COMMITTEE CERTIFICATION OF APPROVED VERSION

The committee for Kendra Leann Stisser certifies that this is the approved version of the following dissertation:

MOLECULAR PROFILING IN EARLY STAGE SQUAMOUS CELL CARCINOMAS OF THE UTERINE CERVIX

Committee:

Concepcion Diaz-Arrastia M.D., Supervisor

Elizabeth R. Unger, Ph.D., M.D.

Golda A. Leonard, Ph.D.

Bruce A. Luxon, Ph.D.

Kathleen O'Connor, Ph.D.

Dean, Graduate School

MOLECULAR PROFILING IN EARLY STAGE IB1 SQUAMOUS CELL CARCINOMAS OF THE UTERINE CERVIX

by
Kendra Leann Stisser, B.S.

Dissertation
Presented to the Faculty of The University of Texas Graduate School of Biomedical Sciences at Galveston
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

Approved by the Supervisory Committee

Concepcion Diaz-Arrastia, M.D.
Bruce A. Luxon, Ph.D.
Elizabeth R. Unger, Ph.D., M.D.
Golda A. Leonard, Ph.D.
Kathleen O'Connor, Ph.D.

May 2007
Galveston, Texas

Key words: Cervical Cancer, Gene Expression

Hidden Text: Ultra-staging, Tumor Heterogeneity

This is dedicated to my loving husband, Steve Stisser. Through sickness, health, the good and the bad, we finally made it.

ACKNOWLEDGEMENTS

The entirety of this research was funded by National Institute of Health's, National Cancer Institute (NCI grant# CON14027 SK22CA108482-02 K22). There are several people who deserve credit for helping me make it through this gratifying, yet arduous process. Please forgive me if I inadvertently miss some of you.

First and foremost, I want to thank my supervising professor, Dr. Arrastia for asking me to join her lab. I am honored to work with someone who is such an inspiration. She is an incredibly talented and dedicated physician AND scientist as well as mother, wife, and mentor to many. Indeed, my graduate experience working on a translational project has been unique in that I often went into the operating room, was in the clinic screening patients, as well as in the laboratory doing basic science research. What an incredible experience! I am inspired and awed at her devotion to helping women with cervical cancer in the clinic and in the laboratory. I am thankful for her insight and wisdom and wish her much success in her endeavors. I am grateful to call her my friend.

I would like to thank the members of my committee. Dr. Unger, from the CDC, for graciously devoted your time and expertise to the project. Dr. Luxon, for your statistical prowess and ability to educate me on the subject matter along the way. Dr. O'Connor for your basic science expertise and willingness to help me develop novel thoughts and ideas for my project. Dr. Leonard for your guidance, understanding, and encouragement through the whole process. Dr. Leonard has been a great friend and advisor.

My utmost gratitude goes to Mala Sinha from the Bioinformatics group at UTMB for her expertise and many long hours working on the statistical analysis for this project. Many, many thanks Mala. I would also like to sincerely thank Dr. Claudia Castro for providing the pathology expertise for reviewing countless slides and helping cutting the red tape to obtain the paraffin blocks. Thanks for making this a truly translational project. Thanks to the UTMB Genomics Core Facility (Dr. Thomas Wood, Director) for

processing the GeneChips and the UTMB Bioinformatics Program headed by Dr. Bruce A. Luxon for the microarray data analysis.

A special thanks goes out to Dr. Tung Van Dinh for nick-naming and always calling me “doctor”. I appreciated the vote of confidence. To Dr. Hannigan for always following my graduate career through newspaper publications and magazines. You always knew what I was doing. Thanks for your interest.

I thank Dr. Coppenhaver, Dr. Saavedra, Jodean Schmiederer, Mike Cromie, Laura Teed, Angie Tropea, Ann Anderson, and Dr. Cooper, for believing in my leadership abilities. You all made me believe I could do great things, and together, I think we did.

Dr. Niesel for believing I could be a successful graduate student even before I ever thought it possible. Thanks to you and Dr. Herzog, I had the ability to pursue a MBA. Thanks for always putting students first.

I would like to acknowledge and thank Dr. Bob Leonard, Dr. Golda Leonard, and Dr. Budelmann for welcoming me in to the Cell Biology Program. I was a late bloomer and you all took very good care of me and I appreciate deeply your friendship and advice.

I would like to offer a special thanks to my dearest friends and fellow students. To Zhenia Tarasenko for being my “Pea” and my closest friend. I will always cherish our crazy adventures. What an awesome person you truly are. To the girls night crew; Tara Ruttley, Liz Warren, Zobeida Cruz-Monserrate, Esther Rodriquez, and Tami Maynard. I treasure and appreciate how much you all are really there for me. Tara, thanks for your down-to-earth advice. Liz, thanks for your sense of humor. Zobeida, thanks for your enthusiasm and grace. Esther, thanks for your kindness, and Tami, thanks for your loving encouragement. You all made sure that the highs were high and the lows were not too low.

I would like to thank my family for not asking me when I was going to graduate too much. Mom, thank you for always believing in me. No matter how tough things were financially or otherwise as a kid, you never let me fall through the cracks. Out of all your children, I just knew I was your favorite (and I think all of us thought that way). Thanks for letting me believe that. Dad, I know you are in heaven watching over me and

hopefully smiling down with pride. I want to thank you for being my inspiration and example. You may have not provided me with your genes, but I am no less your daughter. I know you are there for me still and will walk across that stage to receive my degree with me. I love you and miss you. To my sisters, Kari and Kym: You both have always put me up on a pedestal (undeservingly I might add). I love you both dearly and thank God that he knew what he was doing when he gave you to me. You are both heaven sent gifts and I am blessed to be your sister. To my father, Chuck: I never knew you were so proud of me. I am thankful for the time we have shared together and look forward to sharing many more precious moments (with mom, haha). To Cathy and Terry Stisser: Thank you both for truly loving me more than just a daughter-in-law and welcoming me with open arms into the family. California, here we come!

Finally, I would like to thank Steve, my husband, for sticking it out with me. We have been through more than most couples could endure and yet we made it through. I thank you for the little everyday things that you do to make my life easy and possible. You are my best friend, confidant, and love. I am truly thankful for you babe.

Thank you God for giving me the ability and opportunity and loving me enough to do something as important as cancer research. I pray this research makes a difference.

MOLECULAR PROFILING IN EARLY STAGE IB1 SQUAMOUS CELL CARCINOMAS OF THE UTERINE CERVIX

Publication No._____

Kendra Leann Stisser, PhD
The University of Texas Graduate School of Biomedical Sciences at Galveston, 2007

Supervisor: Concepcion Diaz-Arrastia

Treatment of tumors of the cervix is based on clinical staging of disease. The International Federation of Gynaecology and Obstetrics system (FIGO staging) has established a system whereby cervical cancers are staged on the basis of anatomical extent of the tumor. Microinvasive tumors most often require non-radical surgical resection and pose little risk of recurrence and metastasis. On the other extreme, cancers that are inoperable, due to spread beyond the cervix (Stage IIB to IVB) require concomitant chemotherapy and radiation therapy. Currently, of all women undergoing radical hysterectomies and pelvic lymphadenectomy for FIGO stage IB1 carcinomas, forty to fifty percent are deemed at risk for recurrence and require radiation therapy and chemotherapy [1]. Much research has focused on linking histologic phenotype to outcome. However, tumor heterogeneity has made predictions of recurrence risk difficult. Ten percent of patients deemed having a low risk of recurrence by histological criteria will present with recurrent tumors within 5 years. The problem is that women may not receive optimal radiation therapy. If under treated, the cancer comes back. If over-treated there is a risk of toxicity associated with the adjuvant therapy.

The overall purpose of this research was to clarify the molecular mechanisms of progressive IB1 cervical disease in order to stratify risk to better match multi-modality therapy to minimize morbidity, mortality, and costs. The goal is to improve characterization of the tumor to individualize therapy.

Molecular profiling was used to identify a group of 98 genes that differentiate early cancers with a low risk of recurrence from those with a high risk of recurrence. Tumor heterogeneity is an important consideration when using molecular profiling to characterize outcome. There are location-specific genes that can be used for elucidating the mechanism of disease and individualizing patient care.

GeneChip technology is a powerful tool for teasing out the orchestration of molecular processes leading to progressive disease. As the molecular signature technology develops, we can move beyond improved diagnosis to improved therapy. This preliminary analysis may help to identify improved markers for predicting outcome so we can offer patients more precisely tailored treatment regimens.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

LIST OF ABBREVIATIONS

CIN	Cervical Intraepithelial Neoplasia
DNA	Deoxyribonucleic Acid
EST	Expressed Sequence Tag
FFPE	Formalin-fixed, paraffin-embedded *[tissues]*
FIGO	The International Federation of Gynecology and Obstetrics
GOI	Gene of Interest
H & E	Haemotoxylin and Eosin (slide stain)
HGSIL	High grade squamous intraepithelial lesion
HIV	Human Immunodeficiency Virus
HPV	Human Papillomavirus
HSV	Herpes Simplex Virus
LGSIL	Low grade squamous intraepithelial lesion
MM	Master Mix
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
qRT PCR	Quantitative Real Time Polymerase Chain Reaction
RNA	Ribonucleic Acid
RT	Reverse Transcriptase
SCCA	Squamous Cell Carcinoma
SCA	Squamous Cell Carcinoma Antigen
SCJ	Squamo-Columnar Junction
STI	Sexually Transmitted Infection
TNA	Total Nucleic Acid (DNA and RNA)
UTMB	University of Texas Medical Branch
WHO	World Health Organization

CHAPTER 1: NATURAL HISTORY OF CERVICAL CANCER

CERVICAL CANCER

Cancer of the uterine cervix is the second most prevalent cancer of women in the world. Firm epidemiological data demonstrate that Human Papillomavirus (HPV) is the cause of cervical cancer [2-5]. The virus is commonly transmitted sexually. While certain types of HPV are associated with benign genital warts, males are usually asymptomatic carriers of the HPV types associated with cervical cancer. Infection with HPV typically occurs in the teenage years, but the manifestation of disease is often not until many years later and most women who are infected never develop disease. The mean age of cervical dysplasia (the precancerous form of disease) development is in the reproductive years around the age of thirty four [6]. The gradual development of disease beginning with pre-cancerous changes, allow physicians the opportunity to catch and treat the disease before it ever becomes cancer. Annual routine screening, using the papanicolaou (pap) smear, tests for abnormal cells exfoliated from the cervix. While precancerous lesions are asymptomatic, symptoms of cervical cancer include abnormal vaginal bleeding, abnormally long or heavy menstrual periods, and bleeding during sexual intercourse. Treatment of cervical neoplasia (pre-invasive and invasive disease) is dependent on extent of disease and potential for spread, and ranges from regular follow-up exams to physical ablation of the cervical epithelium to hysterectomy, chemotherapy, and radiation therapy.

EPIDEMIOLOGY

Cancer of the uterine cervix is second only to breast cancer with an annual incidence of 493,000 cases and 274,000 deaths worldwide [7]. In developing countries of south central Africa, southwest Asia, and Latin America, cervical cancer remains the leading cause of death in women and accounts for 83% of all cervical cancer deaths

globally. In developed countries such as the United States, the incidence and mortality rates are significantly less because of early detection programs. In 2007, the National Cancer Institute (NCI) estimated 11,150 new cases and 3, 670 deaths from cervical cancer in the United States [8].

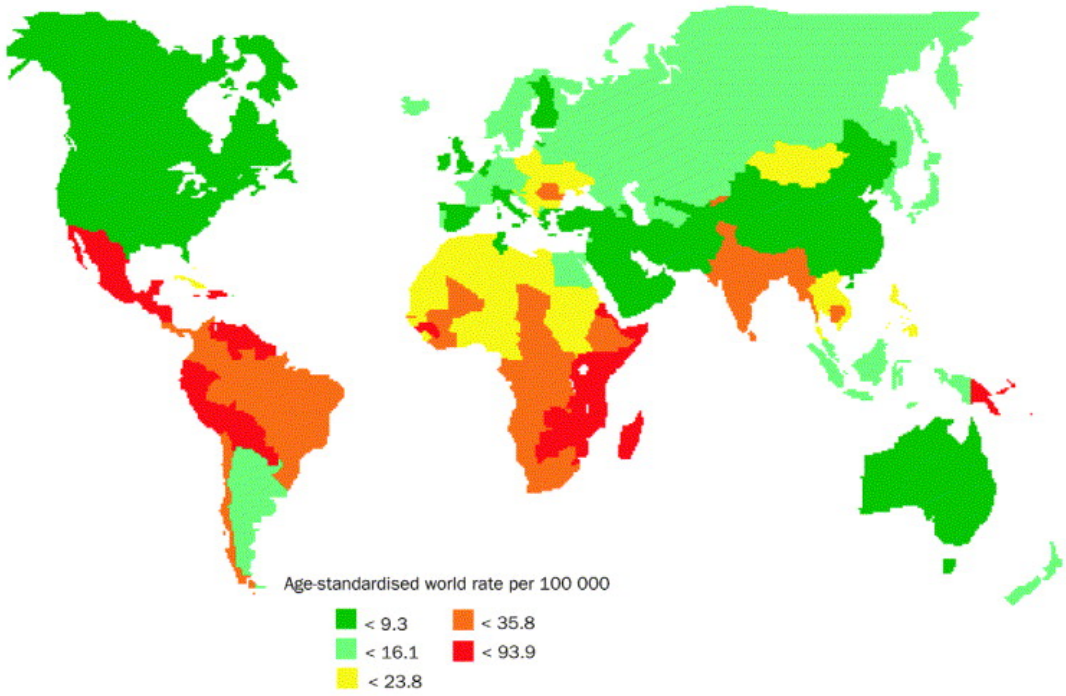

Fig 1.1 Global incidence of cervical cancer. Rates are age-standardized per 100,000 people. The rates range from <9.3 women per 100,000 (in dark green) to the highest incidence of <93.9 women per 100,000 (in red) [A].

EARLY DETECTION

Cervical cancer is one of a very few cancers that is nearly 100% preventable. Early detection methods, such as pap smear screening and visual cervical inspection have decreased the incidence of cervical cancer significantly [9-12]. Pap smear screening allows the detection of dysplastic cells before the cells invade and progress to cancer. Most

[A] Source: GLOBOCAN, International Agency for Research on Cancer. http://www-dep.iarc.fr/

cases of cervical cancer occur in women who have never had or have not had a pap smear within the past five years [13,14]. Women who have access to health care and get annual pap smear screenings are more likely to receive adequate preventative care [15]. The implementation of annual pap Smear screening programs in developed countries has had the greatest impact on diminishing cancer incidence [16-18]. Before preventative screening in the United States, approximately 75-80% of cervical neoplasia was invasive [19]. Currently, with appropriate screening and care, this trend has reversed. In the United States, there has been a steady decrease in the incidence of cervical cancer.

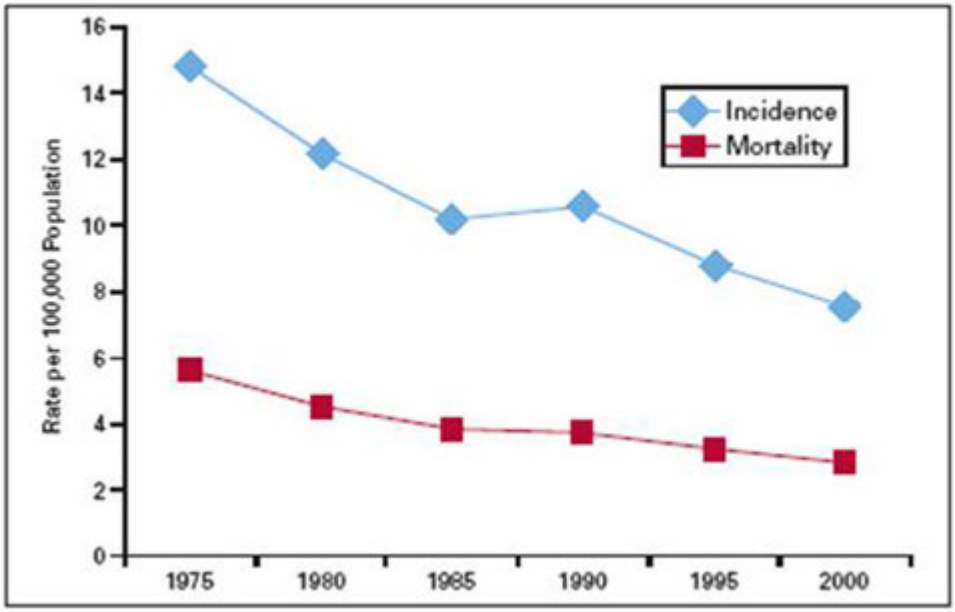

Fig 1.2 2004 National Cancer Institute data on Incidence and Mortality in the United States shows a steady decrease in incidence of cervical cancer and deaths associate with cervical cancer[B].

While the incidence and mortality has declined steadily in the United States due to early detection, in developing countries of Latin America screening programs have failed. Access to healthcare and poverty are a problem. In Mexico, rates of cervical

[B] Graphic representation www.merckmedicus.com/.../clinicalmanifest.jsp. Source: Ries LAG, Eisner MP, Kosary CL, Hankey BF, Miller BA, Clegg L, Mariotto A, Fay MP, Feuer EJ, Edwards BK (eds). SEER Cancer Statistics Review, 1975-2000, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2000/, 2003. Accessed Feb 2007.

cancer correlate with poverty levels [20]. The Pan Health Organization (PAHO) published a report implicating that the failure was often not due to lack of screening, but as a result of treatment failure and lack of an adequate health care infrastructure[21]. Thus, not only is early detection imperative, but also adequate follow-up and treatment is required to reduce cervical cancer related deaths.

HUMAN PAPILLOMAVIRUS (HPV)

HPV is a non-enveloped, double stranded DNA virus belonging to the Papillomaviridae family. It has a circular genome and replicates in the host nucleus. HPV is associated with many benign and malignant diseases including genital warts (*condyloma acuminate*), head and neck cancers, oral cancers, and anogenital carcinomas including cervical cancer. There are over 100 different types of HPV's, but around 40 are found in the genital tract and are associated with cervical disease.

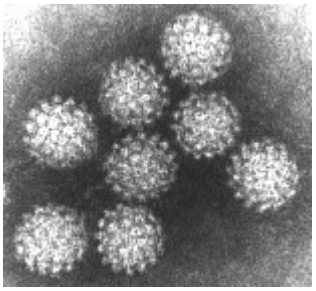

Fig 1.3 HPV virus particles[C].

Because the antigenic portion of the capsid antigens are similar to all HPV types, and may be present in low levels, immunochemical methods lack the sensitivity and type-specificity for HPV identification. HPV detection and typing requires DNA-based methods dependant on DNA sequence analysis. It is one of a few carcinogenic sexually transmitted viruses. Many animals have papillomaviruses, but papillomaviruses are species-specific. HPV infects basal cell keratinocytes at the squamo-columnar junction (SCJ). The virus initially replicates as a plasmid, but virus genotypes with a high

[C] Image provided by Linda M. Stannard, University of Cape Town http://web.uct.ac.za/depts/mmi/stannard/emimages.html. Copyright is retained by Linda M. Stannard.

potential for malignant transformation (high-risk HPV) may integrate into the genome[22]. The completion of the virus life cycle coincides with the terminal differentiation of the keratinocytes. Certain low-risk, HPV genotypes (e.g., HPV 6 and 11) do not integrate into the genome, replicate as a plasmid, and are associated with benign disease (i.e., genital warts).

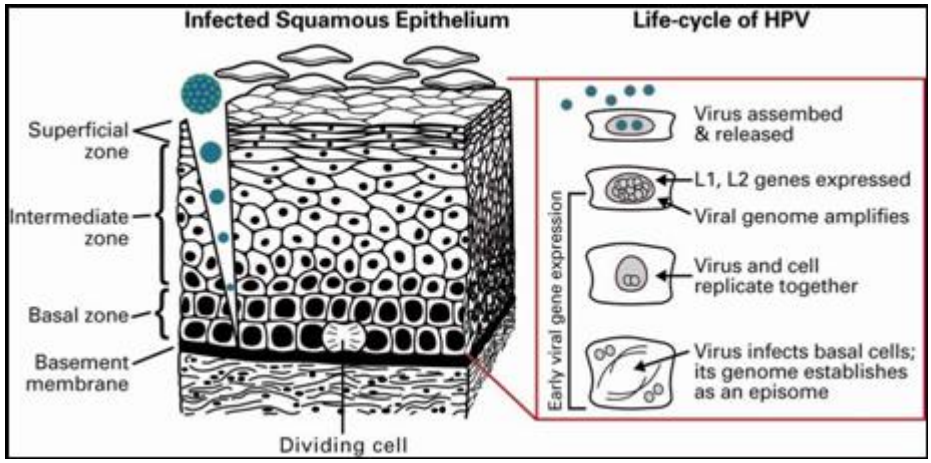

Fig 1.4 HPV virus life cycle[D]

Differentiation of the infected epithelial cells plays an integral role in the HPV life cycle. The cervical epithelium is made up of several distinct layers of cells. Basal epithelial cells line the basement membrane. As these cells mature and differentiate, they slough off from the cervix. While the virus infects the basal layer of cells, the virus cannot complete its life cycle within these cells. Not until the epithelial cells mature, do all of the viral gene products get transcribed.

Viral Gene Products

Viral early genes products (E1- E7) are transcribed in the lower basal cell layer. E2 is a viral transcriptional regulator/repressor. The virus initially replicates as a

[D] Graphic representation www.merckmedicus.com/.../natural-history.jsp Source: Fehrmann F, Laimins LA. Human papillomaviruses: targeting differentiating epithelial cells for malignant transformation. *Oncogene* 2003;22:5201-5207. Copyright reserved by Merck medicus. Accessed Feb 2007.

plasmid. In HPV's episomal state, the E2 protein binds to the E6 and E7 promoter blocking the transcription of the viral oncogenes.

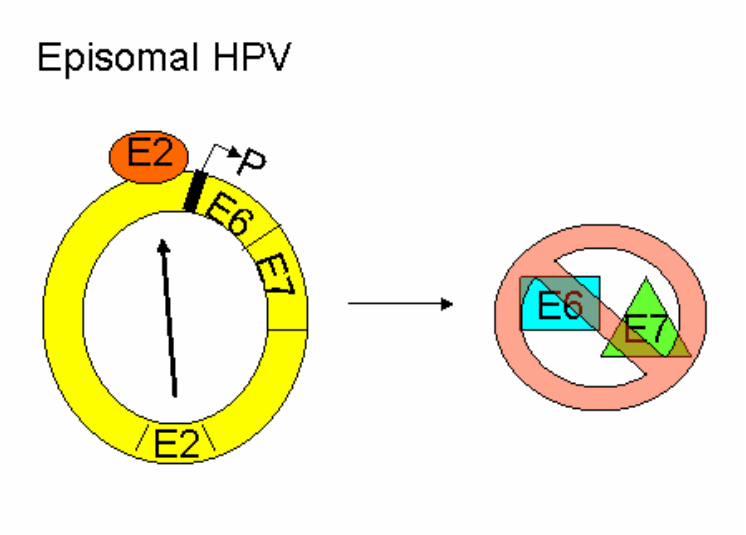

Fig 1.5 Episomal HPV genome. E2 binds to DNA and blocks transcription of E6 and E7.

As keratinocytes mature, high risk HPV genotypes (e.g., HPV 16 and 18) may integrate into the host DNA. Although there are hot-spots of integration into the host genome, integration is a random event. With respect to the virus integration occurs within the E1/E2 region. Once the viral genome integrates into the host DNA, E2 is commonly disrupted so the transcription and translation of functional E2 protein is diminished allowing the de-regulation and overexpression of E6 and E7.

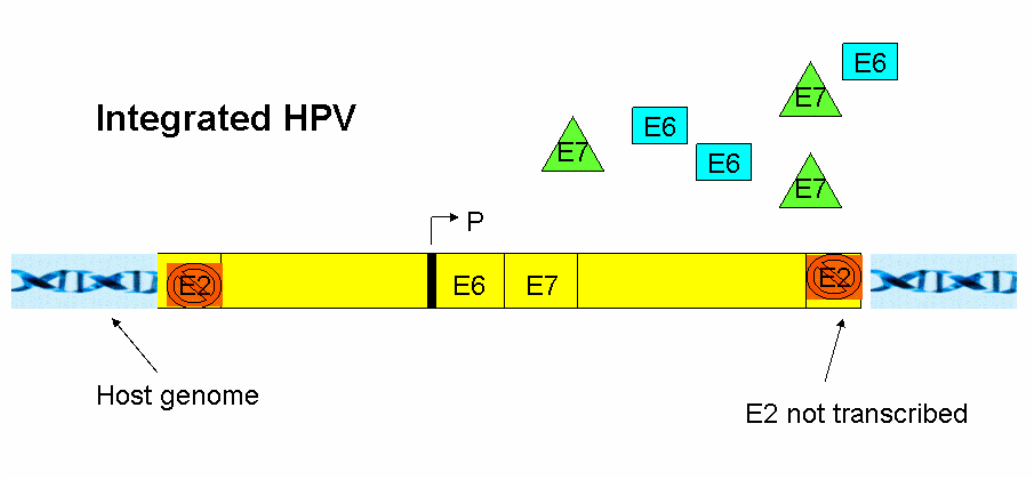

Fig 1.6 Integrated HPV. High-risk HPV types integrate into genome. E2 is disrupted and E6 and E7 oncogenes are over-expressed.

E6 binds to and inactivates tumor suppressor, p53. Normally, p53 activates cip1 (p21) leading to the cell cycle arrest via inactivation of cyclin dependent kinase 2 (cdk2). E6 binds the cellular protein, E6 associated protein (E6AP), and activates the Ubiquitination of p53. Ubiquitination of p53 leads to its degradation and progression of the cell cycle.

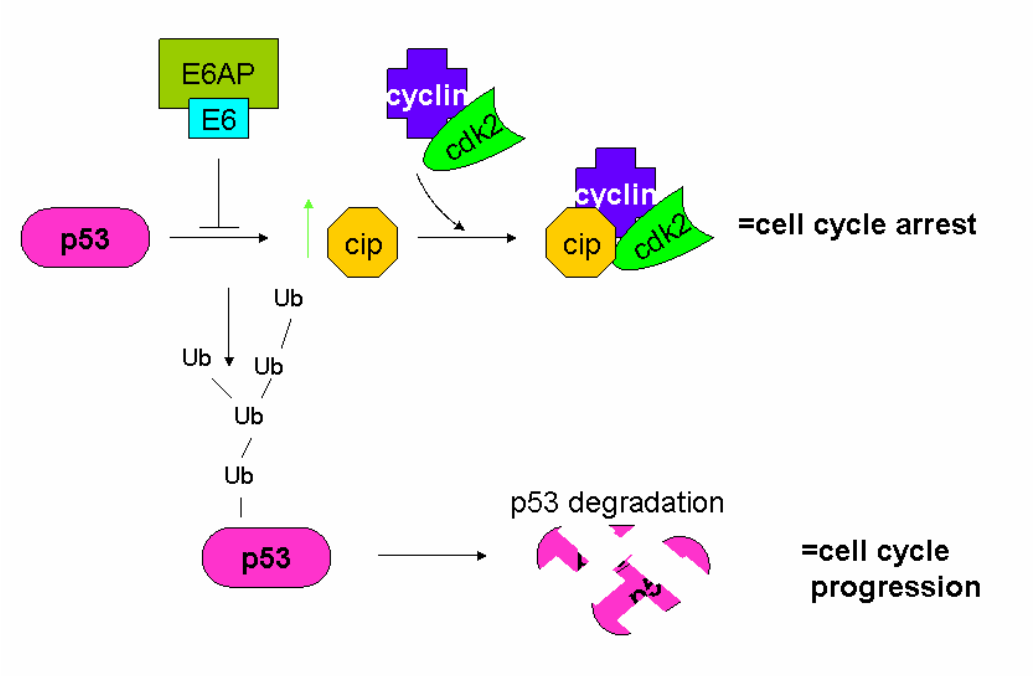

Fig 1.7 E6 blocks cell cycle arrest by targeting p53 for ubiquitin degradation.

The E7 viral oncoprotein binds and inactivates phosphorylated retinoblastoma (pRb). E2F forms an inactive complex when bound to pRB. E2F/pRb dissociates cyclin D from cyclin-dependent kinase 4 (cdk4) thereby causing the cell cycle to arrest in G1. E7 displaces E2F from pRb and allowing progression through the cell cycle.

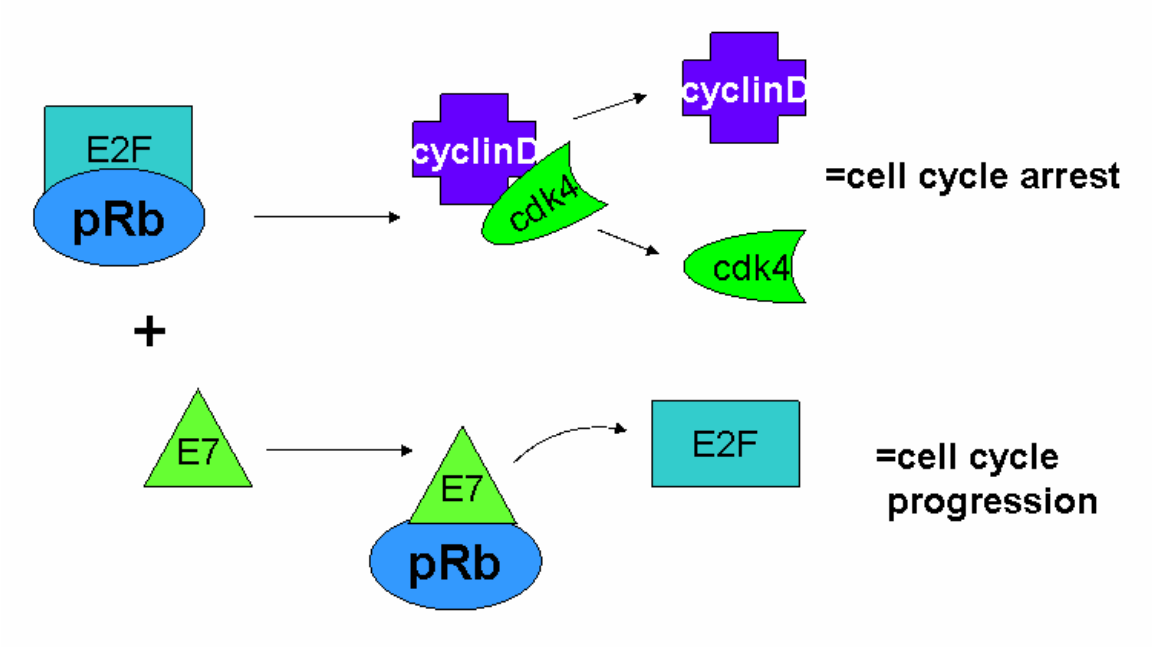

Fig 1.9 E7 blocks cell cycle arrest by binding to pRb and displacing E2F.

Integration of HPV DNA into the hosts genome, with viral persistence and over-expression of the viral E6 and E7 oncoproteins contributes to cervical carcinogenesis [23,24].

The virus produces its capsid proteins (L1 and L2) to encapsulate the viral genome and become an infectious particle. It is important to note that because HPV requires the differentiation of epithelial cells to complete its life cycle and is species-specific, there is no perfect model for the *in vivo* or *in vitro* study of HPV. A lot has been learned from the canine oral papillomavirus system (used in vaccine development), bovine papillomavirus system and cottontail rabbit species. In addition, raft cultures that allow epithelial differentiation have been used to study the viral life cycle and neoplastic induction. Our translational project will use human cervical tissue.

DYSPLASIA

Cervical cancer is one of the few cancers that have an accessible, defined and distinct precancerous form of disease (dysplasia). HPV induces changes in cell cycle and atypical, hyper-proliferative cells can acquire subsequent mutations. The development of progressive disease is a multi-hit process with the progression from dysplasia to cancer often taking many years to develop. The manifestation of disease has been categorized into distinct levels of disease based on histologic findings. Mild dysplasia, also known as cervical intraepithelial neoplasia 1 (CIN 1) is characterized by abnormal cells limited to the lower level of the epithelial layer. A common feature of CIN 1 is the presence of atypical, large nucleated cells termed koilocytes. Moderate dysplasia (CIN 2) to severe dysplasia (CIN 3) is characterized by atypical cells extending up to the upper-most superficial layers of the epithelium. These changes reflect an increasing failure of the normal maturation process such that cell division is retained in cells that have entered the compartment where terminal differentiation should have occurred. Invasive SCCA occurs once the basement membrane is compromised and cells spread into the underlying stroma.

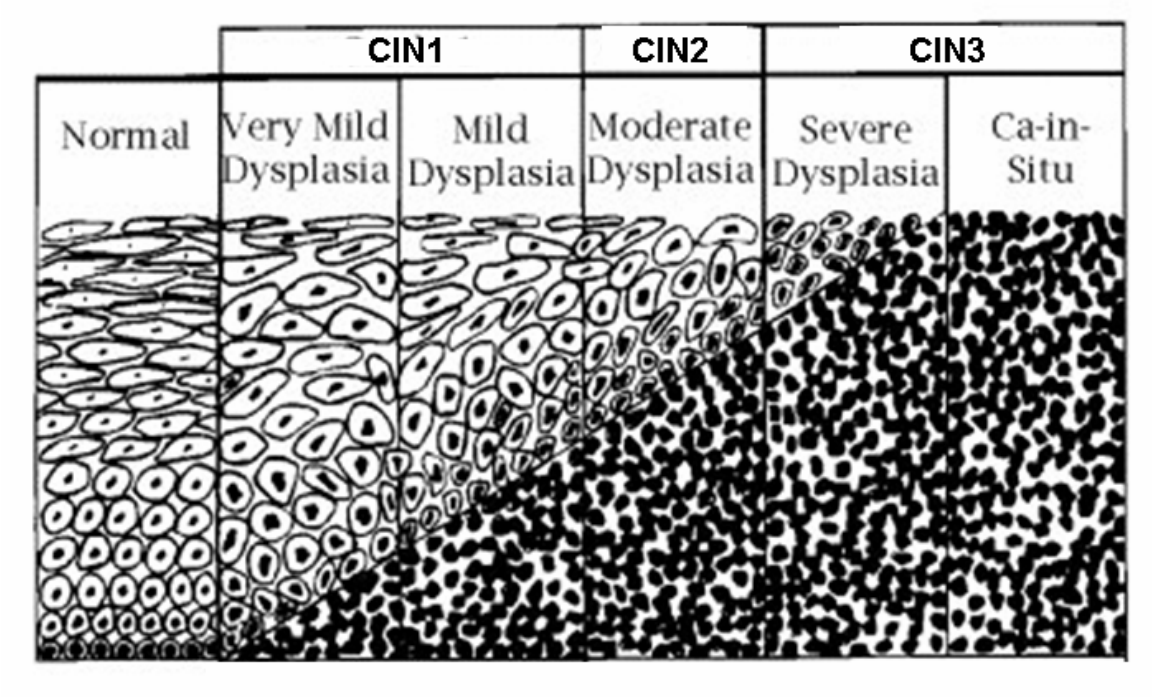

Fig 1.9 The evolution of cervical dysplasia[E].

Interestingly, cervical cytology smears are excellent screening tools for the detection of CIN1, but have 50% false negative rates in detecting CIN2-3 and invasive disease [25]. Perhaps one explanation is because pap smears rely on the exfoliation of epithelial cells. CIN2-3 alter the topography of the surface of the cervix and alter the ability of cells to exfoliate. Keratinization and an increase in cadherins and integrins in dysplastic cells serve as barriers to exfoliation and have the potential to decrease the effectiveness of the pap smear [26,27]. Additionally, small focal lesions and lesions present in the endocervical canal are less likely to be sufficiently sampled by a pap smear. Worldwide, annually there are an estimated ten million cases of CIN2-3 [28]. The World Health Organization estimates approximately 630 million people are infected by HPV at any given time although seventy percent of genital HPV infections are subclinical [29].

[E]Source: Modified from http://www.sh.lsuhsc.edu/fammed/OutpatientManual/PapSmear.htm. Accessed Feb 2007.

Most cervical dysplasias are contained by the hosts' immune system and regress spontaneously (especially the lower grade lesions). In an evaluation of all natural history studies since 1950, 60% of CIN I, 40% of CIN II, and 33% of CIN III regress completely. Only 1% of CIN I, and 5% CIN II, and greater than 12% CIN III progress to cancer [30]. While the exact mechanism of progression remains unclear, infection with oncogenic HPV is only the first step in a series of events that leads to invasive carcinoma. Cofactors play a significant role in cancer development.

COFACTORS

Although oncogenic potential is related to the ability of the virus to persist, it is important to note that HPV infection alone is insufficient for the development of cancer. In fact, one study found over sixty five percent of healthy volunteers have evidence of past or present infection with HPV [31]. Thus, infection with HPV alone does not cause disease. Various cofactors and molecular events influence the transformation of cervical epithelial cells. Important factors such as race, age of first coitus, number of sexual partners, smoking, and socio-economic status have been implicated in the development and progression of disease [32]. Other factors including multi-parity, and diet. Vitamin A and C, and folate deficiencies have been linked with an increased risk of cervical cancer [33,34].

Sexual Behavior

Oncogenic HPV's are mostly asymptomatic in male carriers. Certain sexual behaviors and practices put women at risk for HPV and cervical cancer. The age at which a female has their first sexual experience is one important cofactor. During puberty, the cells of the cervix are undergoing metaplasia and are hyper-proliferative. During metaplasia, the squamous cells are more exposed and at risk for infection with HPV. If the cells are infected with an oncogenic HPV during this time of hyper-proliferation, the virus infection is more extensive. Additionally, the more sex partners a women has, the higher likelihood that she will encounter HPV. Another important risk

factor is the sexual behavior of the partner. A study in Spain linked an increased risk of developing cervical cancer to the sexual behavior of the spouse[35]. A history of other sexually transmitted infections (STI's) in both the female and her partner is another risk factor that may be explained by behaviors that increase the risk for exposure to HPV or to a direct interaction between HPV and other STI's such as *Chlamydia trachomatis*, Herpes Simplex Virus (HSV), or *Neisseria gonorrhoeae* [36-38]. Because uncircumcised men are more likely to carry HPV, having sex with an uncircumcised male increases a woman's risk of getting HPV [39].

Carcinogens

Although the exact mechanism is unknown, women who smoke are twice as likely to get cervical cancer [40,41]. Much research has focused on the tobacco smoke toxins in the carcinogenic transformation of viral oncoproteins. Chemical by-products of tobacco smoke that induce nucleic acid damage are found in the cervix [42-44]. Additionally, environmental exposure from diethylstilbestrol (DES) have been associated with cervical cancer risk [45].

Immunodeficiency

The immune system plays a significant role in the clearance of the virus and disease regression [46]. Women who are on immunosuppressive drugs to prevent transplant organ rejection (e.g., renal transplant patients), have an increased likelihood of developing cervical cancer [47,48]. Also, infection with the Human Immunodeficiency virus (HIV) compromises the immune system by attacking cells of the immune system (T cells and macrophages) required for recognizing and eliminating viruses. Women co-infected with HIV and HPV have a higher rate of cervical neoplasia compared to women infected with HPV alone [49].

Socio-economic status

Social-economic status plays a significant role in cervical cancer. Socio-economic status influences the prevention, early detection, diagnosis, disease treatment,

and quality of life of cancer patients [50,51]. Indicators of social-economic status correlate with education level, race, and cultural factors [50]. These factors are intimately associated and difficult to discern. Cervical cancer incidence and mortality increase proportionately with poverty [52]. Impoverished populations of women do not readily have access to health care and are less likely to go to their doctor for a pap smear. Additionally, once cervical dysplasia is detected (as an abnormal pap smear), these women are less likely to come back for follow-up care until symptomatic [53]. Even within developed countries, like the United States, where access to healthcare is adequate, certain ethnic populations have higher incidence of cervical cancer.

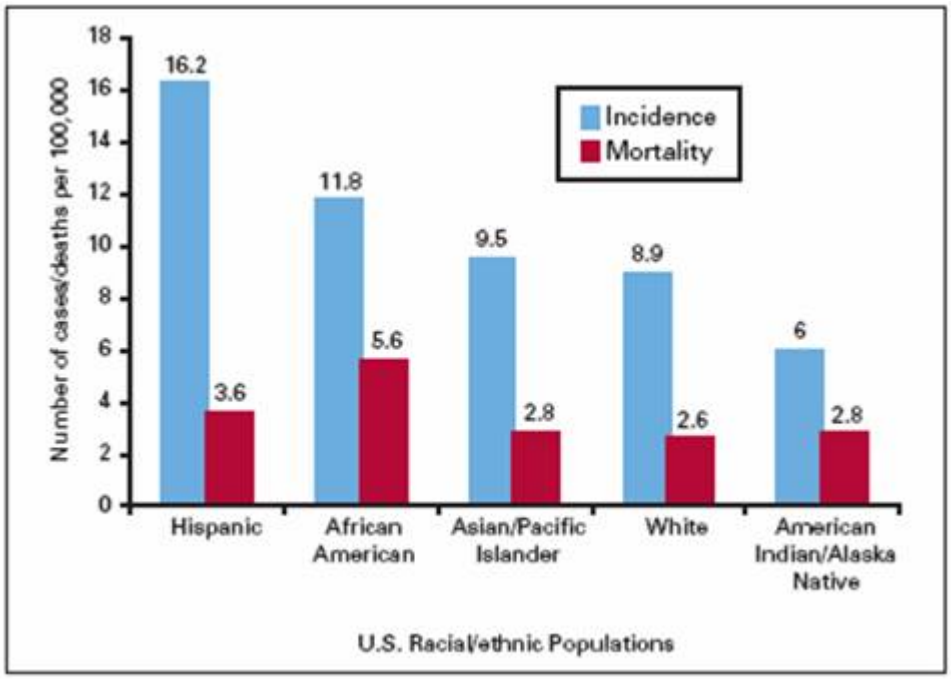

Fig 1.10 U.S. Cervical Cancer Incidence and Mortality Among Racial/ Ethnic Populations (per 100,000 population) 1997-2001[F]

[F] Source: Graphic representation www.merckmedicus.com/.../clinicalmanifest.jsp. Source: Cancer Facts & Figures 2005. http://www.cancer.org/downloads/STT/CAFF2005f4PWSecured. pdf Accessed Feb 2007.

There is increasing literature that there are genetic racial differences in the disparities of cancer [54]. Hispanics and African Americans account for a disproportionately high number of cervical cancer patients and are nearly twice as likely to die of cervical cancer compared to Caucasians [55]. Hispanics make up the highest minority population in the United States and are among the least likely to have access to medical care [56]. Additionally, impoverished women diagnosed with late-stage cervical cancer are 30% less likely to survive five years after diagnosis [57].

Genetic Predisposition

Although race and socio-economic status are closely associated, certain races have higher incidences of cancer. While much of the incidence correlates with lack of access to health care, certain human leukocyte antigens (HLA's) have been found to play a role in cervical cancer risk. HLA's present viral antigens to the immune system allowing the immune system the ability to attack and clear HPV. Certain HLA's, such as HLA DRB1*0401 and DQB1*0301, are more predominant in specific races and are linked to an increase in cervical cancer incidence [58].

CHAPTER 2: PREDICTING OUTCOME IN EARLY SQUAMOUS CELL CARCINOMAS OF THE UTERINE CERVIX

INTRODUCTION

Cervical cancer is an uncommon, but often tragic consequence of infection by the HPV [59]. It is the second most prevalent cancer in women globally, with over 493,000 new cases occurring annually [60]. Radical hysterectomy and pelvic lymphadenectomy may be curative in patients with disease limited to the cervix and upper vagina (Stage I-IIA). Based on the extent of histological spread, patients are grouped into low, intermediate, or high risk for recurrence. Patients stratified into intermediate or high risk categories are advised to undergo post operative whole pelvis radiation therapy and concomitant chemotherapy[61]. At the University of Texas Medical Branch (UTMB), approximately 50% of patients with early stage (IB1) cervical carcinomas are offered post operative chemotherapy and radiation therapy. Post surgical radiation therapy is associated with an increased risk of morbidity compared with primary radiation therapy [62]. Unfortunately, a detailed histological assessment of risk factors for recurrence occurs only with surgical resection of the uterus. Radical surgery followed by concomitant chemotherapy and radiation therapy lacks selectivity, increases morbidity, and is expensive. Patients who have received all three treatment modalities are at increased risk for recto-vaginal and vesico-vaginal fistulas requiring permanent colostomy or urinary diversion. On the other hand, patients with a low risk of recurrence are not offered adjuvant treatments, although 10-15% of these patients ultimately die of recurrent cervical cancer and could have benefited from a more aggressive treatment plan. These data clearly show that standard treatment selection criteria lacks specificity; therefore, there is a need for more sensitive prognostic criteria to guide our management decisions.

Gynecologic oncologists need reliable criteria for the selection of radical surgery, chemotherapy or radiation therapy, based on preoperative tumor characteristics.

Molecular markers have proven beneficial in the diagnosis of haematologic cancers and if applied to cervical cancer, have the potential to allow for more accurate staging and more selective management. Currently, there are no molecular markers for cervical cancer in clinical use.

BACKGROUND AND RATIONALE

Characterizing Risk

Routine pap smear screenings are designed to detect abnormal cells of the cervix. A biopsy of a colposcopic or gross cervical lesion is taken to confirm the diagnosis. Once invasive disease is detected, the extent of tumor invasion must be assessed. The International Federation of Gynecology and Obstetrics (FIGO) created a standard for staging tumors. Staging occurs during a clinical evaluation and is based on the physical characteristics of the tumor. Early stage disease (Stage I) includes tumors that are confined to the uterus. Stage II characterizes tumors that have physically spread beyond the uterus, but not to pelvic wall or lower vagina. Stage III cancers extend to pelvic wall, vagina, and are associated with hydronephrosis that can lead to obstructive uropathy. Stage IV is the most severe and includes tumors that have invaded into the bladder or rectum and beyond the pelvis. Lymph node metastasis is the only significant independent factor that determines survival across all stages of disease [63]. One limitation of FIGO staging is that physical characteristic of tumor size and invasion cannot detect lymph node metastasis. Pre-treatment imaging studies, including computed tomography (CT) scan, magnetic resonance imaging (MRI), and positron emission tomography (PET) scans, are not strong predictors of lymph node involvement [64,65]. Surgical staging is required in order to determine spread of cancerous cells beyond the cervix.

Surgical staging is performed on tissue samples taken during surgical intervention (i.e., hysterectomy). Histopathologic analysis of the tissue samples evaluate factors that are indicative of an aggressive tumor that are high risk for recurrence, thus

treated more aggressively, often with post operative chemotherapy and radiation therapy. These risk factors include the tumor depth of stromal invasion, lymphovascular space invasion (LVSI), marginal involvement and lymph node metastasis [66].

Current Management of Stage 1 Cervical Cancers

Standard therapy for Stage 1 cervical carcinoma includes surgical intervention with a radical hysterectomy and pelvic lymphadenectomy. The Southwest Oncology Group recently conducted a study that randomized patients with pathologic poor prognostic factors including positive margins of resection, or lymph node metastases to post operative cisplatinum and 5-fluorouracil with concomitant pelvic radiation therapy versus radiation therapy alone. Patients in the chemotherapy with radiation arm had an improved recurrence-free survival, compared to the group receiving radiation alone [67]. However, in cases of failures, metastasis occurred both regionally and distantly, indicating that we can still do better by identifying a group of patients who are at risk of distant failure and could benefit from systemic therapy. The Gynecology Oncology Group (GOG) investigated the role of adjuvant radiation therapy for Stage 1 patients with intermediate risk factors [68]. After radical hysterectomy and pelvic lymphadenectomy, patients with bulky cervical lesions, deep stromal invasion, and LVSI with negative lymph nodes and negative margins were randomized to postoperative radiation therapy versus no further therapy.

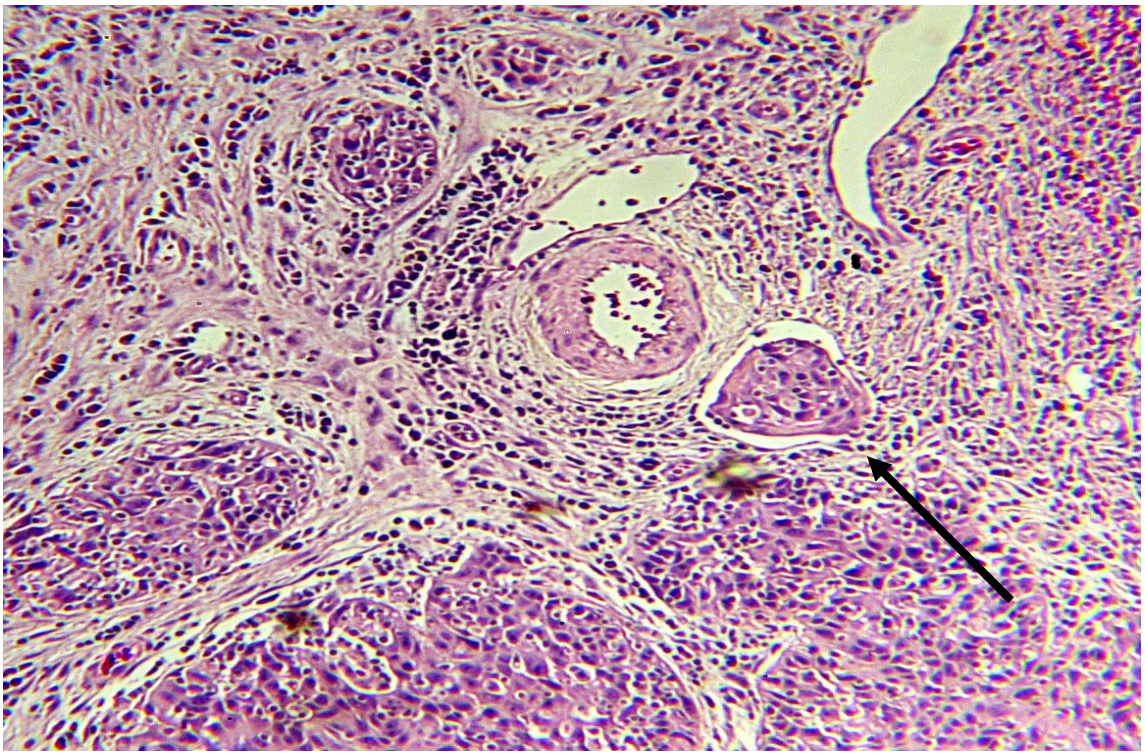

Fig 2.1 H & E slide showing lymphovascular space invasion (LVSI) (arrow)[G].

An improved recurrence-free interval was found in the patients randomized to receive post-operative radiation therapy [69]. Since chemo-radiation has been found to be superior to radiation alone in every tested clinical scenario of cervical carcinoma, many institutions, including ours, treat patients who have intermediate and high risk factors with chemotherapy and radiation therapy, which carries a three fold higher risk of severe toxicity (grade 3-4). Despite this aggressive therapy, 15 % of these women have been found to develop a recurrence and ultimately die of their disease. These results show that the clinicopathological risk factors currently in use are of limited value in predicting prognosis and guiding therapy.

[G] Source: Photo courtesy of Dr. Concepcion Diaz-Arrastia, M.D.

The Effects of Radiation

While a majority of patients with 1B1 SCCA are at risk for metastatic disease, and radiation does play a positive role in preventing recurrence, it is not the standard treatment for all patients. Radiation therapy consists of four to five weeks of an external beam radiation treatment with an internal beam (brachytherapy). In brachytherapy, radioactive material is inserted directly into the vagina and uterus in order for the radiation source to be adjacent to the tumor. External beam radiation consists of a broad band beam aimed at the pelvis (50-55 Gray, Gy) [70]. Radiation therapy is a regional treatment, not specific to the cervix. All organs in the radiation field, including the bladder and the rectum are consequently affected.

Complications range from vaginal dryness, urinary incontinence, vaginal stenosis, premature menopause, to more serious complications of decreased bone density resulting in pelvic fractures, bowel obstructions, abdominal adhesions, fistulas, and even death. Often women undergo multiple follow-up surgeries to repair damage caused by radiation. Serious complications occur in about 30% of patients treated with radiation [71]. The physical bulk of the uterus shields or protects adjacent structures from the non-specific effects of radiation. Unfortunately, histologic characterization occurs only after the uterus is removed, thus increasing radiation sequelae. With an improved technique such as molecular profiling for determining outcome, specifically targeting patients with a high risk of recurrence prior to surgery could lead to decreased morbidity associated with radiation and ultimately decreased cost of treatment.

OBJECTIVE

The objective of this study was to increase our understanding of the processes that lead to an aggressive tumor phenotype with the ultimate goal of stratifying risk so that patients requiring radiation would not have surgery, thereby reducing complications of radiation. The long-term goal is to use this molecular profile to develop and test tailored therapy of patients with cervical carcinoma.

HYPOTHESIS

My general hypothesis is that a molecular diagnosis of early cervical carcinoma is more accurate than the histologic staging in predicting the biological behavior of the tumor. I will establish a genomic signature of early squamous cell carcinoma of the cervix at advancing degrees of invasiveness.

SPECIFIC AIM

To determine a molecular signature that predicts clinical outcome of early stage IB1 squamous cell carcinomas.

METHODS

Affymetrix Microarray GeneChip Arrays

Specimen collection

Women who underwent a radical hysterectomy and pelvic lymphadenectomy for stage IB1 squamous cell carcinoma (SCCA) of the cervix were biopsied and grouped according to histopathologic criteria into superficially or deeply invading tumors(IRB#02-272). Biopsies had to have good quality 18s and 28s ribosomal RNA, and sufficient (24 ug) RNA had to be present in the biopsies for the Affymetrix GeneChips. Tumor biopsies were transported from the operating room to the laboratory in 10 volumes of RNA Later (Cat# 7020, Ambion, Austin, Texas). Following the manufacturers instructions, biopsies were stored at 4^{o}C overnight, decanted and stored at -80^{o}C. We extracted RNA from 12 tumor biopsies. Additionally, women undergoing a hysterectomy for benign reasons were also asked to participate. Normal biopsies were taken from 11 cervices. Four tumor biopsies were degraded and 7 of the normal biopsies had insufficient RNA. We selected two tumors with a low risk of recurrence and two with a high risk of recurrence based on certain histopathic criteria (see Table 2.1). The high and low risk patients were selected because they were very similar. For the superficial group, we wanted to identify tumors that had less than a 20% depth of invasion. Our two superficial tumors in actuality had less than 10% depth of invasion. Likewise, the deeply invading tumors were similar in their histopathic criteria. We wanted to identify patients with greater than 90% depth of invasion. Our actual patients had 85% and 90% depth of invasion with no lymph node metastasis. Two normal tissues were also included in the analysis. Superficial and deep tumors are surrogates for patients having a good or poor prognosis, respectively.

Histopathic Criteria

Recurrence Risk	Depth of Invasion	Lymphovascular Space Invasion	Lymph Node Metastasis	Classification	Prognosis
Low Risk	<20%	No	No	Superficially Invasive	Good
High Risk	>80%	Yes	No	Deeply Invasive	Poor

Table.2.1 Histopathic criteria for risk stratification.

The superficial group had less than 20% depth of invasion, no lymphovascular space invasion (LVSI) and no lymph node metastasis. The deeply invasive group had at least 80% depth of stromal invasion, extensive lymphovascular space invasion (LVSI), and no pelvic lymph node metastasis. Biopsies were taken from the external cervix prior to hysterectomy. Additionally, biopsies were taken from the normal cervix (no cancer) and served as a control to eliminate non-tumor specific genes.

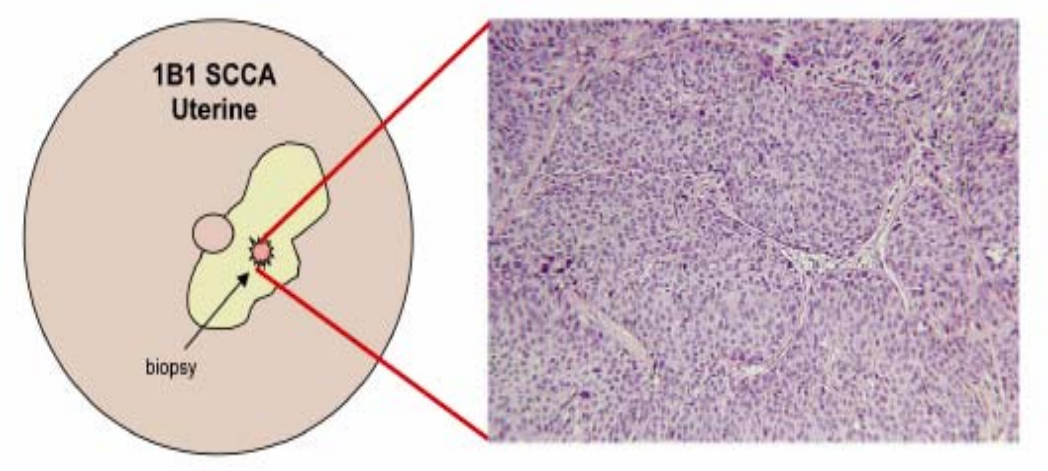

Fig 2.2 Tumor biopsy.

Human tissue is inherently heterogeneous, but we chose tumors such as this representative section, with pushing margins instead of fingerlike projections, to maximize the homogeneity of the tissue sampling. Less than 3% of the cells in this section were stromal cells.

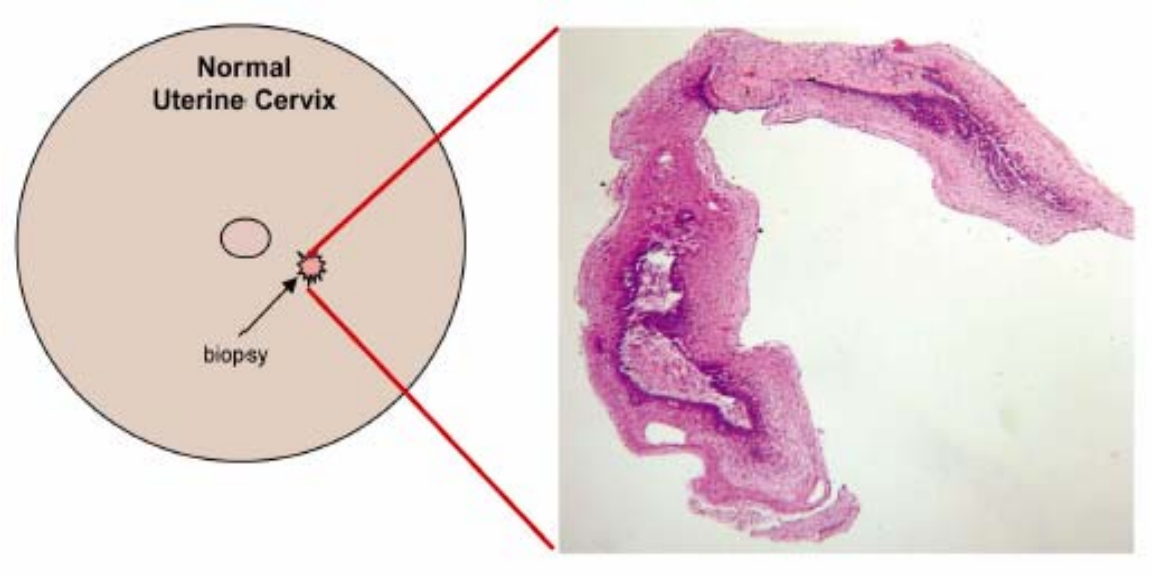

Fig 2.3 Normal biopsy.

We obtained normal cervical epithelium from patients without cancer who underwent hysterectomies. This is a representative section of our normal epithelium; there is primarily epithelium, with minimal underlying stroma. These four specimens were carefully selected to represent the ideal representation of tumors with a high and low risk of recurrence.

RNA Extraction

RNA was extracted using the RNAqueous-4PCR (cat#1914, Ambion, Austin TX) kit following manufacturer's instructions. Frozen biopsies (approximately 15 milligram (mg) biopsies) were crushed in 3 milliliters (ml) of liquid nitrogen in a pre-cooled mortar and pestle (Coors). Biopsies were pulverized into a fine frozen powder and 500 microliters (ul) of lysis/binding solution was added to the bottom of the mortar. The lysed

tissue (lysate) was then transferred to a 1.5ml microfuge tube. An equal volume (500ul) of 64% ethanol was added to the lysate. The ethanol/lysate was then gently mixed by inversion. Five hundred ul of the ethanol/lysate was added to a filter cartridge. The filter cartridge is placed into a collection tube and the sample then centrifuged at 12,000 x *g* for 1 minute (min). This process was repeated per sample until all lysate was filtered (max 2ml per filter cartridge). The RNA, bound to the filter cartridge, was washed with 700ul wash solution #1 and centrifuged at 12,000 x *g* for 1 minute. Five hundred ul wash solution #2/3 was added to the filter cartridge and centrifuged at 12,000 x *g* for 1 minute and then repeated one time. The RNA was eluted using 50ul preheated (80^0C) elution solution. The RNA was DNase treated by adding 5ul 10X DNase 1 buffer and 1ul DNase 1 and incubated 15-30 minutes at 37^0C. Five ul DNase Inactivation Reagent was added. The solution was gently mixed by flicking and incubated 2min at room temperature. The DNase Inactivation Reagent was pelleted by centrifuging the RNA at 10,000 x *g* for 1 min. The DNase-free, RNA was removed from the pellet by pipetting into a fresh, labeled microfuge tube. A 5ul aliquot was removed for RNA quantification and the sample stored at -80^0C.

Assessing RNA Quality/Quantity

The quality and quantity of RNA was assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Walbronn, Germany) following manufacturer's instructions. Using the Agilent RNA 6000 Nano kit (cat#G2938-90030), 1ul of RNA was labeled with a fluorescent gel-dye mix. The RNA was resolved on a gel matrix and RNA quality and quantity was calculated. Only RNA with the highest fidelity (A_{260}/A_{280} ratio between 1.9 and 2.1) was used for microarray analysis.

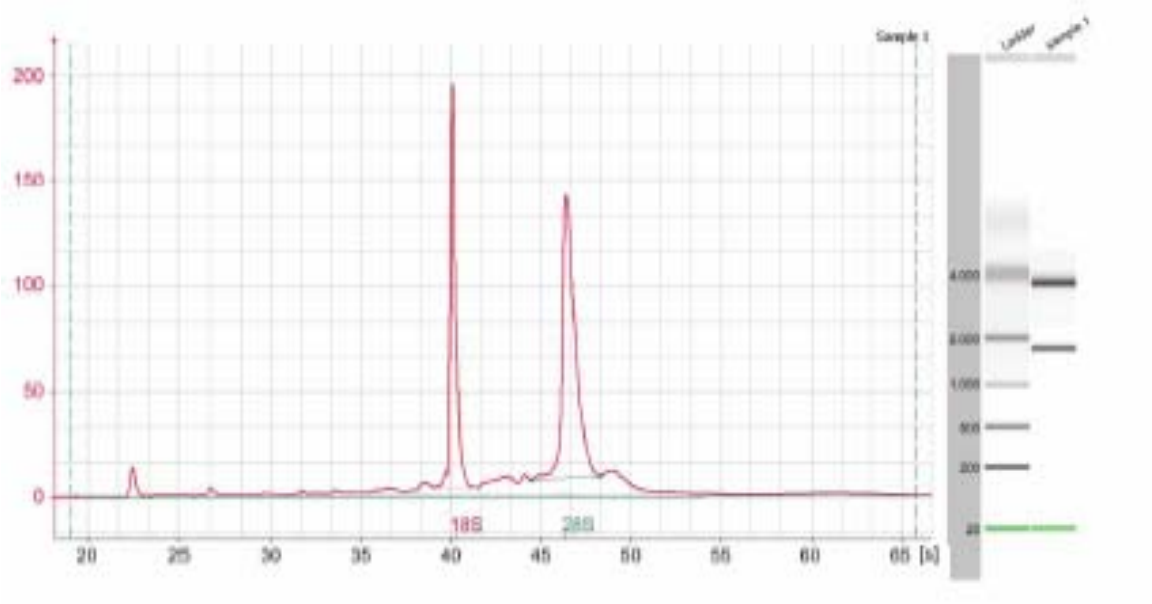

Fig. 2.4 RNA from tumor biopsy. Agilent Bioanalyzer data showing ribosomal 18S and 28S integrity.

Microarray GeneChip

RNA was extracted from cervical biopsy tissue and applied to the Affymetrix microarray GeneChips (Affymetrix, Inc., Santa Clara, CA) to identify differentially expressed genes. Twenty to forty ug of DNA-free, messenger RNA (mRNA) was converted to its double-stranded compliment (cDNA) using the One-Cycle cDNA Synthesis Kit (cat# 900493, Affymetrix, Inc.). *B. subtilis* Poly-A RNA was spiked in with the target RNA to monitor the labeling process. The double-stranded cDNA was then cleaned using the Sample Cleanup Module (contained in the Once-cycle cDNA Synthesis lit)). An *in vitro* transcription (IVT) reaction resulted in a biotin-labeled complementary RNA (cRNA). To remove unincorporated nucleotides, another clean-up step was performed on the biotin-labeled cRNA. The fluorescently labeled targets were allowed to hybridize 16 hours with the Affymetrix GeneChip Human Genome U133 array set (HG-U133A and HG-U133B) (cat# 900444, Affymetrix). In our hands, microarray to microarray variance is very low, with only 0.2% of the probe sets showing a 1 fold change or greater (between chips). Therefore, we used one GeneChip per

sample. Fluorescently-labeled cRNA hybridized to its complementary DNA on the GeneChips. Non-hybridized fragments were then washed from the chips. Fluorescence intensities, representing gene expression levels were captured using the Affymetrix system. The probes were tiled in probe pairs consisting of a perfect match for measuring specific binding to the gene of interest and a mismatch, for non-specific binding. Detection *p*-value cut-offs were used to calculate the values for the detection of transcripts as present, absent or marginal.

Statistical Analysis

The S+ARRAYANALYZER 2.1.1(S+AA) module of S-Plus7.0.4 (Insightful Corporation, Seattle, WA) was used to analyze microarray data. This software has the enhanced data import, QC diagnostics, normalization and differential testing methods for array analysis. The CEL files from the superficially and deeply invading tumors samples were summarized and normalized using GCRMA (GC- Robust Multichi Analysis) method[72]. The GCRMA method estimates the background in the probe level data based on the GC content in the sequences. It further normalizes and summarizes data using quantiles and medianpolish methods. The genes absent across all samples were eliminated from the analysis

Low and high risk tumors were differentiated using the local pooled-error (LPE, test) from S+AA[73]. The LPE test method is designed specifically for the analysis of microarray experiments with low replicates. For experiments performed on chips with a large number of genes (e.g., Affymetrix GeneChips) and few replicates, the estimates of within-genes standard errors are imprecise due to unrealistically high signal-to-noise ratio and a low number of degrees of freedom (one or two), which resulted in increased Type I and Type II errors. In differential expression testing (DET), LPE estimates, which are used like *P* values, are obtained by pooling the variance estimates of genes with similar expression intensities. Further, the Bonferoni adjustments (a family-wise error rate method), was used to reduce the likelihood of false positives. Clustering was performed in Spotfire DecisionSite 8.2 (Spotfire, Somerville, Massachusetts).

Signal intensities were normalized using a *z* score. Genes with an absent call across all samples were eliminated from the analysis. "A" chips were used only because the "B" chips are mostly expressed sequence tags (EST's).

Ingenuity Pathways Analysis (IPA)

Ingenuity Pathways Analysis (IPA) 3.1 software (Ingenuity® Systems, www.ingenuity.com) was used to identify important functional groups and networks to model and analyze those gene relationships that appear to be important in differentiating between the high- and low-risk groups.

Network Generation

The LPE test filtered 98 genes that were differentially expressed between the low and high risk tumors at a *p*-value ≤ 0.01 with Bonferoni adjustment. These 98 genes were analyzed using the Ingenuity Pathways Analysis software to identify the networks, functional groups and pathways that are affected due to the tumors. Each gene identifier was mapped to its corresponding gene object in the Ingenuity Pathways Knowledge Base. These genes, called focus genes, were overlaid onto a global molecular network developed from information contained in the Ingenuity Pathways Knowledge Base. Networks of these focus genes were then algorithmically generated based on their connectivity.

Functional Analysis

The Functional Analysis of a network identified the biological functions and/or diseases that were most significant to the genes in the network. The network genes associated with biological functions and/or diseases in the Ingenuity Pathways Knowledge Base were considered for the analysis. Fischer's exact test was used to calculate a *p*-value determining the probability that each biological function and/or disease assigned to that network is due to chance alone.

Total Nucleic Acid Extraction (TNA)

Total nucleic acid (TNA) was extracted from formalin-fixed, paraffin-embedded (FFPE) archival tissues for HPV genotyping (DNA) and for the quantitative real-time PCR (qRT PCR) applications (RNA after DNAse treatment) using the RecoverAll Total Nucleic Acid Isolation Kit optimized for FFPE tissues (cat#1975 Ambion, Austin, TX). . A full chart review of all IB1 SCCA's from year 2001 to present were compiled from the University of Texas Medical Branch (UTMB) archives. Haemotoxylin and Eosin (H&E) slides were prepared for each block. All microscopic slides were reviewed and tumor samples fit the following criteria; they had a greater than 70% cancer cell: stroma ratio, less than 2% necrosis, and less than 5% lymphocytic infiltrate. Fifty-five patients fit the above microscopic criteria. Four patients were excluded because they had undergone previous chemotherapy. Sixteen patients whose paraffin blocks were not available were excluded. The remaining 35 samples were evaluated in our study. Multiple blocks were tested separately from several patients. Fifty eight samples from thirty five patients were used to look at relative quantification of selected genes. The tumors were characterized microscopically (ratios taken of tumor to stroma, percent necrosis, and lymphocytic infiltrate were noted) and carefully cut using a microtome. Special care was taken to reduce the amount of "normal" tissue from the slivers of FFPE tissue. Using a clean scalpel, the tissues were laid out on a glass slide and cut to remove the "normal" tissue leaving only tumor tissue. We felt this was important for the accurate quantification of gene expression with real time PCR.

Two tubes of tissue per sample was prepared; one for RNA and one for DNA. One ml of 100% xylene was added to the paraffin-embedded tissue curls/sections (≤80um total) for deparaffinization. The samples were incubated at 50^{o}C for 3 minutes and pelleted by centrifugation . The xylene was discarded; the pellet washed twice in 100% ethanol, and briefly air dried. Four hundred ul digestion buffer and 4ul protease were added to each tube. The samples were incubated at 50^{o}C for 3 hours for RNA and 48 hours for DNA. Nucleic acid was isolated by adding 480ul isolation additive and 1.1 ml

100% ethanol. The mixture was passed through a filter cartridge and washed several times. DNase was added to the RNA samples and incubated for 30 min to remove DNA. The RNA samples were washed several more times. Elution solution was added to all columns and the nucleic acid eluted from the columns. The concentration of the nucleic acid was quantified using an Eppendorf Biofotometer spectrophotometer to obtain the A260/280 ratio. A 25 fold dilution was made using 4 ul of nucleic acid and 96 ul Tris EDTA(TE) Buffer (10mM Tris, 1mM EDTA, pH 7.6).

HPV Genotyping

HPV genotyping was performed using the Linear Array HPV Genotyping Test (cat# 04472209 Roche, Indianapolis, IN). The Linear Array method utilizes PCR amplification of target DNA, subsequent hybridization, and a colorimetric reaction to detect 37 different types of anogenital HPV genotypes (HPV 6, 11, 16, 18, 26, 31, 33, 35, 39, 40, 42, 45, 51, 52 (linear array uses a cross-reactive probe to detect HPV 52), 53, 54, 55, 56, 58, 59, 61, 62, 64, 66, 67, 68, 69, 70, 71, 72, 73, 81, 82, 83, 84, IS39, and CP6108)[74]. *B*-globin was used to determine the integrity and presence of the sample DNA.

Sample DNA was amplified using a pool of biotinylated primers to the polymorphic L1 (Late 1) region of the HPV genome per manufacturer's instructions. A no template negative control and an HPV 16 positive control were also included. The PCR reaction conditions were as follows; HOLD 2min 50^{o}C, HOLD 9min 95^{o}C, CYCLE (40 cycles) 30sec 95^{o}C, 1min 55^{o}C, 1min 72^{o}C, HOLD 5min 72^{o}C, HOLD 72^{o}C indefinitely. After PCR was complete, samples were denatured using DN solution (provided in kit). 8ul of amplified DNA was then hybridized to the test strips. The DNA was added to pre-warmed hybridization buffer and incubated at 53^{o}C for 15 minutes. The strips were then washed with a stringent wash buffer. Working horseradish peroxidase (HRP)-conjugate was added to the strips and incubated for 30 minutes at room temperature, washed with an ambient wash buffer. Substrate was added to the strips and allowed to develop until bands were clearly seen on the strips. The reaction

was stopped by washing with distilled water. Strips were read against a key card. Visible bands corresponded to various HPV genotypes.

Reverse Transcription (RT)

Reverse transcription of RNA isolated from cervical biopsies and FFPE tissues was performed using the High-Capacity cDNA Archive Kit (cat# 4322171 Applied Biosystems, Foster City, CA). A 2X Master Mix made up of 10X reverse Transcriptase (RT) buffer, 25X dNTP mix, 10X random hexamer primers, MultiScribe RT (50U/ul), and RNase-free water was added at a 1:1 ratio with up to 10ug of purified RNA in a 25ul total volume. A no-template negative control and HPV 16 positive control was also included. The PCR reaction was as follows; HOLD 10 min at 25^{o}C, HOLD 37^{o}C for 2 hours. cDNA was stored short term at -80^{o}C until processed further.

Endogenous Control Optimization

The TaqMan Human Endogenous Control Plate (cat# 4309199 Applied Biosciences, Foster City, CA) was used to screen RNA from the cervix of patients with stage IB1 SCCA. The TaqMan kit included a pre-coated plate of eleven different human housekeeping genes [18s rRNA (18s), acidic ribosomal protein (huPO), beta-actin (hu*B*A), cyclophilin (huCYC), glyceraldehyde 3-phosphate dehydrogenase (huGAPDH), phosphoglycerokinase (huPGK), *b*2-microgloulin (hu*B*2m), *b*-Glucronidase (huGUS), Hypoxanthine ribosyl transferase (huHPRT), transcription factor IID TATA binding protein (huTBP), and transferring receptor (huTfR)] and an internal positive control (IPC) to detect the presence of inhibitors in our samples.

RNA samples from four patients with stage IB1 SCCA were reverse transcribed (RT) using random hexamers as described above [*under Methods, Reverse Transcription*]. Using the 2X TaqMan Universal PCR Master Mix (cat# 4352042, Applied Biosystems), cDNA was amplified and a VIC reporter dye was incorporated at the 5' end and a quencher at the 3' end of the transcripts. Using the ABI PRISM 7700 Sequence Detection System (Applied Biosystems), the PCR reaction was as follows; HOLD 2 minutes at 50^{o}C, HOLD 95^{o}C for 10 minutes, CYCLE (X40) 95^{o}C for 15

seconds, 60°C for 1 minute. The fluorescence was collected and the standard deviations between all the samples were tabulated from the C_T (cycle number where fluorescence crosses threshold) values [*as described below under Methods, Relative Quantification Calculations*]. The housekeeping gene with the lowest standard deviation between samples was used as our experimental baseline.

Fast Quantitative Real-Time PCR (qRT PCR)

Using the 7500 FAST Real-Time System Instrument (Applied Biosystems), cDNA from the FFPE samples, endogenous controls, and no-template controls, were amplified in a singleplex reaction. Pre-made, unlabeled forward and reverse primers and a FAM-MGB probe were supplied in the 20X Taqman Gene Expression Assay mix (cat# 4331348, Applied Biosystems) for each gene of interest (GOI). Primers were designed to be inclusive of splice variants.

TaqMan Gene Expression Assays-on-Demand (cat# 4331348)

TaqMan Gene Expression Assays	RefSeq	Applied Biosystems Assay ID
CEACAM 7	NM_006890	Hs00185152_m1
Serpin B4	NM_002974.2	Hs00741313_g1
Neurofilament	NM_006158.2	Hs00196245_m1
PTPRF	NM_003622.2	Hs00794171_mH

The 2X TaqMan Fast MM, AmpErase UNG (cat#4352042, Applied Biosystems) was used at a 1:1 concentration in a 20ul final reaction volume per well. The PCR reaction was as follows; HOLD 20 seconds at 95°C, CYCLE (X40) 95°C for 3 seconds, 60°C for 30 seconds. The fluorescence was collected and the relative quantification was calculated relative to an optimized endogenous control (18s).

Relative Quantification (RQ) Calculations

With the use of the Sequence Detection System (SDS) Software (Applied Biosystems), relative quantification (RQ) values were calculated. The change in fluorescence (deltaR$_n$) detection was plotted against cycle number. A threshold (C_T) value was automatically set at the halfway point between the baseline and the maximum fluorescence. The C_T represents the cycle number at which the fluorescence is in exponential phase.

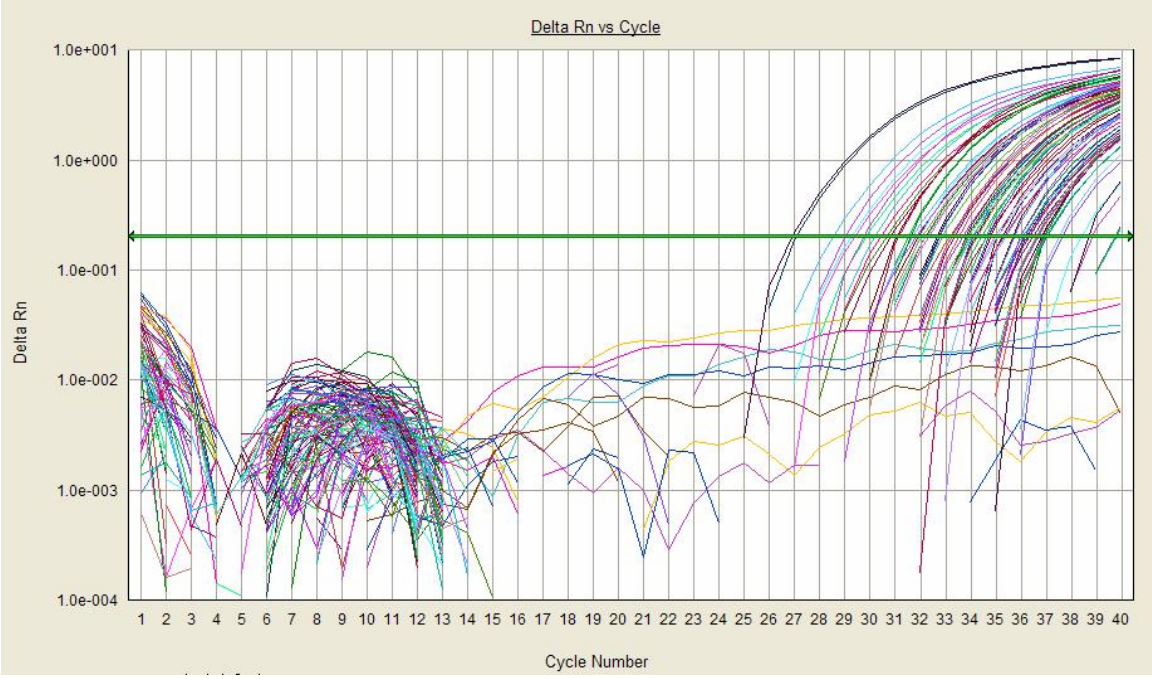

Fig. 2.5 Relative Quantification. Threshold falls within exponential growth phase[H].

Outliers, falling outside of the dynamic range, were removed from replicate samples and C_T values were averaged. RQ values for each sample were quantified based on the GOI expression value to an endogenous control (delta C_T).

Principal Component Analysis (PCA)

The Principal Component Analysis (PCA), part of Spotfire DecisonSite 8.2, transforms non-linear data into linear data. It minimizes the dimensionality of the data

[H] Actual data using SDS Software (Applied Biosciences) obtained from Serpin B4 experiment.

points (patient samples) and scores them according to their variability. The variance along each axis corresponds to the Eigen value. Ninety-eight percent of our variability is contained within the first three principal components; PCA1, PCA2, and PCA3.

RESULTS

GeneChip Array

Women undergoing a radical hysterectomy and pelvic lymphadenectomy for stage IB1 squamous cell carcinoma of the cervix were biopsied. Postoperatively, they were grouped according to histopathologic findings into superficially (good prognosis) or deeply invasive (poor prognosis) groups. Biopsies from normal cervices were also evaluated. RNA was extracted from cervical biopsy tissue of six patients using the RNAqueous 4-polymerase chain reaction kit (Cat#1914, Ambion, Austin, Texas). The fidelity and quantification of RNA was tested using an Agilent Bioanalyzer 2100 (Agilent Technologies, Palo Alto, California). A cDNA copy was made or the transcripts and applied to the Affymetrix microarray human GeneChip HU133 (Affymetrix, Santa Clara, California). Genes differentially expressed with a *P* value of $\leq .01$ were included in the analysis.

Patient Information

	Superficial	Superficial	Deep	Deep
High Risk Factors	Patient 1	Patient 2	Patient 3	Patient 4
AGE (years)	52	51	34	39
Depth of invasion	1 mm/2 cm (5%)	1.5 mm/1.5 cm (10%)	29/32 mm (91%)	17/20 mm (85%)
LVSI	No	No	Yes	Yes
Lymph node metastasis	No	No	No	No

Table.2.2 Patient Information Summary Table (Superficial versus Deep Study)

Statistical Analysis

Analysis of variance (ANOVA) was used to filter genes differentiating between the normal cervices and the deeply and superficially invasive IB1 tumors of the cervix. Spotfire DecisionSite 8.1 was used to generate heat maps and hierarchical clustering (see Fig 2.6). A total of 309 genes were identified by the ANOVA filter at a *P* value of ≤.01 responded to the disease state rather than some random variation. At the 99% confidence level, 309 probe sets (i.e., genes) appear to be clearly differentiated between deep cancers, superficial cancers, and normal cervices.

LPE *t* test

We further compared the GeneChip data from our superficially and deeply invading tumors using the local pooled-error test (LPE test, LPE *t* test) in the S+ Array Analyzer 2.0 module of S-Plus 7.0.4 (Insightful). Bonferoni adjustments were made to decrease Type I errors. Spotfire DecisionSite 8.1 was used to produce a heat map with 2-way hierarchical clustering for visualizing the comparative expression values of the significant genes (see Fig 2.7). A total of 98 genes were differentially expressed at a 99% confidence level between superficially and deeply invasive cervical tumors. Genes, including serpins, S100 genes, CEACAMs, and mucins with molecular function in cell cycle control, signaling, and invasion were significantly altered in the deeply invading group.

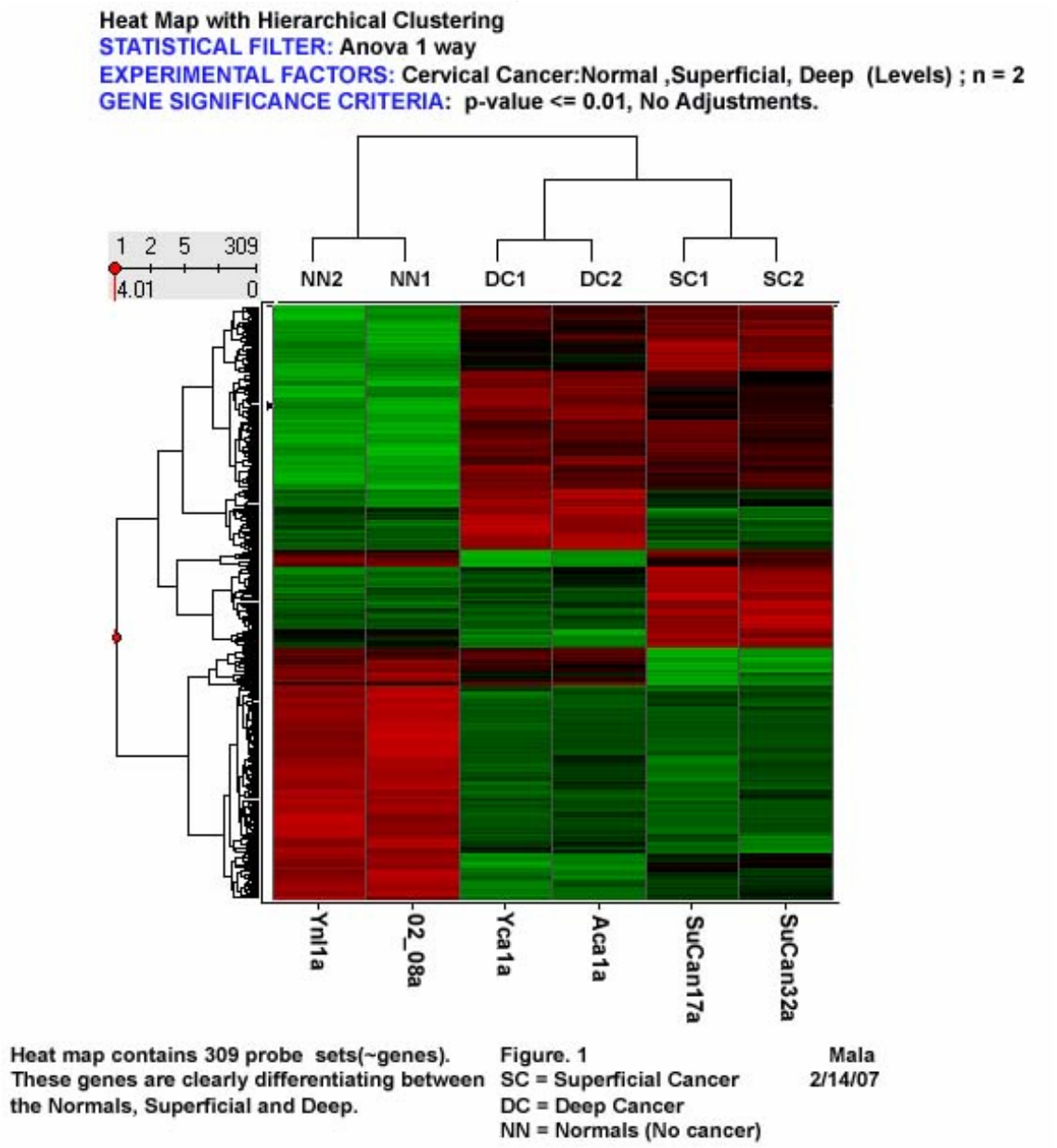

Fig. 2.6 Heat map with hierarchical cluster (99% confidence level, $P \leq .01$) showing disease-specific clustering. The first two columns represent gene expression from the normal cervices with no cancer (NN). The third and fourth columns are from deeply invasive tumors (DC), and the last two columns are biopsies from superficially invasive tumors (SC). Patient identifiers are listed at the bottom of each column (Ynl1a, 02_08a, etc).

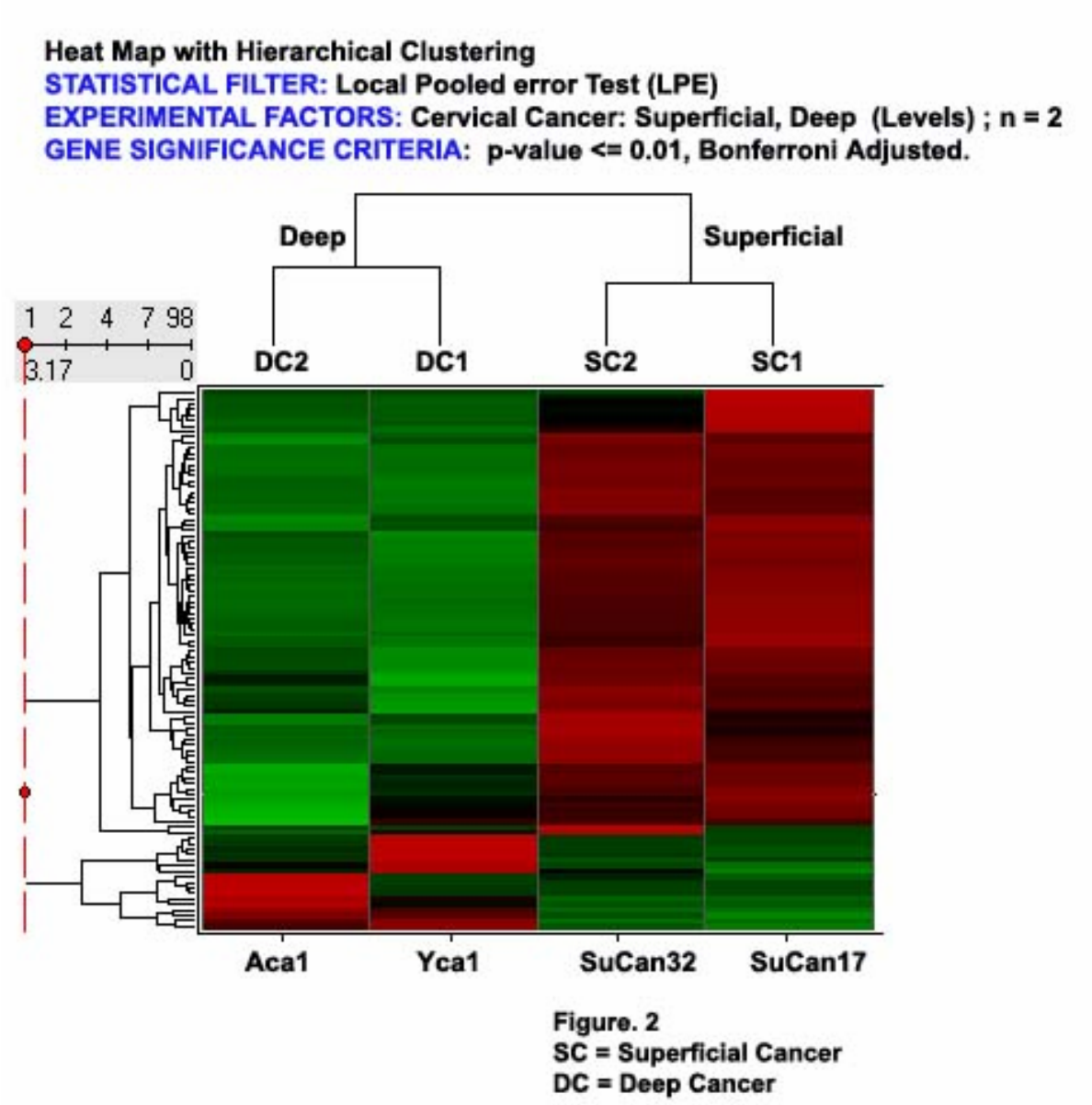

Fig 2.7 Heat map with hierarchical cluster (99% confidence level, P ≤0.01) showing depth of invasion-specific clustering. The first two columns represent two deeply invading tumors while the last two columns represent two superficially invading tumors.

Ingenuity Pathways Analysis

Using IPA, we were able to determine which functions might be most relevant to our dataset. The functions with the highest number of associated genes include cell-to-cell signaling, immune response, and cancer.

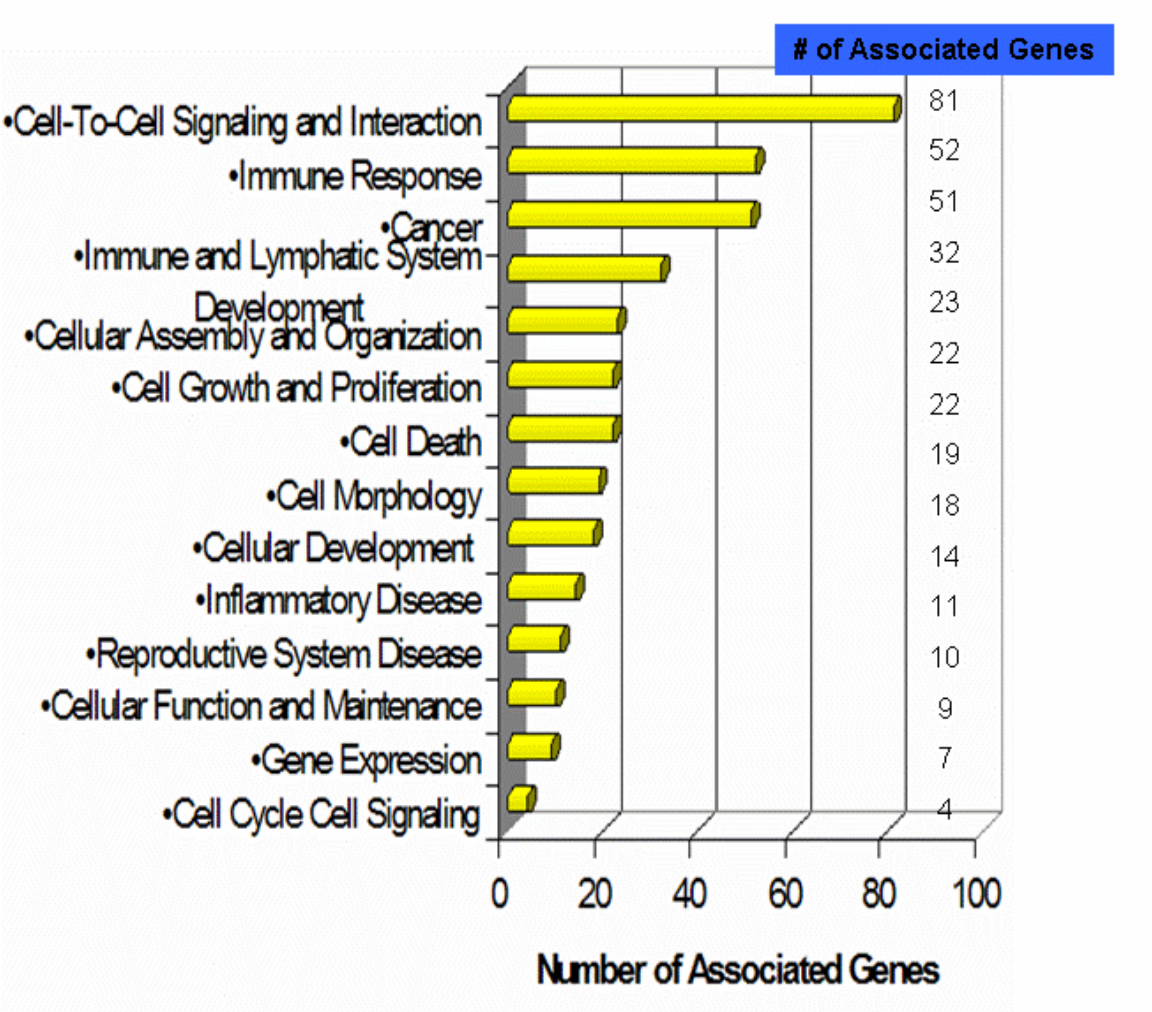

Fig 2.8 Top functional groups and the number of associated genes. Each gene in our dataset may fall within several functional groups and, therefore, will not equal the 98 genes in our dataset.

Next, we used IPA to generate networks that help us identify gene relations within the most significant functional groups. A network is a graphical representation of the molecular relationships between genes/gene products. Genes or gene products are represented as nodes, and the biological relationship between two nodes is represented as an edge (line). Solid lines represent direct interactions and dotted lined represent indirect interactions between the nodes. All edges are supported by at least 1 reference from the literature, from a textbook, or from canonical information stored in the Ingenuity Pathways Knowledge Base. The intensity of the node color indicates the degree of up- (red) or down- (green) regulation. Nodes are displayed using various shapes that represent the functional class of the gene product (i.e., cytokines, enzymes, nuclear receptors, transcriptional regulators, kinases, G-protein coupled receptors, etc).

The top network (with the greatest number of associated genes) was involved in cellular growth and proliferation, cellular movement.

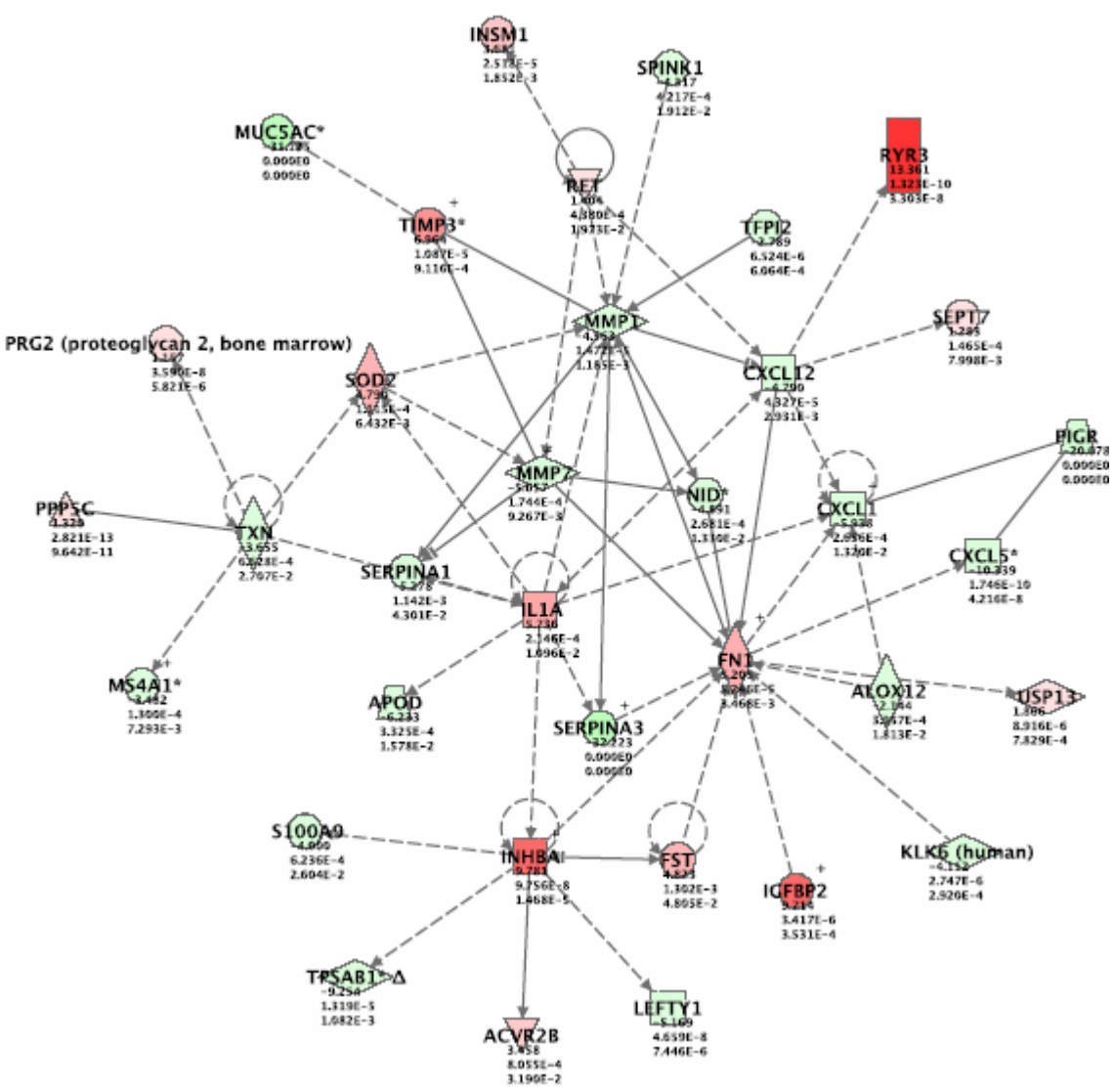

Fig.2.9 Network showing relationships between various genes involved in differentiating superficial and deep cancers.

Based on our stringent statistical methods and IPA analysis as detailed above, we were able to narrow down potential candidates for qRT PCR confirmation. We chose two genes that were down-regulated in deep cancers and two that were up-regulated. Based on our statistical analysis and the IPA analysis, we selected SERPIN B4, CEACAM 7, PTPRF, and Neurofilament for qRT-PCR confirmation.

Serpins

Serpins (serine protease inhibitors) showed up repeatedly in our analysis. There were ten different serpins (A1, A3, A5, B1, B3, B4, B6, B7, B9, and B13) that were all down-regulated in deep cancers compared to the superficial cancers.

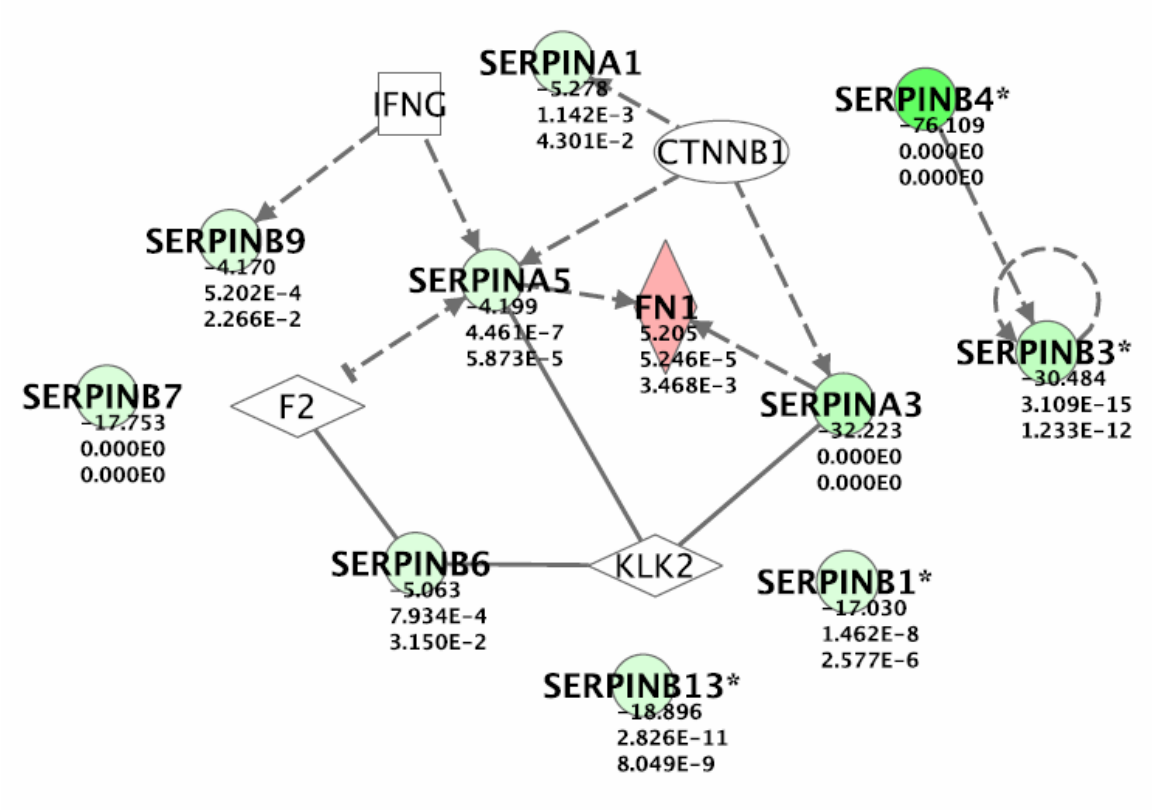

Fig 2.10 IPA Network showing Serpin interactions.

Serpins (serine protease inhibitors) are part of a large super-family of serine and cysteine protease inhibitors found to be involved in the regulation of tumor suppressors and apoptosis. In particular, they have been extensively studied in relation to early stage cervical cancers. Some studies have suggested that serum levels of serpin B3 and B4, also known as squamous cell carcinoma antigen (SCA)-1 and -2, respectively, correlate with tumor bulk [75], and as prognostic indicators of recurrence [76,77]. SCC-2 has been found to regulate cell-to-cell adhesion and growth [78]. Its absence or down-regulation has been associated with a decrease in E-cadherin and an increase in cell migration and

invasion [79]. While serum levels of SCC have shown promise as prognostic indicators of recurrence, there are issues with sensitivity and specificity that have limited its clinical usefulness as a biomarker [80]. Elevated serum levels have been found in cancers besides cervical cancers, renal failure [81], and even in benign diseases such as psoriasis and eczema [82]. Serpin B4 was down-regulated in deep cancers by 76 fold compared to the superficial group at a 99% confidence level.

CEACAM

There were six carcinoembryonic antigen-related cell adhesion molecules (CEACAMS) family members involved in differentiating superficially from deeply invading IB1 SCCA's in our dataset.

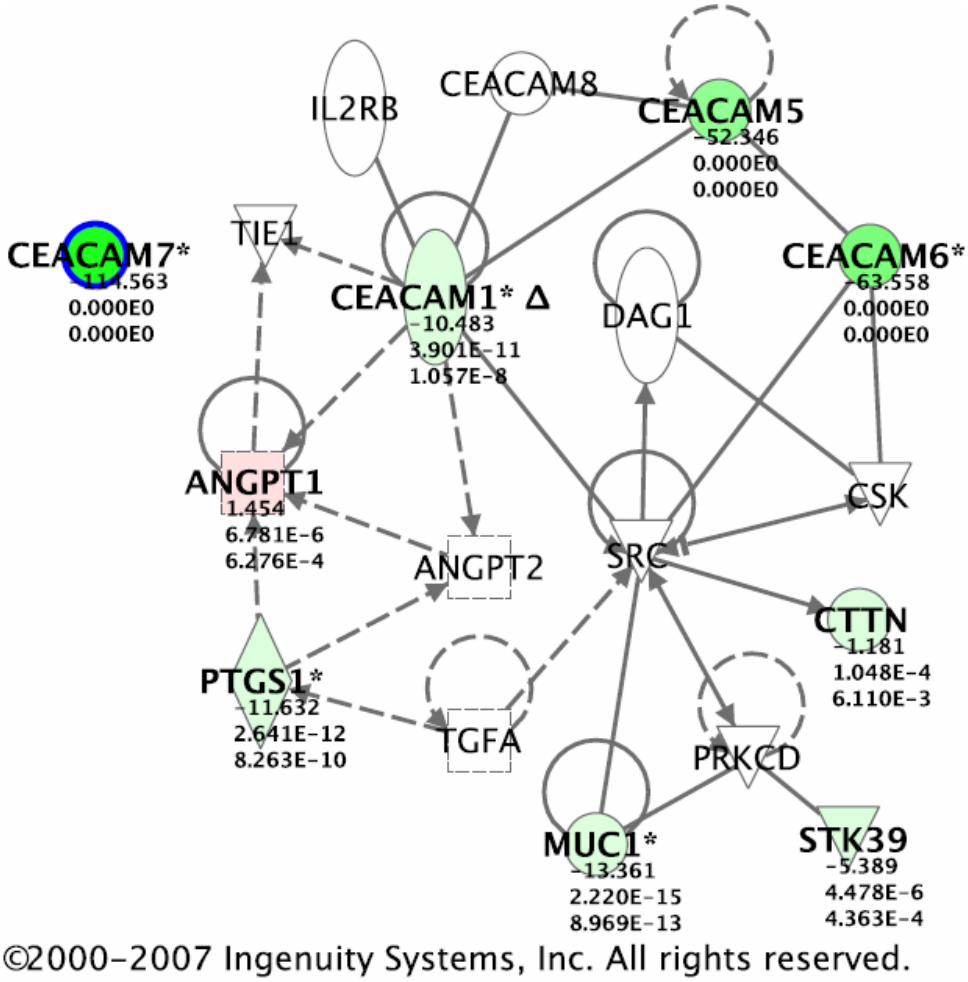

Fig. 2.11 IPA network showing CEACAM interactions. The top functions involved with this network include cellular movement, and growth, proliferation, and cell death.

CEACAMS are epithelial cell, membrane-associated, immuno-globulin-like proteins that have been implicated in cell adhesion and signaling [83]. CEACAMs have been studied as tumor markers for colorectal cancers although they have many of the same specificity and sensitivity issues as found with serpin serum markers [84,85]. CEACAM 7 has been found to be down-regulated in colorectal cancer compared to normal tissue [86], but upregulated in gastric carcinomas [87]. Although the comparison was between superficial and deep cancers, in our study, CEACAM 1, 5, 6, and several variants of CEACAM 7

were all down-regulated in the deep cancers. The down-regulation of CEACAMS may have a role in decreasing cellular adhesion thus facilitate deep stromal invasion.

Neurofilament and Protein Tyrosine Phosphatase, Receptor Type, F (PTPRF)

We also selected genes that were up-regulated in deep cancers. In our ANOVA analysis, we used Bonferoni adjustments to minimize false positives or Type I errors. Neurofilament and PTPRF were both upregulated in deep cancers under this stringent statistical analysis. Neurofilament was found to be upregulated by 22 fold in deep cancers of the cervix. Neurofilament forms the intermediate filaments which provide the tensile strength and rigidity to neurons. While this molecule is associated with certain types of neurological disorders, its role in cancer is not understood. We are the first to describe its presence in deeply invasive carcinomas of the cervix. PTPRF, also known as Liprin beta 1, is a transmembrane tyrosine phosphatase-interacting protein. Like neurofilament, not much is understood about the role this molecule plays in cancer. However, it has been found to play a role in regulating neurotransmitter release. In our analysis, it was upregulated 10 fold in deep cancers. Neither of these molecules is well understood in the context of cancer. For this reason we thought that they might be novel candidates for predicting outcome.

Confirmation with qRT PCR

To validate and confirm our data analysis, we isolated total nucleic acid (TNA) from formalin-fixed, paraffin-embedded (FFPE) archival cervical tissue Eighty percent of patient samples were deeply invasive and twenty percent were superficially invasive according to histologic classification. Patients were grouped as deeply invasive with an unfavorable prognosis or superficially invasive with a favorable prognosis according to their risk of recurrence. Risk factors included LVSI invasion, LN metastasis, depth of stromal invasion, and size of tumor.

qRT PCR Patient Information	Favorable (superficial)	Unfavorable (deep)
Risk Factors	12 Patients	23 Patients
Depth of invasion	(0.4/2.2-0.9/1.8) 18%-50%	(0.9/1.7-3/3) 53%-100%
LVSI	No*	No/Yes
Lymph node metastasis	No	No/Yes

*One patient in this group had LVSI.

Table 2.3 Patient Information for qRT PCR Experiment.

HPV Testing

To determine whether gene clustering varied by HPV type, HPV testing was performed on all samples using the Linear Array Assay (Roche). This assay allowed us to simultaneously test our samples against thirty-seven different types of anogenital HPV DNA genotypes. To test DNA integrity, Beta-globin was used as a positive control.

HPV Testing Results:
35 patients were tested for HPV
31 samples were *B*-globin positive (4 samples were degraded)
24 HPV positive samples (7 samples were HPV negative)
7 HPV 18 positive samples
10 HPV 16 positive samples
7 "other" HPV positive samples
71% of HPV positive samples were HPV 16/18
29% "other" HPV type

Table 2.4 HPV Prevalence

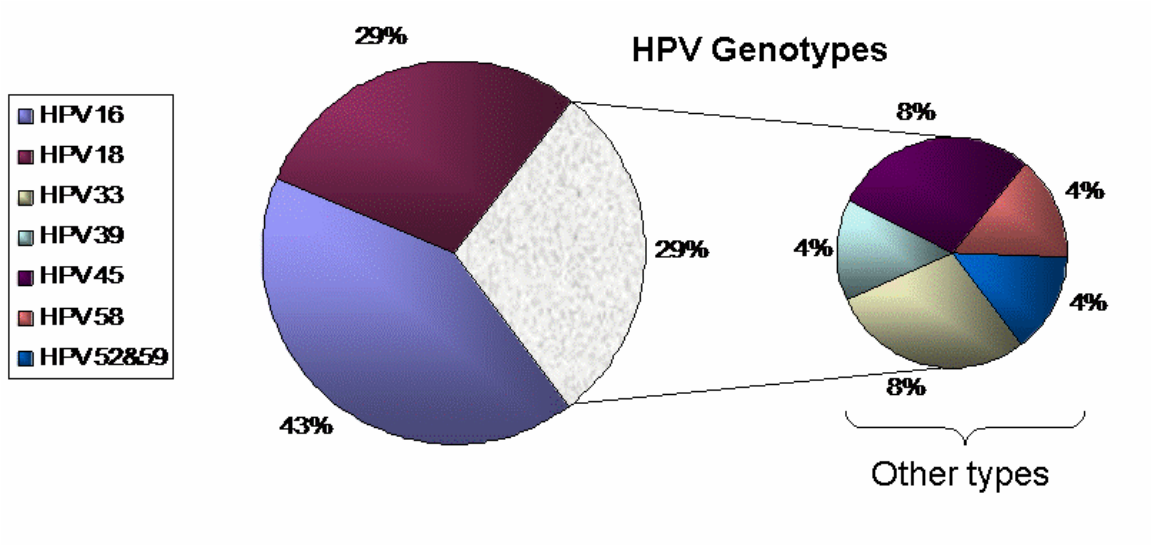

Fig.2.12 HPV genotypes prevalence in our patient population.

Eight different HPV types showed up in our patient group making up 29% of total HPV types. We detected one patient with two different types of HPV (HPV 52 and 59). The prevalence of HPV 16 or 18 was 72% in our patient population.

Endogenous control experiment

In non-cancerous cells "housekeeping" genes are stably expressed and do not vary with the cell cycle or state of differentiation. However, genes from cancer cells, including "housekeeping" genes, are often aberrantly expressed. In order to determine the best endogenous control for our sample population, we isolated RNA from the tumor of 4 patients with stage IB1 SCCA and screened them against 11 different endogenous controls. Quantification of endogenous control genes are based on qRT PCR fluorescent values.

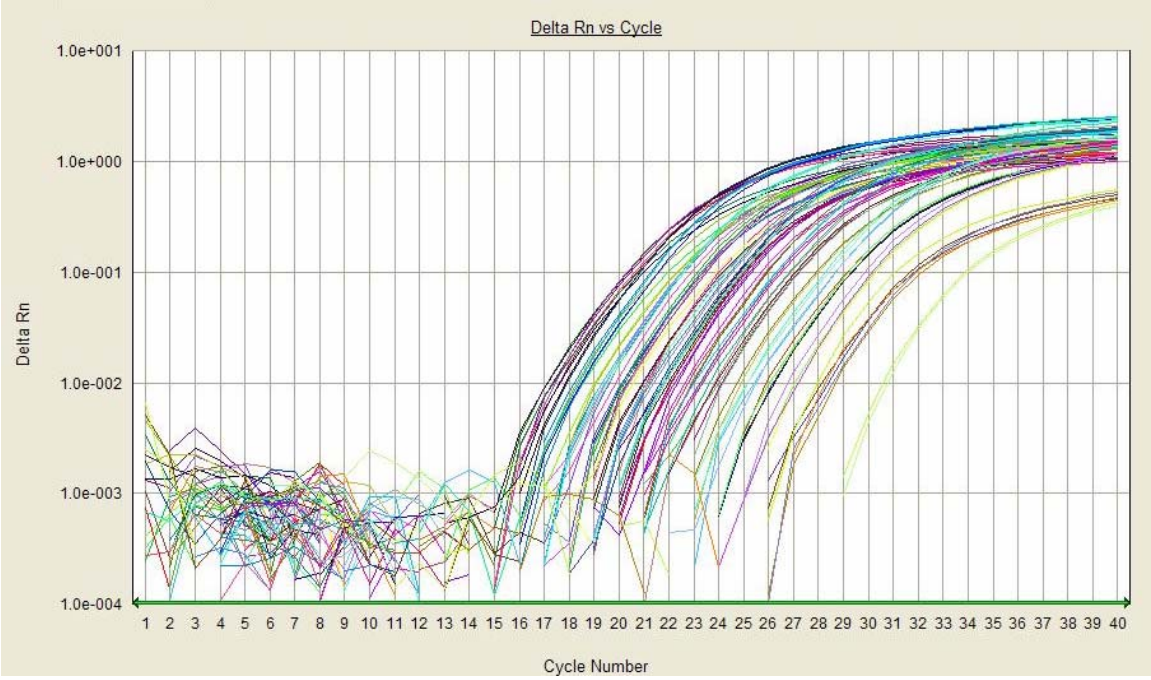

Fig 2.13 Amplification Plot of Endogenous Control Experiment.

The amplification plot represents the raw fluorescent data for all endogenous controls. A threshold is automatically set for each gene and C_T values are calculated for each value. Standard deviations are calculated from all four patient samples.

Endogenous Control	Average CT	Standard Deviation
18s	11.4	+/-0.35
Phosphoglycerokinase (PGK)	24.52	+/-0.40
Glyceraldehyde-3-phosphate dehydrogenase (GAP)	21.69	+/-0.72
Acidic ribosomal protein (PO)	23.03	+/-0.79
Cyclophilin (CYC)	28.81	+/-0.98
Beta-actin (BA)	21.22	+/-1.07
Transcription factor IID, TATA binding protein (TBP)	31.08	+/-1.13
B-Glucronidase (GUS)	27.59	+/-1.20
B2-microglobulin (B2m)	21.63	+/-1.35
Hypoxanthine ribosyl transferase (HPRT)	27.22	+/-1.56
Transferrin Receptor (TfR)	25.85	+/-1.79

Table.2.5 Endogenous Control Optimization. The standard deviations (highlighted in yellow) for each endogenous control are listed from lowest to highest. 18s has the lowest standard deviation, followed by PGK, etc,.

Based on the standard deviation between all of our samples, 18s is the best choice as an endogenous control for use with SCCA's of the uterine cervix.

QRT PCR results

Samples were processed using Applied Biosciences Quantitative Real time Fast PCR system [*See methods section for details*]. Relative quantification values were based on each individual samples expression level versus its endogenous control, 18s.

Based on our GeneChip analysis, we expected that we would be able to differentiate superficially invading IB1 SCCA's from deeply invading IB1 SCCA's. Additionally, we expected CEACAM and Serpin to be lower in aggressive, deeply invading IB1 SCCA's compared to non-aggressive, superficially invading IB1 SCCA's. Additionally, we expected Neurofilament and PTPRF to be expressed at a higher level in

deeply invading IB1 SCCA's compared to superficially invading tumors. Based on our expression values relative to 18s, the following results were obtained.

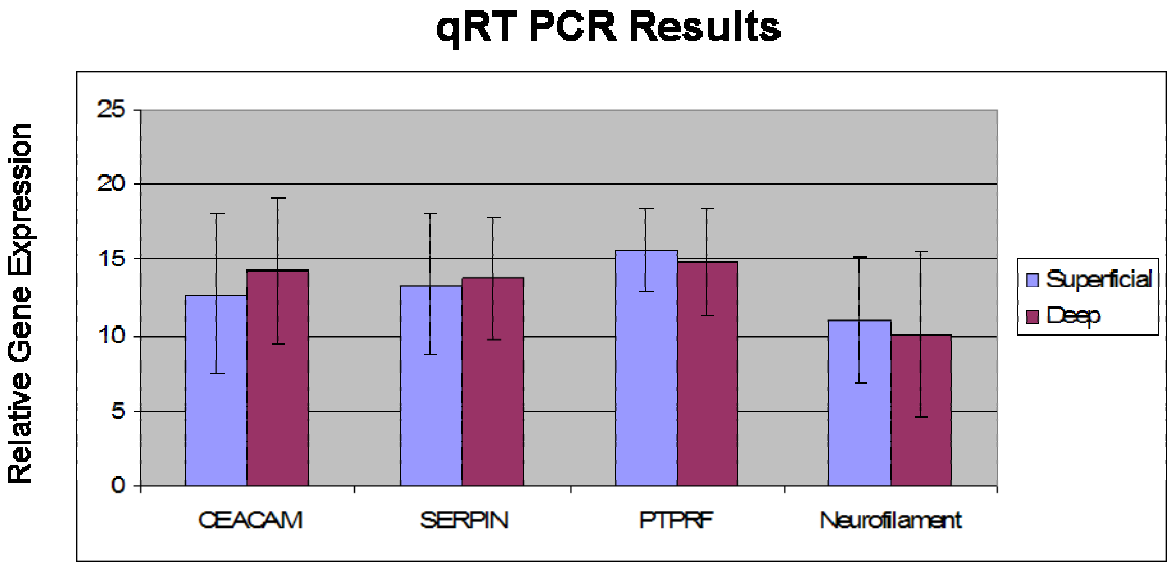

Fig. 2.14 Relative Gene Expression (delta C_T) based on gene of interest versus 18s.

There was no significant differences between in the relative gene expression of CEACAM, Serpin, PTPRF, and Neurofilament in the superficially invasive and deeply invasive IB1 SCCA's. Additionally, CEACAM and Serpin were not expressed at a lower level and over-expressed in PTPRF and Neurofilament in our deeply invading IB1 SCCA's as expected.

Univariate Analysis

CA-125, a serum marker for ovarian cancer in the presence of an adnexal mass, has a sensitivity of 56%-81%, a specificity of 75-99% and a positive predictive value ranging from 58-93% (age dependant) [88-90]. In detecting CIN 2, the pap smear has a 76% sensitivity and 95% specificity [91]. We wanted to determined the specificity, sensitivity, and overall positive predictive values (PPV) for each of our genes of interest.

	CEACAM	SERPIN	NEFL	PTPRF
Sensitivity	50%	50%	41%	67%
Specificity	67%	54%	33%	17%
PPV	29%	22%	70%	76%

Table 2.6 Univariate Analysis. PPV is Positive Predictive Value.

PTPRF has the best PPV value (detection of true positives) and sensitivity (percentage of tumor patients correctly identified by a positive test) out of all the genes, but has the lowest specificity (percentage of tumor patients correctly identified by a positive test).

Mutivariate Analysis

The Principal Component Analysis (PCA) in Spotfire DecisionSite 8.2 was used for clustering algorithms. The PCA method reduces the dimensionality of the data creating linear combinations of the original data. It takes non-linear data and linearizes it so that meaning can be obtained from the dataset. The results are displayed in a scatter plot that maps the principal component score of each patient sample. The position along a certain axis (PCA1, PCA2, or PCA3) represents the score of the component (patient sample value).

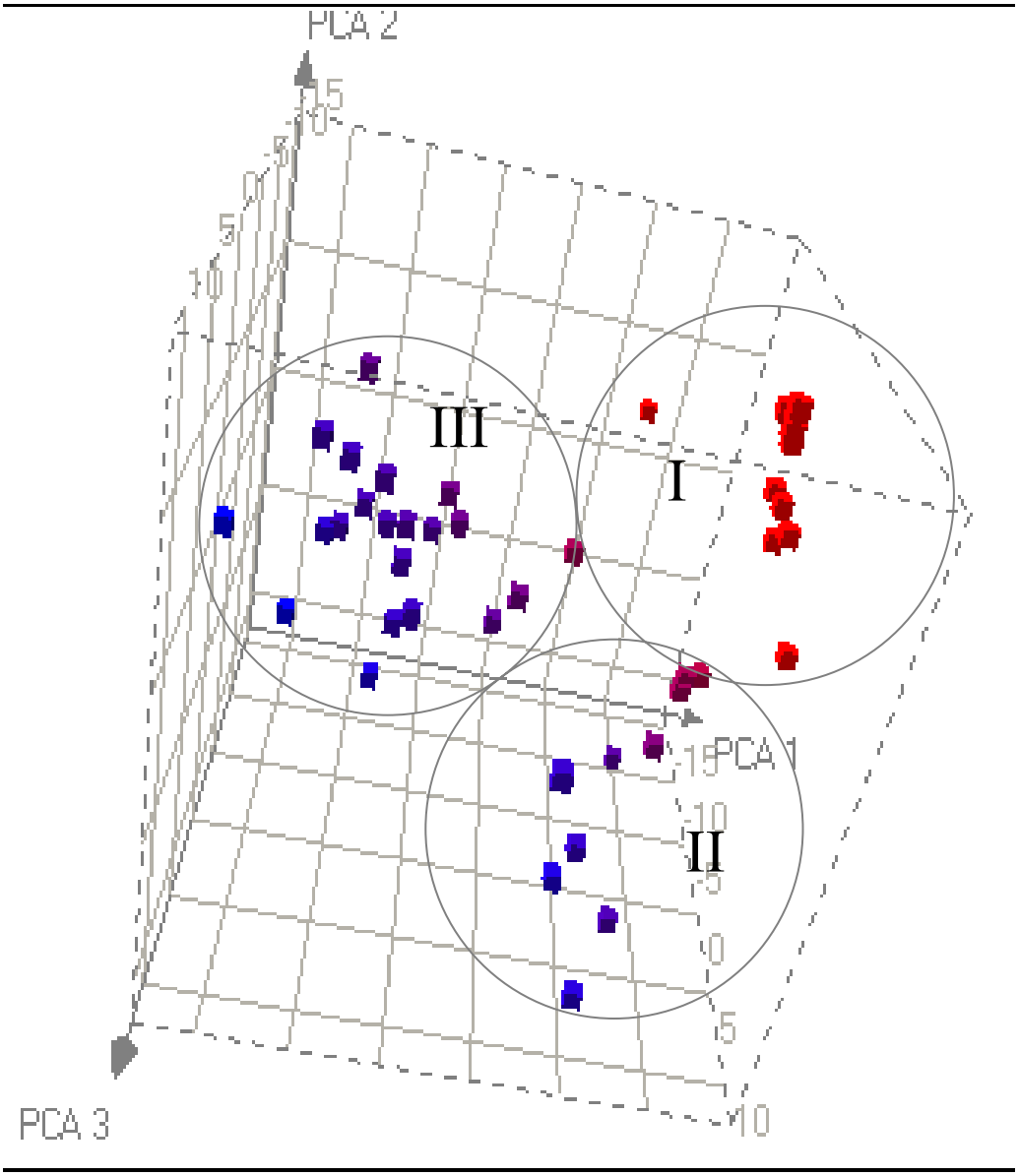

Fig. 2.17 PCA analysis showing clustering of patient samples.

The multivariate PCA analysis resulted in a 3 dimensional view of our dataset. Since the data values did not cluster at zero along the three PCA variability dimensions, the patients were heterogeneous. However, the PCA analysis showed us that our patients did segregate into three distinct clusters. There were a mix of superficially and deeply invasive IB1 SCCA's in groups I and III, but group II contained only deeply invading IB1 SCCA's.

Risk Factor by Group

We wanted to know whether our cluster analysis correlated with any other factors such as depth of stromal invasion, LVSI invasion, lymph node (LN) metastasis, and tumor size. We divided our patients on the basis of superficially or deeply invading patients within each of the clustering groups (I, II, or III).

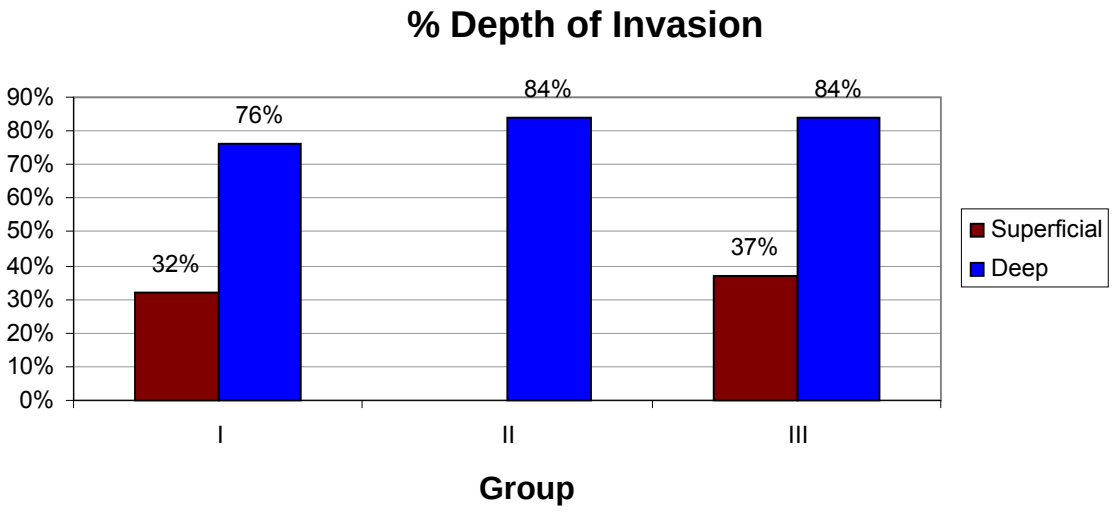

Table 2.7 Average percentage depth of invasion. Calculations are based on the measurement of the tumor from the epithelial-stromal junction to the deepest point of stromal invasion relative to the depth of total stroma. Group II had no superficial patients.

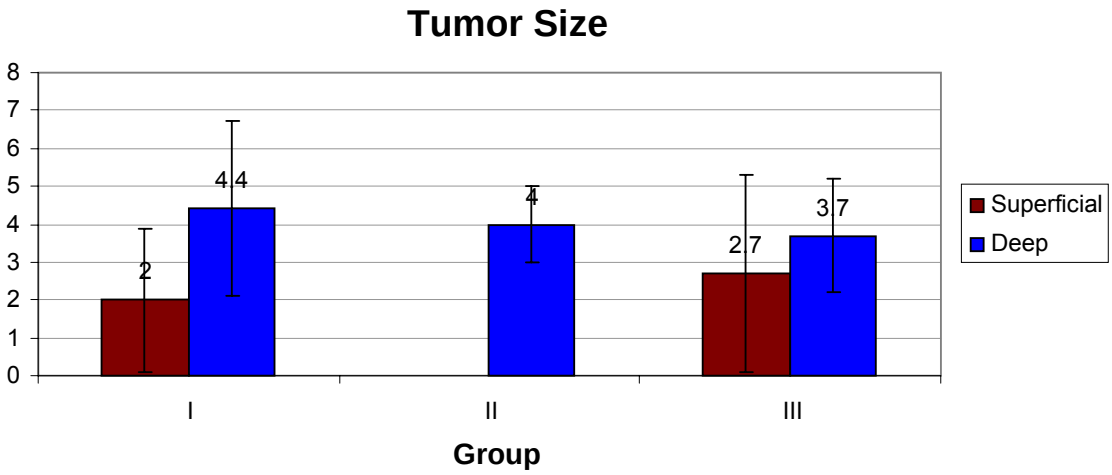

Table 2.8 Average tumor size (in centimeters). Measurements were taken at the largest tumor dimension.

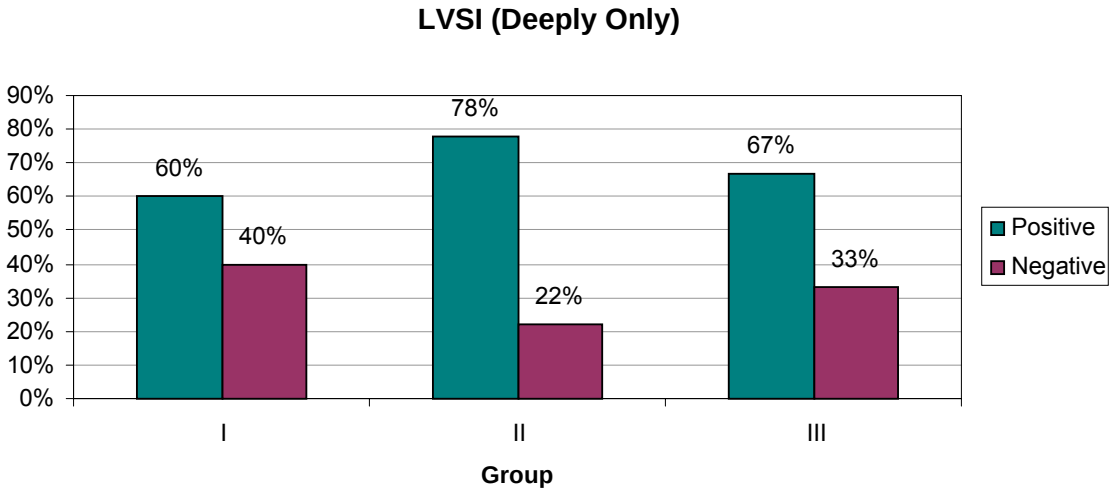

Table 2.9 Lymphovascular Space Invasion (LVSI) in our deeply invading samples. Percentages are based on total deeply invasive patients in group.

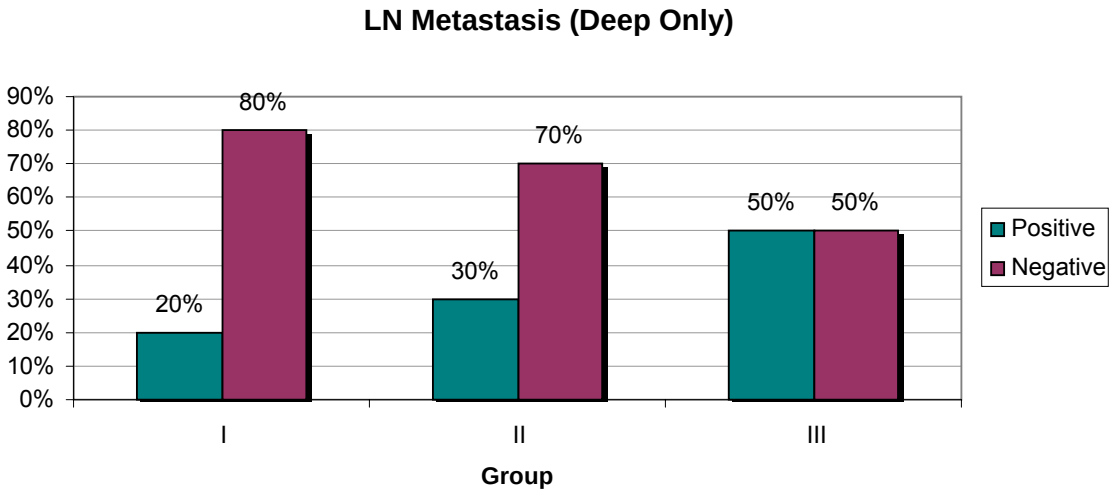

Table 2.10 Lymph node metastasis in our deeply invading samples. Percentages are based on total deeply invasive patients in group.

PCA Results Summary

	42 samples total (33DC:9SC)	Group I N=9 (5DC:4SC)	Group II N=10 (10DC)	Group III N=23 (18DC:5SC)
DOI	Deep	76%	84%	84%
	Superficial	32%	N/A	37%
Size	Deep	4.4+/-2.3	4.0+/-1.0	3.7+/-1.5
	Superficial	2.0+/-1.9	N/A	2.7+/-2.6
LVSI	Deep	60%	78%	67%
LN	Deep	20%	30%	50%

Table 2.10 Breakdown of risk factors by group. DC=Deep cancers; SC=Superficial cancers

These clusters did not appear to correlate with any of our known risk factors. This correlation could be related to some risk factor not yet described.

DISCUSSION

Several studies, have used microarray GeneChip technology to analyze transcript expression in cervical cancer [92,93]. However, these studies have focused on differentiating genes between normal and cancer or dysplasia and cancer. Our approach was to determine risk of recurrence in stage IB1 SCCA's. While the standard of care for cervical cancer is clear for treating Stage I and Stage IIB to IVB cervical cancers, Stage IB1 cervical cancers are not so clear cut. Fifty percent of Stage IB1 SCCA's have a high risk of recurrence and 50% have a low risk of recurrence. Using GeneChip technologies, we distinguished stage IB1 SCCA's with a low-risk of recurrence (superficially invasive) from those with a high risk of recurrence (deeply invasive) in cervical biopsies obtained preoperatively.

While these results have implications in decreasing morbidity and mortality of multi-modality radical treatments, our GeneChip study was limited to patients at the polar opposite of the stage 1B1 SCCA spectrum; superficial stromal invasion with no lymphovascular space invasion compared to deep stromal invasion with lymphovascular space invasion. These patients were highly selected and may not representative of the population as a whole. Additionally, in this preliminary analysis, we used a limited number of patients (n=2 in each group) for our GeneChip analysis. Using microarray GeneChip technology, there is a huge disproportion of variables tested (genes) compared to number of samples. One limitation of having such few patients involves the over-fitting of data in the statistical analysis [94]. This dataset needs to be further expanded and validated in a larger, heterogeneous population of stage IB1 cancers of the cervix.

The platform used in this analysis was Affymetrix GeneChips. GeneChip technologies allow the user to screen thousands of genes simultaneously. Additionally, it

provides a snapshot of the orchestration of molecular events involved in the tumorogenesis process. Identifying a molecular profile of cancer using GeneChip technology has the potential to predict future biological behavior. There are several drawbacks to using this platform. Affymetrics GeneChips are costly, require a large amount of high quality, RNA, and there is no gold standard for the statistical analysis of the data. We used fresh tissues for the GeneChips. At our institution, there are approximately 3 to 4 stage IB1 SCCA's scheduled for surgery each month. Not all patient samples are sufficient for our study because of RNA quality/quantity. Therefore, it is difficult to obtain enough patients in a relatively short time period. The alternative is to use formalin-fixed paraffin-embedded archival tissues. However, for Affymetrix platforms, the RNA quality and quantity from paraffin-embedded tissue is insufficient. Using GeneChips, we were able to identify 98 potential gene candidates that distinguish stage IB1 SCCA's with a low risk of recurrence from those with a high risk of recurrence. For validation purposes, it is advantageous to narrow down the list of gene candidates even further.

The purpose of narrowing the list of potential gene candidates down to just a few putative biomarkers has the advantage of reducing validation time and cost significantly. Currently, there is no standardized method for narrowing down candidate biomarkers. Because fresh biological specimens are limited, for verification, we used paraffin embedded tissues. To verify the GeneChip data, candidates were narrowed down using stringent statistical methods. Genes that were differentially expressed below a 99% confidence level were eliminated. The remaining gene candidates were further filtered using Bonferoni adjustments to minimize false positives. Their biological function was determined using IPA software and candidate markers were picked based on biological significance in the literature. While, we could differentiate patient samples from this limited sample set using GeneChip analysis, we were unable to correlate risk of recurrence using individual genes (CEACAM 7, Serpin B4, Neurofilament, and PTPRF) in a larger set of paraffin-embedded tissues. Based on the univariate analysis of each of

the gene candidates (CEACAM, Serpin, PTPRF, and Neurofilament), our sensitivity, specificity, and positive predictive values were insufficient to apply to large scale validation studies.

One explanation of why our GeneChip results could not be validated may be due to the differences in our patient population. While all tumors were stage IB1 SCCA's, they reflected the entire spectrum of this stage in comparison to the highly selected tumors for the GeneChip experiments. There were patients with mixed risk factors. For example, several patients had deeply invasive tumors with no LVSI and others were superficially invasive with positive LVSI.

Patient Population		Depth of Invasion	LVSI	LN
GeneChip	Superficial	5% - 10%	no	no
	Deep	85% - 91%	yes	no
qRT PCR	Superficial	18%-50%	no	no
	Deep	53%-100%	no/yes	no/yes

Table 2.13 Patient differences between GeneChip and qRT PCR experiments.

Using an unsupervised cluster method (PCA analysis), we found that the patients in this study clustered into three distinct groups contrary to expectations. Stage IB1 SCCA's are currently grouped according to presence of certain risk factors of recurrence (high and low risk). Our PCA data indicate that perhaps there is an additional group with an as yet described risk of recurrence.

Seventy-two percent of the patients in this study were HPV 16 or 18 positive. These results are consistent with the current literature [95]. Using DNA from archival FFPE tissues, 89% of samples were positive for beta-globin. Beta-globin was used as an indictor of DNA fidelity. The DNA was isolated from archival tissues from as far back as 2001 and may have degraded. Thirty–three percent of the beta-globin positive samples were positive for HPV. This is slightly lower than expected according to epidemiology

studies. Several explanations could be possible. The HPV levels might be below the detection limit of the assay. The Roche Linear Assay does use a pre-amplification step, but the transcripts are detected by an enzymatic colorimetric reaction. Secondly, the Roche Linear Array Assay was developed for use with exfoliated cells. We applied the assay to DNA extracted from paraffin embedded archival tissues. Although the assay worked, it may not be optimal for this type of tissue. RNA from the same tissue samples amplified using ribosomal 18s in our real-time experiments suggesting that the 450 base pair target used in the Linear array assay is at the upper limit of what can be expected from this kind of material.

Thus far, there is no one biomarker for cancer predictive of disease or outcome. However, in combination, unique genetic profiles have been used to selectively diagnose early stage disease. One such study has discovered a set of four serum markers that can differentiate patients with ovarian cancer from normal patients. While no one of these proteins were able to distinguish ovarian cancer patients from normal controls, a combination of all four resulted in a 95% sensitivity and specificity, and a 94% positive predictive value [98]. The clinical efficacy of this study is yet to be determined, but it does illustrate the importance of combinations of proteins for biomarker discovery.

Molecularly-specific target based therapies have the potential to guide treatment decisions. Several new target-specific drugs have recently received Food and Drug Administration approval. Iressa is a tyrosine kinase inhibitor that blocks the epidermal growth factor receptor (EGFR) has been approved for use in non-small cell lung carcinomas [96]. Herceptin, an anti-body specifically is currently being used in women with breast cancer over-expressing the HER-2 receptor [97]. While these drugs are advantageous for certain patients over non-specific treatments such as radiation and chemotherapy, they are limited to targeting a single molecule. Because cancers are unstable and mutate quickly, they can become resistant to a single treatment. Using GeneChip technology, it is possible to identify several cancer targets for future multi-

drug development. The age of molecular medicine offers the ability to predict prognosis, drug response, and tailor treatment options to best fit patients on an individual basis.

CHAPTER 3: MOLECULAR ANATOMY OF STAGE IB1 SQUAMOUS CELL CARCINOMAS OF THE UTERINE CERVIX

INTRODUCTION

Little is known about the molecular processes that lead to cervical carcinogenesis. However, transformation does begin at the squamo-columnar junction (SCJ) on the external opening of the cervix. This hotspot of transformation is commonly referred to as the Transformation Zone or "T-zone". Early abnormal, dysplastic cells form at the squamous columnar junction on the exocervix and spread up the columnar cells of the endocervix. Cancer occurs once the basement membrane is compromised and cancer cells start to spread into the stroma.

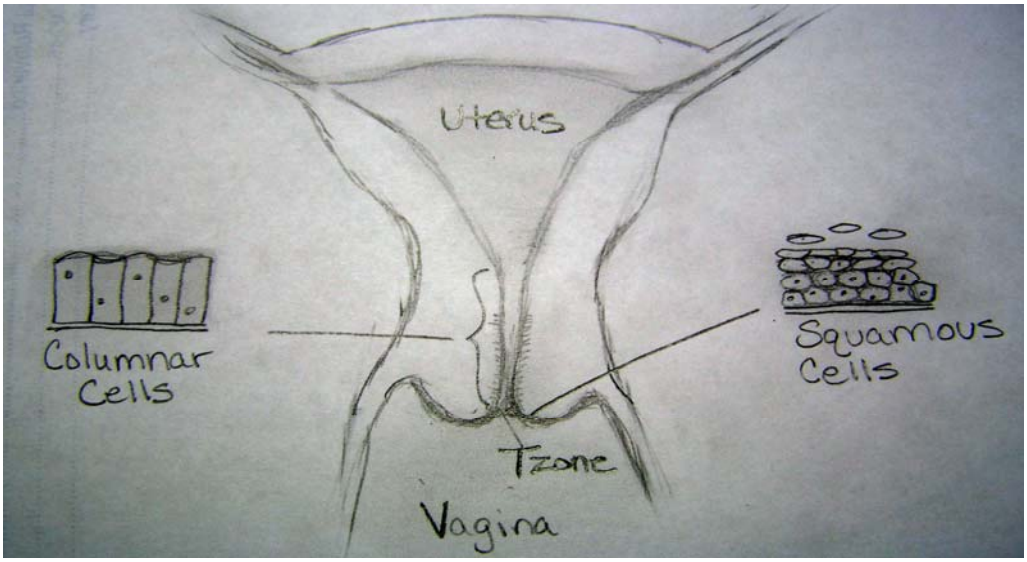

Fig.3.1 Diagram of Cervical Anatomy

BACKGROUND AND RATIONALE

Characterizing the potential for reoccurrence is currently an imprecise science. Clinical staging lacks the precision needed to characterize outcome and thus determine the best treatment option for patients. While microscopic classification is more precise, there are still issues that must be overcome. Molecular profiling is the best tool we have in our repertoire, but we still have a long way to go in order to make sense of the vastness of the genetic information provided by these methods. Molecular markers have proven beneficial in the diagnosis of haematologic cancers [99,100] and solid tumors [101,102] and [*if applied to cervical cancer*], have the potential to allow for more accurate staging and more selective management of early stage cervical cancers.

In chapter two, we were able to identify a set of 98 genes that distinguished low and high risk in IB1 SCCA tumors of the cervix. In order to tease out genes that may typify the metastatic process, a few gene candidates were validated in a larger patient study. While gene patterns were similar, there were distinct expression profiles that varied by patient. Genetic and/or functional differences may account for patient to patient variability. However, in a heterogeneous tumor, differences may also be a result of site of biopsy within a tumor. Microscopically patients within stage IB1 SCCA have very different tumors. Genetically, this is also the case. Several studies have established loss of heterogeneity, chromosome aberrations, and multiple HPV oncoprotein variants that fluctuate throughout cervical carcinomas [103-106]. Thus, one biopsy may not represent the whole tumor. Patient to patient variability may account for some of the variability; however intra-tumor heterogeneity may also be the culprit.

Microscopic Mechanism of Spread

Early stage IB1 cervical cancer presents itself in several different ways microscopically. Typically, abnormal dysplastic cells originating at the SCJ spread in an expansive fashion involving the exocervix and endocervical canal. Invasive carcinoma develops once the basement membrane is compromised. In order to spread to lymph

nodes or distal sites, cancer cells must obtain the ability to intravasate into the lymphovascular space. Clinicians are often faced with a tumor no more than 1mm depth of invasion that has positive LVSI or even lymph node metastasis. By FIGO convention, this size of a tumor should not require adjuvant therapy because of the low risk of recurrence, but any risk of spread beyond the cervix warrants follow-up chemo- and radiation therapy. Thus, early stage IB1 SCCA's have a variety of phenotypic differences and clinical outcomes. Therefore, in order to elucidate the importance of these variables and to delineate clinical mechanism of progression, it is important to characterize these differences at a molecular level.

Multiple Molecular Events

Becoming a Tumor

Multiple molecular events occur in the progression to cancer. Although cell cycle dysregulation is affected by HPV oncoproteins, the metastatic phenotype is dependent upon other genetic alterations. These mechanisms need to be better understood, in order to elucidate the complex processes of tumor cell invasion and metastasis. One of the critical events whose mechanism should be understood in cancer cell development is the transition from a benign to an invasive tumor that facilitates its "escape". An early event in this process is the compromise of the basement membrane. This process involves the interaction between the tumor cell and the basement membrane. Although the exact mechanism is not well understood, the tumor cell attaches to the underlying basement membrane, degrades it with proteases, and migrates through it into the underlying stroma [107].

Escaping the Cervix

A separate molecular hallmark occurs after the basement membrane is compromised. In addition to the cells invading the underlying stroma, an aggressive tumor has the capacity to intravasate into the endothelial-lined lymph-capillary channels known as the lymphovascular space, travel in the circulation to a lymph node or distant organ, and extravasate at the distant site. Tumors that break through the basement

membrane do not necessarily have the ability to invade the lymphovascular space. These tumors will stay localized within the cervix. Intravasation into the lymphovascular space involves molecular events that are distinct from invasion through the basement membrane. Even though tumors may look similar microscopically, it is crucial to determine a molecular signature of genes that regulate metastasis before metastasis ever occurs.

Tumor Heterogeneity

Much like the limitations of culture of a single cell derived from a tumor, one biopsy from a heterogeneous tumor may not be a sufficient representation to infer about the molecular processes of the whole tumor. Although the majority of cervical tumors are moderately differentiated squamous cell carcinomas, different patterns can be identified at the phenotypic level.

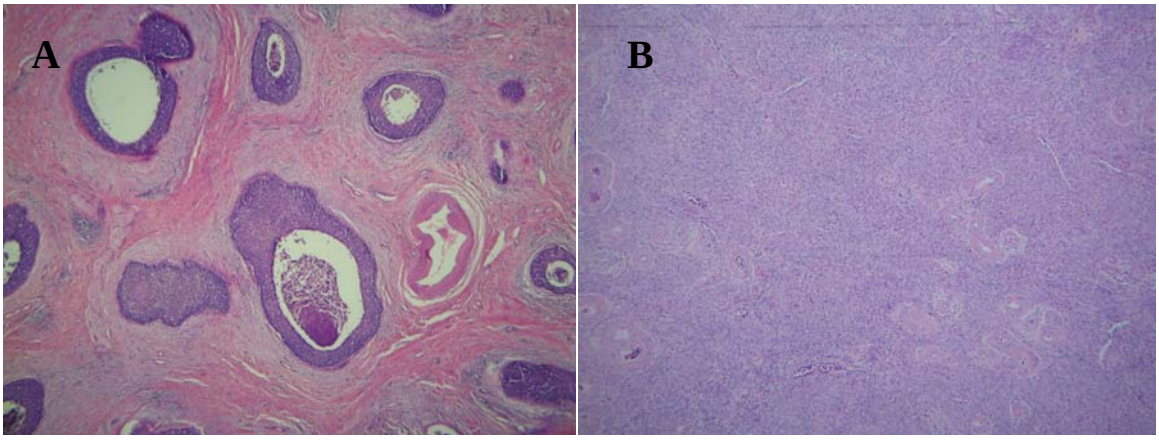

Fig.3.2 H & E slide from two different IB1 SCCA's[I]. Slide "A" contains focal islands of tumor (purple) among atypical stroma. Slide "B" is a more homogeneous tumor with little stroma.

Additionally, each individual tumor can be quite heterogeneous microscopically (intra-tumor heterogeneity). Yet, these differences seen at the microscopic level are grouped under a single stage of disease (IB1) and, based on clinical outcome, are not

[I] Photographs kindly provided by Dr. Concepcion Arrastia, MD.

representative of the biology within various areas of the tumor. This is reasonable considering tumors originate from a common parental cell. Through a multi-step process, cancer cells incur genetic aberrations through each concurrent cell division resulting in a genetically related, but heterogeneous mix of cells. For example, at the leading edge of a tumor, the cells, although derived from a common lineage, are quite different from the parental cancer cell and presumably posses different abilities to grow, evade, and invade. Oxygen deprivation, cells of the immune system, etc., provide selective pressures that change genetically unstable cancer cells. Pro-angiogenic proteins are up-regulated and pro-apoptotic genes are down-regulated ultimately leading to progeny cancer cells with a growth/survival advantage. Thus, when using a single tumor biopsy for molecular profiling and tumor characterization, the site of biopsy may be an important variable.

OBJECTIVE

In order to develop a molecular signature of early cervical cancer, the optimal biopsy site and the genetic differences within the tumor must be clarified. Our objective is to identify the genetic variation within a cervical tumor, and to evaluate the effect of sampling of gross tumor in a genomic analysis of cervical carcinoma.

HYPOTHESIS

When using molecular profiling to characterize tumor function and behavior, the site of biopsy is an important variable.

SPECIFIC AIM

To define the molecular profile of heterogeneity in early stage IB1 SCCA's.

METHODS

Radical hysterectomy specimens from patients with stage IB1 squamous cell carcinomas of the cervix were evaluated. The specimens were sectioned and grossly examined.

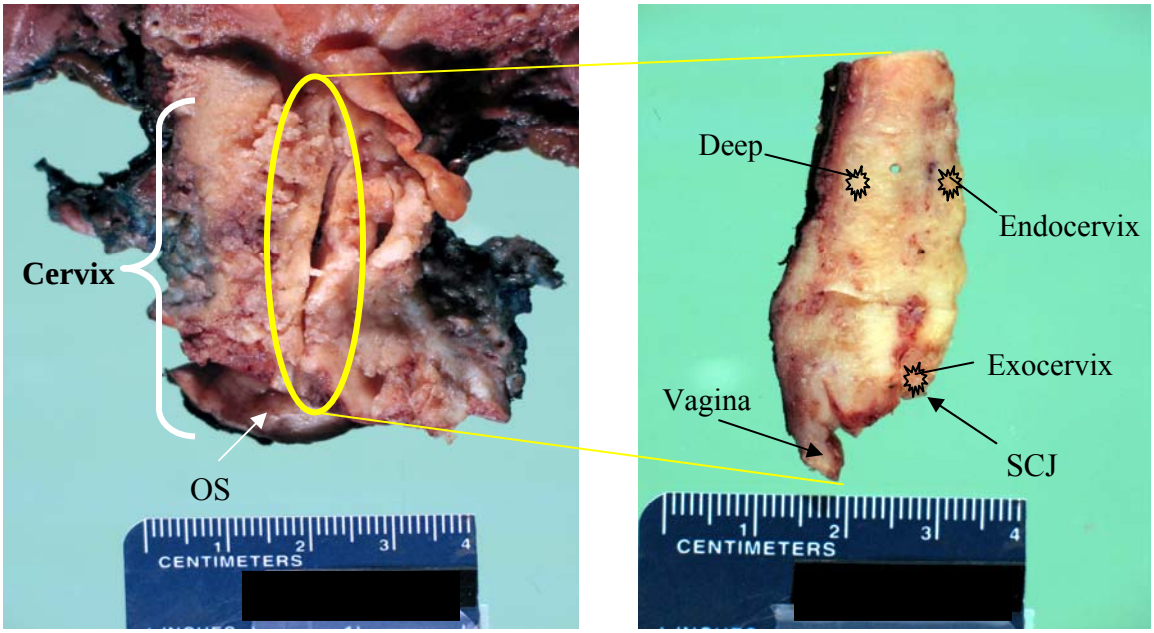

Fig.3.3 Uterine cervix with IB1 SCCA. Left side shows the cancerous cervix with a slice removed. Right side is the tissue slice showing locations of the biopsy sites[J].

Representative biopsies were obtained from 3 sites within the tumor: 1) site of presumed carcinogenesis at the squamo-columnar junction (SCJ), at the exocervix; 2) site of tumor extension from the endocervical epithelium; and 3) site of deepest area of cervical stromal invasion. Half of each biopsy specimen was snap frozen in liquid nitrogen for molecular analysis and half submitted for histopathologic evaluation. Three biopsies were taken from 3 patients for this analysis.

[J] Photographs kindly provided by Dr. Claudia Castro, M.D.

Patient #	Depth of Invasion	LVSI	Margins	Nodes
103	1.1/1.1cm (100%)	+	+	+
327	1.2/1.8cm (67%)	+	-	-
340	1.5/1.5cm (100%)	+	+	-

Table 3.1 Patient information (Heterogeneity study).

Upon histopathologic review, we selected tissues with greater than a seventy percent carcinoma: stroma ratio, less than 5 percent a minimal amount of lymphocytic infiltrate, and less than two percent necrosis.

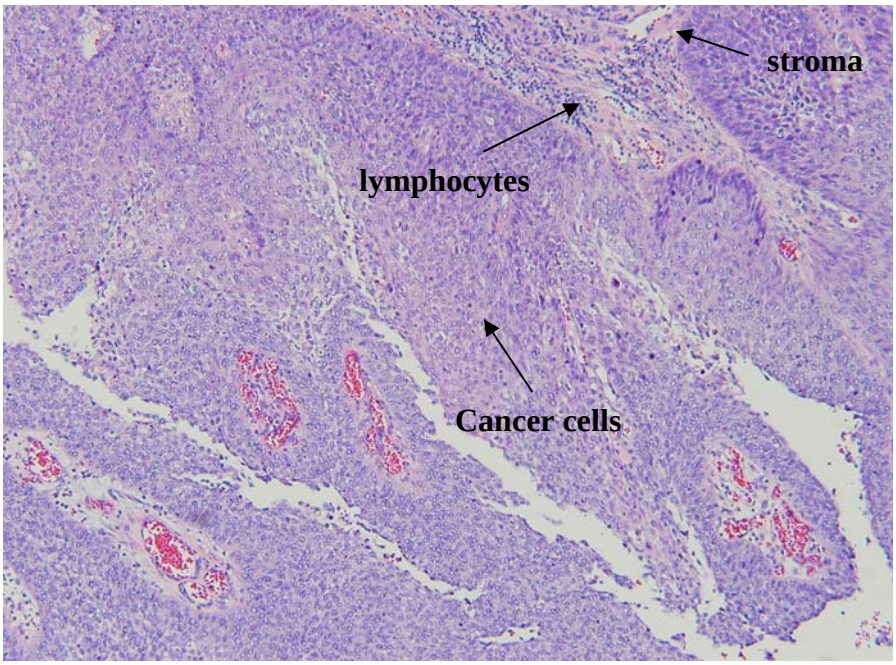

Fig.3.4 Haematoxylin and Eosin (H&E) stain of cervical section showing stroma (light pink areas), lymphocytes (dark blue spots), and cancer cells(light purple areas)[K].

RNA Extraction

Snap frozen biopsies were stored in the gas phase of liquid nitrogen until nucleic acid was extracted. DNA-free, RNA was extracted using the RNAqueous-4PCR (Ambion, Austin TX) kit. Briefly, RNA was crushed in liquid nitrogen, lysed, washed, and treated with DNase. The quality and quantity of RNA was assessed using an Agilent 2100 Bioanalyzer. ,One ul of RNA was labeled with a fluorescent gel-dye mix (Agilent RNA 6000 Nano kit). The RNA was resolved on a gel matrix and RNA quality and quantity was calculated. Only RNA with the highest fidelity was used for microarray analysis. Twenty ug of DNA-free RNA from each biopsy was applied to the GeneChip Human HG_U133 plus 2.0 (Affymetrix) [*see chapter 2 "Methods" for a detailed overview]*.

[K] Photograph courtesy of Dr. Concepcion Arrastia, MD.

Statistical Analysis

The CEL files were imported in S+AA for probe level analysis. The Robust Multi-chip Analysis along with quantile and medianpolish method was performed for background correction, normalization, and summarization respectively. The probe sets absent across all the chips were also filtered out. One way analysis of variance (ANOVA) was used to filter genes that were differentially expressed between the 3 locations (exocervix, endocervix, and deep stromal invasion) within the tumor from 3 study subjects. Additionally, we performed one way ANOVA to filter genes by patient. Spotfire DecisionSite 8.2 was used for clustering algorithms. Ingenuity Pathways Analysis Software was used to identify important networks, functional groups and pathways that are affected due to these genes.

No Change Controls

Perilesional biopsies were obtained from women who underwent a radical hysterectomy and pelvic lymphadenectomy for stage IB1 squamous cell carcinoma (SCCA) of the cervix (IRB#02-272).

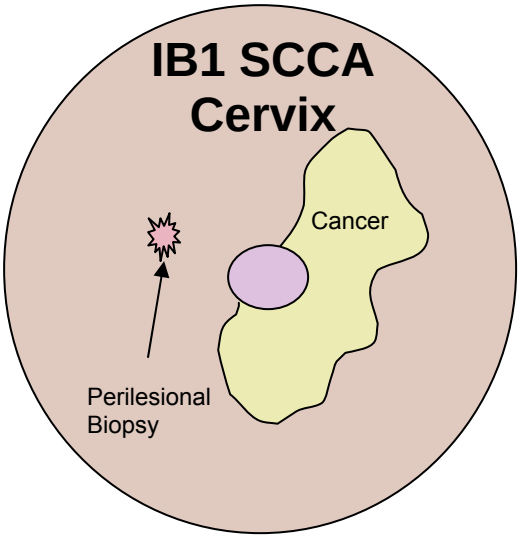

Fig 3.5 Perilesional Biopsy

Biopsies had to have good quality 18s and 28s ribosomal RNA, and sufficient (24 ug) RNA had to be present in the biopsies for the Affymetrix GeneChips. Tumor biopsies were transported from the operating room to the laboratory in 10 volumes of RNA Later (Cat# 7020, Ambion, Austin, Texas). Following the manufacturers instructions, biopsies were stored at 4^{o}C overnight, decanted and stored at -80^{o}C. RNA was extracted from 4 perilesional biopsies. Perilesional tissues were processed the same as the corresponding biopsies from the tumor tissues [*Chapter 2, methods*]. Additionally, data obtained from the 4 normal biopsies probe sets from women undergoing a hysterectomy for benign reasons (chapter 2) were used to compare probe sets from the exocervical biopsies.

RESULTS

Expression Patterns in a Heterogeneous Tumor

We used one-way ANOVA to determine whether genes varied by location within the tumor. The one way ANOVA filtered 157 genes at p-value$\leq$0.05 that were differentially expressed between the 3 locations in a tumor (exocervix, endocervix, and deep). Hierarchical clustering helped to visualize the changing gene patterns in a heterogeneous tumor.

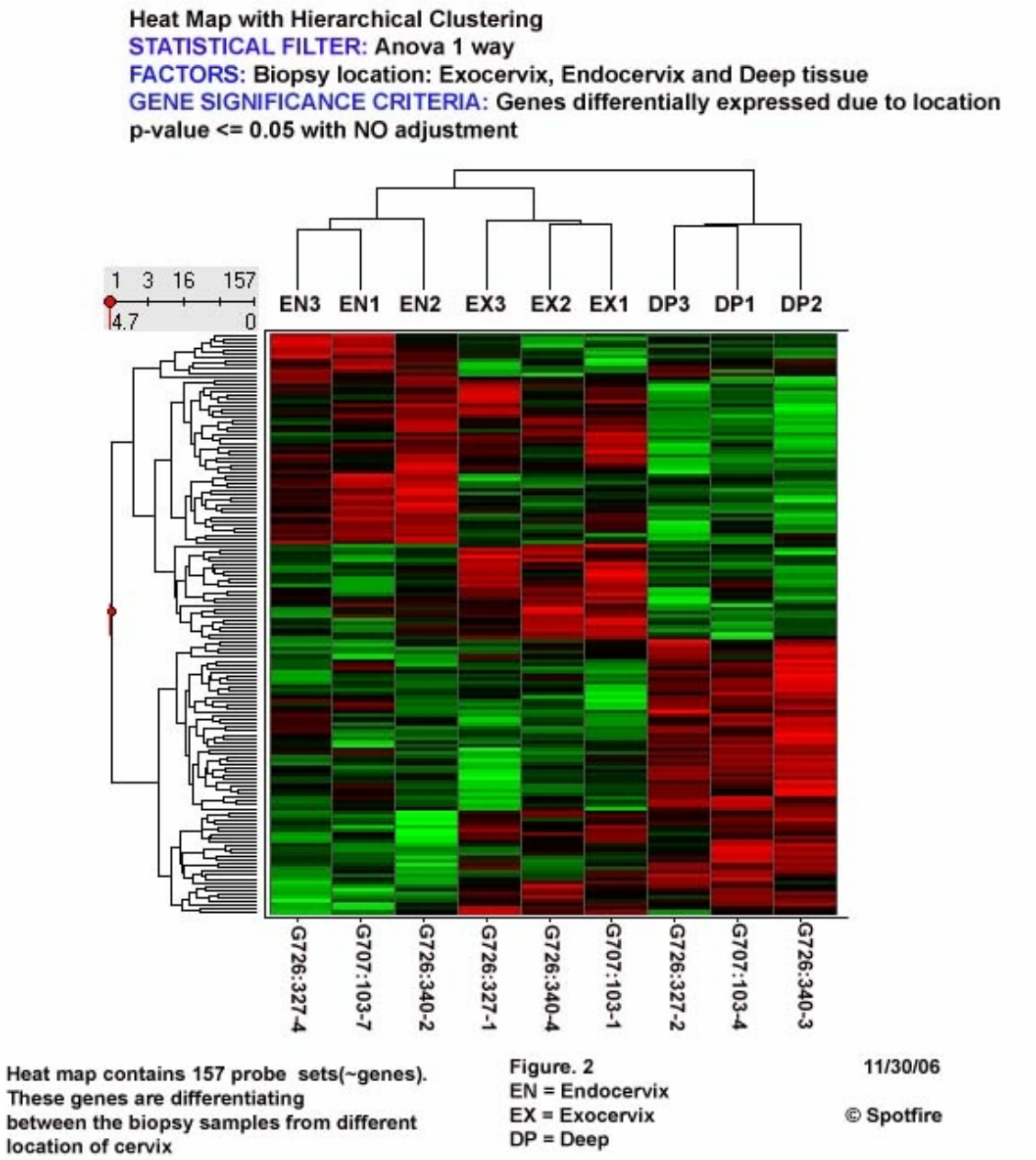

Fig.3.6 Heat Map with Hierarchical Clustering showing 157 genes differentially expressed due to location p-value ≤0.05 with no adjustments.

The three biopsies from the endocervix, exocervix and deep biopsies from all three patients clustered together. Based on the dendogram, the exocervical and endocervical

profiles are more closely clustered together compared to the biopsies at the deepest edge of the tumor.

Inter-tumor Heterogeneity

We wanted to validate whether the 157 location-specific genes were varying by location or whether it could be due to patient to patient variability. To test this idea, using 1 Way ANOVA, we identified 8893 genes that distinguished patients. We then compared the 157 location-specific genes to the 8893 genes identified by patient.

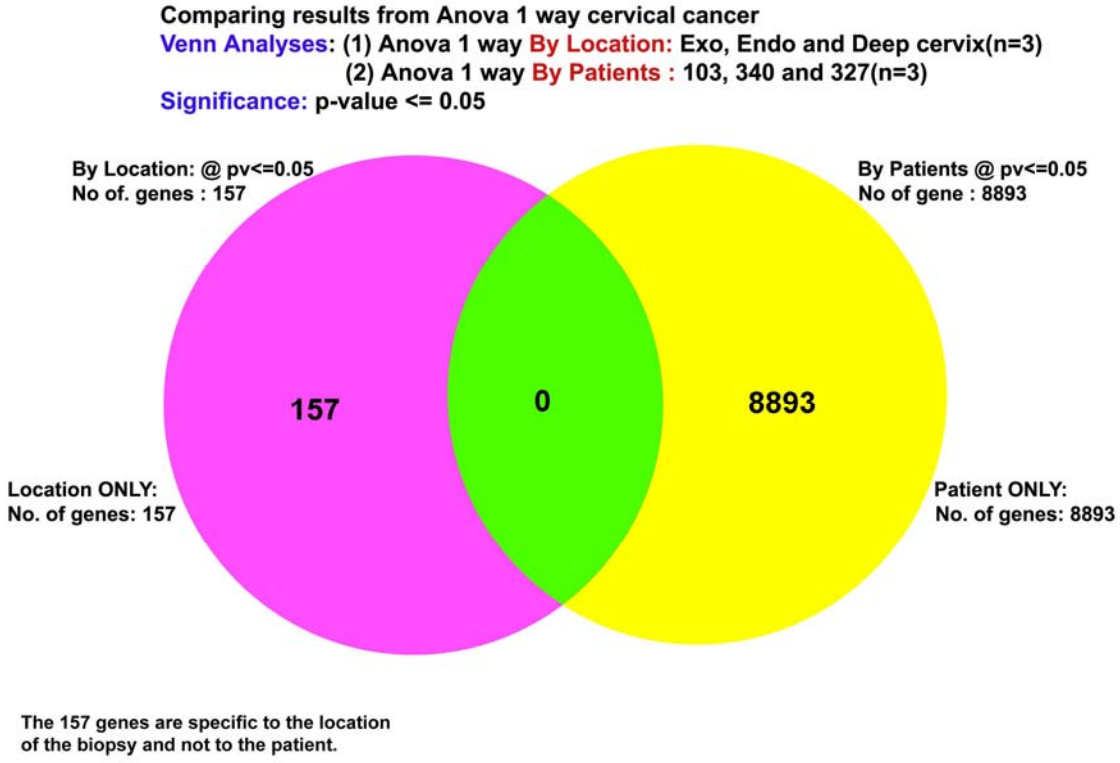

Fig.3.7 Venn diagram showing location specific genes are different from patient variability genes.

There was no overlap of genes in the two data sets. This suggests that the 157 genes truly vary by location and are not due to patient to patient variability. Although

sampling location is important when comparing gene profiles, it is possible to identify genes that do not vary by patient, but only by location.

Clinical Relevance

From a biological stand-point, molecular differences are important to study the mechanism of invasion, growth, and immune evasion. However, from a clinical stand-point, it is useful to take a pre-operative biopsy without removing the uterus for clinical staging. Therefore, we wanted to know whether we can take a pre-operative biopsy from the exocervix and gain some insight into the behavior of the tumor at the deepest leading edge of the tumor. Using ANOVA 1 way analysis, we identified 32 genes showing similar expression values in both the exocervical and deep biopsies.

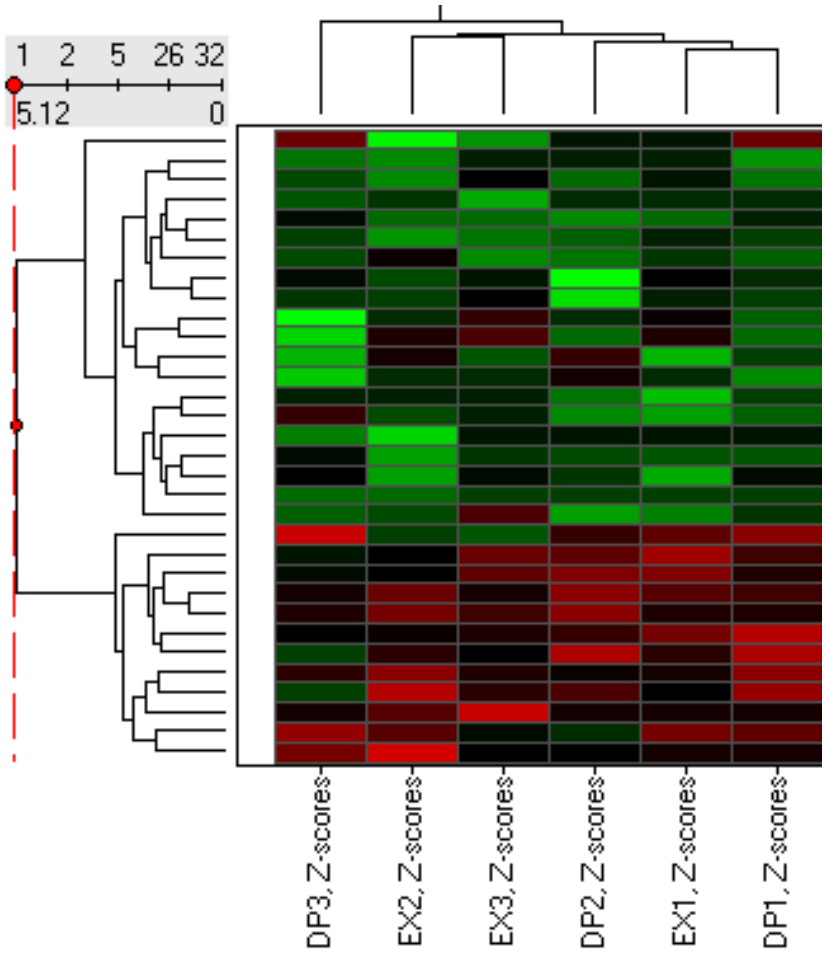

Fig.3.8 Heat Map with Hierarchical Clustering showing 32 genes whose expression profiles showed no change between the exocervical and deep biopsies ($p \leq 0.05$).

Listed in the following table., 32 genes showed an absolute fold change of 1 and thus show no difference in expression between the exocervix and deep biopsies in all 3 patients.

Probe set ID	Genbank ID	Gene	Description
241366_at	BE618393	---	CDNA FLJ35956 fis, clone TESTI2012289
205283_at	NM_006731	FCMD	Fukuyama type congenital muscular dystrophy (fukutin)
1558166_at	BM824870	MGC16275	hypothetical protein MGC16275
230194_at	AI341076	---	Transcribed locus
210138_at	AF074979	RGS20	regulator of G-protein signalling 20
229418_at	AV709958	---	Transcribed locus
228660_x_at	AA523537	SEMA4F	sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4F
215378_at	AU148255	SRA1	Steroid receptor RNA activator 1
231008_at	AI733001	UNC5CL	Unc-5 homolog C (C. elegans)-like
240776_at	AI378893	PGR	Progesterone receptor
1556436_at	BC043205	C8orf50	chromosome 8 open reading frame 50
238020_at	BG166796	PSMC2	Proteasome (prosome, macropain) 26S subunit, ATPase, 2
239302_s_at	AA931539	---	MRNA; cDNA DKFZp686P18215 (from clone DKFZp686P18215)
216863_s_at	AC004542	MORC2	MORC family CW-type zinc finger 2
218500_at	NM_016647	C8orf55	chromosome 8 open reading frame 55
210458_s_at	BC003388	TANK	TRAF family member-associated NFKB activator
218904_s_at	NM_017998	C9orf40	chromosome 9 open reading frame 40

1557113_at	AK095276	LOC283588	hypothetical protein LOC283588
232749_at	AU145533	---	CDNA FLJ11701 fis, clone HEMBA1005062
203702_s_at	AL043927	TTLL4	tubulin tyrosine ligase-like family, member 4
240292_x_at	N50412	ANKS1B	ankyrin repeat and sterile alpha motif domain containing 1B
210050_at	M10036	---	---
212350_at	AB029031	TBC1D1	TBC1 (tre-2/USP6, BUB2, cdc16) domain family, member 1
200684_s_at	AI819709	UBE2L3	ubiquitin-conjugating enzyme E2L 3
232369_at	AF339768	MBNL2	Muscleblind-like 2 (Drosophila)
44563_at	AI858000	WDR79	WD repeat domain 79
205767_at	NM_001432	EREG	epiregulin
218681_s_at	NM_022044	SDF2L1	stromal cell-derived factor 2-like 1
201465_s_at	BC002646	JUN	v-jun sarcoma virus 17 oncogene homolog (avian)
221995_s_at	BF195165	---	---
1553575_at	NM_173714	---	---
1565876_x_at	H50649	NUP153	Nucleoporin 153kDa

Table 3.2 Table showing genes that did not change from the exocervix to the deep biopsy in all three patients.

Of the 32 genes, IPA mapped 26 genes in their database. IPA does not include EST's and genes with no known identity/function. While these genes may play a significant role in cervical cancer, not enough is known about them in the literature. Out of the 26 genes, 15 of these genes fell within one IPA network.

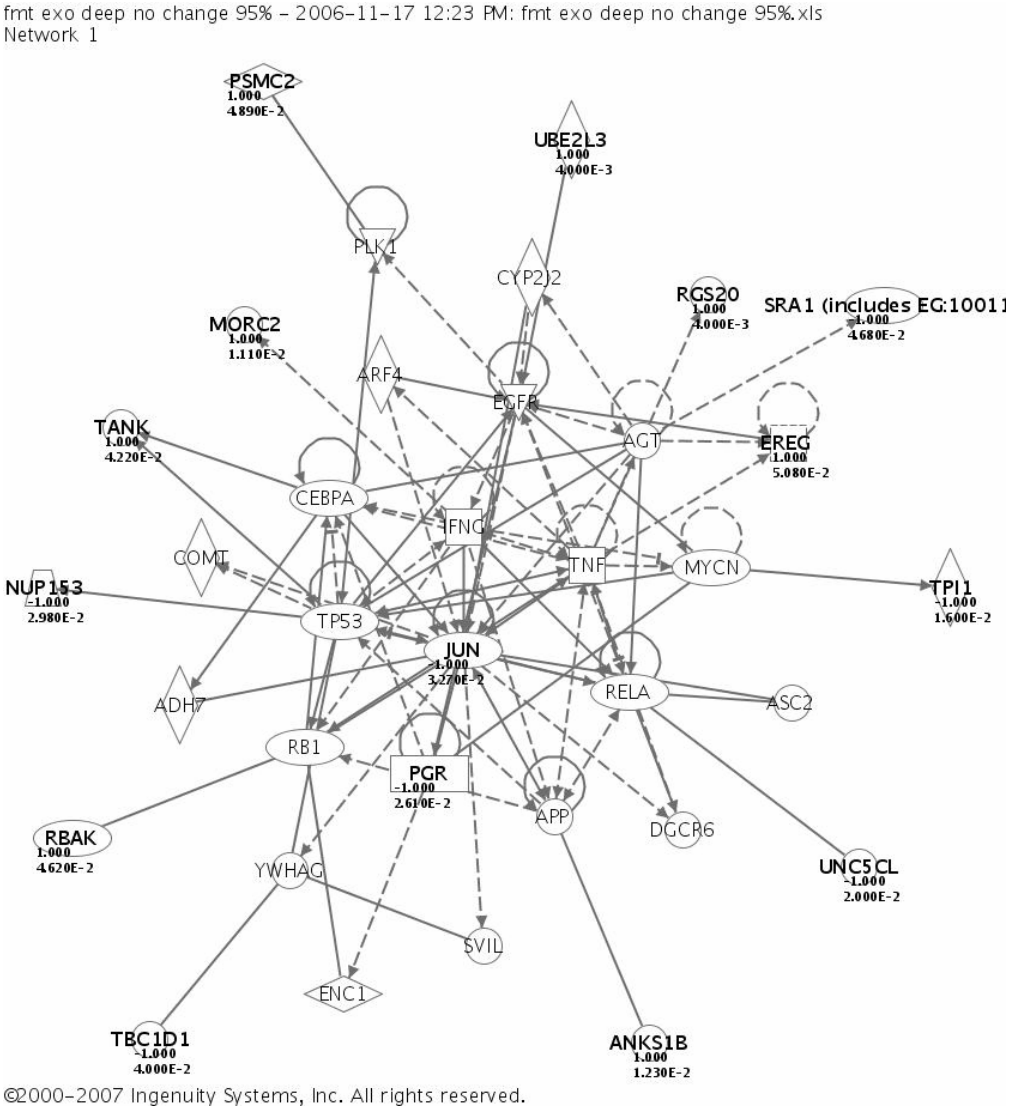

Fig 3.9 IPA functional network identifying the key relationships of genes whose expression values are common to both deep and exocervical biopsies (in bold). Solid lines represent direct relationships while dotted indicate indirect relationships.

The following is a list of the 15 genes in network 1 above.

Affymetrix	Gene	Description	Location	Family
240292_x_at	ANKS1B	ankyrin repeat and sterile alpha motif domain containing 1B	Nucleus	other
205767_at	EREG	epiregulin	Extracellular Space	growth factor
201465_s_at	JUN	v-jun sarcoma virus 17 oncogene homolog (avian)	Nucleus	transcription regulator
216863_s_at	MORC2	MORC family CW-type zinc finger 2	Unknown	other
1565876_x_at	NUP153	nucleoporin 153kDa	Nucleus	transporter
240776_at	PGR	progesterone receptor	Nucleus	ligand-dependent nuclear receptor
238020_at	PSMC2	proteasome (prosome, macropain) 26S subunit, ATPase, 2	Nucleus	peptidase
241366_at	RBAK	RB-associated KRAB zinc finger	Nucleus	transcription regulator
210138_at	RGS20	regulator of G-protein signalling 20	Cytoplasm	other
215378_at	SRA1 (includes EG:10011)	steroid receptor RNA activator 1	Nucleus	transcription regulator
210458_s_at	TANK	TRAF family member-associated NFKB activator	Cytoplasm	other
212350_at	TBC1D1	TBC1 (tre-2/USP6, BUB2, cdc16) domain family, member 1	Nucleus	other
210050_at	TPI1	triosephosphate isomerase 1	Cytoplasm	enzyme
200684_s_at	UBE2L3	ubiquitin-conjugating enzyme E2L 3	Cytoplasm	enzyme
231008_at	UNC5CL	unc-5 homolog C (C. elegans)-like	Cytoplasm	other

Table 3.3 No change genes in network #1.

The genes within this list are most likely to be involved in viral function, gene expression, cancer, the cell cycle, etc. The IPA table below represents the top functions that genes from the above table are involved.

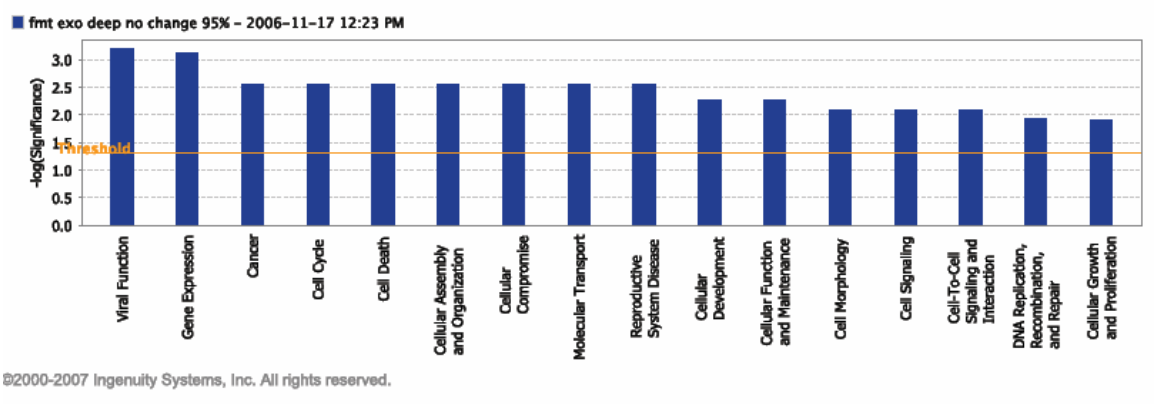

Fig 3.10 Bar chart representing the probability that genes are involved in a particular function. For example, genes of this list are most likely to be involved in viral function.

The following network was generated from the cancer functional group from the above bar graph. Many of the genes in our dataset fall into this pathway and are involved in cancer development.

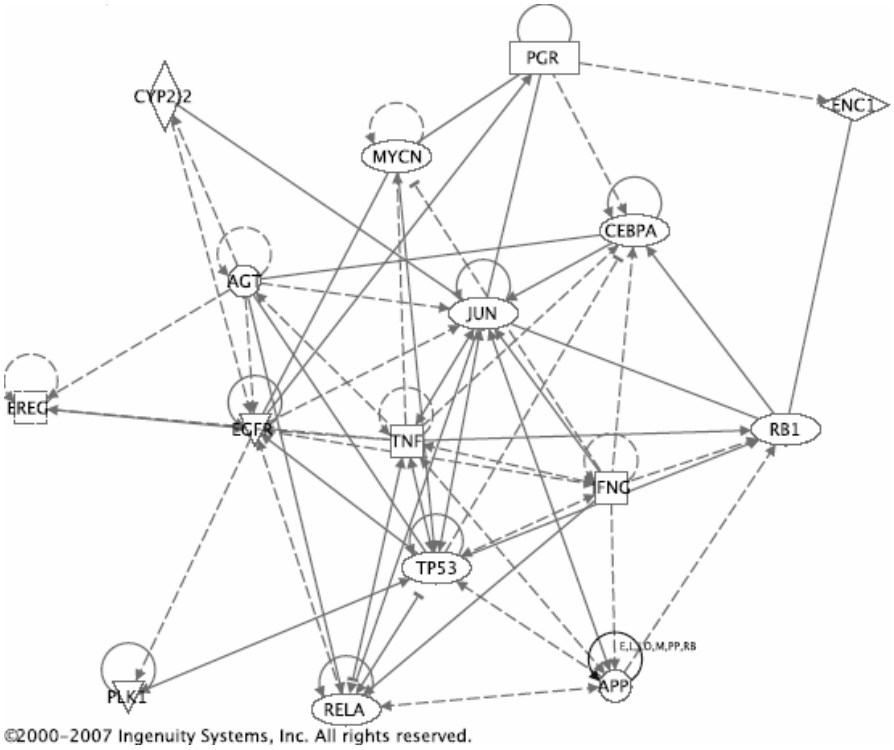

Fig 3.11 IPA functional network identifying the key relationships of genes whose function is involved in cancer.

Because we were only interested in tumor specific genes and not "normal" or HPV-regulated genes, we compared our dataset to genes expressed in biopsies taken from cancer-free normal controls and from HPV-infected, perilesional controls. Out of the 32 genes identified, 16 were absent in the normal (taken from the cervices of patients undergoing a hysterectomy for reasons other than cancer) and perilesional biopsies (taken from patients with IB1 SCCA).

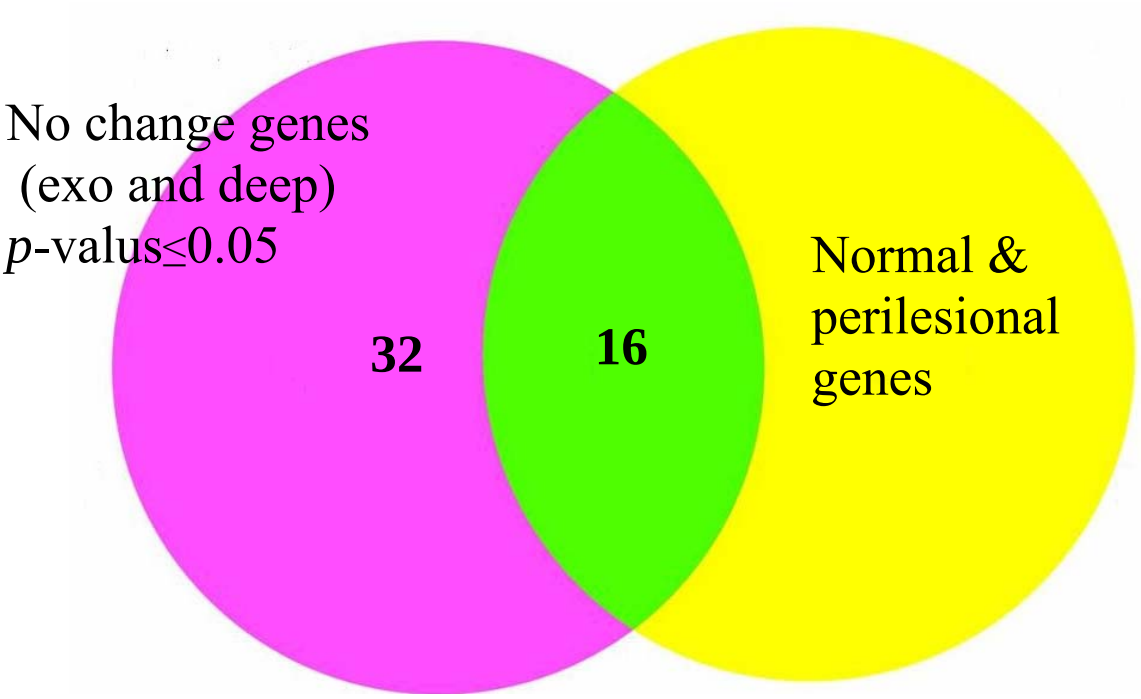

Fig 3.12 No change genes unique to exocervical (exo) and deep biopsies.

The following is a list of genes that were specific to the tumor and not perilesional or normal biopsies.

Probe Set	Gene	Gene Title
210458_s_at	TANK	TRAF family member-associated NFKB activator
203702_s_at	TTLL4	tubulin tyrosine ligase-like family, member 4
205283_at	FCMD	Fukuyama type congenital muscular dystrophy (fukutin)
210050_at	M10036	---
212350_at	TBC1D1	TBC1 (tre-2/USP6, BUB2, cdc16) domain family, member 1
215378_at	SRA1	Steroid receptor RNA activator 1
218500_at	C8orf55	chromosome 8 open reading frame 55
218904_s_at	C9orf40	chromosome 9 open reading frame 40
216863_s_at	MORC2	MORC family CW-type zinc finger 2
221995_s_at	BF195165	---
44563_at	WDR79	WD repeat domain 79
205767_at	EREG	epiregulin
218681_s_at	SDF24	stromal cell-derived factor 2-like 1
201465_s_at	JUN	v-jun sarcoma virus 17 oncogene homolog (avian)
200684_s_at	UBE2L3	ubiquitin-conjugating enzyme E2L 3
210138_at	RGS20	regulator of G-protein signalling 20

Table 3.4 Tumor-specific genes that are in common to exocervical and deep biopsies.

In a biopsy taken at the exocervix, these 16 genes are representative of gene expression at the deepest part of the tumor and are not found in the normal or perilesional tissues.

DISCUSSION

In this study we used molecular profiling as a tool to analyze intra- and inter-tumor heterogeneity in squamous cell carcinomas of the cervix. We wanted to clarify whether site of biopsy matters when characterizing squamous cell carcinomas of the uterine cervix. While the precision of molecular profiling is useful in characterizing disease, we have discovered limitations as well as advantages to this technique.

Intra-tumor Heterogeneity

Much like the limitations found using monoclonal cancer cells in culture and as histological evidence suggests, biopsy location does matter in molecular profiling in squamous cell carcinomas of the cervix. Despite the histologic appearance, we identified a set of location-specific genes that differentiate the site of biopsy. Our data suggests that not all genes are expressed consistently throughout a tumor, but that there are genes that are expressed in specific locations. Thus, assumptions about the behavior of solid tumors may be inadequate using only a singe biopsy. Multiple samples need to be taken and pooled in order to get a more precise snapshot of a tumor as a whole. This finding may also be relevant in other solid tumors such as those found in the prostate, colon, breast, etc,. A big advantage of intra-tumor heterogeneity is that it allows for the identification of tumor cell subpopulations with varying metastatic potential and prognostic significance.

Inter-tumor Heterogeneity

Although we discovered a set of genes that are location specific, we found significant similarities between all study patients in the expression patterns at those locations. Out of 30,000 genes, 157 location-specific genes were in common to all three of our patients. In other words, the expression profile of certain genes taken from a biopsy of the exocervix from one patient can be compared to that same set of genes on the exocervix of another patient. This suggests that there are genes from the exocervix

that might be reflective of the biology at the deepest, leading edge of the tumor. The clinical implication is that a pre-surgical biopsy from the exocervix could be used to determine tumor biology at the deepest edge of a tumor. Even though genes are variable from different locations, we are still able to use genetic profiling to make comparisons between patients as long as the biopsy is taken at the same site. Although other studies, as well as ours, have suggested tumor heterogeneity, we are the first to identify location-specific genes that do not change from patient to patient. The implication from this finding is that although heterogeneity is present, using location-specific genes allow comparisons to be made between patients.

Clinical Relevance

We also identified gene expression values that remain unchanged in exocervical and deep cervical biopsies suggesting that genetic information from the exocervix is reflective of the biological activity in the deeply invading and inaccessible portion of the tumor. Certain genes from the exocervix are consistent with genes of the endocervix. IPA identified Jun at the center of many of the key pathways of the genes with similar expression values in both exocervical and deep biopsies. Jun, while repeatedly appearing in our analysis of cervical cancer, may not be specific for cervical cancer, but may play a major role in its genesis. Jun mediates DNA binding, dimerization, and transcriptional activation. Jun forms homodimers or heterodimers with Fos and binds to the activating protein 1 (AP-1) sites in the promoter for transactivation. HPV E7 binds to c-Jun and increases transcriptional activation and transformation of HPV-infected cells [108,109]. Additionally, the HPV genome has an AP-1 site in its promoter region and early transcriptional regulation of early viral gene products are under control of jun [110]. V-Jun is the viral mutated form of cellular Jun (c-Jun) and has been found to induce tumors in animal models [111]. Jun is involved in many cellular processes including apoptosis, differentiation, transformation, proliferation, mitosis, survival, transcriptional activation, etc.

The molecular signature of a tumor, obtained from an accessible exocervical biopsy, may be useful in individualizing therapy to optimize tumor control while minimizing treatment morbidity. Currently, tumor characterization occurs after surgical intervention. Overall, our data support our hypothesis that site of biopsy is an important factor for molecular profiling studies.

Further studies comparing the GeneChip data from the deeply invasive patient samples (chapter 2) to the no change genes in our exocervical biopsies from chapter 3 will narrow down the potential gene candidates which are expressed on the exocervix that have implications in differentiating superficial from deeply invading tumors.

These studies will provide a molecular signature of cervical cancer that can be used for mechanistic studies of the unique genetic changes of families of genes that define the stages of cervical cancer. As the molecular signature technology develops, we can move beyond improved diagnosis to improved therapy. Our preliminary analysis may help to identify novel potential targets for new therapies, so we can offer patients with different molecular signatures more precisely tailored treatment regimens.

CHAPTER 4: FUTURE STUDIES

EXPANDING THE DATA

The current trend in biomarker discovery is shifting away from the "Holy Grail" ideology of biomarker discovery to multi-arrays with many genes to profile patients. The development of cancer is a complex process whereby multiple carcinogenic events ensure adaptive cells that can respond to a changing microenvironment. Moreover, tumor heterogeneity and genetic variability make the discovery of one molecule to target or characterize all people with a certain tumor less likely. For this preliminary study, we limited our patient selection to tumors of squamous origin. Eighty percent of all cervical carcinomas originate in the squamous epithelium, however, there are other cell types that are affected by cervical cancer including adenocarcinomas of the cervix, neuron-endocrine cancers of the cervix, or mixed cancers of adenatomous and squamous origin. Additionally, we selected tumor tissues with greater than 70% stroma to cancer cell ratios, less than 2% necrosis, and less than 5% lymphocytic infiltration. Not all tumors fit these criteria and may lead to increased variability when applied to a larger sample population. Thus, targeting many genes is probably a better option compared to targeting a single molecule.

Commercially available arrays to detect breast cancer are becoming readily available for research and diagnostic purposes. Various platforms are now available that overcome the issues of sample quantity, but it is unknown whether these separate platforms are comparable. Even within this study, the first generation Affymetrix GeneChips (HG_U133A and HG_U133B set) used in differentiating tumors by risk of recurrence (chapter 2), were different from the newer generation Affymetrix GeneChips (HG_U133 2.0) used in the heterogeneity study (chapter 3) and are not directly comparable. Other real-time PCR platforms for large scale validation have entered the market, but validation needs to be performed before large scale implementation for this

study. MammaPrint (The Molecular Profiling Institute, Inc.) is the first diagnostic microarray kit that is 97% accurate in predicting which lymph node negative breast cancer patients are at risk for metastasis based on a 70 gene profile [112]. While multi-arrays offer the added advantage of screening many genes simultaneously, they still require high quality and quantity of mRNA from freshly obtained, frozen biopsies. The advantage of a PCR-based approach is that archival FFPE tissues can be used.

In this study, we used qRT PCR to validate a few gene candidates. Currently, there are new qRT PCR platforms that allow the screening of many genes simultaneously. Oncotype DX (Genomic Health, Inc.) is one such recently-introduced diagnostic technology that scores breast cancer patients on the basis of risk of recurrence using real time PCR to screen 21 genes. With further validation of our 98 gene dataset on an expanded, heterogeneous population, we believe this technology can be applied in a similar manner to cervical cancer.

STATISTICAL ANALYSIS

Since this study was designed, there have been vast improvements in the statistical analysis of GeneChip data. Support vector machines (SVMs) are based on simple ideas that originated in the area of statistical learning theory [113]. The simplicity arises from the fact that SVMs apply a transformation to highly dimensional data to enable researchers to linearly separate the various features and classes. Luckily, the data transformation allows researchers to avoid calculations in the high dimension space. The popularity of SVMs owes much to the simplicity of the transformation as well as their ability to handle complex classification and regression problems. SVMs are trained with a learning algorithm from optimization theory and tested on the remainder of the available data that were not part of the training dataset [114]. The main aim of support vector machines is to devise a computationally effective way of learning optimal separating parameters for two classes of data. SVMs use an implicit mapping of the input data, commonly referred to as Φ, into a highly dimensional feature space defined by some kernel function. The learning then occurs in the feature space, and the data points appear

in dot products with other data points [115]. One particularly nice property of SVMs is that once a kernel function has been selected and validated, it is possible to work in spaces of any dimension. Thus, it is easy to add new data into the formulation since the complexity of the problem will not be increased by doing so.

Some of the advantages of SVMs versus other techniques are that SVMs produce a unique solution, they can deal with large quantities of seemingly dissimilar information, and the computations are quicker than other machine learning techniques because the discriminant function is characterized by only a small subset of the training data. Thus, SVMs are highly promising for use as a biomarker discovery tool. They are gaining popularity because of their good performance in real-world applications. SVMs are also robust in high dimension which is essential for creating accurate biomarkers. SVMs are also based on sound theoretical foundations which means they are a reliable choice for building biomarkers. Non-traditional data such as trees which are the output of hierarchical clustering can be used as input to SVM which makes them ideal candidates for application with large sets of GeneChip data.

PROTEOMICS

Ultimately, the goal is to develop a clinically applicable diagnostic test to determine risk of recurrence in stage IB1 carcinomas of the cervix. Gene expression platforms have allowed the monitoring of thousands of mRNA's simultaneously. However, messenger RNA is quite unstable. Handling of tissue biopsies is an important variable in obtaining good quality RNA requiring a timely transition from the operating room (or clinic) to the laboratory. This is often an impractical scenario in many laboratories. An alternative to RNA is to identify highly stable downstream metabolites and proteins indicative of risk. Unfortunately, the correlation coefficients between RNA and protein expression are less that 0.5 [116]. Complementing the field of genomics, proteomic analysis can be used to elucidate both protein pool levels and post-translational modifications in cancer development and progression. Changes in protein levels may, therefore, be correlated to changes observed in corresponding mRNA levels that are

detected in the GeneChip arrays. Further, post translational modifications such as methylation and glycosylation have been found to be involved in cancer diagnosis and prognosis [117]. For example, protein modification by glycosylation is found in several cancers including colorectal cancer[118]. This finding is important because glycosylation is a post-translational change that may influence cancer cell behavior, but not necessarily affect the mRNA or protein level and cannot be studied genetically. Moreover, cells can alter their proteome in response to changes on the tissue microenvironment, a phenomenon that makes proteomic analysis an essential component of oncology research. In esophageal carcinoma cell lines, proteomic differences have been reported between transformed (from normal epithelium) and immortalized cell lines suggesting that these differentially expressed proteins might play a role during malignant transformation.[119] An interesting observation is that in all these normal to cancer transformation studies, certain changes are typical of all cancers, while others are unique to each cancer, the latter representing a true cancer signature. Thus, we propose to validate the genomic findings at the protein level.

We propose obtaining protein from superficially and deeply invasive biopsy tissues. These proteins can be positively identified with the peptide mass fingerprinting technique called Matrix-Assisted Laser Desorption Ionization Time of Flight (MALDI-TOF). Using 2-D gel electrophoresis, we can identify proteins unique to the IB1 SCCA's at high risk of recurrence. Two-dimension gel electrophoresis allows the screening of 1200 to 1500 genes simultaneously [120]. Already, our laboratory has been able to identify a unique pattern of proteins expressed in the carcinoma tissues in comparison with the adjacent normal epithelium.

To further expand this validation to multiple readily available samples, we plan on utilizing core samples from FFPE tissues to generate tissue micrarrays [121]. Multiple sections can be combined made from one block containing multiple cancer specimens. Tissue microarrays will allow us to validate our GeneChip data with the use of labeled RNA and DNA probes or antibodies to screen multiple samples simultaneously.

While biopsy material is accessible from the cervix in the clinical setting, a less invasive procedure for determining risk would be optimal. Serum proteins from a blood draw or metabolites found in urine would be a better option for obtaining risk information. We have identified unique proteomic profiles in serum from women with dysplasia compared to normal control serum obtained from women with normal pap smear. Once our genomic dataset is validated in a larger, heterogeneous population of IB1 carcinomas, we will be able to identify secreted protein candidates using IPA software that might be present in a more accessible body fluid such as serum, saliva, or urine.

CLINICAL APPLICATION

Currently, adjuvant therapies such as radiation and chemotherapy do not differentiate between normal and cancer cells. The collateral damage resulting from the non-specific effects of such therapies can lead to multiple adverse effects. Bioinformatics is a promising field for cancer diagnostics. The potential to individualize treatment and decrease non-specific effects make the applicability of this technique monumental. However, the transition from basic science research and patient management must overcome several obstacles. Cross study meta-analysis is difficult due to incompatibilities of study design, incompatibilities in array platforms, and statistical analysis. Study design is an important factor in comparing studies. For example, two studies done at the same hospital looking at the disease progression markers in breast cancer found different results. One study enrolled women under the age of 55 and the other study enrolled all available patients [122,123]. Additionally, retrospective studies are difficult because detailed clinical information is often incomplete [124]. There are new initiatives aimed at standardizing gene expression datasets. The MIAME (Minimum information About Microarray Experiment) allows complete array datasets to be uploaded and publicly available for combined analysis [125]. Future advances in sample processing, microarray technology, and statistical analysis will aid in the practical application of gene array data to individualize patient care in the clinic.

LITERATURE CITED

1. Shimada, M., Kigawa, J., Takahashi, M., Minagawa, Y., Okada, M., Kanamori, Y., Itamochi, H., Oishi, T., Iba, T., Terakawa, N. *Stromal invasion of the cervix can be excluded from the criteria for using adjuvant radiotherapy following radical surgery for patients with cervical cancer.* Gynecol.Oncol. **93**[3], 628-631. 2004.

2. DiMaio, D., and Liao, J. B. *Human papillomaviruses and cervical cancer*. Adv.Virus Res. **66**, 125-159. 2006.

3. Walboomers, J. M., Jacobs, M. V., Manos, M. M., Bosch, F. X., Kummer, J. A., Shah, K. V., Snijders, P. J., Peto, J., Meijer, C. J., and Munoz, N. *Human papillomavirus is a necessary cause of invasive cervical cancer worldwide*. J.Pathol. **189**[1], 12-19. 1999.

4. Pfister, H. *The role of human papillomavirus in anogenital cancer*. Obstet Gynecol.Clin.North Am. **23**[3], 579-595. 1996.

5. Zur Hausen, H. *Papillomaviruses causing cancer: evasion from host-cell control in early events in carcinogenesis*. J.Natl.Cancer Inst. **92**[9], 690-698. 5-3-2000.

6. Reagan, J. W., Hicks, D. J., and Scott, R. B. *Atypical hyperplasia of uterine cervix*. Cancer **8**[1], 42-52. 1955.

7. Parkin, D. M., Bray, F., Ferlay, J., and Pisani, P. *Global cancer statistics, 2002*. CA Cancer J.Clin. **55**[2], 74-108. 2005.

8. http://www.cancer.gov/cancertopics/types/cervical/. *What You Need To Know About Cancer of the Cervix.* Accessed 3-3-2007

9. Clarke, E. A., and Anderson, T. W. *Does screening by "Pap" smears help prevent cervical cancer? A case-control study*. Lancet **2**[8132], 1-4. 7-7-1979.

10. Hakama, M., Chamberlain, J., Day, N. E., Miller, A. B., and Prorok, P. C. *Evaluation of screening programmes for gynaecological cancer*. Br.J.Cancer **52**[4], 669-673. 1985.

11. Miller, A. B., Chamberlain, J., Day, N. E., Hakama, M., and Prorok, P. C. *Report on a Workshop of the UICC Project on Evaluation of Screening for Cancer*. Int.J.Cancer **46**[5], 761-769. 11-15-1990.

12. *Control of cancer of the cervix uteri: A WHO meeting. Bulletin of the World Health Organization*. A WHO meeting.Bulletin of the World Health Organization **64**, 607-618. 1986.

13. Womack, C., and Warren, A. Y. *Achievable laboratory standards; a review of cytology of 99 women with cervical cancer*. Cytopathology **9**[3], 171-177. 1998.

14. Janerich, D. T., Hadjimichael, O., Schwartz, P. E., Lowell, D. M., Meigs, J. W., Merino, M. J., Flannery, J. T., and Polednak, A. P. *The screening histories of women with invasive cervical cancer, Connecticut*. Am.J.Public Health **85**[6], 791-794. 1995.

15. Leyden, W. A., Manos, M. M., Geiger, A. M., Weinmann, S., Mouchawar, J., Bischoff, K., Yood, M. U., Gilbert, J., and Taplin, S. H. *Cervical cancer in women with comprehensive health care access: attributable factors in the screening process*. J.Natl.Cancer Inst. **97**[9], 675-683. 5-4-2005.

16. Adami, H. O., Ponten, J., Sparen, P., Bergstrom, R., Gustafsson, L., and Friberg, L. G. *Survival trend after invasive cervical cancer diagnosis in Sweden before and after cytologic screening. 1960-1984*. Cancer **73**[1], 140-147. 1-1-1994.

17. Stenkvist, B., Bergstrom, R., Eklund, G., and Fox, C. H. *Papanicolaou smear screening and cervical cancer. What can you expect?* JAMA **252**[11], 1423-1426. 9-21-1984.

18. Sasieni, P., and Adams, J. *Effect of screening on cervical cancer mortality in England and Wales: analysis of trends with an age period cohort model*. BMJ **318**[7193], 1244-1245. 5-8-1999.

19. Miller, B.A., Kolonel, L.N., Bernstein. L., Young Jr., J.L., Swanson, G.M., West, D., Key, C.R., Liff, J.M., Glover, C.S., Alexander, G.A., et al. eds. *Racial/Ethnic Patterns of Cancer in the United States 1988-1992.* National Cancer Institute. NIH Pub.No.96-4104.Bethesda, MD. 1996.

20. http://www.paho.org/English/DD/PIN/ptoday04_nov04.htm. *Scant Progress on Cervical Cancer.* Accessed 3-3-2007

21. http://www.paho.org/English/DD/PIN/ptoday04_nov04.htm. *Scant Progress on Cervical Cancer.* Accessed 3-3-2007

22. Dell, G., and Gaston, K. *Human papillomaviruses and their role in cervical cancer*. Cell Mol.Life Sci. **58**[12-13], 1923-1942. 2001.

23. Dalstein, V., Riethmuller, D., Pretet, J. L., Le Bail, Carval K., Sautiere, J. L., Carbillet, J. P., Kantelip, B., Schaal, J. P., and Mougin, C. *Persistence and load of high-risk HPV are predictors for development of high-grade cervical lesions: a longitudinal French cohort study*. Int.J.Cancer **106**[3], 396-403. 9-1-2003.

24. Walboomers, J. M., Meijer, C. J., Steenbergen, R. D., van, Duin M., Helmerhorst, T. J., and Snijders, P. J. *Human papillomavirus and the development of cervical cancer: concept of carcinogenesis*. Ned.Tijdschr.Geneeskd. **144**[35], 1671-1674. 8-26-2000.

25. Sasieni, P. D., Cuzick, J., and Lynch-Farmery, E. *Estimating the efficacy of screening by auditing smear histories of women with and without cervical cancer. The National Co-ordinating Network for Cervical Screening Working Group*. Br.J.Cancer. **73**[8], 1001-1005. 1996.

26. de Boer, C. J., van Dorst, E., van Krieken, H., Jansen-van Rhijn, C. M., Warnaar, S. O., Fleuren, G. J., and Litvinov, S. V. *Changing roles of cadherins and catenins during progression of squamous intraepithelial lesions in the uterine cervix*. Am.J.Pathol. **155**[2], 505-515. 1999.

27. Vessey, C. J., Wilding, J., Folarin, N., Hirano, S., Takeichi, M., Soutter, P., Stamp, G. W., and Pignatelli, M. *Altered expression and function of E-cadherin in cervical intraepithelial neoplasia and invasive squamous cell carcinoma*. J.Pathol. **176**[2], 151-159. 1995.

28. *World Health Organization. Geneva, Switzerland: World Health Organization*. WHO, 1-22. 1999.

29. Sonya Pagliusi. http://www.who.int/vaccines/en/hpvrd.shtml/shtml. *Vaccines Against Human Papillomavirus.* Accessed 3-03-2007

30. Ostor, A. G. *Natural history of cervical intraepithelial neoplasia: a critical review*. Int.J.Gynecol.Pathol. **12**[2], 186-192. 1993.

31. Ramaswamy, S., Ross, K. N., Lander, E. S., and Golub, T. R. *A molecular signature of metastasis in primary solid tumors*. Nature Genet. **33**, 49-54. 2003.

32. Bosch, F. X., Lorincz, A., Munoz, N., Meijer, C. J., and Shah, K. V. *The causal relation between human papillomavirus and cervical cancer*. J.Clin.Pathol [55], 244-265. 1-22-2002.

33. Potischman, N. and Brinton, L. A. *Nutrition and cervical neoplasia*. Cancer Causes Control **7**[1], 113-126. 1996.

34. Butterworth, C. E., Jr., Hatch, K. D., Macaluso, M., Cole, P., Sauberlich, H. E., Soong, S. J., Borst, M., and Baker, V. V. *Folate deficiency and cervical dysplasia*. JAMA **267**[4], 528-533. 1-22-1992.

35. Bosch, F. X., Castellsague, X., Munoz, N., de, Sanjose S., Ghaffari, A. M., Gonzalez, L. C., Gili, M., Izarzugaza, I., Viladiu, P., Navarro, C., Vergara, A., Ascunce, N., Guerrero, E., and Shah, K. V. *Male sexual behavior and human papillomavirus DNA: key risk factors for cervical cancer in Spain.* J.Natl.Cancer Inst. **88**[15], 1060-1067. 8-7-1996.

36. Anttila, T., Saikku, P., Koskela, P., Bloigu, A., Dillner, J., Ikaheimo, I., Jellum, E., Lehtinen, M., Lenner, P., Hakulinen, T., Narvanen, A., Pukkala, E., Thoresen, S., Youngman, L., and Paavonen, J. *Serotypes of Chlamydia trachomatis and risk for development of cervical squamous cell carcinoma*. JAMA. **285**[1], 47-51. 1-3-2001.

37. Hawes, S. E., and Kiviat, N. B. *Are genital infections and inflammation cofactors in the pathogenesis of invasive cervical cancer?* J.Natl.Cancer Inst. **94**[21], 1592-1593. 11-6-2002.

38. Johansen, C., Mellemkjaer, L., Frisch, M., Kjaer, S. K., Gridley, G., and Olsen, J. H. *Risk for anogenital cancer and other cancer among women hospitalized with gonorrhea*. Acta.Obstet.Gynecol.Scand. **80**[8], 757-761. 2001.

39. Castellsague, X., Bosch, F. X., Munoz, N., Meijer, C. J., Shah, K. V., de, Sanjose S., Eluf-Neto, J., Ngelangel, C. A., Chichareon, S., Smith, J. S., Herrero, R., Moreno, V., and Franceschi, S. *Male circumcision, penile human papillomavirus infection, and cervical cancer in female partners*. N.Engl.J.Med. **346**[15], 1105-1112. 4-11-2002.

40. Castellsague, X., Bosch, F. X., and Munoz, N. *Environmental co-factors in HPV carcinogenesis*. Virus Res. **89**[2], 191-199. 2002.

41. Slattery, M. L., Robison, L. M., Schuman, K. L., French, T. K., Abbott, T. M., Overall, J. C., Jr., and Gardner, J. W. *Cigarette smoking and exposure to passive smoke are risk factors for cervical cancer*. JAMA **261**[11], 1593-1598. 3-17-1989.

42. Schiffman, M., Brinton, L., Holly, E., Lannom, L., Kurman, R., Lancaster, W., Andrews, A. W., and Felton, J. *Regarding mutagenic mucus in the cervix of smokers*. J.Natl.Cancer Inst. **78**[3], 590-591. 1987.

43. Schiffman, M. H., Haley, N. J., Felton, J. S., Andrews, A. W., Kaslow, R. A., Lancaster, W. D., Kurman, R. J., Brinton, L. A., Lannom, L. B., and Hoffmann, D. *Biochemical epidemiology of cervical neoplasia: measuring cigarette smoke constituents in the cervix*. Cancer Res. **47**[14], 3886-3888. 7-15-1987.

44. Prokopczyk, B., Cox, J. E., Hoffmann, D., and Waggoner, S. E. *Identification of tobacco-specific carcinogen in the cervical mucus of smokers and nonsmokers*. J.Natl.Cancer Inst. **89**[12], 868-873. 6-18-1997.

45. Reagan, J. W. and Herbst, A. L. *A correlative nuclear DNA and histologic study of genital squamous lesions in DES exposed progeny*. Obstet.Gynecol.Surv. **34**[11], 849-850. 1979.

46. Konya, J. and Dillner, J. *Immunity to oncogenic human papillomaviruses*. Adv.Cancer Res. **82**, 205-238. 2001.

47. Kessler, M., Jay, N., Molle, R., and Guillemin, F. *Excess risk of cancer in renal transplant patients*. Transpl.Int. **19**[11], 908-914. 2006.

48. Fairley, C. K., Chen, S., Tabrizi, S. N., McNeil, J., Becker, G., Walker, R., Atkins, R. C., Thomson, N., Allan, P., and Woodburn, C. *Prevalence of HPV DNA in cervical specimens in women with renal transplants: a comparison with dialysis-dependent patients and patients with renal impairment*. Nephrol.Dial.Transplant. **9**[4], 416-420. 1994.

49. Hawes, S. E., Critchlow, C. W., Faye Niang, M. A., Diouf, M. B., Diop, A., Toure, P., Aziz, Kasse A., Dembele, B., Salif, Sow P., Coll-Seck, A. M., Kuypers, J. M., and Kiviat, N. B. *Increased risk of high-grade cervical squamous intraepithelial lesions and invasive cervical cancer among*

African women with human immunodeficiency virus type 1 and 2 infections. J.Infect.Dis. **188**[4], 555-563. 8-15-2003.

50. Institute of Medicine. *Unequal Treatment: Confronting Racial and Ethnic Disparities in Healthcare.* 2003. Washington D.C., The National Academic Press.

51. Freeman, H. P. *Commentary on the meaning of race in science and society.* Cancer Epidemiol.Biomarkers Prev. **12**[3], 232s-236s. 2003.

52. Ward, E., Jemal, A., Cokkinides, V., Singh, G. K., Cardinez, C., Ghafoor, A., and Thun, M. *Cancer disparities by race/ethnicity and socioeconomic status*. CA Cancer J.Clin. **54**[2], 78-93. 2004.

53. http://www.paho.org/English/DD/PIN/ptoday04_nov04.htm. *Scant Progress on Cervical Cancer.* Accessed 3-3-2007

54. Couzin, J. *Cancer research. Probing the roots of race and cancer.* Science **315**[5812], 592-594. 2-2-2007.

55. Ries LA, Eisner M, Kosary C, and et al. http://seer.cancer.gov/csr/1975_2003/. *SEER Cancer Statistics Review, 1975-2003.* Accessed 3-3-2007

56. Ward, E., Jemal, A., Cokkinides, V., Singh, G. K., Cardinez, C., Ghafoor, A., and Thun, M. *Cancer disparities by race/ethnicity and socioeconomic status*. CA Cancer J.Clin. **54**[2], 78-93. 2004.

57. Singh, G. K., Miller, B. A., Hankey, B. F., and Edwards, B. K. *Persistent area socioeconomic disparities in U.S. incidence of cervical cancer, mortality, stage, and survival, 1975-2000*. Cancer **101**[5], 1051-1057. 9-1-2004.

58. Madeleine, M. M., Brumback, B., Cushing-Haugen, K. L., Schwartz, S. M., Daling, J. R., Smith, A. G., Nelson, J. L., Porter, P., Shera, K. A., McDougall, J. K., and Galloway, D. A. *Human leukocyte antigen class II and cervical cancer risk: a population-based study*. J.Infect.Dis. **186**[11], 1565-1574. 12-1-2002.

59. Bosch, F. X., Lorincz, A., Munoz, N., Meijer, C. J., and Shah, K. V. *The causal relation between human papillomavirus and cervical cancer*. J.Clin.Pathol. **55**[4], 244-265. 2002.

60. http://www-dep.iarc.fr/ . *GLOBOCAN, International Agency for Research on Cancer (IARC) CANCERMondial.*

61. Sedlis, A., Bundy, B. N., Rotman, M. Z., Lentz, S. S., Muderspach, L. I., and Zaino, R. J. *A randomized trial of pelvic radiation therapy versus no further therapy in selected patients with stage IB carcinoma of the cervix after radical hysterectomy and pelvic lymphadenectomy: a gynecologic oncolgy group study*. Gynecol.Oncol. **73**[2], 177-183. 1999.

62. Classe, J. M., Rauch, P., Rodier, J. F., Morice, P., Stoeckle, E., Lasry, S., and Houvenaeghel, G. *Surgery after concurrent chemoradiotherapy and brachytherapy for the treatment of advanced cervical cancer: morbidity and outcome: results of a multicenter study of the GCCLCC (Groupe des Chirurgiens de Centre de Lutte Contre le Cancer)*. Gynecol Oncol. **102**[3], 523-529. 2006.

63. Denschlag, D., Gabriel, B., Mueller-Lantzsch, C., Tempfer, C., Henne, K., Gitsch, G., and Hasenburg, A. *Evaluation of patients after extraperitoneal lymph node dissection for cervical cancer*. Gynecol.Oncol. **96**[3], 658-664. 2005.

64. Weber, T. M., Sostman, H. D., Spritzer, C. E., Ballard, R. L., Meyer, G. A., Clark-Pearson, D. L., and Soper, J. T. *Cervical carcinoma: determination of recurrent tumor extent versus radiation changes with MR imaging*. Radiology **194**, 135-139. 1995.

65. Grigsby, P. W., Siegel B. A., and Dehdashti, F. *Lymph node staging by positron emission tomography in patients with carcinoma of the cervix*. J.Clin.Oncol **19**, 3745-3749. 2001.

66. Holtz, D. O., and Dunton, C. *Traditional management of invasive cervical cancer*. Obstet.Gynecol Clin.North Am. **29**[4], 645-657. 2002.

67. Monk, B. J., Wang, J., Im, S., Stock, R. J., Peters, W. A., III, Liu, P. Y., Barrett, R. J., Berek, J. S., Souhami, L., Grigsby, P. W., Gordon, W., Jr., and Alberts, D. S. *Rethinking the use of radiation and chemotherapy after radical hysterectomy: a clinical-pathologic analysis of a Gynecologic Oncology Group/Southwest Oncology Group/Radiation Therapy Oncology Group trial*. Gynecol.Oncol. **96**[3], 721-728. 2005.

68. Sedlis, A., Bundy, B. N., Rotman, M. Z., Lentz, S. S., Muderspach, L. I., and Zaino, R. J. *A randomized trial of pelvic radiation therapy versus no further therapy in selected patients with stage IB carcinoma of the cervix after radical hysterectomy and pelvic lymphadenectomy: a gynecologic oncolgy group study*. Gynecol.Oncol. **73**[2], 177-183. 1999.

69. Rotman, M., Sedlis, A., Piedmonte, M. R., Bundy, B., Lentz, S. S., Muderspach, L. I., and Zaino, R. J. *A phase III randomized trial of postoperative pelvic irradiation in Stage IB cervical carcinoma with poor prognostic features: follow-up of a gynecologic oncology group study*. Int.J.Radiat.Oncol.Biol.Phys. **65**[1], 169-176. 5-1-2006.

70. Greer, B. E., Weu-Jin, K., Nedeem, A., Bookman, M. A., and *et al*. www.nccn.org. *National Comprehensive Cancer Network (NCCN) Guidelines in Oncology, Cervical Cancer.* Accessed 3-03-2007

71. Classe, J. M., Rauch, P., Rodier, J. F., Morice, P., Stoeckle, E., Lasry, S., and Houvenaeghel, G. *Surgery after concurrent chemoradiotherapy and brachytherapy for the treatment of advanced cervical cancer: morbidity and outcome: results of a multicenter study of the GCCLCC (Groupe des Chirurgiens de Centre de Lutte Contre le Cancer).* Gynecol.Oncol. **102**[3], 523-529. 2006.

72. Katz, S., Irizarry, R. A., Lin, X., Tripputi, M., and Porter, M. W. *A summarization approach for Affymetrix GeneChip data using a reference training set from a large, biologically diverse database.* BMC Bioinformatics **7**, 464. 2006.

73. Jain, N., Thatte, J., Braciale, T., Ley, K., O'Connell, M., and Lee, J. K. *Local-pooled-error test for identifying differentially expressed genes with a small number of replicated microarrays.* Bioinformatics **19**[15], 1945-1951. 10-12-2003.

74. Gravitt, P. E., Peyton, C. L., Apple, R. J., and Wheeler, C. M. *Genotyping of 27 human papillomavirus types by using L1 consensus PCR products by a single-hybridization, reverse line blot detection method.* J.Clin.Microbiol. **36**[10], 3020-3027. 1998.

75. Maiman, M., Feuer, G., Fruchter, R. G., Shaw, N., and Boyce, J. *Value of squamous cell carcinoma antigen levels in invasive cervical carcinoma*. Gynecol.Oncol. **34**[3], 312-316. 1989.

76. Gaarenstroom, K. N., Kenter, G. G., Bonfrer, J. M., Korse, C. M., Gallee, M. P., Hart, A. A., Muller, M., Trimbos, J. B., and Helmerhorst, T. J. *Prognostic significance of serum antibodies to human papillomavirus-16 E4 and E7 peptides in cervical cancer.* Cancer **74**[8], 2307-2313. 10-15-1994.

77. Patsner, B., Orr, J. W., Jr., and Allmen, T. *Does preoperative serum squamous cell carcinoma antigen level predict occult extracervical disease in*

patients with stage Ib invasive squamous cell carcinoma of the cervix? Obstet.Gynecol. **74**[5], 786-788. 1989.

78. Suminami, Y., Nagashima, S., Murakami, A., Nawata, S., Gondo, T., Hirakawa, H., Numa, F., Silverman, G. A., and Kato, H. *Suppression of a squamous cell carcinoma (SCC)-related serpin, SCC antigen, inhibits tumor growth with increased intratumor infiltration of natural killer cells*. Cancer Res. **61**[5], 1776-1780. 3-1-2001.

79. Murakami, A., Nakagawa, T., Kaneko, M., Nawata, S., Takeda, O., Kato, H., and Sugino, N. *Suppression of SCC antigen promotes cancer cell invasion and migration through the decrease in E-cadherin expression*. Int.J.Oncol. **29**[5], 1231-1235. 2006.

80. Duk, J. M., de Bruijn, H. W., Groenier, K. H., Hollema, H., ten Hoor, K. A., Krans, M., and Aalders, J. G. *Cancer of the uterine cervix: sensitivity and specificity of serum squamous cell carcinoma antigen determinations*. Gynecol.Oncol. **39**[2], 186-194. 1990.

81. Molina, R., Filella, X., Torres, M. D., Ballesta, A. M., Mengual, P., Cases, A., and Balaque, A. *SCC antigen measured in malignant and nonmalignant diseases*. Clin.Chem. **36**[2], 251-254. 1990.

82. Duk, J. M., van, Voorst, V, ten Hoor, K. A., Hollema, H., Doeglas, H. M., and de Bruijn, H. W. *Elevated levels of squamous cell carcinoma antigen in patients with a benign disease of the skin*. Cancer **64**[8], 1652-1656. 10-15-1989.

83. Hammarstrom, S. *The carcinoembryonic antigen (CEA) family: structures, suggested functions and expression in normal and malignant tissues*. Semin.Cancer.Biol. **9**[2], 67-81. 1999.

84. Ychou, M., Pelegrin, A., Faurous, P., Robert, B., Saccavini, J. C., Guerreau, D., Rossi, J. F., Fabbro, M., Buchegger, F., Mach, J. P., and Artus, J. C. *Phase-I/II radio-immunotherapy study with Iodine-131-labeled anti-CEA monoclonal antibody F6 F(ab')2 in patients with non-resectable liver metastases from colorectal cancer*. Int.J.Cancer **75**[4], 615-619. 2-9-1998.

85. Delgado, C., Pedley, R. B., Herraez, A., Boden, R., Boden, J. A., Keep, P. A., Chester, K. A., Fisher, D., Begent, R. H., and Francis, G. E. *Enhanced tumour specificity of an anti-carcinoembrionic antigen Fab' fragment by poly(ethylene glycol) (PEG) modification*. Br.J.Cancer **73**[2], 175-182. 1996.

86. Thompson, J., Zimmermann, W., Nollau, P., Neumaier, M., Weber-Arden, J., Schrewe, H., Craig, I., and Willcocks, T. *CGM2, a member of the carcinoembryonic antigen gene family is down-regulated in colorectal carcinomas*. J.Biol.Chem. **269**[52], 32924-32931. 12-30-1994.

87. Kinugasa, T., Kuroki, M., Takeo, H., Matsuo, Y., Ohshima, K., Yamashita, Y., Shirakusa, T., and Matsuoka, Y. *Expression of four CEA family antigens (CEA, NCA, BGP and CGM2) in normal and cancerous gastric epithelial cells: up-regulation of BGP and CGM2 in carcinomas*. Int.J.Cancer **76**[1], 148-153. 3-30-1998.

88. Jacobs, I., and Bast, R. C., Jr. *The CA 125 tumour-associated antigen: a review of the literature*. Hum.Reprod. **4**[1], 1-12. 1989.

89. Einhorn, N., Bast, R. C., Jr., Knapp, R. C., Tjernberg, B., and Zurawski, V. R., Jr. *Preoperative evaluation of serum CA 125 levels in patients with primary epithelial ovarian cancer*. Obstet.Gynecol **67**[3], 414-416. 1986.

90. Maggino, T., Gadducci, A., D'Addario, V., Pecorelli, S., Lissoni, A., Stella, M., Romagnolo, C., Federghini, M., Zucca, S., Trio, D., and . *Prospective multicenter study on CA 125 in postmenopausal pelvic masses*. Gynecol. Oncol. **54**[2], 117-123. 1994.

91. Cuzick, J., Szarewski, A., Cubie, H., Hulman, G., Kitchener, H., Luesley, D., McGoogan, E., Menon, U., Terry, G., Edwards, R., Brooks, C., Desai, M., Gie, C., Ho, L., Jacobs, I., Pickles, C., and Sasieni, P. *Management of women who test positive for high-risk types of human papillomavirus: the HART study*. Lancet **362**[9399], 1871-1876. 12-6-2003.

92. Sopov, I., Sorensen, T., Magbagbeolu, M., Jansen, L., Beer, K., Kuhne-Heid, R., Kirchmayr, R., Schneider, A., and Durst, M. *Detection of cancer-related gene expression profiles in severe cervical neoplasia*. Int.J.Cancer **112**[1], 33-43. 10-20-2004.

93. Santin, A. D., Zhan, F., Bignotti, E., Siegel, E. R., Cane, S., Bellone, S., Palmieri, M., Anfossi, S., Thomas, M., Burnett, A., Kay, H. H., Roman, J. J., O'Brien, T. J., Tian, E., Cannon, M. J., Shaughnessy, J., Jr., and Pecorelli, S. *Gene expression profiles of primary HPV16- and HPV18-infected early stage cervical cancers and normal cervical epithelium:identification of novel candidate molecular markers for cervical cancer diagnosis and therapy.* Virology **331**[2], 269-291. 1-20-2005.

94. Ahmed, A. A., and Brenton, J. D. *Microarrays and breast cancer clinical studies: forgetting what we have not yet learnt.* Breast Cancer Res. **7**[3], 96-99. 2005.

95. Del Mistro A., Salamanca, H. F., Trevisan, R., Bertorelle, R., Parenti, A., Bonoldi, E., Zambon, P., and Minucci, D. *Human papillomavirus typing of invasive cervical cancers in Italy*. Infect.Agent.Cancer **1**, 9. 2006.

96. Paez, J. G., Janne, P. A., Lee, J. C., Tracy, S., Greulich, H., Gabriel, S., Herman, P., Kaye, F. J., Lindeman, N., Boggon, T. J., Naoki, K., Sasaki, H., Fujii, Y., Eck, M. J., Sellers, W. R., Johnson, B. E., and Meyerson, M. *EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy.* Science **304**[5676], 1497-1500. 6-4-2004.

97. Cobleigh, M. and Frame, D. *Is trastuzumab every three weeks ready for prime time?* J.Clin.Oncol **21**[21], 3900-3901. 11-1-2003.

98. Mor, G., Visintin, I., Lai, Y., Zhao, H., Schwartz, P., Rutherford, T., Yue, L., Bray-Ward, P., and Ward, D. C. *Serum protein markers for early detection of ovarian cancer*. Proc.Natl.Acad.Sci.U.S.A **102**[21], 7677-7682. 5-24-2005.

99. Zvara, A., Hackler, L., Jr., Nagy, Z. B., Micsik, T., and Puskas, L. G. *New molecular methods for classification, diagnosis and therapy prediction of hematological malignancies*. Pathol.Oncol.Res. **8**[4], 231-240. 2002.

100. Levene, A. P., Morgan, G. J., and Davies, F. E. *The use of genetic microarray analysis to classify and predict prognosis in haematological malignancies*. Clin.Lab Haematol. **25**[4], 209-220. 2003.

101. Mor, G., Visintin, I., Lai, Y., Zhao, H., Schwartz, P., Rutherford, T., Yue, L., Bray-Ward, P., and Ward, D. C. *Serum protein markers for early detection of ovarian cancer*. Proc.Natl.Acad.Sci.U.S.A **102**[21], 7677-7682. 5-24-2005.

102. Tonini, G. P. and Pistoia, V. *Molecularly guided therapy of neuroblastoma: a review of different approaches.* Curr.Pharm.Des **12**[18], 2303-2317. 2006.

103. Kirchhoff, M., Rose, H., Petersen, B. L., Maahr, J., Gerdes, T., Lundsteen, C., Bryndorf, T., Kryger-Baggesen, N., Christensen, L., Engelholm, S. A., and Philip, J. *Comparative genomic hybridization reveals a recurrent pattern of chromosomal aberrations in severe dysplasia/carcinoma in situ of the*

cervix and in advanced-stage cervical carcinoma. Genes Chromo.Cancer **24**[2], 144-150. 1999.

104. Hu, X., Pang, T., Asplund, A., Ponten, J., and Nister, M. *Clonality analysis of synchronous lesions of cervical carcinoma based on X chromosome inactivation polymorphism, human papillomavirus type 16 genome mutations, and loss of heterozygosity*. J.Exp.Med. **195**[7], 845-854. 4-1-2002.

105. Lyng, H., Beigi, M., Svendsrud, D. H., Brustugun, O. T., Stokke, T., Kristensen, G. B., Sundfor, K., Skjonsberg, A., and De Angelis, P. M. *Intratumor chromosomal heterogeneity in advanced carcinomas of the uterine cervix*. Int.J.Cancer **111**[3], 358-366. 9-1-2004.

106. Pang, T., Hu, X., Mazurenko, N., Kisseljov, F., and Ponten, J. *Multiple variants of HPV16 E6 gene in cervical invasive squamous cell carcinoma*. Anticancer Res. **22**[2A], 1011-1016. 2002.

107. Devita, V. T., Hellman, S., and Rosenberg, S. A. *Cancer: Principles and Practice of Oncology*. 123-134. 2001. Philadelphia, PA, Lippincott Williams & Wilkins.

108. Martin, M. L., Lieberman, P. M., and Curran, T. *Fos-Jun dimerization promotes interaction of the basic region with TFIIE-34 and TFIIF*. Mol.Cell Biol. **16**[5], 2110-2118. 1996.

109. Antinore, M. J., Birrer, M. J., Patel, D., Nader, L., and McCance, D. J. *The human papillomavirus type 16 E7 gene product interacts with and trans-activates the AP1 family of transcription factors*. EMBO J. **15**[8], 1950-1960. 4-15-1996.

110. Li, J. J., Rhim, J. S., Schlegel, R., Vousden, K. H., and Colburn, N. H. *Expression of dominant negative Jun inhibits elevated AP-1 and NF-kappaB transactivation and suppresses anchorage independent growth of HPV immortalized human keratinocytes.* Oncogene **16**[21], 2711-2721. 5-28-1998.

111. Wong, W. Y., Havarstein, L. S., Morgan, I. M., and Vogt, P. K. *c-Jun causes focus formation and anchorage-independent growth in culture but is non-tumorigenic*. Oncogene **7**[10], 2077-2080. 1992.

112. van, de, V, He, Y. D., van't Veer, L. J., Dai, H., Hart, A. A., Voskuil, D. W., Schreiber, G. J., Peterse, J. L., Roberts, C., Marton, M. J., Parrish, M.,

Atsma, D., Witteveen, A., Glas, A., Delahaye, L., van, der, V, Bartelink, H., Rodenhuis, S., Rutgers, E. T., Friend, S. H., and Bernards, R. *A gene-expression signature as a predictor of survival in breast cancer*. N.Engl.J.Med. **347**[25], 1999-2009. 12-19-2002.

113. Karatzoglou, A., Meyer, D., and Hornik, K. *Support Vector Machines in R.* J.of Stat.Software **15**[9]. 2006.

114. Furey, T. S., Cristianini, N., Duffy, N., Bednarski, D. W., Schummer, M., and Haussler, D. *Support vector machine classification and validation of cancer tissue samples using microarray expression data* Bioinformatics **16**[10], 906-914. 2000.

115. Scholkopf, B., Platt, J. C., Shawe-Taylor, J., Smola, A. J., and Williamson, R. C. *Estimating the support of a high-dimensional distribution.* Neural.Comput. **13**[7], 1443-1471. 2001.

116. Yeatman, T. J. *The future of clinical cancer management: one tumor, one chip.* Am.Surg. **69**[1], 41-44. 2003.

117. Herranz, M. and Esteller, M. *DNA methylation and histone modifications in patients with cancer: potential prognostic and therapeutic targets.* Methods Mol.Biol. **361**, 25-62. 2007.

118. Chandrasekaran, E. V., Xue, J., Neelamegham, S., and Matta, K. L. *The pattern of glycosyl- and sulfotransferase activities in cancer cell lines: a predictor of individual cancer-associated distinct carbohydrate structures for the structural identification of signature glycans.* Carbohydr.Res. **341**[8], 983-994. 6-12-2006.

119. Xiong, X. D., Cai, W. J., Luo, J. M., Han, Y. L., and Li, E. M. *Identification of differentially expressed proteins between human esophageal immortalized and carcinomatous cell lines by two-dimensional electorphoresis and MALDI-TOF-mass spectrometry*. World J.Gastroent. **8**[5], 777-781. 2002.

120. Smith, R. D., Anderson, G. A., Lipton, M. S., Pasa-Tolic, L., Shen, Y., Conrads, T. P., Veenstra, T. D., and Udseth, H. R. *An accurate mass tag strategy for quantitative and high-throughput proteome measurements.* Proteomics **2**[5], 513-523. 2002.

121. Kononen, J., Bubendorf, L., Kallioniemi, A., Barlund, M., Schraml, P., Leighton, S., Torhorst, J., Mihatsch, M. J., Sauter, G., and Kallioniemi, O. P. *Tissue*

microarrays for high-throughput molecular profiling of tumor specimens. Nat.Med. **4**[7], 844-847. 1998.

122. van't, Veer, L. J., Dai, H., van de, V, He, Y. D., Hart, A. A., Mao, M., Peterse, H. L., van der., Kooy K., Marton, M. J., Witteveen, A. T., Schreiber, G. J., Kerkhoven, R. M., Roberts, C., Linsley, P. S., Bernards, R., and Friend, S. H. *Gene expression profiling predicts clinical outcome of breast cancer.* Nature **415**[6871], 530-536. 1-31-2002.

123. van de, V., He, Y. D., van't Veer, L. J., Dai, H., Hart, A. A., Voskuil, D. W., Schreiber, G. J., Peterse, J. L., Roberts, C., Marton, M. J., Parrish, M., Atsma, D., Witteveen, A., Glas, A., Delahaye, L., van der, V., Bartelink, H., Rodenhuis, S., Rutgers, E. T., Friend, S. H., and Bernards, R. *A gene-expression signature as a predictor of survival in breast cancer*. N.Engl.J.Med. **347**[25], 1999-2009. 12-19-2002.

124. Riley, R. D., Abrams, K. R., Sutton, A. J., Lambert, P. C., Jones, D. R., Heney, D., and Burchill, S. A. *Reporting of prognostic markers: current problems and development of guidelines for evidence-based practice in the future.* Br.J.Cancer **88**[8], 1191-1198. 4-22-2003.

125. Microarray Gene Expression Data Society: MGED Society. www.mged.org. Accessed 4-01-2007

VITA

Kendra Stisser was born December 8th, 1971 in Oklahoma City, Oklahoma. In 1995 she received her undergraduate degree in Microbiology (Chemistry minor) from the University of Arizona, Tucson, Arizona. Kendra won the "Senior Research Award" from the University of Arizona for her studies in molecular diagnostics of virus infections in the commercial fishery industry.

After graduating, Kendra worked as a "Special Projects Manager" for an industrial water testing company in Houston, Texas. In 1997, she was hired by Dr. Concepcion Diaz-Arrastia, M.D., as a "Senior Research Assistant" at the University of Texas, Medical Branch (UTMB), Galveston, Texas. In 2000, she matriculated into the Graduate School of Biomedical Sciences (GSBS) and currently under the tutelage of Dr. Arrastia.

While a student at UTMB, Kendra has earned the "Excellence in Clinical Sciences Research Award" by the Center for Interdisciplinary Research in Women's Health (CIRWH) and received the Who's Who Among Students in American Colleges. She has also been actively involved in student government and held simultaneous positions as the Chair of the Student Government Association (SGA) and President of the Graduate Student Organization (GSO). She is the first person to enroll in the concurrent Ph.D/ Masters in Business Association (MBA) Program.

Publications

Feussel, A. L., He, Q., Rady, P. L., Zhang, L., Grady, J., Hughes, T.K., Stisser, K. L., Konig, R., Tyring, S. K. *Nested PCR with the PGMY09/11 and GP5(+)/6(+) primer sets improves detection of HPV DNA in cervical samples*. J.Virol.Methods **122**[1]:87-93. Dec 2004.

Abstracts

Arrastia, C. D., Stisser, K. L. *Molecular Anatomy of Squamous Cell Carcinomas of the Uterine Cervix.* The New Frontiers in Cancer Detection and Diagnosis. January 23, 2007, Ventura, California

Stisser, K. L., Sinha, M., Luxon, B. A., Papaconstantinou, J., and Arrastia, C D. *Differential Gene Expression in Early Cervical Carcinomas Predictive of Risk of Recurrence.* The 11th Annual Structural Biology Symposium, May 19-20th, 2006, Galveston, Texas

Stisser, K. L., Sinha, M., Luxon, B. A., Papaconstantinou, J., and Arrastia, C D. *Differential Gene Expression in Early Cervical Carcinomas Predictive of Risk of Recurrence.* The National Women's Health Symposium**,** May 15-19th, 2006, Galveston, Texas

Stisser, K. L., Sinha, M., Luxon, B. A., Papaconstantinou, J., and Arrastia, C D. *Differential Gene Expression in Early Cervical Carcinomas Predictive of Risk of Recurrence.* Society of Gynecologic Oncologists**,** March 22-26th, 2006 San Diego, California

Stisser, K. L., Sinha, M., Luxon, B. A., Papaconstantinou, J., and Arrastia, C D. *Differential Gene Expression in Early Cervical Carcinomas Predictive of Risk of Recurrence.* Clinical Biomarker Summit, March 29-31st, 2006 Coronado, California

Stisser, K. L., Galindo, C., Papaconstantinou, J., and Arrastia, C D. *Differential Gene Expression in Early Cervical Carcinomas Predictive of Risk of Recurrence.* Gynec.Oncol. 101 (2006) p.S133.

Tenney, K. L., Arrastia, C D., Arany, I., Tyring, S. K. *A Better Consensus Primer for the Detection of Human Papillomavirus (HPV) in Cervical Biopsies.* 1st International Conference on Human Papillomavirus Infections & Cervical Cancer, July 7-11, 1998, Montreal, Quebec, Canada

Tenney, K. L., Nunan, L. M., Lightner, D. V., *Measurement Using polymerase Chain Reaction (PCR) of the Survival of Infectious Hypodermal and Hematopoietic Necrosis Virus (IHHNV) Subjected to Shrimp Culture Discinfection Techniques.* World Aquaculture Society Annual Meting, February 1-4, 1995, San Diego, California