-21.

Lacey J.V. Jr, Devesa S.S., Brinton L.A., 2002. Recent trends in breast cancer incidence and mortality. Environ. Mol. Mutagen. 39(2-3), 82-8.

Lee P.P., Hwang J.J., Murphy G., Ip M.M., 2000. Functional significance of MMP-9 in tumor necrosis factor-induced proliferation and branching morphogenesis of mammary epithelial cells. Endocrinology 141, 3764-73.

Levi E., Fridman R., Miao H.Q., Ma Y.S., Yayon A., Vlodavsky I., 1996. Matrix metalloproteinase 2 releases active soluble ectodomain of fibroblast growth factor receptor 1. Proc. Natl. Acad. Sci. USA 93(14), 7069-74.

Liaw D., Marsh D.J., Li J., Dahia P.L., Wang S.I., Zheng Z., Bose S., Call K.M., Tsou H.C., Peacocke M., Eng C., Parsons R., 1997. Germline mutations of the PTEN gene in Cowden disease, an inherited breast and thyroid cancer syndrome. Nat. Genet. 16, 64-7.

Liu B.Y., Kim Y., Leatherberry V., Cowin P., Alexander C.M., 2003. Mammary gland development requires syndecan-1 to create a b-catenin/TCF-responsive mammary epithelial subpopulation. Oncogene 22, 9243–9253.

Lux M.P., Fasching P.A., Beckmann M.W., 2006. Hereditary breast and ovarian cancer: review and future perspectives. J. Mol. Med. 84(1), 16-28.

Lyle R., Gehrig C., Neergaard-Henrichsen C., Deutsch S., Antonarakis S.E., 2004. Gene expression from the aneuploid chromosome in a trisomy mouse model of Down syndrome. Genome Res. 14, 1268-1274.

Magnusson C., Baron J., Persson I., Wolk A., Bergstrom R., Trichopoulos D., Adami H.O., 1998. Body size in different periods of life and breast cancer risk in post-menopausal women. Int. J. Cancer 76(1), 29-34.

Malkin D., 1994. Germline p53 mutations and heritable cancer. Annu. Rev. Genet. 28, 443-465

Mao R., Zielke C.L., Zielke H.R., Pevsner J., 2003. Global up-regulation of chromosome 21 gene expression in the developing Down syndrome brain. Genomics 81, 457-467.

Marti A., Lazar H., Ritter P., Jaggi R., 1999. Transcription factor activities and gene expression during mouse mammary gland involution. J. Mam. Gland Biol. Neoplasia 4(2), 145-152.

Martin M., 2006. Molecular biology of breast cancer. Clin. Transl. Oncol. 8(1), 7-14.

McKay L.M., Carpenter B., Roberts S.G., 1999. Evolutionary conserved mechanism of transcriptional repression by even-skipped. Nucleic Acids Res. 27(15), 3064-3070.

McTiernan A., Rajan K.B., Tworoger S.S., Irwin M., Bernstein L., Baumgartner R., Gilliland F., Stanczyk F.Z., Yasui Y., Ballard-Barbash R., 2003. Adiposity and sex hormones in postmenopausal breast cancer survivors. J. Clin. Oncol. 21, 1961-1966.

Meng X., Shi J., Peng B., Zou X., Zhang C., 2006. Effect of mouse Sim2 gene on the cell cycle of PC12 cells. Cell Biol. Int. 30(4), 349-53.

Metz R.P., Kwak H.I., Gustafson T., Laffin B., Porter W.W., 2006. Differential transcriptional regulation by mouse single minded 2S. J. Biol. Chem. 281(16):10839-48.

Michaud J., Fan C., 1997. Single minded—two genes, three chromosomes. Genom. Res. 7, 569–571

Mikkelsen M., 1977. Down's syndrome cytogenetic epidemiology. Hereditas 86, 45-59.

Minami T., Horiuchi K., Miura M., Abid M.R., Takabe W., Noguchi N., Kohro T., Ge X., Aburatani H., Hamakubo T., Kodama T., Aird W.C., 2004. Vascular endothelial growth factor- and thrombin-induced termination factor, Down syndrome critical region-1, attenuates endothelial cell proliferation and angiogenesis. J. Biol. Chem. 279, 50537-54.

Minchenko O, Opentanova I, Caro J., 2003. Hypoxic regulation of the 6-phosphofructo-2-kinase/fructose-2, 6-bisphosphatase gene family (PFKFB-1-4) expression in vivo. FEBS Lett. 554, 264–270

Miyoshi A., Kitajima Y., Kido S., Shimonishi T., Matsuyama S., Kitahara K., Miyazaki K., 2005. Snail accelerates cancer invasion by upregulating MMP expression and is associated with poor prognosis of hepatocellular carcinoma. Br. J. Cancer 92, 252-8.

Miyoshi A., Kitajima Y., Sumi K., Sato K., Hagiwara A., Koga Y., Miyazaki K., 2004. Snail and SIP1 increase cancer invasion by upregulating MMP family in hepatocellular carcinoma cells. Br. J. Cancer 90, 1265-73.

Moffett P.P., 2000. Different transcriptional properties of mSim-1 and mSim-2. J. FEBS Lett. 466(1), 80-86.

Moffett P., Dayo M., Reece M., McCormick M.K., Pelletier J., 1996. Characterization of msim, a murine homologue of the *Drosophila* sim transcription factor. Genomics 35, 144-55.

Moffett P., Reece M., Pelletier J., 1997. The murine Sim-2 gene product inhibits transcription by active repression and functional interference. Mol. Cell. Biol. 17, 4933-47.

Morrow D., Qin C., Smith R. 3rd, Safe S., 2004. Aryl hydrocarbon receptor-mediated inhibition of LNCaP prostate cancer cell growth and hormone-induced transactivation. J. Steroid Biochem. Mol. Biol. 88(1), 27-36.

Moynahan M.E., 2002. The cancer connection: BRCA1 and BRCA2 tumor suppression in mice and humans. Oncogene 21, 8994-9007.

Muenke M., Bone L.J., Mitchell H.F., Hart I., Walton K., Hall-Johnson K., Ippel E.F., Dietz-Band J., Kvaloy K., Fan, C-M., Tessier-Lavigne M., Patterson D., 1995. Physical mapping of the holoprosencephaly critical region in 21 q22.3, exclusion of SIM2 as a candidate gene for holoprosencephaly, and mapping of SIM2 to a region of chromosome 21 important for Down syndrome. Am. J. Hum. Genet. 57, 1074–1079.

Muller W.J., Sinn E., Pattengale P.K., Wallace R., Leder P., 1988. Single-step induction of mammary adenocarcinoma in transgenic mice bearing the activated c-neu oncogene. Cell 54, 105-115.

Nambu J.R., Lewis J.O., Wharton K.J., Crews S.T., 1991. The *Drosophila* single minded gene encodes a helix-loop-helix protein that acts as a master regulator of CNS midline development. Cell 67, 1157-1167.

Nambu J.R., Lewis J.O., Crews S.T., 1993. The development and function of the *Drosophila* CNS midline cells. Comp. Biochem. Physiol. Comp. Physiol. 104(3), 399-409.

Nicholson R.C., Mader S., Nagpal S., Leid M., Rochette-Egly C., Chambon P., 1990. Negative regulation of the rat stromelysin gene promoter by retinoic acid is mediated by an AP1 binding site. Embo. J. 9, 4443-54.

Osborne C., Wilson P., Tripathy D., 2004. Oncogenes and tumor suppressor genes in breast cancer: potential diagnostic and therapeutic applications. Oncologist 9(4), 361-77.

Osborne M.P., 2000. Breast anatomy and development. In: Harris JR, Lippman ME, Morrow M, Osborne CK, eds. Diseases of the Breast, 2nd edition. Philadelphia: Lippincott, Williams and Wilkins 1-14.

Parkin D.M., Bray F., Ferlay J., Pisani P., 2005. Global cancer statistics, 2002. CA Cancer J. Clin. 55(2), 74-108.

Parkin D.M., Fernandez L.M., 2006. Use of statistics to assess the global burden of breast cancer. Breast J. 12 Suppl 1, S70-80.

Parmar H., Cunha G.R., 2004. Epithelial–stromal interactions in the mouse and human mammary gland in vivo. Endocrine-Related Cancer 11(3) 437-458.

Patterson D., Costa A.C.S., 2005. Down syndrome and genetics - a case of linked histories. Nat. Rev. Gen. 6, 137-147.

Penrose L.S., 1933. The relative effects of paternal and maternal age in mongolism. J. Genet. 27, 219 only.

Pflug B.R., Reiter R.E., Nelson J.B., 1999. Caveolin expression is decreased following androgen deprivation in human prostate cancer cell lines. Prostate 40, 269-73.

Pielage J., Steffes G., Lau D.C., Parente B.A., Crews S.T., Strauss R., Klambt C., 2002. Novel behavioral and developmental defects associated with *Drosophila* single minded, Dev. Biol. 249, 283–299.

Poschl G., Seitz H.K., 2004. Alcohol and cancer. Alcohol Alcohol. 39, 155-165.

Pueschel S.M., Rynders J.E., 1982. Down Syndrome. Advances in Biomedicine and the Behavioral Sciences. Ware Press, Cambridge, UK 133 only.

Quarrie L.H., Addey C.V.P., Wilde C.J., 1996. Programmed cell death during mammary tissue involution induced by weaning, litter removal, and milk stasis. J. Cell. Physiol. 168, 559–569.

Rachidi M., Lopes C., Charron G., Delezoide A., Paly E., Bloch B., Delabar J., 2005. Spatial and temporal localization during embryonic and fetal human development of the transcription factor SIM2 in brain regions altered in Down syndrome. Int. J. Devl. Neuroscience. 23, 475–484.

Radice G.L., Ferreira-Cornwell M.C., Robinson S.D., Rayburn H., Chodosh L.A., Takeichi M., Hynes R.O., 1997. Precocious mammary gland development in P-cadherin–deficient mice. J. Cell Biol. 139, 1025-32.

Reeves R.H., Irving N.G., Moran T.H., Wohn A., Kitt C., Sisodia S.S., Schmidt C., Bronson R.T., Davisson M.T., 1995. A mouse model for Down syndrome exhibits learning and behaviour deficits. Nature Genet. 11, 177.

Roizen N.J., Patterson D., 2003. Down's syndrome. Lancet. 361, 1281–1289.

Ross R.K., Paganini-Hill A., Wan P.C., Pike M.C., 2000. Effect of hormone replacement

therapy on breast cancer risk: estrogen versus estrogen plus progestin, J. Natl. Cancer Inst. 92, 328-332.

Rudolph-Owen L.A., Cannon P., Lynn M., 1998. The matrix metalloproteinase matrilysin influences early stage mammary tumorigenesis. Cancer Res. 58, 5500-5506.

Russo I., Russo I.H., 2004. Development of the human breast. Maturitas 49(1), 2-15.

Russo J., Russo I.H., In: Neville M.C., Daniel C.W., editors. 1987. The Mammary Gland. Plenum Publishing Corporation, New York. 67–93.

Sago H., Carlson E.J., Smith D.J., Kilbridge J., Rubin E.M., Mobley W.C., Epstein C.J., Huang T.T., 1998. Ts1Cje, a artial trisomy 16 mouse model for Down syndrome, exhibits learning and behavioral abnormalities. Proc. Natl. Acad. Sci. USA 95, 6256-6261.

Sakakura T., 1991.New aspects of stroma–parenchyma relations in mammary gland differentiation. Int. Rev. Cytol. 125, 165–202.

Satge D., Sasco A.J., Cure H., Leduc B., Sommelet D., Vekemans M.J., 1997. An excess of testicular germ cell tumors in Down's syndrome: three case reports and a review of the literature. Cancer 80, 929-35.

Scully K.M., Jacobson E.M., Jepsen K., Lunyak V., Viadiu H., Carrie C., Ross D.W., Hooshmand F., Aggarwal A.K., Rosenfeld M.G., 2000. Allosteric effects of Pit-1 DNA sites on long-term repression in cell type specification. Science 290(5494), 1127-1131.

Sellers T.A., Vierkant R.A., Cerhan J.R., Gapstur S.M., Vachon C.M., Olson J.E., Pankratz V.S., Kushi L.H., Folsom A.R., 2002. Interaction of dietary folate intake, alcohol, and risk of hormone receptor-defined breast cancer in a prospective study of postmenopausal women, Cancer Epidemiol. Biomarkers Prev. 11, 1104-1107.

Semenza G.L., 1998. Hypoxia-inducible factor 1: master regulator of O2 homeostasis. Curr. Opin. Genet. Dev. 8, 588-594.

Shiloh Y., 2003. ATM and related protein kinases: safeguarding genome integrity. Nat. Rev. Cancer. 3, 155–168.

Silberstein G.B., 2001. Postnatal mammary gland morphogenesis. Microsc. Res. Tech., 52(2), 155-62.

Silberstein G.B., Strickland P., Coleman S., Daniel C.W., 1990. Epithelium-dependent

extracellular matrix synthesis in transforming growth factor-beta 1-growth-inhibited mouse mammary gland. J. Cell Biol., 110, 2209–2219.

Singletary K.W., Gapstur S.M., 2001. Alcohol and breast cancer: review of epidemiologic and experimental evidence and potential mechanisms, JAMA 286, 2143-2151.

Sledge G.W. Jr, Qulali M., Goulet R., Bone E.A., Fife R., 1995. Effect of matrix metalloproteinase inhibitor batimastat on breast cancer regrowth and metastasis in athymic mice. J. Natl. Cancer Inst. 87, 1546-1550.

Stampfer M.R., Yaswen P., Alhadeff M., Hosoda J., 1993. TGF beta induction of extracellular matrix associated proteins in normal and transformed human mammary epithelial cells in culture is independent of growth effects. J. Cell. Physiol. 155, 210–221.

Sternlicht M.D., Lochter A., Sympson C.J., Huey B., Rougier J.P., Gray J.W., Pinkel D., Bissell M.J., Werb Z., 1999. The stromal proteinase MMP3/stromelysin-1 promotes mammary carcinogenesis. Cell 98, 137-146.

Stetler-Stevenenson W., 1999. Matrix metalloproteinases in angiogenesis: a moving target for therapeutic intervention. J. Clin. Invest. 103, 1237-1241.

Stewart T.A., Pattengale P.K., Leder P., 1984. Spontaneous mammary adenocarcinomas in transgenic mice that carry and express MTV/myc fusion genes. Cell 38, 627-637.

Storey A., Thomas M., Kalita A., Harwood C., Gardiol D., Mantovani F., Breuer J., Leigh I.M., Matlashewski G., Banks L., 1998. Role of a p53 polymorphism in the development of human papillomavirus-associated cancer. Nature 393, 229-234.

Stroka D.M., Burkhardt T., Desbaillets I., Wenger R.H., Neil D.A., Bauer C., Gassmann M., Candinas D., 2001. HIF-1 is expressed in normoxic tissue and displays an organ-specific regulation under systemic hypoxia. FASEB J. 15; 2445–2453.

Sympson C.J., Bissell M.J., Werb Z., 1995. Mammary gland tumor formation in transgenic mice overexpressing stromelysin-1. Semin. Cancer Biol. 6, 159-63.

Sympson C.J., Talhouk R.S., Alexander C.M., Chin J.R., Clift S.M., Bissell M.J., Werb Z., 1994. Targeted expression of stromelysin-1 in mammary gland provides evidence for a role of proteinases in branching morphogenesis and the requirement for an intact basement membrane for tissue-specific gene expression. J Cell Biol. 125, 681-93.

Thiery J.P., 2002. Epithelial–mesenchymal transitions in tumour progression. Nat. Rev.

Cancer 2, 442-454.

Titus-Ernstoff L., Longnecker M.P., Newcomb P.A., Dain B., Greenberg E.R., Mittendorf R., Stampfer M., Willett W., 1998. Menstrual factors in relation to breast cancer risk, Cancer Epidemiol. Biomarkers Prev. 7, 783-789.

Vincent-Salomon A., Thiery J.P., 2003. Host microenvironment in breast cancer development: Epithelial–mesenchymal transition in breast cancer development. Breast Cancer Res. 5, 101-106.

Vogel W.F., Aszodi A., Alves F., Pawson T., 2001. Discoidin domain receptor 1 tyrosine kinase has an essential role in mammary gland development. Mol. Cell. Biol. 21, 2906–2917.

Vokes E.E., Golomb H.M., 2003. Oncologic Therapies 2nd Ed. Springer, New York.

Vu T., Werb Z., 2000. Matrix metalloproteinases: effectors of development and normal physiology. Genes Dev. 14, 2123–2133.

Wechsler J., Greene M., McDevitt M.A., Anastasi J., Karp J.E., Le Beau M.M., Crispino J.D., 2002. Acquired mutations in GATA1 in the megakaryoblastic leukemia of Down syndrome. Nat. Genet. 32, 148-52.

Whitelaw M.L., Gustafsson J.A., Poellinger L., 1994. Identification of transactivation and repression functions of the dioxin receptor and its basic helix-loop-helix/PAS partner factor Arnt: inducible versus constitutive modes of regulation. Mol. Cell Biol. 14(12), 8343-8355.

Wiesen J.F., Werb Z., 1996. The role of stromelysin-1 in stromal-epithelial interactions and cancer. Enzyme Protein 49, 174-181.

Williams T.M., Lisanti M.P., 2005. Caveolin-1 in oncogenic transformation, cancer, and metastasis. Am. J. Physiol. Cell Physiol. 288, C494-506.

Williams T.M., Hassan G.S., Li J., Cohen A.W., Medina F., Frank P.G., Pestell R.G., Di Vizio D., Loda M., Lisanti M.P., 2005. Caveolin-1 promotes tumor progression in an autochthonous mouse model of prostate cancer: genetic ablation of Cav-1 delays advanced prostate tumor development in tramp mice. J. Biol. Chem. 280, 25134-45.

Williams T.M., Medina F., Badano I., Hazan R.B., Hutchinson J., Muller W.J., Chopra N.G., Scherer P.E., Pestell R.G., Lisanti M.P., 2004. Caveolin-1 gene disruption promotes mammary tumorigenesis and dramatically enhances lung metastasis in vivo. Role of Cav-1 in cell invasiveness and matrix metalloproteinase (MMP-2/9) secretion. J. Biol. Chem. 279, 51630-46.

Woods S.L., Whitelaw M.L., 2002. Differential activities of murine single minded 1 (SIM1) and SIM2 on a hypoxic response element. J. Biol. Chem. 277(12); 10236-10243.

Zetter B., 1998. Angiogenesis and tumor metastasis. Annu. Rev. Med. 49, 407-424.

Ziegler R.G., Hoover R.N., Pike M.C., Hildesheim A., Nomura A.M., West D.W., Wu-Williams A.H., Kolonel L.N., Horn-Ross P.L., Rosenthal J.F., Hyer M.B., 1993. Migration patterns and breast cancer risk in Asian-American women. J. Natl. Cancer Inst. 85(22), 1819-27.

Zorick T.S., Mustacchi Z., Bando S.Y., Zatz M., Moreira-Filho C.A., Olsen B., 2001. Passos-Bueno MR. High serum endostatin levels in Down syndrome: implications for improved treatment and prevention of solid tumours. Eur. J. Hum. Genet. 9, 811-4.

VITA

NAME:	Hyeong-il Kwak
DATE OF BIRTH:	12/25/1970
EDUCATION:	Ph.D. Toxicology,Texas A&M University, College Station, Texas. December 2006. M.S. Toxicology, Seoul National University, Seoul, Korea. February 1999. B.S. Veterinary Medicine, Chonnam National University, Kwangju, Korea, February 1996.
PERMANENT ADDRESS:	368 Chopyungri Eusinmyun Chindogun Chonnam, Korea.