

NOVEL BIOMARKERS FOR EARLY CANCER DETECTION AND SCREENING

by

GERALD LI

B.Sc., The University of British Columbia, 2005

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Interdisciplinary Oncology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

December 2012

© Gerald Li, 2012

Abstract

Early detection and screening have reduced mortality from many cancers, but there remains a need for improved biomarkers of risk. Cytometric DNA ploidy analysis has been used for the detection, treatment, and management of many cancers, but greater clinical utility would come with increased accuracy. Improvements to ploidy-based screening might come from adding complementary biological information.

The first aim combined ploidy with the additional biological information provided by malignancy associated changes as detected by automated nuclear morphometry. In 2249 sputum samples, the resultant biomarker, the Combined Score (CS), correlated with lung cancer risk factors like dysplasia grade, age, smoking status, and p53 and Ki-67 immunostaining. CS is a minimally invasive tool for risk assessment for the presence of precancerous lung lesions and could enrich chemoprevention trials with subjects likely to have high-risk dysplasias.

The second aim complemented ploidy with biological information provided by immunocytochemistry in a double staining procedure. Testing 49 cervical cytology brushings showed addition of Ki-67 immunostaining to distinguish abnormal cells from normal cycling cells did not improve ploidy's ability to separate high- and low-grade dysplasias. Nevertheless, double staining with Feulgen thionin and immunocytochemistry was shown to be technically feasible, even with antigen retrieval, and might be applicable to other immunocytochemical stains.

Motivated by the ability to combine ploidy with immunocytochemistry, the third aim investigated techniques for biomarker discovery pertinent to cervical dysplasia development. Cervical squamous epithelium consists of a continuum of differentiating cells and carcinogenesis

disrupts this cell maturation program. Gene expression differences between the basal and superficial epithelial layers and across various grades of dysplasia could catalyze the discovery of novel biomarkers through a better understanding of carcinogenesis. Microdissection and expression microarray analysis of molecular fixative preserved cervical biopsies resulted in the immunohistochemistry validation of four candidate targets showing correlation with dysplasia grade. This work underscores the importance and potential of accounting for heterogeneity within stratified squamous epithelium and constitutes the first report of successful gene expression microarray analysis of microdissected epithelial layers from molecular fixative preserved paraffin-embedded cervical specimens.

Ploidy combined with digital morphometry and immunocytochemistry can generate useful biomarkers of early squamous cell carcinomas.

Preface

Portions of this thesis have been published, as detailed below.

Section 1.3 is excerpted from: Cecic IK, Li G, & MacAulay C (2012) Technologies supporting analytical cytology: clinical, research and drug discovery applications. *J Biophotonics* 5(4):313-326. The publication was an invited review article and the section excerpted was written almost entirely by me, with editorial contributions by the other co-authors. In the original published manuscript, I wrote about 35% of the text, designed and created the figure, and helped with the formatting of the manuscript. IKC wrote most of the remainder and compiled the finished manuscript, while CM contributed a few excerpts and edited the manuscript. All co-authors contributed to the proof-reading and editing of this manuscript.

Chapter 2 is from: Li G, Guillaud M, leRiche J, McWilliams A, Gazdar A, Lam S, & MacAulay C (2012) Automated sputum cytometry for detection of intraepithelial neoplasias in the lung. *Anal Cell Pathol (Amst)* 35(3):187-201. I was responsible for data analysis and the preparation of the manuscript. I made all the figures (including performing the associated statistical analyses) and wrote essentially all the text for the manuscript. I also implemented the data adjustment to account for the effect of cell count on each slide. MG contributed image cytometry analysis, JIR was a study pathologist, AM was a bronchoscopist in the chemoprevention trials, AG was a pathologist and assisted in the design of the study, SL contributed to the design of this study and the chemoprevention trials cited therein, CM contributed to the design of the study and data analysis.

Chapter 3 is from: Li G, Guillaud M, Follen M, & MacAulay C (2012) Double staining cytologic samples with quantitative Feulgen-thionin and anti-Ki-67 immunocytochemistry as a method of distinguishing cells with abnormal DNA content from normal cycling cells. *Anal Quant*

Cytopathol Histopathol 34(5):273-284. I conducted all experiments described and wrote the manuscript. MG conducted the original study for which the specimens used here were collected and to which some of the present results are compared. MF contributed to the design of the original study for which the specimens used were collected. CM contributed to the design of the study.

In all of the above, edits have been made to the published version to provide extra clarity where needed and in order to integrate the material into the thesis. Many are minor and stylistic in nature. Of note, however, substantial changes have been made to the introductory sections of each manuscript, including incorporating material into Chapter 1 of this thesis.

Chapters 4 and 5 contain work from our collaborators. Fixation and embedding of molecular fixative samples was performed by staff at the Centre for Translational and Applied Genomics. All other tissue handling up to and including sectioning was performed by Dorothy Hwang. All gel electrophoresis experiments and microarray analysis of the frozen sample were performed by members of the laboratory of Dr. Cathie Garnis (Shevaun Hughes and Danielle Truong). Some of the microdissection of the frozen specimens (about 50% of the laser microdissection and all of the manually microdissected case 5065) was performed by Lorena Barclay. The Bioanalyzer assays were run by Bruce Woolcock. The study pathologist was Dr. Dirk van Niekerk. All other experiments and data analysis not listed above were performed by me.

The work described in this thesis was approved by the Research Ethics Board of The University of British Columbia and the BC Cancer Agency. The Ethics Certificate Numbers are: C98-0411, C02-0298, and C96-0426 for Chapter 2; H02-61476 and C99-0441 for Chapters 3 and 5; H03-61235 for Chapters 4 and 5; and R04-0112 for Chapter 5.

Table of Contents

Abstract	ii
Preface	iv
Table of Contents	vi
List of Tables	xi
List of Figures	xii
List of Abbreviations	xv
Acknowledgements	xvii
Dedication	xix
1 Introduction.....	1
1.1 Early cancer detection and screening.....	1
1.1.1 Lung cancer screening overview and challenges	4
1.1.2 Cervical cancer screening programs overview and challenges	6
1.2 Multistep carcinogenesis.....	7
1.2.1 Lung cancer biology and carcinogenesis.....	11
1.2.2 Cervical cancer biology and carcinogenesis.....	12
1.3 Image cytometry and quantitative image analysis	15
1.3.1 Clinical applications of image cytometry systems	16
1.3.2 Ploidy analysis.....	18
1.3.3 Other quantitative image analyses and applications.....	20

1.4	Gene expression microarrays.....	22
1.4.1	Data analysis and validation.....	24
1.5	Immunohistochemistry and immunocytochemistry.....	26
1.6	Hypothesis and aims	30
1.6.1	Aim 1: A novel sputum biomarker for bronchial dysplasia	31
1.6.2	Aim 2: Double staining as a potential improvement over thionin alone	31
1.6.3	Aim 3: Biomarker discovery via microdissection of cervical epithelial layers...	32
2	Automated Sputum Cytometry for Detection of Intraepithelial Neoplasias in the Lung	33
2.1	Introduction.....	33
2.2	Materials and methods.....	35
2.2.1	Chemoprevention trial subject recruitment and eligibility	35
2.2.2	Semi-automated quantitative sputum analysis	36
2.2.3	Ploidy measures.....	37
2.2.4	MAC measures.....	38
2.2.5	Biopsy collection and analysis	38
2.3	Results.....	40
2.4	Discussion.....	55
2.5	Conclusion	63
3	Double Staining Cytologic Samples with Quantitative Feulgen-Thionin and Anti-Ki-67 Immunocytochemistry as a Method of Distinguishing Cells with Abnormal DNA Content from Normal Cycling Cells.....	65

3.1	Introduction.....	65
3.2	Materials and methods.....	67
3.2.1	Cell culture.....	67
3.2.2	Patient samples.....	68
3.2.3	Immunocytochemistry.....	68
3.2.4	Thionin staining.....	69
3.2.5	Double staining with thionin and immunocytochemistry	70
3.2.6	Imaging and analysis.....	71
3.3	Results.....	72
3.3.1	Destaining of thionin by antigen retrieval.....	73
3.3.2	Reduction of thionin staining intensity after immunocytochemical staining due to antigen retrieval.....	73
3.3.3	Optimization of hydrolysis time.....	75
3.3.4	Image cytometry.....	76
3.3.5	Cervical cytology samples	77
3.4	Discussion.....	81
3.4.1	A procedure for optimizing double staining conditions	82
3.4.2	Ability to discriminate HGSIL and LGSIL samples	84
3.4.3	Limitations of double staining.....	86
3.5	Conclusion.....	89

4	Microarray Analysis of Microdissected Molecular Fixative Cervical Dysplasias: Technical Aspects	90
4.1	Molecular fixative.....	91
4.2	Materials and methods	93
4.2.1	Samples	93
4.2.2	Sample preparation.....	94
4.2.3	Microdissection	94
4.2.4	RNA extraction and purification	95
4.2.5	Microarrays	96
4.2.6	Data analysis	96
4.3	Results.....	98
4.3.1	Laser microdissection.....	99
4.3.2	Manual microdissection	100
4.3.3	Frozen samples.....	101
4.3.4	Molecular fixative samples	103
4.3.5	False discovery rates	107
4.3.6	Cluster analysis	108
4.4	Discussion.....	109
5	Microarray Analysis of Epithelial Layers in Cervical Dysplasia: Biological Validation	115
5.1	Introduction.....	115
5.2	Materials and methods	115

5.2.1	Data analysis	115
5.2.2	Human Protein Atlas	117
5.2.3	Validation samples	118
5.2.4	Immunohistochemistry	119
5.2.5	Data analysis of immunohistochemistry	121
5.3	Results.....	122
5.3.1	Target lists	125
5.3.2	Human Protein Atlas	145
5.3.3	Immunohistochemistry	145
5.4	Discussion.....	158
6	Conclusion	171
6.1	Recapitulation of aims	171
6.2	Impact and final remarks	172
	References	176
	Appendix: Microdissection of Molecular Fixative Samples	204

List of Tables

Table 1.1: Applications of quantitative image analysis in early cancer detection.....	20
Table 3.1: Patient specimens classified by conventional cytology and histopathology.	72
Table 4.1: Summary of all cases used in this study.....	99
Table 4.2: False discovery rates when differential expression criteria based on the scatter of duplicate data are applied to the three sets of duplicate data.....	108
Table 5.1: Summary table of all the LEEP specimens used in the validation of microarray targets by IHC.	119
Table 5.2: Target list found by selecting only those probes that were overexpressed in the top layers of at least half of the high-grade squamous intraepithelial dysplasia samples and less than half of the low-grade squamous intraepithelial dysplasias	128
Table 5.3: Target list found by selecting only those probes that were overexpressed in the bottom layers of at least half of the low-grade squamous intraepithelial dysplasias but less than half of the high-grade squamous intraepithelial dysplasias	131
Table 5.4: Target list of negative markers for high-grade squamous intraepithelial dysplasia, showing probes that are overexpressed in the upper layers of a majority of low-grade but not high-grade squamous intraepithelial dysplasia regions	137
Table 5.5: Target list from t-test analysis comparing only the top layers of high- and low-grade squamous intraepithelial dysplasias.	143
Table 5.6: List of genes found in the present analysis that have previously been associated with cervical cancer or CIN.....	145

List of Figures

Figure 2.1: Dependence of the Combined Score on the number of identifiable cells.	42
Figure 2.2: Box plots of Combined Score for sputum samples containing more than 500 identifiable cells, grouped according to the highest histopathological grade in that patient.	43
Figure 2.3: Analysis of Combined Scores for samples from patients in each of the four histopathological groups created by combining similar grades.	45
Figure 2.4: Box plots of Combined Score for sputum samples, sorted according to m-risk.	47
Figure 2.5: ROC curves showing the ability of the Combined Score to distinguish between high- and low-m-risk patients.	48
Figure 2.6: Box plots comparing p53 (A) and Ki-67 (B) staining score to histopathological grade for the 159 biopsy samples from sites that had a biopsy grading of dysplasia or worse at baseline or follow-up.	51
Figure 2.7: Box plots comparing Combined Score to the maximum p53 (A, B) and Ki-67 (C, D) staining score in that patient at that point in time.	53
Figure 3.1: Successful double staining of HL-60 cells.	74
Figure 3.2: Aborted double staining test.	75
Figure 3.3: A typical DNA histogram of Ki-67-negative cells taken from a patient who was negative for dysplasia.	77
Figure 3.4: ROC curve of the performance of the diploid-exceeding rate in detecting high-grade cervical dysplasias.	79
Figure 3.5: ROC curve comparing the use of Ki-67-positivity rate and double staining for detecting high-grade cervical dysplasias.	80

Figure 3.6: Microspectrophotometry data measured from spots of oxidized DAB chromogen deposited on a slide, before and after thionin staining, compared with thionin-stained HL-60 nuclei.	87
Figure 4.1: Example composite figure documenting microdissection of case 0028.	101
Figure 4.2: Plot of expression data from two aliquots of frozen sample case 5065 stroma.	102
Figure 4.3: 200 ng of basal layer from case 0028 in the middle lane of a 1% agarose gel.	104
Figure 4.4: Plot of expression data from two aliquots of case 0043 basal layer.	105
Figure 4.5: Plot of expression data from two aliquots of case 0028 basal layer.	106
Figure 4.6: Overlaid M-A plots of duplicate data for the frozen and two MFPE samples.	107
Figure 4.7: Cluster analysis of all log-transformed microarray data.	109
Figure 5.1: M-A plots of representative CIN I (0030) (A) and CIN III (0053B) (B) cases.	124
Figure 5.2: Summary of KLK7 staining in LEEP specimens	148
Figure 5.3: Sample LEEP sections (cases 0018 (A, B) and 0030 (C, D)) stained with KLK7 IHC (A, C), along with nearby reference H&E sections (B, D).	149
Figure 5.4: Summary of IFITM3 staining in LEEP specimens.	151
Figure 5.5: Sample LEEP sections (cases 0018 (A, B) and 0047 (C, D)) stained with IFITM3 IHC (A, C), along with adjacent reference H&E sections (B, D).	152
Figure 5.6: Summary of CRNN staining in LEEP specimens.	154
Figure 5.7: Sample LEEP sections (cases 0028 (A, B) and 0047 (C, D)) stained with CRNN IHC (A, C), along with adjacent reference H&E sections (B, D).	155
Figure 5.8: Summary of CLDN1 staining in LEEP specimens.	156

Figure 5.9: Sample LEEP sections (cases 0018 (A, B) and 0030 (C, D)) stained with CLDN1 IHC (A, C), along with adjacent reference H&E sections (B, D).	157
Appendix Figure 1: Case 0027.	204
Appendix Figure 2: Case 0028.	205
Appendix Figure 3: Case 0030.	205
Appendix Figure 4: Case 0033A.	206
Appendix Figure 5: Case 0033B.	206
Appendix Figure 6: Case 0043.	207
Appendix Figure 7: Case 0044.	207
Appendix Figure 8: Case 0053.	208

List of Abbreviations

The following is a list of abbreviations used in this thesis.

5cER	5c exceeding rate
AP	Alkaline phosphatase
AR	Antigen retrieval
AUC	Area under the curve
BCCA	British Columbia Cancer Agency
BSA	Bovine serum albumin
CIN	Cervical intraepithelial neoplasia
CIS	Carcinoma <i>in situ</i>
CS	Combined Score
CT	Computed tomography
CV	Coefficient of variation
DAB	3,3'-Diaminobenzidine (tetrahydrochloride)
DNA	Deoxyribonucleic acid
EGFR	Epidermal growth factor receptor
FFPE	Formalin-fixed paraffin embedded
FN	False negative
FP	False positive
H&E	Hematoxylin and eosin
HC II	Hybrid Capture II
HGSIL	High-grade squamous intraepithelial lesion
HPA	Human Protein Atlas
HPV	Human papillomavirus
HRP	Horseradish peroxidase
HSD	Honestly Significant Difference
ICC	Immunocytochemistry
IHC	Immunohistochemistry
IOD	Integrated optical density
LEEP	Loop electrosurgical excision procedure
LGSIL	Low-grade squamous intraepithelial lesion
MAC	Malignancy associated change

MFPE	Molecular fixative (preserved) paraffin-embedded
MI	Morphometry Index
NLST	National Lung Screening Trial
NPV	Negative predictive value
NSCLC	Non-small cell lung carcinoma
Pap smear	Papanicolaou smear
PBS	Phosphate-buffered saline
PPV	Positive predictive value
PSA	Prostate-specific antigen
RNA	Ribonucleic acid
ROC	Receiver operating characteristic
RT-PCR	Reverse transcription polymerase chain reaction
SAGE	Serial analysis of gene expression
SD	Standard deviation
TN	True negative
TP	True positive
Tris	Tris(hydroxymethyl)aminomethane

Acknowledgements

First off, a special thank you goes to my supervisor, Dr. Calum MacAulay, for the privilege of working in your laboratory, as well as your guidance and mentorship through this whole experience. I've learned a lot and I can say I am a better scientist for having worked with you.

I would like to thank my Supervisory Committee, Drs. Stephen Lam, Torsten Nielsen, and Haishan Zeng, for their helpful advice and insights on my work.

I would like to acknowledge the scholarship support provided by the Canadian Institutes of Health Research, the Natural Sciences and Engineering Council of Canada, and the Michael Smith Foundation for Health Research.

More specifically, the work described in the research chapters would not have been possible without the assistance and technical expertise of many people:

Chapter 2: I would like to thank Jagoda Korbelik, Anita Carraro, Priscilla Fung, Deanna Ceron, Anne Dy Buncio, and Sukhinder Khattri for technical assistance and database management. I would like to thank my co-authors on this project, Drs. Martial Guillaud, Jean leRiche, Annette McWilliams, Adi Gazdar, and Stephen Lam, for their contributions as detailed in the Preface. Funding for this project was provided by the National Institutes of Health (project numbers 1P01-CA96964-05 and 1U01-CA96109-05).

Chapter 3: I would like to acknowledge the assistance of Priscilla Fung with HL-60 cells and thionin staining, Jagoda Korbelik and Anita Carraro with imaging, as well as Dr. Haishan Zeng and Jianhua Zhao with microspectrophotometry. I would like to thank my co-authors on this project, Drs. Martial Guillaud and Michele Follen, for their contributions as detailed in the

Preface. Funding for this project was provided by the National Institutes of Health (R01-CA-103830 and P01-CA-82710).

Chapters 4 and 5: I would like to thank Dr. Dirk van Niekerk for access to patient samples and pathology review, and Drs. Martial Guillaud, Michele Follen, Dianne Miller, and Tom Ehlen for obtaining patient samples. As well, I would like to thank Dorothy Hwang for sectioning, Sylvia Lam, Judit Banáth, Lorena Barclay, May Zhang, and members of the laboratory of Dr. Cathie Garnis (especially Shevaun Hughes, Danielle Truong, Rebecca Towle, and Angie Chu) for technical assistance. I would further like to acknowledge Dr. Catherine Poh for access to the slide imager, Dr. Wan Lam for access to the NanoDrop and microarray scanner, and the Centre for Translational and Applied Genomics for access to the Bioanalyzer (and to Bruce Woolcock for helping run the machine). Funding for this project was provided by the National Institutes of Health (project numbers 1P01-CA96964-05 and 1U01-CA96109-05).

Finally, I want to thank all my other colleagues in the Integrative Oncology Department and the BC Cancer Research Centre who have helped me along the way but whose names I cannot all list here. Thank you for your support and friendship.

Dedication

To my family: my parents and my brother. It has been a long journey but the end is finally here. Thank you for all your support and for being behind me every step of the way.

1 Introduction

Cancer is a major public health problem. It is the leading cause of death in Canada (1) and one of the leading causes of death worldwide (2). In the global fight against cancer, multiple approaches are being taken, including improvements in prevention, screening, early detection, and treatment.

1.1 Early cancer detection and screening

Early detection is important for the successful management and treatment of many cancers. Lung cancer, for example, has an overall 5-year survival rate of just 16% (3), but this number jumps to about 70% for resected early-stage tumours (4). Meanwhile, mortality rates of cervical cancer vary widely between developed and developing countries, a discrepancy that has been attributed to the success of screening programs in the industrialized countries (2, 5, 6). If cancer could be detected and treated earlier, significant numbers of late-stage cancer cases might be prevented. There is potential for early detection and treatment to greatly reduce mortality and also morbidity associated with the symptom burden of late-stage cancer and the harmful side-effects of the treatments used to combat this disease.

The ability of a potential test to separate cancers from non-cancers can be quantified by several key measures. Any new test is compared against a widely accepted test, the gold standard, performed independently on the same test subjects. Based on the results of the two tests, each subject can be placed into one of four categories: true positives (TP) are those for which the new test and the gold standard agreed were positive (e.g., cancer was present), true negatives (TN) are negative by both tests, false positives (FP) are those identified as positive by the new test but

negative by the gold standard, and false negatives (FN) are negative by the new test but positive by the gold standard. The sensitivity of a test is its ability to detect correctly the positive cases (e.g., cancers) in a test population: $Sensitivity = TP / (TP + FN)$. The specificity is the ability to accurately classify all the negative cases as negative: $Specificity = TN / (TN + FP)$. While an ideal test would have 100% sensitivity and 100% specificity, in practice, improved sensitivity in a test often comes at the cost of decreased specificity. Hence, a good test for a given disease must consider the costs and benefits of improving sensitivity or specificity to strike the right balance between the two. For tests with an adjustable threshold for positivity, the interplay between sensitivity and specificity when setting different threshold scores can be represented graphically on a receiver operating characteristic (ROC) curve. A screening test that is not sensitive enough would cause many cancers to go undetected, providing patients with a false sense of security. On the other hand, a test that is not specific enough would subject too many patients to unnecessary investigations or treatment that may be associated with such adverse effects as discomfort, anxiety, or time lost from work, or may lead to complications or even death.

Another way to measure the effectiveness of a screening test is to look at the positive and negative predictive values (PPV and NPV, respectively). These refer to the probabilities that a positive or negative screening test result accurately reflects the presence or absence of disease, respectively, as determined independently using the gold standard test. In a test sample in which the prevalence of the disease mirrors that of the target population, $PPV = TP / (TP + FP)$ and $NPV = TN / (TN + FN)$. Whereas the sensitivity and specificity measure the discriminating power of the test itself, PPV and NPV also take into account the prevalence of the disease being detected. This can make PPV and NPV more meaningful than sensitivity and specificity when interpreting a test result. A test with a high NPV, for example, could provide an extra measure of reassurance to those with a negative test result, allowing for a lengthening of the test interval for

subjects with a negative result. This could make the overall screening program more cost-effective by reducing the number of tests required to be performed. However, as cancers in a typical screening population can be quite rare, it may be more costly to conduct a clinical trial to measure PPV and NPV directly. Alternatively, prevalence can be determined in the target population separately and used to calculate PPV and NPV from the measured sensitivity and specificity:

$$PPV = \frac{sensitivity \times prevalence}{(sensitivity \times prevalence) + (1 - specificity)(1 - prevalence)} \text{ and}$$

$$NPV = \frac{specificity \times (1 - prevalence)}{(specificity \times (1 - prevalence)) + ((1 - sensitivity) \times prevalence)}.$$

The success of a screening program depends on many factors (7, 8). The impact of screening is maximized in more common cancers with high rates of mortality and those where earlier stage disease has better treatment outcomes. The disease or screening endpoint should be well-defined and the test itself should clearly discriminate those with and without the disease, while remaining acceptable to patients and affordable by the health care delivery system. Those with a positive test result should have timely access to the appropriate care. Finally, the overall screening program should be cost-effective and ultimately reduce mortality.

Cervical screening programs in industrialized countries are often held up as examples of successful cancer screening. Participation rates are around 70% and mortality rates have come down significantly as a result of screening (9). Other screening tests like mammography for breast cancer (10, 11) and colonoscopy, fecal occult blood testing, or sigmoidoscopy for colorectal cancer (12, 13) have also been shown to lead to a reduction in mortality.

However, increased cancer screening is not always better. Prostate cancer, for example, is emerging as a case study on the potential problem of over-diagnosis. The discovery that serum levels of prostate-specific antigen (PSA) correlated with the stage of prostate cancer (14) led to the widespread adoption of PSA testing as a screen for prostate cancer. Incidence rates of prostate

cancer shot up, especially for early stage disease as more men were being diagnosed earlier with prostate cancer (15). More recent studies, however, suggest that while PSA screening might have reduced mortality (16), so many more men were being over-diagnosed and subjected to unnecessary and potentially harmful treatment that the risks outweighed the benefits in otherwise healthy individuals (17, 18).

The overall success of cancer screening is dependent on the nature of the cancer being screened for, the accuracy and costs of the test(s) being proposed to detect it, as well as the availability and effectiveness of downstream treatment options. A variety of early detection and screening technologies have been developed over the years, many of which have been adopted into routine clinical practice. Some cancer screening methods rely on directly visualizing the tumour mass (e.g., mammography for breast cancer) or abnormal cells (e.g., Pap smear for cervical cancer). Others look for secondary signs that might indicate the presence of cancerous cells, such as elevated prostate specific antigen (PSA) levels for prostate cancer or fecal occult blood tests for colorectal cancer. Ultimately, the unique biology of each type of cancer to be screened for dictates the type of test that would best be applied. Within this broad field of early cancer detection and screening, the work presented in this thesis will focus on lung and cervical cancers.

1.1.1 Lung cancer screening overview and challenges

Lung cancer is the leading cause of cancer death worldwide (2). Despite decades of work by clinicians and research scientists, lung cancer still has a bleak 5-year survival rate of just 16% (19). While the 5-year survival rate is about 70% for resected early-stage lung cancer (4), the majority of lung cancers are detected at an advanced stage with metastases already present (3, 4). Historical efforts to find an effective lung cancer screening strategy using chest X-rays to detect cancerous nodules and standard sputum cytology to find abnormal cells (4) unfortunately found

no decrease in mortality (20). Many promising techniques are currently being studied, including low-dose spiral computed tomography (CT) (21, 22), fluorescence bronchoscopy (23, 24), and sputum and blood biomarkers (25, 26). More recently, the National Lung Screening Trial (NLST) reported a 20% reduction in lung cancer mortality in high-risk smokers with low-dose CT screening (27).

Advances in imaging technology have led to faster and safer devices for detecting lung cancer. CT scans use multiple X-ray scanners and detectors to generate a three-dimensional image of a patient's lungs. Today's scanners can generate an image on one breath-hold, minimizing motion artifacts (28). More sensitive devices and improvements in computer-based image acquisition and analysis mean less radiation exposure for patients. Unlike traditional chest X-rays, spiral CT is able to detect many early-stage cancers, especially those in the peripheral lung. In contrast, bronchoscopy and sputum analysis are more sensitive to early-stage central airway cancers. In particular, autofluorescence bronchoscopy (29, 30) has been shown to improve sensitivity over white light bronchoscopy, but at the cost of a slight decrease in specificity (31). High-grade preinvasive cancerous tissues fluoresce a weak (brownish) red (30), while normal tissue autofluorescence is a more intense green.

While both spiral CT and fluorescence bronchoscopy are sensitive technologies to detect early lung cancer, both methods suffer from limitations, including low specificity (25, 32), and are costly if applied to high-risk subjects defined by age and smoking history alone. Consequently, a critical part of any lung cancer screening strategy is a means to assess a participant's risk of having cancer or precancerous lesions at high risk of progressing to cancer. This could, for example, be used to narrow the screening pool so that only the highest risk patients need to be screened, further improving the cost-effectiveness of a test like low-dose CT (26, 33).

As the NLST is the first large randomized controlled trial to show success with any lung cancer screening method, there is still no broad consensus on who should be screened for lung cancer. In the meantime, the American College of Chest Physicians and the American Society of Clinical Oncology have recommended that only smokers and former smokers between 55 and 74 years of age who have at least 30 pack-years of smoking history and who are still smoking or quit in the past 15 years should receive annual low-dose CT screening, and only in settings where patients can receive the same level of care as those provided in the NLST trial (34). In other words, only those patients fitting the description of those enrolled in the NLST trial should be offered annual low-dose CT screening for lung cancer.

1.1.2 Cervical cancer screening programs overview and challenges

Cervical cancer is the third most commonly diagnosed cancer in females globally (2). Screening programs based on the Papanicolaou (Pap) smear have significantly reduced mortality due to cervical cancer in industrialized nations (2, 5, 6). In Canada, for example, mortality from cervical cancer dropped almost 50% from 1973 to 1998 (9), and has continued to decline since then at a rate of 2.9% per year (35). In this procedure, a sample of cells is scraped from the cervix and deposited on a slide. In many countries that rely on liquid-based cytology, the cells are first immersed in a fixative solution before being transferred to a slide. The cells are stained with a series of dyes: hematoxylin stains nuclei dark blue, Orange G 6 stains keratin orange, and Eosin Azure (composed of Eosin Y, Light Green SF yellowish, and Bismarck brown Y) stains cytoplasm. The slide is then visualized under a microscope and interpreted by a trained cytotechnologist or pathologist.

Today, more than 85% of cervical cancer cases arise in low-resource settings, making cervical cancer the second-leading cause of cancer death among women in developing countries (2, 5). This presents a distinct challenge to establishing cervical screening programs where they

are needed most, as screening programs based on the Pap smear require an extensive and costly infrastructure, in addition to significant training and skill to interpret the patient slides. Potential alternatives include visual inspection with acetic acid (36) and testing for human papillomavirus (HPV) DNA (37-41). Unfortunately, visual inspection methods, even with the application of dilute acetic acid to cause abnormal regions to appear white (42), continue to rely on adequate training of practitioners while rollout of HPV testing programs in low-resource settings has been hindered by cost and logistics (43). Moreover, high-risk HPV testing has a higher sensitivity but lower specificity for detecting high-grade cervical precancers than conventional cytology (i.e., Pap stain with a human analysis process) (44) and low-resource settings are particularly sensitive to the follow-up costs of false positive cases.

Recently updated recommendations in the United States call for regular Pap tests for women between 21 and 65 years of age (45, 46). HPV tests are recommended for some women and some studies have called for a greater role for HPV testing or even HPV testing alone (41, 47), but cytology currently remains the foundation for most screening programs.

1.2 Multistep carcinogenesis

Carcinogenesis, the process by which normal tissue transforms into a malignant cancer, is a long and complex process that usually takes many decades. The prevailing hypothesis of multistep carcinogenesis suggests that cancer is the end result of a series of genetic and/or epigenetic alterations in normal cells that causes them to become cancerous (48-51). Colorectal cancer, for example, is believed to require at least 7 independent genetic alterations (50) while lung cancer might require 10 to 20 (52). In some cases, these changes might occur on a background of some degree of inherited predisposition to cancer. Through eons of evolution,

humans have developed innate defences against cells harbouring mutations. Meanwhile, the constant need to renew certain cells in our bodies and exposure to various mutagens in our environment continue to introduce potentially harmful mutations into our genes. For the most part, our bodies do an excellent job of repairing the damage. However, over time, mutations may be introduced that escape the natural repair mechanisms of our bodies. These mutations may confer new capabilities upon the transformed cells, eventually enabling them to become malignant (53).

There are many mechanisms by which normal cells can be transformed into abnormal cells. Environmental, physical, and chemical mutagens, for example, can introduce genetic mutations by damaging or chemically altering DNA. Smoking is known to introduce countless chemical mutagens to the lungs and is the major risk factor for lung cancer. Carcinogenesis may also be initiated by viruses, such as the human papillomavirus (HPV), implicated in cervical and a few other cancers. In the case of hereditary cancers, individuals may inherit a defective copy of a gene that predisposes them to developing cancer later on in life.

The search for specific genetic mutations driving carcinogenesis has traditionally focussed on oncogenes and tumour suppressors. Oncogenes are those that, once mutated, lead to the development of cancer. Tumour suppressors, on the other hand, normally function to prevent cancer. Once mutated, this suppression function is lost and the cell can progress one step further along the path of carcinogenesis. Typically, both copies of a tumour suppressor gene must be altered to render it ineffective, either by genetic damage to both copies or by inheriting a defective copy with subsequent damage to the other one. In some cases, a mutation to one copy may inhibit the function of the wild type copy. Such a mutation is known as a dominant negative mutation.

Mutations come in many forms. Individual nucleotides can be altered, deleted, or repeated, potentially leading to gain or loss of function, or even a change of function of the gene. Sometimes, larger sections of a gene can be mutated or two genes can be spliced together, forming a fusion gene. In many cancers, genomic instability is observed, such that large sections of chromosomes, or even entire chromosomes, can be duplicated, deleted, or even recombined. Normal non-dividing human cells have two homologous copies of each chromosome, which is referred to as diploid. Mitotic cells (and those that have replicated their genome in preparation for mitosis) are termed tetraploid. Abnormal cells harbouring gross chromosomal mutations may have DNA content outside the normal range between diploid and tetraploid or they may have a non-diploid resting DNA amount. This condition is called aneuploidy. Aneuploidy has been found in many cancers and is often associated with a worse prognosis (54). It has even been suggested that abnormal gene dosages caused by aneuploidy might be the critical event in carcinogenesis (55). As such, the quantitative assessment of cellular DNA content, also termed ploidy analysis, has been found to be effective in aiding the diagnosis, prognosis, and management of most solid tumours (56).

In addition to mutations to the coding regions of genes, which could directly affect the amino acid sequence of the gene product, mutations can occur in regulatory regions such as promoters, affecting the regulation of gene expression. The product of such a mutated gene might be over- or underexpressed relative to the wild type, which in turn may lead to potentially oncogenic downstream effects. An emerging field in the study of altered gene expression is epigenetics, where changes do not even occur to the DNA sequence itself. Changes to levels of DNA methylation, histone modifications, or even the secondary structures of nucleic acids can impact expression levels of many genes. Silencing RNAs, short pieces of RNA that lead to the degradation of complementary strands of mRNA, are another mechanism by which normal cells regulate gene expression and hence another route by which alterations can be introduced into the

genetic code to enable carcinogenic changes. Multistep carcinogenesis, then, can exploit any or all of the above mechanisms to disrupt the genetic program of normal cells in order to put them on the path towards malignancy.

The paradigm of multistep carcinogenesis lends itself well to the concept of “field cancerization,” first proposed by Slaughter *et al* (57) after observing that many individuals with one malignancy tend to develop second primary tumours. A mutation or genetic alteration in one cell may be passed on to its progeny cells, establishing a region of increased malignant potential. Under the two-hit hypothesis (58), only one more “hit” or mutation to any of the cells in this cancer field could potentially transform it into a malignant cell. More generally, in cancers where multiple “hits” are required to complete the transformation to malignancy, the cancer field represents an intermediate state where some but not all of the required genetic alterations have been acquired.

The long latency between the initial genetic insults and the clinical manifestation of invasive cancer presents an excellent opportunity for early detection and possibly early intervention. Chemoprevention approaches the neoplastic process as the disease, with invasive cancer merely being the final manifestation of this disease process (49). Consequently, the goal is to slow down, stop, or even reverse the process of carcinogenesis (59-62). Much work has gone into pursuing this strategy, with various levels of success, and excellent reviews of chemoprevention studies can be found in (59, 63, 64). Past work in our group and others, for example, has shown promise in lung cancer (65-68). Separate small studies using either *myo*-inositol or folate and vitamin B₁₂ found significantly increased rates of regression of dysplasias in the treated groups versus controls (67, 68). However, few trials have been successful. Many trials only show modest if any benefit from using chemopreventive agents and some even show negative effects. There remains a need for improved biomarkers both as intermediate endpoints

for evaluating efficacy of chemopreventive agents and for better risk stratification to identify patients who would benefit most from chemoprevention (64). Designing the optimal strategy for cancer screening and chemoprevention will require an understanding of the biological mechanisms underpinning carcinogenesis.

1.2.1 Lung cancer biology and carcinogenesis

Lung cancer is broadly classified according to histology into small cell and non-small cell lung carcinoma. Small cell carcinomas tend to be found in the central airways. They account for fewer than 15% of lung cancer cases (69), but are more aggressive and have a poorer prognosis. Non-small cell lung carcinomas (NSCLC) are further subdivided into adenocarcinomas (arising from Type II pneumocytes, Clara cells, or glandular cells), squamous cell carcinomas (arising from bronchial squamous epithelial cells), and large cell carcinomas. Of the NSCLCs, adenocarcinomas are the most common, accounting for over 50% of lung cancers, and are found primarily in the peripheral lung. Squamous cell carcinomas tend to be more centrally located and account for about 20% of all cases. Over the last few decades, there has been a general trend towards more adenocarcinomas relative to squamous cell carcinomas (70). NSCLC tumours are staged according to the traditional TNM (tumour, node, metastasis) system (71, 72). Due to the rapid progression and spread of the disease, small cell carcinomas are instead classified as either limited stage or extensive stage.

The progression of normal lung tissue to invasive cancer is a complex process, involving multiple stages of increasing genetic and molecular insult. Precancerous squamous cell lesions are classified according to a system laid out by the World Health Organization. In order of increasing severity, the histopathological grades are normal, hyperplasia, metaplasia, mild dysplasia, moderate dysplasia, severe dysplasia, carcinoma *in situ* (CIS), and cancer (73, 74). Of these, only the dysplasias and CIS are considered preneoplastic lesions (73). In the central airway,

Saccomanno *et al* showed that squamous cell carcinomas arise from a series of distinct pathological “stages” (75). Moreover, lower grade dysplasias have a considerably lower rate of progression than severe dysplasias and CIS (4). Besides those associated with squamous cell carcinoma, the other recognized classes of preinvasive lesions of the lung are atypical adenomatous hyperplasia and adenocarcinomas *in situ* (both believed to be precursor lesions to adenocarcinomas) (76) and diffuse idiopathic pulmonary neuroendocrine cell hyperplasia (74).

Smoking is the primary cause and risk factor for lung cancer, accounting for 80% of cases in males and 50% of cases in females globally (2). Other environmental risk factors include exposure to second-hand smoke, air pollution, radon, and asbestos. Tobacco smoke carries hundreds of chemicals, many of them known carcinogens, into the smoker’s lungs. These chemicals, along with other environmental carcinogens, are responsible for introducing the genetic damage that ultimately leads to lung cancer. Much work by many researchers has gone into identifying the specific changes responsible for lung carcinogenesis. Sato *et al* provide an excellent review of the many genetic and epigenetic alterations that have been implicated in lung cancer (77). Among these is p53, a tumour suppressor that is also the most mutated gene in human cancers. Inactivating mutations of p53 are found in approximately 90% of small cell lung cancers and 50% of NSCLCs (77). Mutant p53 is often detected at elevated levels in cells because while wild type p53 is constantly degraded, mutant p53 (in the absence of wild type p53) is not, causing it to accumulate in the cell (78). Meanwhile, mutant p53 can inhibit wild type p53 in a dominant negative phenotype.

1.2.2 Cervical cancer biology and carcinogenesis

The great majority of cervical cancers are squamous cell carcinomas (80%), with about 15% more being adenocarcinomas. In countries with well-established screening programs, a slightly higher proportion of adenocarcinomas is seen, perhaps because they arise from the glands

that are poorly sampled in conventional forms of screening (44). Staging is done according to the International Federation of Gynecology and Obstetrics system (79). In most developed countries, however, regular screening programs mean that most potential cases are detected as precancerous dysplasias or intraepithelial neoplasias. Dysplasias of the squamous epithelium are graded as mild, moderate, or severe, corresponding to cervical intraepithelial neoplasia (CIN) I, II, or III, respectively. Abnormal cells are confined to the bottom third of the epithelium in CIN I, spreading to include the bottom two-thirds in CIN II, and the full thickness of epithelium in CIN III. Carcinomas *in situ* (CIS) were previously considered the immediate precursor lesion to invasive cancer, being more severe than CIN III. However, the distinction between CIN III and CIS is no longer currently made. Furthermore, it has been suggested that CIN II and CIN III be combined as CIN II/III (80). A parallel grading system, known as the Bethesda system, is used for grading cytology (Pap) smears (81). This system divides dysplasias into low-grade squamous intraepithelial lesions and high-grade squamous intraepithelial lesions, along with various other classifications for other abnormal results. The risk that a precancerous lesion will progress to invasive cancer increases with the severity of the lesion.

Virtually all cervical cancers are caused by infection with HPV. Hence, it should come as no surprise that the story of cervical cancer carcinogenesis is closely intertwined with the biology of HPV infection. This process is wonderfully summarized in a seminar by Schiffman *et al* (44) and a review by Doorbar (82) but some of the highlights will be presented here.

There are many known types of HPV, but HPV 16 and HPV 18 are by far the most carcinogenic types, accounting for over 70% of all cervical cancer cases between them (83). After initial infection, viral particles reach the basal layer of the cervical epithelium. Cells in the normal squamous epithelium are arranged as layers, with stem-like cells believed to reside in the basal layer and more differentiated cells closer to surface.

The HPV genome consists of 8 genes, denoted E1, E2, E4, E5, E6, E7, L1, and L2. These are divided into early (E) and late (L) genes, based on when they are expressed in the host cell's differentiation program. Consequently, E genes are expressed closer to the basal layer while L genes are expressed more superficially. E4 is expressed throughout. While all the HPV genes play a role in acute infection (82), E6 and E7 are the most important drivers of carcinogenesis. Both interfere with the host cell's normal biochemical machinery in many ways, including inhibition of p53 by E6 (84) and disruption of retinoblastoma protein (pRb) activity by E7 (85).

The vast majority of HPV infections are cleared or suppressed by the host immune system within the first two years (44). However, a small proportion develops a persistent infection and is at greatest risk of progressing to cervical cancer. It is unclear what ultimately triggers some infections to become cancerous. It has been suggested this may be related to deregulation of E6 and E7 expression caused by the integration of HPV DNA into the host cell's genome (82). p16^{INK4A}, a marker for elevated E7 expression, might be a marker for HPV DNA integration (82, 86) and is strongly expressed in many high-grade lesions (87).

Work by countless researchers has greatly contributed to our understanding of cancer. To truly translate this knowledge into improved patient outcomes, however, requires the development of assays and imaging tools that inform us about how far a patient has progressed down the road of carcinogenesis. The next few sections will discuss a number of these tools, starting with image cytometry and quantitative image analysis.

1.3 Image cytometry and quantitative image analysis

A version of this section has been published. This section is modified and excerpted from Cecic IK, Li G, & MacAulay C (2012) Technologies supporting analytical cytology: clinical, research and drug discovery applications. J Biophotonics 5(4):313-326.

Ever since the invention of the microscope and the subsequent rise of cell biology, researchers and clinicians alike have sought to understand the biology behind the cellular structures they saw. However, early observations were of a qualitative nature and thus challenging to reproducibly teach and convey to others. Eventually, computer technology allowed these observations to become more objectively defined. Today, we can broadly define analytical cytology as the combination of single-cell analysis of cellular features for a defined analytical outcome. In response to the needs of researchers in the fields of cell biology, immunology, molecular biology, microbiology, and medicine, a variety of analytical cytology technologies have been developed over the past few decades, including flow cytometry, laser scanning cytometry, and image cytometry. Of these, this thesis will focus on image cytometry.

Compared to flow cytometry, image cytometry provides better structural resolution that allows operators to visualize cellular and subcellular morphology and requires less sample (88). Some image cytometers are also compatible with chromatic (absorbance-based) dyes, avoiding the problems of photobleaching and higher costs associated with fluorescence-based imaging systems. Fundamentally, image cytometry allows information about each cell to be derived from the image of the cell, instead of reducing the signal from each cell to a single data point. This also allows images to be reviewed to ascertain that signals are derived from cells and not other non-cellular debris that may be present in the sample. While imaging flow cytometers to some extent duplicate this function (89, 90), most image cytometers are also slide-based, which, unlike flow cytometers, gives operators the ability to revisit and track individual cells. This allows the same

individual cells to be measured before and after fixation (91) or to be repeatedly re-imaged after re-staining with different stains (92-95). As the cells remain on the slide after analysis, they can be reanalyzed by numerous methods, allowing more information to be derived from each individual cell.

Slide-based cytometry can be performed in fluorescence or absorbance (bright field) modes. Fluorescence-based cytometers are often (but not exclusively) laser scanning cytometers. This helps to minimize photobleaching as light exposure to each cell is kept to a minimum. Bright field image cytometry typically employs a white light source instead of lasers like flow and laser scanning cytometry systems. Transmitted light is usually collected by a charge-coupled device from a wider field of view, allowing many cells to be imaged at once. Although this allows for faster image acquisition, corrections have to be made to account for uneven illumination across the field of view, especially if quantitative analysis of image intensity is desired. Filters can also be applied to the source and/or the detector, depending on the desired analysis.

1.3.1 Clinical applications of image cytometry systems

Bright field image cytometry has been applied to a wide range of sample types. Early work was driven in large part by the fields of hematology, cytogenetics, and cervical smear analysis (96). Systems were built to automatically distinguish blood cell types (97-99) and to automatically identify metaphase nuclei for cytogenetic analysis (100-102). With increasing acceptance of Papanicolaou smear testing, there was considerable interest in developing prescreening machines, as the smears were tedious and time-consuming to grade manually and the false negative rate was high. Despite the many early successes of Pap smear screening programs, there remained many challenges associated with the test (103), so over the course of several decades, multiple groups from around the world designed, built, and in some cases commercialized, various machines designed to assist in the interpretation of not only cervical

cytology tests (104-125), but other cytological specimens as well, such as breast fine-needle aspirates (126-128), and urinary sediment (128-131).

PAPNET (formerly marketed by now-defunct Neuromedical Systems Inc, Suffern, NY, USA, but no longer commercially available) was developed as an automated system using neural networks to identify 128 of the most abnormal-appearing cells on each slide to be reviewed by a pathologist (132). PAPNET was found to be useful for prescreening triage (133) as well as rescreening of negative smears from women with a history of cervical abnormalities (134). NeoPath Inc (Redmond, WA, USA) marketed a pair of competing products, the AutoPap 300 QC and AutoPap Primary Screener systems. These systems complemented human review by being used to reassess negative smears or identify potentially malignant cells, respectively (135). However, the introduction of liquid-based cytology by Cytoc Corp with their ThinPrep method led to a major shift in the way cervical cancer screening was performed in the United States.

The ThinPrep Imaging system was approved by the US Food and Drug Administration in 2003 and today, the vast majority of cervical screening laboratories in the United States use some form of liquid-based cytology (136). The ThinPrep imaging system consists of a fully integrated workflow platform with a processor and imaging system utilizing a proprietary nuclear stain (137). Moving away from the traditional cervical smear interpretation, the development of semi-automated scanning of liquid-based cytology specimens emerged, utilizing the quantification of DNA content in abnormal cells being imaged on a slide highlighting cells of potential malignancy. The cervical brushings are placed in a specimen wash and deposited on a slide in a thin layer after which the cells are stained with proprietary DNA stain, the slide is scanned and imaged, and the DNA content of each cell quantified. Abnormalities in DNA content are highlighted in 22 fields of interest for a cytotechnologist to assess intensity of nuclear staining, size of the nuclei/cells, and morphological inconsistencies. The imaging system includes an

image processor, a computer storage device, and a review microscope. The image processor identifies multiple fields of diagnostic interest and the X and Y coordinates are recorded by the computer attached. The cytotechnologist can then refer back to areas of interest previously identified by the software algorithm. The areas are also highlighted for review by a pathologist. This is a semi-automatic diagnostic system that has potential for non-gynecological use, including fine needle aspirates, urines, and mucoid specimens.

Several companies manufacture and have commercialized similar liquid-based cytology prep and analysis systems, including Tri-Path (formed from the merger of NeoPath and AutoCyte and now a part of BD Diagnostics, Franklin Lakes, NJ, USA, this product line includes FocalPoint, SurePath, and PrepStain products) and Cytoc (developers of ThinPrep, now part of Hologic). While the performance advantages of liquid-based cytology over traditional Pap smears have been called into question (138, 139), most laboratories in the United States continue to use liquid-based cytology because the workflow is more efficient, there are fewer unsatisfactory samples, and the residual liquid sample can be used for the detection of oncogenic strains of human papillomavirus (HPV) by the Digene test (Qiagen, Hilden, Germany). Overall, imaging has improved the accuracy of the Pap test for cervical cancer screening, reducing human error, and increasing quality control. By combining manual screen by skilled cytotechnologists with imaged data, clinical studies have shown that imaged cohorts have a significant reduction in false negative rates, improved detection of low-grade squamous intraepithelial lesions and high-grade squamous intraepithelial lesions, and a significant decrease in the proportion of atypical squamous cells of undetermined significance (140-142).

1.3.2 Ploidy analysis

Many slide-based cytometry systems today rely on a quantitative DNA stain such as Feulgen staining (Table 1.1). In this method, a sample is subjected to acid hydrolysis, removing a

fraction of the nucleotide bases from cellular DNA. In the second step, a stain combines with these reactive abasic sites (143-145). Two stains commonly used with the Feulgen method, pararosaniline and thionin, have different spectral profiles (the former is red, while the latter is blue), but both perform favourably compared with other stains used to quantify DNA (146). Using a quantitative DNA stain, image cytometry can be used to study ploidy (i.e., the frequency distribution of total DNA within cells) and its implications for cancer detection and prognosis. Clinical guidelines for DNA cytometry have been standardized (147) as a steady stream of studies over the past few decades continues to show the benefits of ploidy analysis in the early detection, prognosis, and management of various types of cancers (56, 148, 149), including breast (148, 150), cervical (148, 151-159), lung (160-163), oral (164-166), and prostate (148, 167, 168) cancers. In cervical cancer, for example, ploidy analysis was found to perform comparably with high-risk HPV testing and conventional cytology (152) and is being used in China for screening (151). Ploidy analysis for cancer detection has already been standardized for clinical use (147). Whereas the need for accurate ploidy measurements in clinical samples typically precludes the use of paraffin-embedded tissue due to concerns of how nuclear overlap and sectioning truncation may impact accurate segmentation and quantification, the introduction of the Hedley method of isolating intact nuclei from archival paraffin blocks (169) has expanded the range of samples that can be used for ploidy analyses, though usually in combination with flow cytometry (155-157, 170-176).

Feulgen stain for DNA ploidy analysis	
Breast	(148, 150)
Cervical	(148, 151-159)
Lung	(160-163)
Oral	(164-166)
Prostate	(148, 167, 168)
Feulgen stain for nuclear morphometry and malignancy associated changes associated with cancers	
Breast	(182, 183)
Cervical	(184-188)
Colon	(189, 190)
Lung	(177, 180, 182, 191, 192)
Oral	(193)
Hematoxylin and Eosin (H&E)	
Oral lesions	(198)
Skin cancers	(199)
Thyroid lesions	(200)

Table 1.1: Applications of quantitative image analysis in early cancer detection.

1.3.3 Other quantitative image analyses and applications

Besides ploidy analysis, image cytometry on cells stained with a stoichiometric DNA stain like Feulgen can be used for measuring nuclear morphometry. One application of this technique is for detecting malignancy associated changes (MACs). These are subtle morphological changes in seemingly histologically normal cells in patients with cancer and are hypothesized to arise from reactions in the non-malignant cells to factors released by cancer cells (177). MACs were first observed by Gruner in 1916 (178) and subsequently expanded upon by

Nieburgs *et al* (179, 180), but their practical utility for cancer detection was limited until computer-assisted image analysis allowed these changes to be quantitatively measured by Klawe and Rowinski in 1974 (181). MAC analysis has been used in patients with breast (182, 183), cervical (184-188), colon (189, 190), and lung (177, 180, 182, 191, 192) cancers, among others. By combining ploidy measures and nuclear morphometry for MACs, novel biomarkers have also been developed for detecting lung (191) and oral cancer (193). These tests are marketed by Perceptronix Medical Inc (Vancouver, Canada) as LungSign and OralAdvance, respectively, and have both received Health Canada and CE Mark approval. Chromatin texture features in cervical histology (194) and cytology (195) specimens have also been found to be indicative of high-risk HPV. Meanwhile, image cytometry has shown promise as a tool for risk assessment of precancerous lesions and for assisting clinical trials of chemopreventive drugs. Risk assessment of precancerous lesions by histopathology has been hampered by the low level of agreement between pathologists (196, 197), while chemoprevention trials are held back by the low risk of developing cancer even in untreated cohorts and the lack of suitable surrogate end points that would allow trials to be completed in a timely and cost-effective manner (64). Image cytometry provides an objective measure of risk in bronchial (197) and cervical dysplasias (152), for example. Image cytometry can be used as a secondary end point for chemoprevention trials (66), while ploidy measurements can be used to enrich such trials with higher risk patients more likely to benefit from chemoprevention (65, 66).

Image cytometry has been applied to other types of cellular stains. Using standard hematoxylin and eosin staining, basic nuclear morphological features like the area, diameter, and perimeter can be easily measured and have been investigated for use in assessing oral lesions (198), skin cancers (199), and thyroid lesions (200). Immunocytochemical staining can also be analyzed quantitatively with image cytometry. For example, staining of 5-bromo-2'-deoxyuridine has been used to study cell proliferation (201), while quantification of synapses (202) and

measurements of hormone levels in the hypothalami and pituitaries of mice (203) have further been demonstrated by image cytometry. Using a consumer-model digital camera and public-domain image analysis software, a group in India also demonstrated how image cytometry could be used to predict metastasis in oral cancer (204), underscoring the potential of image cytometry as a low-cost approach to various clinical problems in low-resource countries. Studies have investigated, for example, the use of ploidy analysis via image cytometry to screen for cervical cancer in China (151, 153).

While cytometric techniques like ploidy analysis can detect large-scale chromosomal abnormalities that might be indicative of cancer, a more molecular approach toward cancer biomarker research can be taken. Beyond simply identifying which cells might have dysregulated genomes, one could probe how such genetic changes manifest themselves within the biochemical machinery of the cell to produce a malignant phenotype.

1.4 Gene expression microarrays

The completion of the Human Genome Project ushered in a new era of biomarker discovery. No longer were we confined to studying individual proteins or molecules, but the entire human genome could be studied at once. Advances in computer technology have also enabled us to process the vast amounts of data that such approaches generate.

One important tool for studying the genome is the cDNA or oligonucleotide microarray, an excellent overview of which is presented in (205). Earlier methods of exploiting complementary base pairing to interrogate nucleic acids present in a complex sample often involved the application of a single, pure, labelled probe on to an immobilized analyte (e.g., Southern blotting or fluorescence *in situ* hybridization). In a microarray, short DNA sequences

complementary to the genes of interest are immobilized directly on a solid support (often a glass slide) and the complex analyte is labelled and applied on to the bound probes. Each spot on the array contains identical copies of a particular oligonucleotide sequence, known as probes. The earliest microarrays consisted of only a few select genes of interest (206, 207), but technological improvements and the availability of the complete sequence of the human genome today allow the reliable construction of microarrays spanning the entire human genome. Alternatively, microarrays can be constructed to study specific mutations or polymorphisms. Microarrays can be used to study DNA, looking for gene mutations or copy number variations, for example. RNA can also be hybridized to microarrays, providing a snapshot of gene expression or, more recently, of regulatory RNA such as siRNAs or microRNAs. Often, one sample is hybridized to a microarray, but two samples can be simultaneously hybridized, such as with comparative genomic hybridization, where the genomes of two samples can be compared directly.

A typical experiment using microarrays to study gene expression starts with purifying mRNA or total RNA from the sample. mRNA is then amplified and labelled. Reverse transcriptases generate cDNA from the RNA sample. Many commercial kits use poly-deoxythymidine primers to specifically amplify mRNA by recognizing the poly-A tail, but some use random primers to amplify all RNA, avoiding a well-known bias of the former kits in which the 3' ends of transcripts are preferentially amplified. The cDNA is then used as a template for transcription by RNA polymerases. A fraction of the substrate nucleotide triphosphates is labelled with a fluorescent probe so that this probe becomes incorporated into the final cRNA. In comparative hybridization experiments, each sample would be labelled with a different fluorophore. The labelled cRNA is purified and applied to the microarray, where complementary strands hybridize. After washing off unbound material, the microarray is placed in a scanner where the signal from each spot is quantified. The fluorescence intensity of each spot correlates with the concentration of the corresponding oligonucleotide sequence in the original sample.

1.4.1 Data analysis and validation

Data from a microarray experiment must be processed, analyzed, and validated. The data must be background corrected and normalized so that the data itself is meaningful and comparable between samples. Then, it is analyzed to determine genes or targets of interest. Finally, these targets must be validated to confirm that the analysis has produced a result of biological meaning and significance. Many strategies have been developed over the years to handle the analysis challenges presented by microarray data, with many of the common ones now implemented in free and commercially available software packages. For a good overview of some of the common data analysis methods, the reader is referred to (208).

The first step of microarray data analysis is background correction to separate the signal from measurement noise, i.e., samples for which the signal is too weak. Often, this can be as simple as subtracting a background level from all the data and discarding negative values. To account for spatial variations across a microarray slide, many array scanners use a background level determined from the pixels surrounding each probe spot. For added stringency, data within one or two standard deviations from the background level can also be discarded.

A major step in microarray data analysis is normalization. Data from one sample must be made comparable to data from another sample, accounting for such factors as different quantities of starting material, variations in labelling or detection efficiencies, or even systematic biases. Again, many approaches are used in the literature and a few of the more common ones have been summarized previously (209-211). The ultimate goal of microarray analysis is typically to compare between two samples, so data at this point is often transformed into intensity ratios. This allows the construction of M-A plots, in which $M = \log_2(S_1/S_2)$ is plotted against $A = \log_2 \sqrt{S_1 \cdot S_2}$, where S_1 and S_2 are the signal intensity values of the two samples (212). The null hypothesis is typically that there is no difference between the expression levels of the majority of

genes in the two samples. Consequently, a fit of the data in the M-A domain can be performed and subtracted from the data set to achieve normalization. Historically, a global average M was often used as the correction factor, but observations of an intensity-dependent deviation of M values away from zero (213) led to the use of more complex fits such as linear or locally weighted scatterplot smoothing (LOESS or LOWESS) fits.

Once microarray data has been normalized, the data is analyzed for genes of interest. As with other stages of microarray analysis, different groups have developed different strategies. A simple approach is to simply call all genes exhibiting greater than a fixed threshold fold-change as differentially expressed. Two-fold change is a common threshold, but doesn't account for intensity-dependence of signal noise. An alternative approach, then, is to use the local standard deviation in the data on an M-A plot to set the appropriate thresholds. For each data point, the standard deviation of all data with similar A values is calculated. If the data point exceeds a threshold multiple of standard deviations from its neighbours, it is considered significant. Another approach is to compare the different groups on a gene-by-gene basis using t-tests.

Regardless of the method used to generate a list of genes of interest, any useful target must be biologically relevant. The final step, validation, attempts to demonstrate this. Targets can be validated by showing that the gene is in fact differentially expressed in the compared samples via an independent test such as reverse transcriptase polymerase chain reaction or immunohistochemistry. Alternatively, validation may involve verifying that the observed differences agree with previous literature, either on an individual gene basis or as part of a biochemical pathway. Many pathway analysis tools (e.g., IPA (Ingenuity Systems, www.ingenuity.com, Redwood City, CA, USA), Cytoscape (214), GSEA (215, 216), DAVID (217, 218), and PANTHER (219)) are currently available for this.

1.5 Immunohistochemistry and immunocytochemistry

Interest in molecular biology has led to the development of many techniques to probe the intracellular distribution of specific biomolecules. Two such techniques are immunohistochemistry (IHC) and immunocytochemistry (ICC), sometimes collectively referred to as immunostaining or immunolabelling. IHC and ICC differ primarily in the types of samples they are applied to: IHC is used on tissue sections, while ICC is used on cytology samples where most of the extracellular matrix has been removed and information on tissue architecture and physical intercellular interactions is mostly lost. Both techniques are well-established in the literature and rely on specific binding of antibodies to a target of interest.

Both IHC and ICC are typically performed on fixed samples. Samples need to be rehydrated, meaning paraffin-embedded IHC samples also need to be dewaxed. The staining procedure can be divided into four major steps: antigen retrieval, blocking steps, primary antibody incubation, and visualization.

The first major step of immunostaining is antigen retrieval (AR). This step reverses some of the cross-links formed during aldehyde fixation, thereby exposing epitopes that may have been masked by structural changes that resulted from fixation. Aldehyde-based fixatives such as formalin (aqueous formaldehyde) are believed to fix tissue by forming cross-links between proteins and other biomolecules (220). While these cross-links increase the durability of the tissue in preparation for the harsh chemical treatments it may be subject to, they can conceal antigenic binding sites recognized by antibodies. Early researchers found IHC to often be unreliable, with discrepancies in the literature as to the presence and intensity of observed staining (221). Attempts were made to use enzymatic digestion with proteases to improve IHC results (221, 222). Although better than without pre-treatment, this was still unreliable and not broadly applicable. The discovery of heat-mediated antigen retrieval opened up immunostaining

to a wider range of possible targets (223). An excellent review of heat-mediated antigen retrieval can be found in (224). Microwave heating, in particular, allowed immunostaining procedures to be completed quickly and inexpensively. Heat-mediated AR involves immersing the sample slides in a buffer solution, most commonly citrate at pH 6, but occasionally tris(hydroxymethyl)aminomethane (Tris) and/or ethylenediaminetetraacetic acid (EDTA) at alkaline pH. Heat, from a microwave oven, pressure cooker, or even a simple water bath, supplies the thermal energy required to hydrolyze the cross-linking bonds and appears to be the crucial element of AR (224). Microwave irradiation can cleave proteins at aspartyl residues, but this seems to have only a very limited effect in AR (225). Secondary effects of the procedure that might also play a role in improving immunostaining effectiveness include rehydration (and consequently renaturation) of proteins as well as chelation of metal ions (notably calcium) that were incorporated into proteins during fixation (220, 226) (all the above-mentioned commonly used buffering agents double as good chelators), although the latter has been called into question (227) and the exact mechanism of heat-mediated AR is not yet fully understood. Some antibodies do not require any form of antigen retrieval, even when used on formalin-fixed samples and some antibodies require AR even when used with non-formalin-fixed samples. The optimal antigen retrieval protocol for any given antibody on any given sample type must be determined empirically.

The second major step of immunostaining is a series of blocking steps. The purpose of this step is to increase the specificity of the immunostain by preventing endogenous factors from generating a signal through mechanisms other than the intended antibody-antigen interaction and its associated visualization reactions. The decision of which blocking agents to use is based on the antibody and the chosen visualization method. Hence, blocking agents will be detailed below with the corresponding visualization method.

Primary antibody incubation is straightforward: after antigen retrieval and the necessary blocking steps, the antibody is applied to the sample, allowing it to bind to its antigen. Unbound antibody is then washed off with a fresh buffer solution (often containing a small amount of detergent).

Visualization of immunostaining typically involves the enzyme-catalyzed reaction of a chromogen precursor to form the final insoluble chromogen that can be seen under a regular light microscope. When a fluorescent probe is used instead, the immunostaining procedure is typically referred to as immunofluorescence. For IHC and ICC, the enzyme can be bound directly to the primary antibody or indirectly via a secondary antibody. The most commonly used enzymes for immunostaining are horseradish peroxidase (HRP) and alkaline phosphatase (AP). HRP catalyzes the oxidation of the chromogen precursor, commonly 3,3'-diaminobenzidine tetrahydrochloride (DAB), by hydrogen peroxide. DAB polymerizes upon oxidation and appears as a brown precipitate, insoluble in water, alcohol, and xylene. A less commonly used HRP chromogen is aminoethyl carbazole, which forms a red precipitate, but is alcohol-soluble. When HRP is used, the blocking step must include a way to quench any peroxidase-like activity in the sample. This is usually achieved by incubation with aqueous or methanolic hydrogen peroxide solution. Methanol is a strong inhibitor of peroxidase (228). AP, on the other hand, catalyzes the cleavage of a phosphate group from its substrate. The other reaction product, besides inorganic phosphate, is now activated and undergoes further reaction to form the visible chromogen. Endogenous AP can be blocked by levamisole. AP is also inhibited by inorganic phosphate, so Tris-buffered saline should be used in place of phosphate buffered saline with AP.

When the visualization enzyme is not bound directly to the primary antibody, it is bound to a secondary antibody, which is applied to the sample after washing off the primary antibody. Secondary antibodies are chosen to recognize all antibodies raised in the same species as the

primary antibody. However, some samples might bind to the secondary antibody directly (bypassing any interaction with the primary antibody). A block for non-specific binding of antibodies is therefore needed. One option is to incubate the sample with serum from the species in which the secondary antibody was raised. This would competitively fill any binding sites in the sample that recognize the secondary antibody. Alternatively a universal, serum-free block can be used. This block works by reducing non-specific hydrophobic interactions between the antibodies and cellular components. It contains components like casein to block hydrophobic binding sites.

In addition to the common visualization methods described above, there are a number of variations that have been developed in an attempt to improve sensitivity and specificity of immunostaining. These include, for example, exploiting the highly specific and tight binding of the avidin-biotin complex to bring more copies of the visualization enzyme/probe to each bound primary antibody. Meanwhile, the EnVision system (and its successors) from Dako (Glostrup, Denmark) utilize a dextran polymer to bind more enzyme molecules to each secondary antibody.

When performing an IHC or ICC analysis, it is important to run positive and negative controls at the same time. This ensures the reliability of the observed positive or negative staining in the test slides by exposing any systematic errors that may have occurred during any given batch. Controls are typically selected on the basis of existing literature or information provided by the antibody vendor and are ideally handled in the same manner as the test samples, including fixation and processing methods. Alternatively (or additionally), a sample that should otherwise stain positively can be run without a primary antibody incubation (replaced with only the antibody diluent). Staining in such no-primary controls might indicate, for instance, inadequate washing, secondary antibody that is too concentrated, or inadequate quenching of endogenous peroxidase (when HRP is being used).

Despite their many similarities, a crucial difference between IHC and ICC is the tissue morphology information that can be derived from IHC. IHC samples are typically taken from localized biopsies, whereas cytology specimens for ICC might include a mixture of various cell types. Tissue architecture information enables IHC analyses regarding intercellular relationships that would be difficult if not impossible with ICC. On the other hand, the requirement for tissue sections for IHC restricts it to fixed samples. ICC can be performed on live cells, opening other avenues of inquiry not accessible to IHC. In a clinical screening setting, cytology specimens may also be preferable to biopsies because they represent sample from a wider region, obviating the need to find and sample a suspect lesion directly. Moreover, cytology specimens can sometimes be obtained in a less invasive manner compared to biopsies. Methodologically, ICC also requires an additional permeabilization step, typically achieved by incubating the sample with a detergent solution. Cells in IHC samples are already permeable on account of having had a microtome blade pass through them.

1.6 Hypothesis and aims

The work of countless researchers over the past few decades has greatly enhanced both our understanding of cancer and our ability to fight this disease. Looking ahead, it is believed that the development of novel biomarkers will aid and play a crucial role in the early detection, screening, and management of cancers and precancerous lesions. In particular, cytology appears promising as a screening tool for lung and cervical cancer. Cytology sample collection is simple and minimally invasive. DNA ploidy cytometry has already been used effectively for cervical screening on over 3 million women in China. This thesis seeks to improve cytometric ploidy-based cancer screening by layering independent biological information on to the DNA ploidy information. The hypothesis examined in this work is that a test combining additional biological

information with ploidy information will perform better than ploidy alone as a screening test for early preinvasive neoplasias.

1.6.1 Aim 1: A novel sputum biomarker for bronchial dysplasia

This aim is to develop a novel sputum biomarker based on ploidy and malignancy associated changes (MAC) features that can improve identification of subjects harbouring dysplasia or cancer for detection of preinvasive neoplasias and chemoprevention. A combined ploidy and MAC biomarker, applied to patients at high-risk of lung cancer, should correlate with other known lung cancer risk factors and act as a more effective means of stratifying lung cancer risk than age and smoking status alone. As such, the addition of MAC features represents additional biological information that can be gleaned with minimal additional effort (no changes required to sample collection or processing).

1.6.2 Aim 2: Double staining as a potential improvement over thionin alone

Ploidy analysis alone cannot always distinguish between cells with abnormal DNA content and normal cycling cells. Aim 2 is to develop and test a double staining strategy combining immunocytochemical staining for proliferation marker Ki-67 with thionin. The additional biological information provided by the immunocytochemical stain should improve the sensitivity and specificity of double staining for detecting high-grade dysplasias over thionin staining alone by improving the recognition of abnormal cells. As such, we are attempting to add biological information from a marker known to correlate with cancer risk and progression in the cervix, at least as seen in cervical tissue samples.

1.6.3 Aim 3: Biomarker discovery via microdissection of cervical epithelial layers

Technologically, Aim 2 unlocks the possibility of combining any immunostain with absorbance-based ploidy analysis. Normal human cervical squamous epithelium consists of a differentiating continuum of cell layers. The basal layer is believed to consist of stem cells, with cells maturing and differentiating as they move towards the surface. Carcinogenesis upsets this regulated program of cell maturation. I aim to use microdissection of cell layers and gene expression analysis to better understand these oncogenic processes. It is intended that this understanding can be translated into the development of novel biomarkers that can be assayed on cervical cytology specimens, either alone or in combination with thionin to improve cytometric ploidy analysis. As such, we are attempting to uncover biological information by identifying a marker that will correlate with progression in the cervix and likely be differentially expressed in cells of conventional cytological samples.

2 Automated Sputum Cytometry for Detection of Intraepithelial Neoplasias in the Lung

A version of this chapter has been published as: Li G, Guillaud M, leRiche J, McWilliams A, Gazdar A, Lam S, & MacAulay C (2012) Automated sputum cytometry for detection of intraepithelial neoplasias in the lung. Anal Cell Pathol (Amst) 35(3):187-201. Edits have been made throughout for additional clarity and integration into the flow of the thesis.

2.1 Introduction

Lung cancer is a major health problem worldwide. Its low survival rates can be attributable to the fact that lung cancers are seldom detected at an early stage where curative treatment is more likely (4, 19). To decrease lung cancer mortality, a strategy is needed to identify both patients with early disease, for treatment, and those with precancerous disease at risk of cancer development, for chemoprevention (229). Some screening technologies like spiral CT and fluorescence bronchoscopy, while sensitive to early lung cancer, are invasive methods that are costly if applied to a subject population defined as high-risk based on age and smoking history alone. Hence, risk assessment would also be useful in selecting only the very highest risk patients to receive more costly or invasive screening methods. Furthermore, if the risk methodology is based on molecular alterations in the lung's genetic material, then changes in the assessment can be used to monitor the effectiveness of treatment or chemoprevention. The process of carcinogenesis is a complex one, transforming normal lung tissue to invasive cancer via a series of steps involving increasing genetic and molecular insult. Ascertaining the degree to which areas of the lung have progressed down this path – and the corresponding increased risk of cancer – is clinically important, as it should guide screening and chemopreventative therapy decisions.

The internationally accepted standard prognostic factor for lung cancer risk is the histopathological grade of a bronchial biopsy based on the World Health Organization classifications (230). Biopsy, taken either endoscopically or surgically, is an invasive procedure and so other attempts to quantify risk have focused on patient factors, or biomarkers in sputum or blood samples (25, 26, 231). Blood screen methodologies have included circulating DNA and RNA markers and proteomic profiling (26, 232, 233). A study employing the detection of nanoarchitectural changes in buccal cells to detect lung cancer gave promising results, but their analysis was based on a small number of manually selected cells from a small number of patients (234). In sputum, markers such as Ras and hnRNP B1 and the aberrant methylation of tumour suppressor genes have all been investigated (25). Recent studies have also studied the presence of specific chromosomal abnormalities in sputum using fluorescence *in situ* hybridization (235, 236) and pulmonary function (160) as possible risk factors for lung cancer.

Sputum biomarkers are promising because they are relatively quick and inexpensive while being adaptable to large-scale population screening (160), making them practical tools to guide both subsequent screening and chemoprevention trials. Studies into the diagnostic utility of conventional sputum cytology (summarized in (237)) have reported widely varying results, likely due to differences in methodologies between studies and significant intra- and inter-observer variations in identifying abnormal cells (31). However, most studies have been directed at the detection of tumours and a lack of research remains into the utility of sputum cytology as a risk assessment tool for precancerous lesions.

As the percentage of bronchial epithelial cells in sputum can be quite low, reported sensitivities of sputum cytology for lung cancer detection also tend to be quite low (31). Malignancy associated changes (MAC) are subtle morphological and physiological changes that have been observed in non-malignant cells when cancer is present in a patient (179). These

changes may be due to soluble factors secreted by the malignant cells and can be measured using image cytometry (177). Due to the larger number of non-malignant cells expected to exhibit MACs, we expect techniques based on MACs to be more sensitive than conventional cytology. Previous work in our group and by others has shown that automated image cytometry based on MACs can be used to detect lung cancer (191, 192), although care may be needed to account for the possible confounding effect of non-malignant pulmonary diseases (238).

In this study, we correlate a number of published lung cancer risk factors – histopathological grade of biopsies from the bronchial tree, age, smoking status, quantitative morphometry, p53 and Ki-67 biopsy status – to a novel sputum biomarker assay based on cell population ploidy status (i.e., the presence or absence of cells with abnormal amounts of DNA) and malignancy associated changes (180, 192).

2.2 Materials and methods

Cell samples and data were drawn from several National Cancer Institute-sponsored lung cancer chemoprevention trials in high-risk smokers, as defined by age and smoking history, (65-67) and from patients undergoing investigation for suspected lung cancer. A total of 2249 sputum samples were obtained between 2000 and 2006 from 1795 participants.

2.2.1 Chemoprevention trial subject recruitment and eligibility

For this study, a former smoker is defined as someone who has not smoked in the previous 12 months. A current smoker has smoked in the previous 12 months. Former and current smokers between 40 to 74 years of age with a smoking history of ≥ 30 pack-years were recruited for the chemoprevention studies through the community outreach network of the public relations

department of the British Columbia Cancer Agency (BCCA) using television programs, radio broadcasts, and through local newspapers. Following an initial interview during which study subjects completed a questionnaire to document their smoking history, we obtained a sputum sample from each subject using simultaneous high-frequency chest wall oscillation with an ABI Vest (Advanced Respiratory Inc., St. Paul, MN) and inhalation of 3% hypertonic saline from an ultrasonic nebulizer for 12 minutes (65, 66). The subjects were instructed to cough intermittently during the induction procedure and for at least 2 hours afterwards to produce sputum samples. This procedure was found to be well tolerated by patients. The sputum samples were fixed in 50% ethanol and each sample was cytopun onto a glass slide and DNA was stained with Feulgen-thionin.

Some patients who volunteered for the chemoprevention studies did not meet the eligibility requirements for continuing on to participate in the bronchoscopy examination phase of those studies after a sputum sample was collected. A total of 1312 sputum samples in the present study were from such patients. They were followed through the Cancer Registry to determine if they developed lung cancer. Approval was granted by the Clinical Investigations Committees of the BCCA and The University of British Columbia. Written informed consent was obtained from all participants.

2.2.2 Semi-automated quantitative sputum analysis

An automated, high-resolution image cytometer (Cyto-Savant™ system from Oncometrics Inc., Vancouver, Canada) was programmed to attempt to measure the DNA content of at least 3000 objects per sample (186, 192, 239). For slides with fewer than 3000 objects, all objects were collected and the sample adequacy was determined on the basis of the criteria described in Section 2.3. The image cytometer was subjected to the daily, weekly, monthly, and yearly quality assurance standard operating procedures described in Chiu *et al* (240) and Guillaud

et al (241) to ensure that the system's components (i.e., device and sample staining) were operating within their expected performance parameters. Each object detected on the slide was individually focused and scanned. Each object was then subjected to a discriminating function, in the form of a classification tree, which separated bronchial epithelial cells from other materials such as food particles, macrophages, lymphocytes, and other inflammatory cells (242, 243). All cells were then reviewed by a trained cytotechnologist (certified by the Canadian Society of Laboratory Technologists). About 90% of all collected objects were identified to be epithelial cells after this procedure, which we have previously demonstrated yields comparable results to manually selecting nuclei (192).

For each epithelial cell, 110 nuclear features that will be used for ploidy and MAC analysis were calculated. These features can be divided into 6 categories: morphology (size and shape); densitometric properties (absorption amount and distribution); discrete texture features (euchromatin/heterochromatin); Markovian texture features (co-occurrence based); fractal texture features; and run-length texture features (239, 242). For each slide, the average, standard deviation, skewness, and kurtosis were calculated for each feature from all epithelial cells found on the slide.

2.2.3 Ploidy measures

We have previously shown that Feulgen-thionin staining with our system is quantitative for DNA (186, 192). Each cell's ploidy status was assessed by measuring the integrated optical density (IOD) of the nucleus and normalizing this against the mean IOD of the sample's diploid cell population, as determined from a frequency histogram of the nuclear IODs (241). Diploid cells were assigned a DNA index of 1.0. A ploidy score for each slide was calculated by examining the frequency of cells falling within a series of DNA index ranges and then finding the range which had the most discriminating performance between normal and abnormal cases,

where abnormal cases were defined as carcinomas *in situ* (CIS) and cancers. The ranges used were: <0.95, 0.95-1.00, 1.00-1.60, 1.60-1.85, 1.85-1.95, 1.95-2.09, 2.09-2.15, >2.15.

2.2.4 MAC measures

All MAC feature calculations were based on the feature set calculated in Section 2.2.2 and considered only cells with DNA indices between 0.7 and 1.3. A training set was constructed by randomly sampling 100 cells from each of the 36 normal, 6 CIS, and 36 cancer samples, as defined by the histopathological grading of their matching bronchial biopsies (see Section 2.2.5). These 78 samples were the same ones used to train the ploidy score in Section 2.2.3. The sampled cells from the CIS and cancer samples were then pooled together and compared against the sampled normal cells. A forward-stepping linear discriminant function analysis (180, 197) was performed on these two sets of about 4000 cells each, resulting in seven features selected to be indicative of malignancy associated changes.

The combined cytometric score for each slide was calculated from a linear combination of the 7 selected MAC features and the ploidy score. This combined cytometric score created a sputum-based biomarker that was used in the subsequent comparative analysis.

2.2.5 Biopsy collection and analysis

Atypia in a sample was defined as the presence of at least five cells which had DNA indices greater than 1.2 (65, 66). All volunteers with atypical sputum were recalled and invited to be examined using autofluorescence bronchoscopy; they had an average of 7-8 bronchial biopsies taken per visit. Each of the 7934 biopsies collected was fixed in buffered formalin, embedded in paraffin, and serially sectioned. H&E-stained sections from each biopsy were systematically reviewed by two experienced lung pathologists (J leRiche, A Gazdar), as previously described (65, 66). Each biopsy was classified into one of the categories (normal, basal cell hyperplasia,

metaplasia, mild/moderate/severe dysplasia, carcinoma *in situ*, cancer) in the histopathological system established by the World Health Organization (244). Minor (i.e., one grade) differences in sample classification were resolved by telephone consultation between the two pathologists. If the diagnosis differed by two or more grades, both pathologists reviewed the slides again and reached a consensus diagnosis after communication by phone, email, or in person. The biopsies were matched by patient and date of collection to 1233 distinct sputum samples. For each of these sputum samples, the most severe consensus biopsy diagnosis associated with that sputum sample was recorded. In addition, all samples taken from patients who were subsequently diagnosed with CIS or cancer by non-bronchoscopic means (e.g., CT scans) within 8 months after sputum collection were also classified as CIS or cancer, as appropriate. All subjects who received, for any biopsy, a biopsy grade of dysplasia, or worse, then had a Morphometry Index (MI) calculated for all their biopsies, according to the procedure set out in (197). A total of 5060 biopsies had MIs calculated for this study.

178 biopsy samples taken from sites with a biopsy grading of at least dysplasia at baseline or follow-up were subjected to immunohistochemical analysis using 4 markers: p53, Ki-67, bcl2, and cleaved caspase 3, as previously described (65). They were graded visually on a 0 – 4 scale (with 0 indicating no stain and 4 indicating more than 75% of the nuclei staining positive) by experienced cytotechnologists. Of these biopsies, 159 corresponded to one of the sputum samples within the data set for the present analysis, matching both patient and time.

Some of the volunteers in this study either developed resectable lung cancer during the trial process or were discovered to have cancer upon enrolment. From 40 of these subjects who developed lung cancer, 73 sputum samples were collected either before or after surgical treatment. These 40 cases included patients with squamous cell carcinomas, adenocarcinomas, large cell lung cancers, and small cell lung cancers.

Except where otherwise noted, statistical significance in the present analysis was assessed using unpaired t-tests and ANOVA performed using STATISTICA software (StatSoft Inc., Tulsa, OK). $P < 0.05$ was considered statistically significant.

2.3 Results

The average age of the 1795 volunteers when samples were taken was 59.7 years (ranging from 39 to 83), and the average pack-years smoked was 48 (ranging from 8 to 221 amongst all current and former smokers). 57% of the samples were taken from male participants, 43% from females. 60.2% of the samples came from former smokers, 38.5% from current smokers, and 0.3% from non-smokers. The age distributions were similar between sexes: male average was 60 (range 39-83) and female average was 59 (range 39-81). However, there was some difference in their smoking history, with the male average pack-year exposure being 50, and the female average pack-year exposure being 44.

Upon comparing the ploidy characteristics of the normal and CIS/cancer training sets, the most discriminating ploidy feature was found to be the frequency of epithelial cells with a ploidy amount between 1.6 and 1.85, denoted here as v_4 . Hence, this was used in the ploidy score that will be combined with malignancy associated changes features and then compared with bronchial biopsy histology: $Ploidyscore = 324 \cdot v_4 - 0.3548$. Seven features were found to be most indicative of malignancy associated changes: 1) the standard deviation (SD) of a nuclear morphology feature, `harmon05_fft` (a measure of nuclear roundness) (242), across all the epithelial cells measured for the sample, 2) the SD of 3 nuclear discrete texture features, `high_DNA_area`, `medium_DNA_amount` and `medium_average_distance` (239), across all the epithelial cells measured for the sample, 3) the SD of a Markovian texture feature, `correlation`

(239), across all the epithelial cells measured for the sample, 4) the mean of a fractal texture feature, *fractal_dimension* (239), across all the epithelial cells measured for the sample and 5) the SD of a run-length texture feature, *maximum_run_length* (adapted from (239)), across all the epithelial cells measured for the sample. These 7 MAC-based features were used to generate a MAC score:

$$\begin{aligned} MACscore = & 11.6186578 \cdot SD_{harmon05_fft} + 102.10314446 \cdot SD_{medium_average_distance} - \\ & 0.22874972 \cdot SD_{maximum_run_length} + 35.60912366 \cdot Mean_{fractal_dimension} - \\ & 36.69652294 \cdot SD_{high_DNA_area} + 0.17210604 \cdot SD_{correlation} - 30.68769782 \cdot \\ & SD_{medium_DNA_amount} - 90.8891872. \end{aligned}$$

Finally, the MAC score was combined with the ploidy score as a weighted sum to create a Raw Combined Score (Raw CS): $Raw\ CS = 1.008354 \cdot MACscore + 0.404192 \cdot Ploidyscore + 0.509748$.

A plot of the Raw Combined Scores showed a dependency on the number of identifiable cells on each slide (Figure 2.1A). To correct for this, we subtracted from each Raw CS the value predicted by the distance-weighted least squares fit as a function of the number of cells measured on the slide. There were insufficient samples with more than 6000 cells to reliably estimate the trend, so the adjustment for samples with more than 6000 cells was set to zero. Except where otherwise specified, the adjusted Combined Score will be denoted simply as the Combined Score or CS for the remainder of the present analysis.

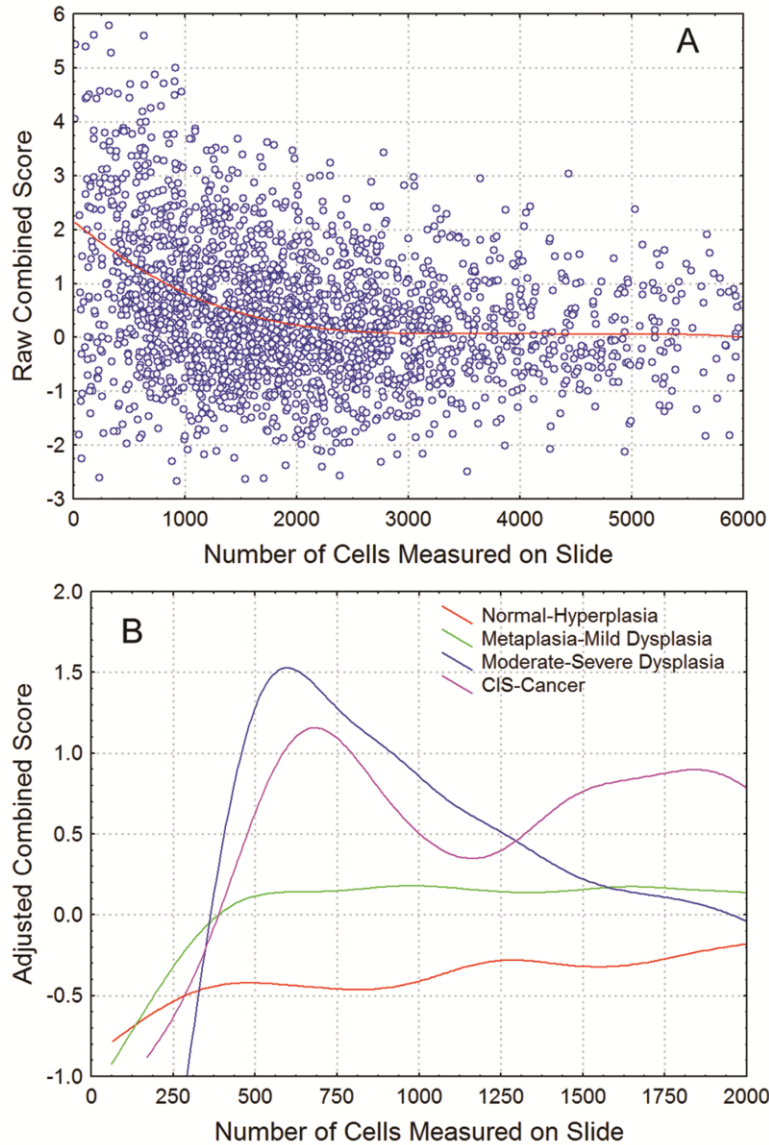


Figure 2.1: Dependence of the Combined Score on the number of identifiable cells. Trend lines are distance-weighted least squares fits. **A:** As the number of identifiable cells increases, the scatter decreases and the Raw Combined Score becomes a more consistent measurement. The distribution of histopathological grades, meanwhile, is quite consistent across the range of identifiable cell counts. However, a distinct trend towards higher Raw Combined Scores at lower cell counts necessitated a cell count normalization procedure. **B:** Adjusted Combined Scores, with data categorized according to the highest grade of abnormality found in that patient's biopsies. Data points have been removed to highlight the trends. Below a count of 500 cells per slide, the consistent patterns that the various histological categories exhibit break down, as seen in the rapid changes and convergence of the four running average curve lines.

As with any cytological test, we must set a sample adequacy threshold that minimizes the scatter from measuring too few cells without excluding so many samples that it causes undue

stress on patients and reduces the test's overall utility in a clinical setting. We chose 500 cells per slide as a threshold because below this level, the somewhat consistent patterns that the various histological categories exhibit break down (Figure 2.1B). Meanwhile, approximately 10% of the sample slides are excluded at this level, which was felt to be an acceptable rate. Hence, only sputum samples with at least 500 identifiable cells were used in the subsequent analysis.

A comparison of the sputum-derived Combined Score (CS) with the maximum histopathological grade of all the bronchial biopsies of the test subject at the corresponding time point is shown in Figure 2.2. There is a clear trend that as pathological severity increases, so does the CS (F-test, $P < 10^{-5}$). Post hoc analysis using the Tukey Unequal N Honestly Significant Difference (HSD) test showed that the CS of the normal and hyperplasia groups were statistically significant from those of the cancers ($P = 0.003$ and 0.009 , respectively).

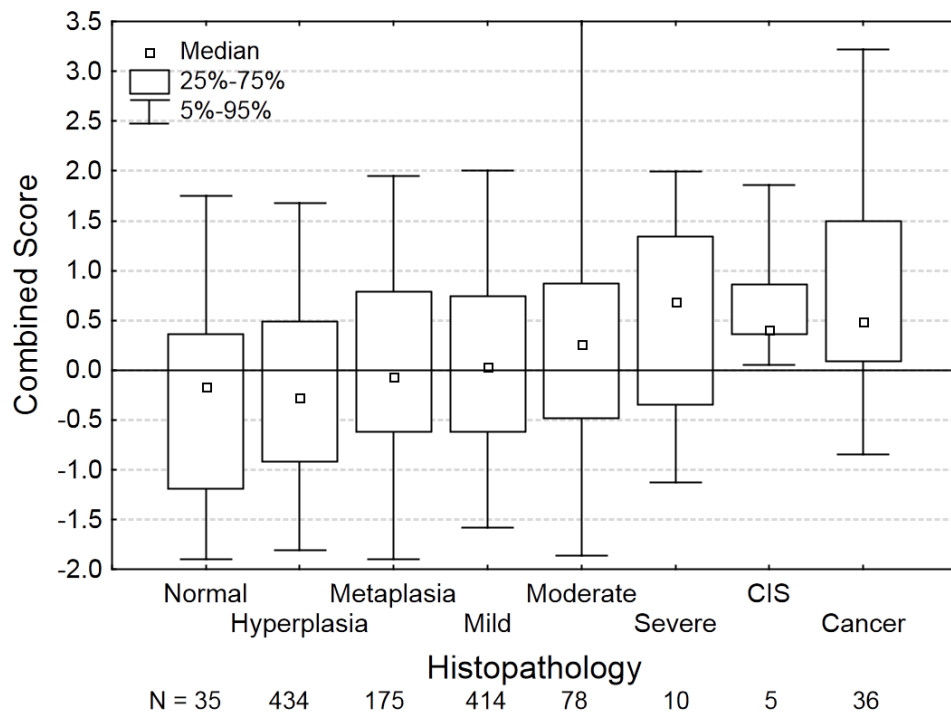


Figure 2.2: Box plots of Combined Score for sputum samples containing more than 500 identifiable cells, grouped according to the highest histopathological grade in that patient. There is a general increase in the median Combined Score in samples taken from patients harbouring more pathologically severe lesions.

As similar histopathological groups can often be difficult to distinguish, we created four new groups: normal/hyperplasia, metaplasia/mild dysplasia, moderate dysplasia to CIS, invasive cancer. When these groups are used, the trend between CS and pathological severity becomes even more evident, as shown in Figure 2.3A. Post hoc analysis shows that the normal/hyperplasia group is significantly different from all other groups and the cancers are significantly different from the metaplasia/mild group (summarized in Figure 2.3B).

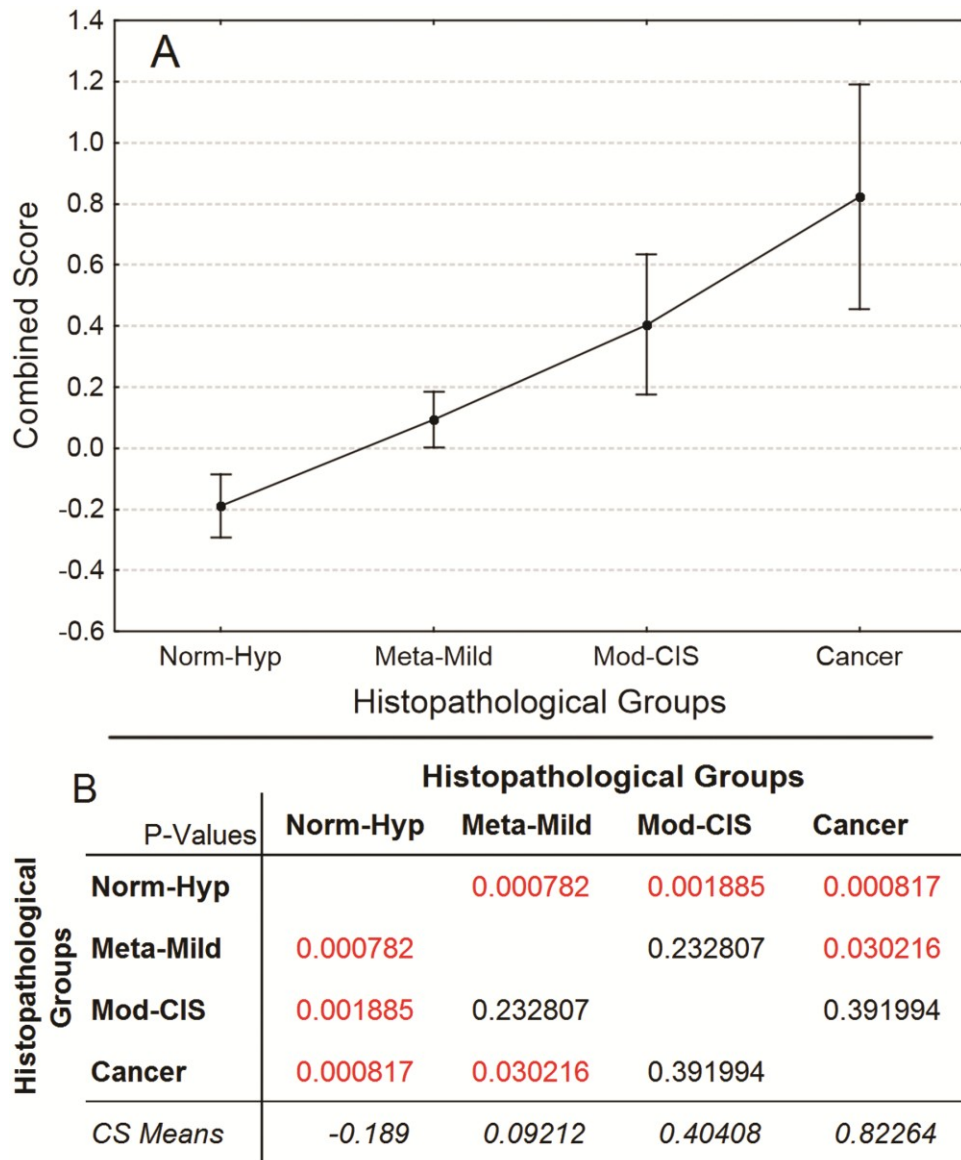


Figure 2.3: Analysis of Combined Scores for samples from patients in each of the four histopathological groups created by combining similar grades: normal/hyperplasia, metaplasia/mild dysplasia, moderate dysplasia to CIS, invasive cancer. **A:** Plot of mean CS. Error bars denote 95% confidence intervals. **B:** Summary P-value matrix of Tukey Unequal N HSD Post Hoc analysis. Significant P-values are highlighted in red and the italicized row shows mean Combined Scores in each group.

The ideal criteria for assessing a novel lung cancer risk biomarker would be reductions in mortality and/or progression to invasive cancer. In the absence of data on whether or not our study subjects progressed, we attempted to estimate the degree to which the Combined Score can be used to ascertain lung cancer risk by comparing the CS to other known risk factors and

biomarkers of lung cancer. Using the Morphometry Index in conjunction with histopathological grading, we created high- and low-risk subject groups, which we will denote m-risk. A given patient was considered low-m-risk if he or she had a histopathological grading of hyperplasia, or less, and a maximum MI < 1.36 , as described in (66). High-m-risk subjects had a maximum MI > 1.36 and a histopathological grading of moderate dysplasia, or worse. Additionally, all CIS and cancer patients were denoted high-m-risk, regardless of MI. Given the strong correlation between the Morphological Index and cancer risk (197), we feel that this combination of dysplasia grade and MI, i.e., the m-risk, represents a more accurate approximation of lung cancer risk than does a system that relies on dysplasia grade alone. If all the subjects that fit into the high- or low-m-risk categories are grouped together, there is a significant correlation between the Combined Score and the m-risk groups ($P = 0.00004$) (Figure 2.4). Removing the samples used for training from this analysis, this correlation between CS and m-risk groups still holds ($P = 0.008$).

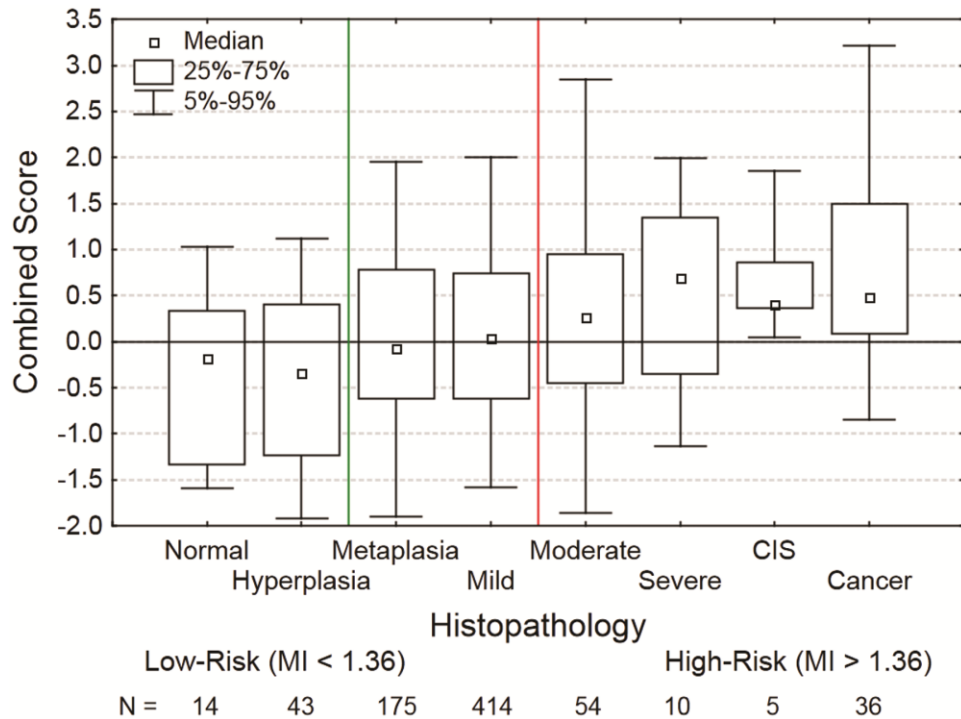


Figure 2.4: Box plots of Combined Score for sputum samples, sorted according to m-risk. Normal and hyperplasia groups are low-m-risk and only data for which the low-m-risk MI criterion is also met (max MI < 1.36) is shown here. Moderate and severe dysplasia are considered high-m-risk and the data shown here only includes cases where the high-m-risk MI criterion is also met (max MI > 1.36). Additionally, all CIS and cancer cases were counted as high-m-risk, regardless of MI. Metaplasia and mild dysplasia are neutral m-risk and all data in these groups is shown. The numbers at the bottom indicate the number of sputum samples in each group. The Combined Scores for the low- and high-m-risk groups are significantly different ($P = 0.00004$).

Using the m-risk groups, we can construct a receiver operating characteristic (ROC) curve representing the ability of the Combined Score to distinguish high-m-risk patients from low-m-risk patients (Figure 2.5). Patients with metaplasia and mild dysplasia are considered neutral m-risk. As we are unsure whether to consider them high- or low-m-risk, we excluded them from this analysis. In this manner, we are assessing the performance of CS on only the patients for whom we are most certain of their risk of progression to invasive lung cancer. Since many patients received multiple biopsies over the course of the study, MI can be used to determine m-risk either by using the maximum MI at a given time point or the average MI at that time point. In either case, the worst histopathological diagnosis was used to determine the m-risk

group. CS performed well by both definitions of m-risk, although using the average MI resulted in a noticeably better area under the curve than using the maximum MI (AUC by trapezoidal rule, 0.766 and 0.711, respectively). If all samples used in training the CS are removed from this analysis, the areas under the curve become 0.752 and 0.677, respectively.

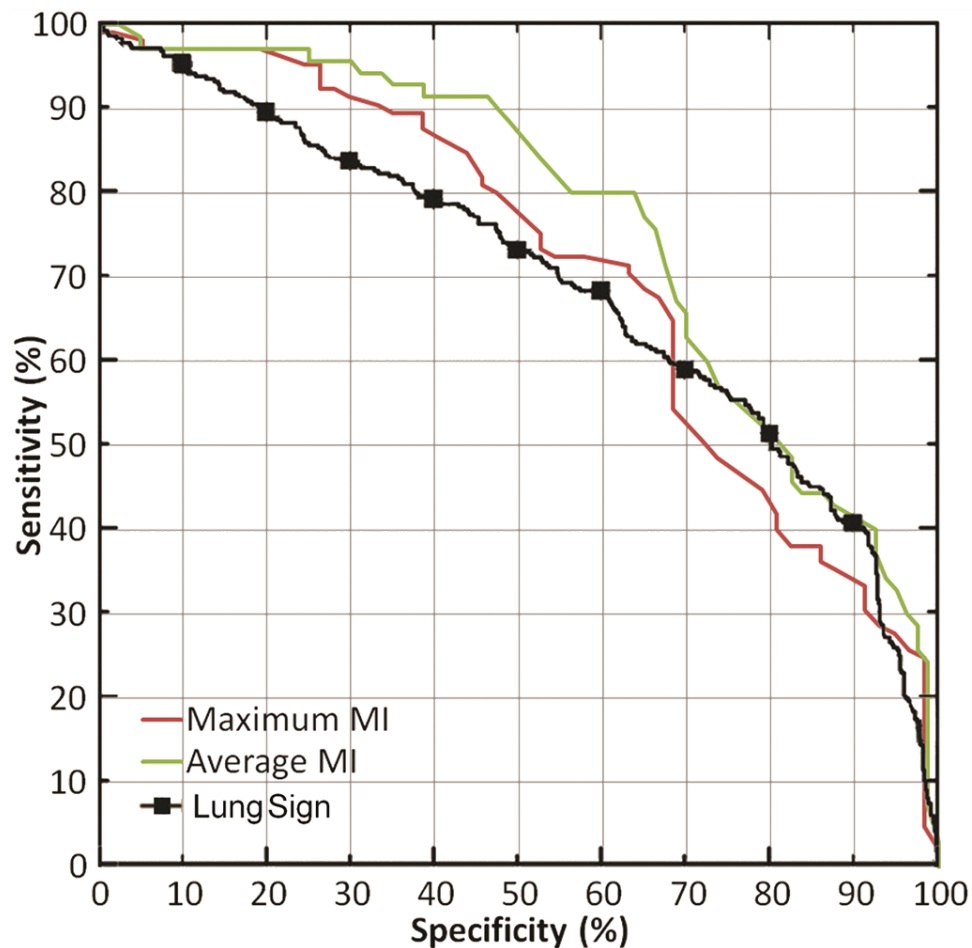


Figure 2.5: ROC curves showing the ability of the Combined Score to distinguish between high- and low-m-risk patients. Patient m-risk groups were defined as described in the text. Patient m-risk can be defined using the maximum MI of the biopsies from the subject or the average MI from the biopsies taken from the subject, both cases for which are shown. For comparison, the ROC curve for LungSign, a test developed to detect more advanced neoplastic lesions, i.e., cancer, is shown as well (adapted from Figure 2 in (191)). Areas under the curve are 0.711, 0.766, and 0.692 for maximum MI, average MI, and LungSign, respectively.

Age and pack-years smoked are the most widely studied epidemiological lung cancer risk factors. Both have been shown to be key predictors of lung cancer risk (245). Plotting each

sample's Combined Score against age and pack-years smoked shows a positive correlation in each case (data not shown). A linear regression with a statistically significant positive slope can be calculated in each case which confirms that CS increases with age and smoking history. Furthermore, these trends remain even if the subject population is broken down into male/female and current smoker/former smoker groups. The subject population was divided into these groups because these subpopulations may exhibit differences in cancer risk or progression (25, 246).

There has been considerable work to find immunohistochemical markers – in blood, sputum, or biopsies – that correlate with lung cancer risk (230, 247). Two of the more successful tissue-based lung cancer immunohistochemical markers are p53 and Ki-67. p53, which is the most studied marker for all cancers including lung cancer (230), has been shown to be overexpressed in many premalignant bronchial lesions (248). Furthermore, overexpression of p53 in a lesion correlates with an increased risk of a lesion progressing to invasive cancer (248). Immunostaining with the proliferation marker Ki-67, which is expressed in the G₁, S, G₂, and M phases of the cell cycle (249), has been shown to be of prognostic value in a number of cancers, including lung cancer (250, 251). Ki-67 expression has further been demonstrated to increase as preneoplastic lung lesions progress from mild dysplasia to CIS (252).

Figure 2.6A and Figure 2.6B plot the immunohistochemical staining score of p53 and Ki-67 respectively, for each of the 159 biopsies selected (see Materials and methods) against its histopathological grading. Overall, there was a positive correlation between p53 staining and histological grade. There was also a statistically significant difference between hyperplasia and all of the more severe grades (Figure 2.6A). The proliferation marker Ki-67 shows an even more pronounced progression of increasing staining with increasing severity of pathological grading (Figure 2.6B). Immunohistochemical staining with bcl2 and cleaved caspase 3 were also performed, but no correlation was observed between these stains and histopathological grade.

This appeared to be partly due to poor staining quality. Hence, bcl2 and cleaved caspase 3 immunostaining were considered poor markers for lung cancer risk and no further analysis with these markers was pursued.

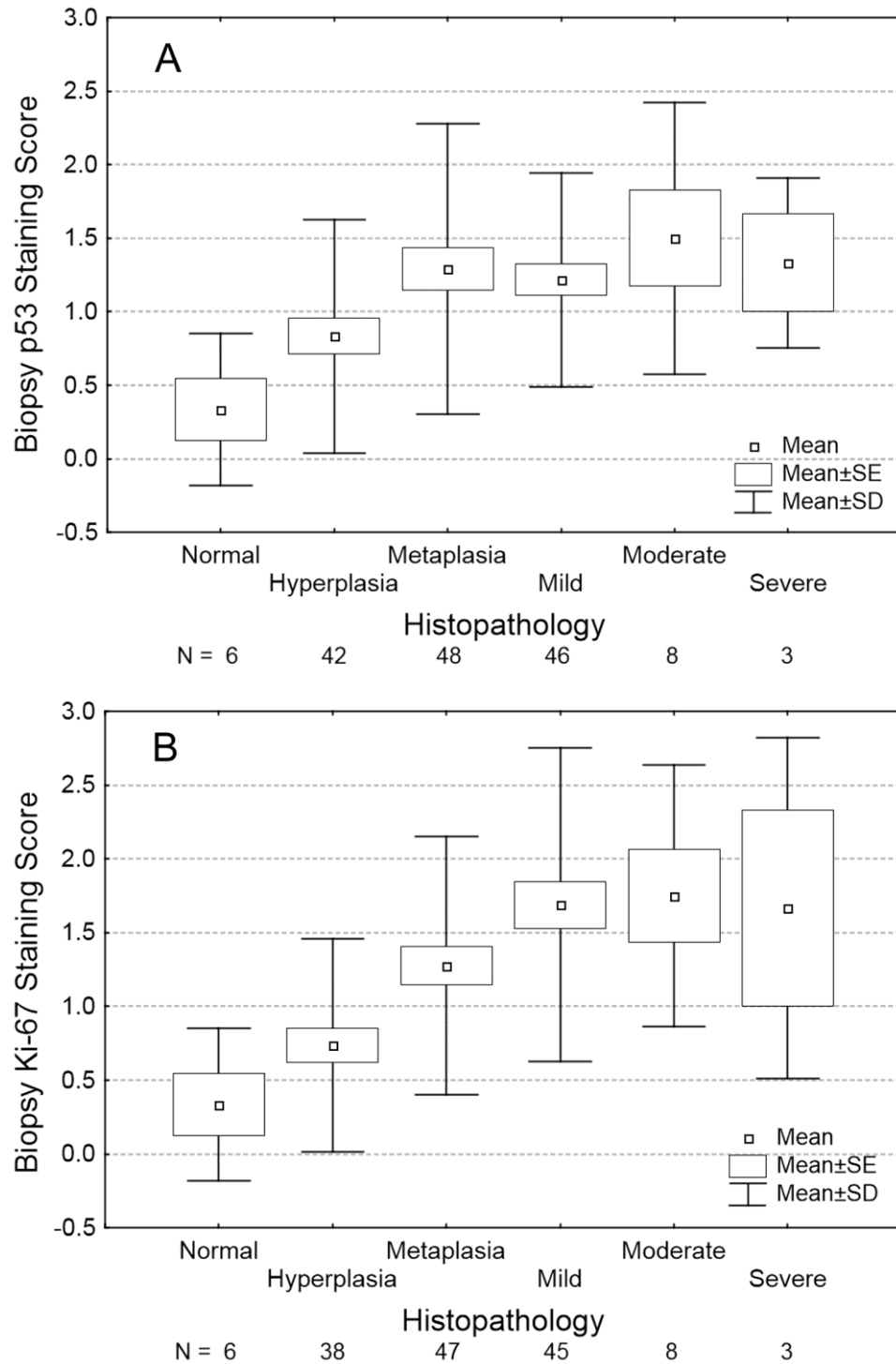


Figure 2.6: Box plots comparing p53 (A) and Ki-67 (B) staining score to histopathological grade for the 159 biopsy samples from sites that had a biopsy grading of dysplasia or worse at baseline or follow-up.

To compare p53 and Ki-67 immunohistochemical staining to the sputum-based Combined Score, the sample populations were separated into current and former smoker subgroups. To more clearly illustrate the trends in the data, the five immunohistochemical staining scores were combined into two groups: no/weak staining (scores 0-1) and stronger staining (scores 2-4). The CS correlates with the p53 immunohistochemical staining for both former smokers and current smokers (Figure 2.7A and Figure 2.7B, respectively). Ki-67 staining correlated with increasing Combined Score values for the former smoker subgroup (Figure 2.7C). However, there was no discernable pattern in the plot comparing Ki-67 immunostaining to the Combined Score values for current smokers (Figure 2.7D). This is mainly because the CS for the current smoker cases with weaker Ki-67 staining is as high as the scores for all cases with stronger Ki-67 staining regardless of smoking status. For the comparisons shown in each of the four panels (Figure 2.7A-D), $P = 0.08, 0.008, 0.1, 0.6$, respectively, although we can likely attribute the findings of insignificance in the former smoker comparisons to an insufficient number of cases.

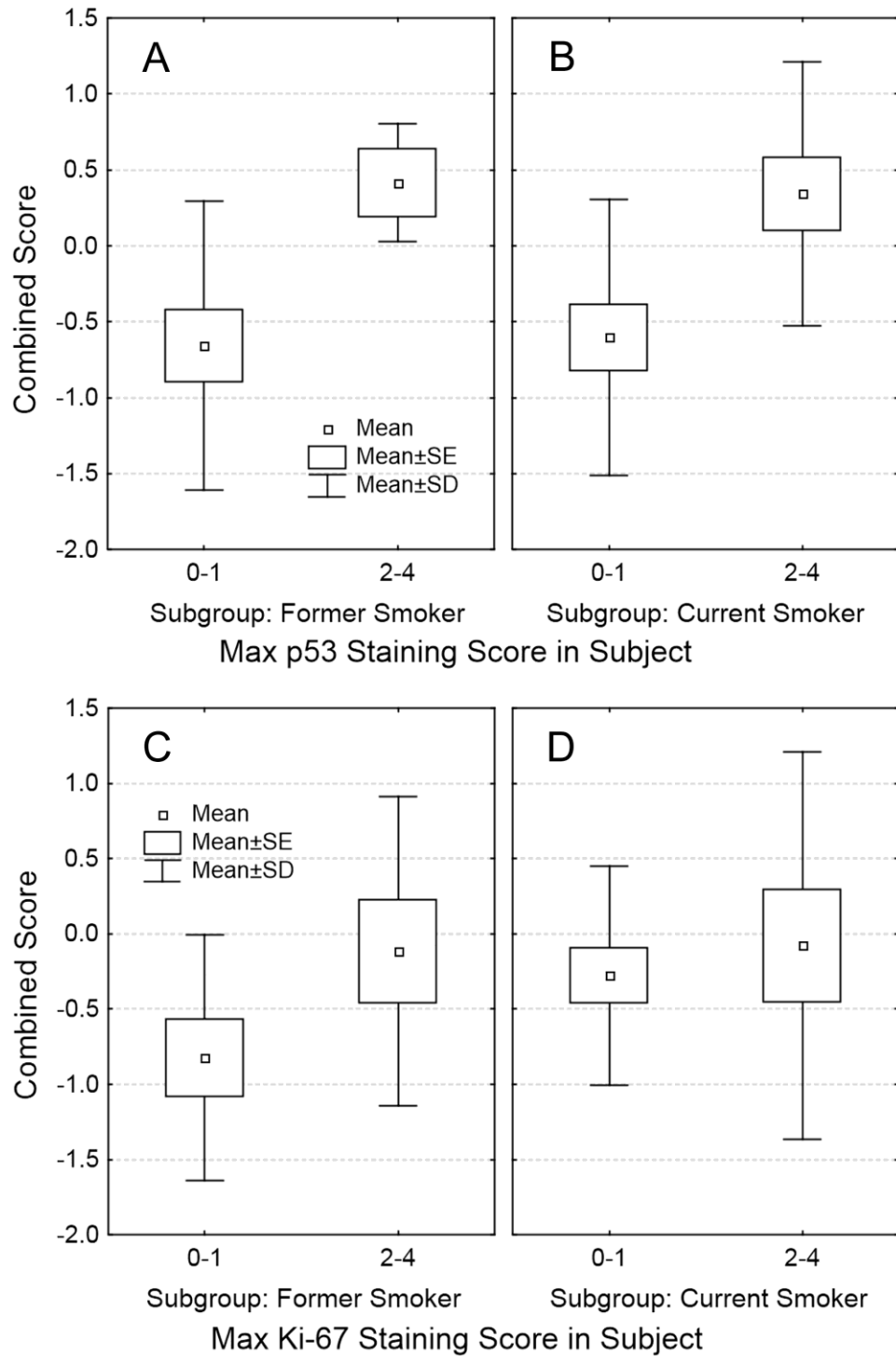


Figure 2.7: Box plots comparing Combined Score to the maximum p53 (A, B) and Ki-67 (C, D) staining score in that patient at that point in time. Immunostaining scores were grouped into two categories: no/weak staining (scores 0-1) and stronger staining (scores 2-4). Cases were further subdivided into former smoker (A, C) and current smoker (B, D) groups. For the comparisons shown in each of the four panels (A-D), $P = 0.08, 0.008, 0.1, 0.6$, respectively.

Given the impact of smoking status on the interaction between the Ki-67 measurements and Combined Score, we turned our attention to other risk factors which might have a confounding effect on our analysis of CS. We found that there were small but not significant differences between the high- and low-m-risk groups in terms of age, smoking history, and sex (t-tests, sex by Pearson χ^2 , $P = 0.1, 0.2, 0.3$, respectively), but high-m-risk patients were significantly more likely to be current smokers (Pearson χ^2 , $P = 0.00009$). Analyzing current and former smokers separately, we found that in both cases, high-m-risk patients had significantly higher CS than low-m-risk patients ($P = 0.01, 0.002$, respectively). Furthermore, Figure 2.2 replotted with current and former smokers separately shows the same general trends in each subgroup as the original figure, demonstrating that smoking status does not have a confounding effect on our analysis of CS overall.

The most important feature of any surrogate biomarker is its correlation with cancer risk or progression. While our participant criteria were not designed to find lung cancer patients, a number of study participants developed lung cancer over the course of the study. Additionally, some patients recruited on account of receiving a bronchoscopy for other clinical indications were found to have lung cancer upon enrolment. We compared sputum samples taken within eight months before surgery with sputum samples collected at least six months after the surgical resection treatment protocol. Samples taken after surgery had significantly lower ($P = 0.003$) Combined Scores than the samples taken before surgery. The t-test was unpaired because not enough patients had data both before and after surgery for a pair-wise test to be statistically meaningful.

Among the sputum samples linked to a positive cancer diagnosis, there was no significant difference in CS between distal and proximal tumours ($P = 0.9$). When broken down by cancer subtype, there was no significant difference between adenocarcinomas and squamous cell

carcinomas ($P = 0.1$). There were insufficient samples of small cell and other non-small cell lung cancers to make any other statistically meaningful comparisons.

2.4 Discussion

It has been suggested that the traditional view that cancer begins when invasive disease is first detected should be replaced by one in which carcinogenesis itself is the disease, with invasive or symptomatic cancer being merely the final outcome (49). Consequently, treatments should aim to “reverse, suppress, or prevent the process of carcinogenesis” (247). This is the goal of chemoprevention, with past work in our group and others showing promise (65-68).

However, many early chemoprevention studies for lung cancer have actually shown neutral or even negative effect from chemopreventative agents. Disappointing results from these early studies may be due to the fact that many of these studies used smoking status as the primary selection criterion, resulting in a study population with an insufficiently high risk to benefit from chemoprevention (64). Many precancerous lesions never progress even without treatment (59) and so chemoprevention will offer these patients no additional benefit. If we can remove these patients from a study population and only study those who are likely to progress to invasive disease without treatment, any effect from chemoprevention should become more evident.

End points for chemoprevention studies are typically the incidence of invasive cancer or mortality (64). Since many pre-cancers never develop into invasive disease, regardless of whether chemopreventative agents are used, this makes such trials long and costly. Governments and pharmaceutical companies may be reluctant to invest in the development of cancer chemoprevention drugs and strategies due to the immense research cost, especially while lung cancer chemoprevention is still not a universally accepted approach to the management of the

disease. In order to accelerate the development and verification of new chemopreventative agents, intermediate end points need to be identified and validated.

Computer analysis of sputum samples has previously been used to detect lung cancer, using criteria based on ploidy (161), MACs (180), or both (191). While CS sought to detect pre-cancers, the LungSign test combined ploidy and MAC analysis and was effective in detecting 40% of all lung cancers with 91% specificity consistently across all subtypes and stages, far better than the results from conventional cytology (191). MAC analysis is appealing because it can be measured on non-malignant cells, which typically greatly outnumber malignant ones in sputum samples. The features we used describe various aspects of the nuclear architecture. Changes in the chromatin distribution and organization may be indicative of changes in activation and expression of genes. Genetic and epigenetic alterations, which may be related to cell cycle, metabolic, or differentiation status of the cell, are reflected in these MAC features (239). Using a similar approach to LungSign, we have devised a novel biomarker combining ploidy and MAC analysis. Unlike LungSign, which was optimized for the detection of invasive cancers with high specificity, the Combined Score presented in this paper is designed to detect dysplasias. By detecting pre-cancerous lesions before they become invasive cancers, the CS could allow the highest-risk patients to be enrolled in chemopreventative therapy trials in an effort to reduce their risk of progression to cancer. Since the Combined Score correlates with dysplasia grade, the effectiveness of any such intervention can also be safely and easily monitored over time.

As a biomarker for lung cancer risk, our analysis shows that the Combined Score correlates with a number of other known lung cancer risk factors. CS is better able to distinguish patients with moderate dysplasia or worse from those with normal histology or hyperplasia than either age or smoking history alone (trapezoidal rule areas under the ROC curves of 0.661, 0.569, and 0.537 for CS, age, and smoking history, respectively). When compared to histopathological

grade as shown in Figure 2.2, there is a clear trend towards higher Combined Scores with increasing disease severity. This trend is apparent even if we were to remove all the normal, CIS, and cancer cases, the sample sets from which the training set was derived. Since Morphometry Indices for biopsies from patients who progressed to cancer were significantly higher than non-progressing lesions of the same histopathological grade (197), MI can supplement histopathology. By combining histopathology and MI, we can get a better assessment of cancer risk (which we denoted m-risk). Adding the MI to our analysis, we found that CS correlated even better with m-risk (Figure 2.4) than with cancer risk defined by dysplasia grade alone. In the context of chemoprevention trials, then, we would aim to enrol patients with high m-risk. CS is a non-invasive test that could potentially identify subjects harbouring high-grade dysplasia and cancer without biopsying everyone.

The fact that the correlation between CS and histopathological grading is not as strong as that between MI and pathological grades in previous studies (197) reflects the difference between the subtle malignancy associated changes that occur lung-wide and the more pronounced changes found in the diagnostic cells of biopsies. However, collecting biopsies (upon which both histopathological grading and MI are based) is still an invasive technique. The correlation between CS and the combination of histopathology and MI suggest that CS could be used as a rapid, non-invasive, and relatively inexpensive alternative to these techniques for both risk assessment and the conduction of chemoprevention studies.

As we lacked data on actual cancer progression, we used MI and histopathology as a gold standard to assess the performance of CS in identifying those patients at highest risk to progress to invasive cancer. As a previous study found that patients who progressed to cancer had significantly higher MIs than patients with non-progressing lesions (197), we believe the combination of MI and histopathology embodied by the m-risk provides a better estimate of lung

cancer progression risk than histopathology alone. Figure 2.5 shows ROC curves using either maximum or average MI to define m-risk. Both these methods of determining m-risk generate noticeably different ROC curves. Clinically, a physician may be interested in determining the risk of progression of the most severe lesion and so a risk assessment using the maximum MI is most appropriate. However, a sputum biomarker is based on a sampling of cells from throughout the lungs. As expected, then, when the average MI is used as the criterion for determining m-risk, the ROC curve for CS looks improved over the maximum MI case. Nonetheless, even when using the maximum MI as the m-risk criterion, the ROC curve for CS compares very well with that for LungSign. This is despite the fact that LungSign seeks to distinguish between cancerous (CIS or worse) and non-cancerous samples, whereas the Combined Score is able to separate high-grade dysplasias from normals, an arguably much more challenging task. While our samples were not routinely screened with conventional cytology, a subset of our samples overlaps with those used in the LungSign study, where they reported a sensitivity of 16% and a specificity of 99.1% for detecting lung cancer with cytology. The CS showed a similar level of sensitivity to high-grade dysplasias at that level of specificity.

The ideal analysis for any novel biomarker would be to see which patients ultimately develop cancer. This requires extensive follow-up and even then, only a small number ever progress. In the absence of data on actual cancer progression, the next best alternative is to ensure that the novel biomarker correlates with known biomarkers. We found that Combined Score correlated with age, smoking history, and immunohistochemical staining of p53 and Ki-67, all of which have been previously found to correlate with lung cancer risk (245, 247, 248, 252). Except for Ki-67 staining in current smokers, these trends were further found to be applicable to both current and former smokers.

In former smokers, the Combined Score shows a particularly strong correlation with p53 staining. This may be due in part to “field cancerization,” a concept first proposed by Slaughter *et al* (57) to explain the propensity of individuals with one malignancy to develop second primary tumours. Mutations and changes in expression level of the p53 gene across wide areas of the lung have previously been reported in subjects with dysplasia or preinvasive lesions (248). Since CS is based on a sampling of cells from throughout the lungs, we might expect a better correlation with an immunohistochemical marker whose expression has likewise been altered over a large region of the pulmonary mucosa.

Ongoing exposure to cigarette smoke causes inflammation in the lungs and has been shown to be associated with an increased expression of not only Ki-67 (253) but also proliferating cell nuclear antigen (254), another important proliferation marker. This confounds our analysis and may help explain why we do not observe a trend between CS and Ki-67 staining in current smokers, as any correlation between CS and Ki-67 may be dwarfed by the impact of smoking on proliferation across the lung. Further, smoking is known to alter the expression not only of a large number of genes (49) but the chromatin structure as well (255) and these changes are different in current and former smokers. Our Ki-67 staining results, when compared to the Combined Score, illustrate one more example of the difference between the lungs of current and former smokers, underscoring the necessity of taking smoking status into consideration for any proliferation-based diagnosis or treatment.

To address the issue of potential confounding effects in our analyses of CS, we compared the age, smoking history, sex, and smoking status of patients in our two m-risk groups. As these are all documented to affect lung cancer risk, we expected to see some differences between the groups. Except in the case of smoking status, the differences we observed were too small to be considered potential confounders. A follow-up analysis showed that among both current and

former smokers, the general trends we observed in Figure 2.2 still hold and high-m-risk patients have higher CS, so smoking status does not impart any additional confounding effect on our analyses. However, as current smokers generally had higher CS than former smokers ($P = 0.00001$), different thresholds may need to be set if CS were to be used in a clinical setting.

Since the Combined Score is presented on a continuous numeric scale, it allows smaller changes in lung health to be detected. The use of automated image analysis also means that it should be more objective than standard histopathology. We've shown that the Combined Score correlates with m-risk, which combines the dysplasia grade and the Morphometry Index. We believe that CS can be used to monitor chemoprevention trials. Unlike the MI, however, the CS is a sputum-based biomarker, which is less invasive and more likely to be tolerated by patients. This further allows CS to be measured repeatedly over the course of the trial.

In chemoprevention trials, these advantages mean that trials can be designed to use a reduction in CS by a certain threshold amount as an alternative end point, instead of waiting for invasive disease to develop. We can consider our analysis of surgically resected lung cancer cases to be an example of this, as we can think of surgery and chemoprevention as two different interventions and CS as a common scale by which to assess their effectiveness. In cases where lung cancer has been treated by surgery, the Combined Scores before and after surgery are significantly different. We are further encouraged by the observation that our sample contains a mix of squamous cell carcinomas and adenocarcinomas. With the small sample size, however, it is difficult to properly assess the ability of CS to detect successful surgery. There is also insufficient data to assess whether CS performs better with squamous cell carcinomas or adenocarcinomas. The present study was not designed to test CS in this setting, but the initial results suggest this is another potential application of CS that merits further study. Similar to how CS can be tuned for optimal detection of precancerous lesions, the continuous scale of the

Combined Score allows us in the future to select a good threshold for detecting successful surgery once data from a larger and more comprehensive study is available.

The observation that post-surgery Combined Scores are lower than scores from before surgery suggests that CS is sensitive to MACs, which was the intention in training the CS. While the correlative evidence is weak, an advantage of using MACs as a pre-screening test is the extra sensitivity inherent in being able to detect malignancy even when a sputum sample consists primarily of non-malignant cells, as is often the case. In addition to being able to detect MACs, the Combined Score appears to be able to detect the effects of field cancerization. The correlation of CS with p53 staining is suggestive of this, as is the observation that CS is better able to assess m-risk when m-risk is calculated on the basis of average MI as opposed to maximum MI (Figure 2.5). The magnitude of a MAC effect might be expected to correlate with the most severe lesion present releasing soluble factors to which the surrounding cells respond, but the CS appears to correlate more with the severity of the overall “cancer field” as reflected in the average MI. This is of benefit to the design of future chemoprevention studies as it would be informative to be able to monitor the overall level of field cancerization in response to a candidate chemopreventive therapy. Our data weakly suggests that the CS can act as a surrogate biomarker in this regard.

We view MACs and field cancerization as separate but possibly related phenomena. While the prevailing field cancerization hypothesis suggests that cancers arise from a field of altered cells, previous work with pre- and post-surgery patients suggests that cancer cells themselves influence histologically normal cells (177). As these effects can be reversed by removing the tumour, it has been hypothesized that such effects may be a response in histologically normal cells to autocrine signals released by malignant cells (177). Although our data cannot provide insight into the mechanisms underlying the morphological changes detected

by the Combined Score, our results weakly suggest that CS may correlate with both MACs and field cancerization.

Despite decades of work, there remains no widely accepted screening test for early lung cancer detection. Studies using spiral CT, for example, showed high sensitivity for detecting non-calcified pulmonary nodules, but had a low specificity, which, coupled with a low overall prevalence of lung cancer even amongst heavy smokers, led to a low positive predictive value (25) and consequently increased costs due to follow-up testing and unnecessary surgical interventions. To address these shortcomings, it has been suggested that automated sputum cytometry could be used as an initial screening test, thereby increasing the disease prevalence amongst those subsequently screened by CT and autofluorescence bronchoscopy (75). While our intent was not to design a novel pre-screening tool, our analysis of the Combined Score as a pre-screen for patients most likely to benefit from chemoprevention suggests that CS could potentially be used to pre-screen for patients most likely to benefit from secondary lung cancer screening with CT and autofluorescence bronchoscopy. Our study population was at high risk of developing lung cancer on the basis of demographic risk factors (i.e., age and smoking history). We envision any potential use of CS in a pre-screening setting would also focus on such a subset of patients as these are patients most likely to benefit from additional screening. Moreover, patients at high risk due to age and smoking history are readily identified by the use of a patient questionnaire.

Like the LungSign test, the Combined Score is a sputum biomarker that has an adjustable classification threshold. This allows the performance to be optimized to best complement other early lung cancer detection methods (191). In such a pre-screening scenario where positive pre-screening tests would be followed up with more (perhaps more costly and/or invasive) screening, we would like a test with a high sensitivity, while tolerating a lower specificity. The performance

scores of the Combined Score for detecting pre-cancers match very well those of the LungSign test for detecting cancers. At a specificity of 50%, for example, we can achieve 78% sensitivity for high-grade dysplasias, which is slightly better than LungSign's ability to detect cancers at that level of specificity. This means we could reduce the number of CTs by half and still catch roughly three-quarters of all high-m-risk pre-cancers. This would have significant cost savings and mean less risk of increased cancer incidence caused by radiation exposure due to unnecessary CT scans (256). However, one must keep in mind that our analysis of the ability of CS to identify high-m-risk lesions excludes metaplasias and mild dysplasias, which may result in better perceived performance. More study will be needed to validate the use of CS as a pre-screener in conjunction with more invasive screening tools in a clinical setting.

2.5 Conclusion

Attempts to develop effective screening tools for lung cancer have faced many challenges. Just as importantly, where patients have been found to harbour precancerous lesions, there remain no widely accepted interventions as research into chemoprevention is currently hampered by a lack of effective surrogate biomarkers to serve as end points for trials. We have presented evidence that the Combined Score, a novel automated sputum image cytometry biomarker based on ploidy and MAC analysis, correlates with other known lung cancer risk factors like histopathology, age, smoking status, and immunohistochemistry of p53 and Ki-67. Compared to LungSign, a similar sputum biomarker, Combined Score achieves a similar performance separating high- and low-m-risk pre-cancers to that of LungSign separating cancers and non-cancers. Patients with high Combined Scores are prime candidates for enrolment in chemoprevention studies, where the Combined Score may be most useful as a method of monitoring response and screening for a higher risk study population more likely to benefit from

treatment. This will hopefully spur more interest in investigating chemopreventative therapies that will treat the carcinogenic process before invasive disease appears, saving money and patient lives in the long run.

3 Double Staining Cytologic Samples with Quantitative Feulgen-Thionin and Anti-Ki-67 Immunocytochemistry as a Method of Distinguishing Cells with Abnormal DNA Content from Normal Cycling Cells

A version of this chapter has been published as: Li G, Guillaud M, Follen M, & MacAulay C (2012) Double staining cytologic samples with quantitative Feulgen-thionin and anti-Ki-67 immunocytochemistry as a method of distinguishing cells with abnormal DNA content from normal cycling cells. Anal Quant Cytopathol Histopathol 34(5):273-284. Significant edits have been made to the Introduction and minor stylistic and grammatical changes have been made throughout in order to integrate this content into the flow of the thesis.

3.1 Introduction

We have shown that ploidy analysis combined with MAC features may be a useful biomarker for risk assessment in lung dysplasias. However, ploidy analysis itself has room for improvement because the amount of DNA present within the nucleus of a normal cycling cell changes as it progresses through the cell cycle. A normal cycling cell can be diploid, tetraploid, or somewhere in between. Frankly abnormal cells (>2.5 times the normal complement of DNA) are rare and occur in widely disparate and very low frequencies, even in high-grade dysplasias such as high-grade squamous intraepithelial lesions (HGSIL) of the cervix (106, 158, 257). Hence, ploidy might be an improved biomarker for cancer screening if normal dividing cells could be distinguished from abnormal non-cycling cells by using an immunostain for Ki-67 as a marker of

cell proliferation. Ki-67 is an antigen expressed in the nuclei or on chromosome surfaces during all active phases of the cell cycle (i.e., all except G₀) (249). As such, it has been used for many years as a proliferation marker and in the assessment of many cancers (249, 250, 258-260), including cervical cancer (261-264).

Previous attempts to simultaneously determine DNA content and assess proliferation status in the same cell have relied heavily on fluorescent labels detected by flow cytometry, in which abnormal cell identification is hindered by the uncertainty of only individual cell passage through the flow cytometer, which can mistakenly detect signals from non-cellular material, especially when trying to detect a relatively rare event (265, 266). We instead propose to use absorbance stains on slide-mounted samples. Absorbance stains are permanent and less costly to image, as a simple light microscope will suffice. A slide-based assay would enable the study of a wider range of sample types without picking up signals from non-cellular material. By double staining cytological specimens, normal cycling cells can be removed from the analysis, focussing on those cells whose abnormal DNA content might be indicative of large-scale chromosomal mutations associated with precancerous changes. Hence, we hypothesize that by studying Ki-67-negative cells only, ploidy analysis can be a better indicator of high-risk dysplastic lesions than ploidy analysis on cycling and non-cycling cells combined.

Double staining will be attempted on cervical cytological specimens. Cervical cancer is a relatively commonly diagnosed cancer on a global scale, ranking third amongst females (2). Screening programs in industrialized nations have successfully reduced cervical cancer mortality, but the vast majority of cases today arise in low-resource settings. Hence, there is a great need for simple and effective screening programs in developing countries. The success of ploidy analysis in other cancers has prompted interest in its potential application to cervical cancer screening (154, 267). Indeed, ploidy analysis is already in use successfully in China for cervical cancer

screening (151, 153). We have previously shown that ploidy analysis using Feulgen-thionin staining performs comparably with conventional cytology and HPV testing for detecting cervical high-grade dysplastic lesions (152). Further study suggests that ploidy and HPV mRNA may be independent predictors of cervical dysplasia (195). Despite its past successes, ploidy still has room for improvement, as noted above. Hence, we believe that double staining will prove to be a better indicator of high-grade cervical dysplasias than ploidy analysis alone.

3.2 Materials and methods

3.2.1 Cell culture

Cell culture was performed to generate large numbers of cytology slides for protocol optimization. HL-60 acute promyelocytic leukemia and H460 large cell lung cancer cell lines were obtained from the American Type Culture Collection (Manassas, VA, USA) and maintained in Iscove's Modified Dulbecco's Medium and RPMI, respectively, supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin. Cells were incubated at 37°C in an atmosphere of 95% air and 5% carbon dioxide. To generate HL-60 slides, autoclaved, uncharged, pre-cleaned glass slides were placed in square culture dishes, 3 per dish, and covered with 15 mL of cell suspension at 5×10^5 cells/mL in growth medium. To each dish was added 15 µL of 1 mg/mL phorbol 12-myristate 13-acetate (Sigma-Aldrich Canada, Oakville, ON, Canada) solution in ethanol, causing the cells to adhere to the slides (268). After an additional 48 hours of growth, the slides were rinsed, fixed in Sed-Fix® (Surgipath, Richmond, IL, USA) for 40 minutes, and allowed to dry overnight. Before use in any staining procedures, slides were cleared of dried fixative by immersion in ethanol for 20 minutes at room temperature followed by thorough air-drying. H460 slides were made by fixing a cell suspension (via trypsinization) of 4×10^4 cells/mL

in 10% buffered formalin for 10 minutes at room temperature. 250 μ L of this suspension was cytopun directly to each slide.

Cell lines were chosen for convenience and because HL-60 slides prepared as described were used routinely in our laboratory as a control for batch-to-batch variation in Feulgen-thionin staining. Staining of cell lines was used to optimize staining and imaging protocols only and no attempt was made to glean information about cancer biology from these results.

3.2.2 Patient samples

Specimens collected from forty-nine cervical cytology brushings representing a range of dysplastic grades from a previous study (152) were used in this work. Approval was granted by the Internal Review Boards at M. D. Anderson Cancer Center, the University of Texas Health Science Center, the Lyndon Baines Johnson Hospital Health District, British Columbia Cancer Agency (BCCA), and the University of British Columbia. In the previous study, brushings were collected and fixed in PreservCyt (Hologic Inc, Bedford, MA, USA) and used to generate slides using the ThinPrep method (Hologic Inc). The residual material was stored at 1°C in a cold room before being used for the present study. As many of the vials contained pieces of tissue and other debris that might confound cytological analysis, the specimens were vortexed and allowed to settle for 15 minutes on ice before use. Samples were taken from the supernatant, post-fixed 10 minutes at room temperature with 10% buffered formalin, and cytopun on to new slides in duplicate. Within a day, one slide was stained with Feulgen-thionin only as a control, while the other was double stained.

3.2.3 Immunocytochemistry

Concentrate buffer solutions for antigen retrieval (pH 6 and 9) were obtained from Vector Laboratories (Burlingame, CA, USA) and used at 1:100 dilution. Bovine serum albumin (BSA),

SIGMAFAST™ 3,3'-diaminobenzidine (DAB) chromogen tablets, and HRP-conjugated rabbit anti-mouse secondary antibody were obtained from Sigma-Aldrich Canada. Anti-Ki-67 monoclonal antibody (clone MIB-1) and serum-free protein block were purchased from Dako Canada (Mississauga, ON, Canada). All antibodies were diluted in 1% BSA in phosphate-buffered saline (PBS) just prior to use.

Antigen retrieval (AR) was performed using the microwave method, followed by cooling for 20 minutes. For patient samples, pH 9 buffer was used for 22.5 minutes, consistent with the vendor's recommendations for the anti-Ki-67 primary antibody; various conditions were tried in the optimization experiments with cell lines. Blocking steps were 15 minutes with 3% v/v H₂O₂ in methanol for endogenous peroxidase, 5 minutes with 1% BSA and 0.1% Triton® X-100 in PBS for permeabilization, and 30 minutes with protein block for non-specific binding. Antibody incubations were one hour at room temperature for primary and 30 minutes for secondary (diluted 1:800), followed by 7 minutes with the DAB chromogen solution. After a thorough rinse, slides were dehydrated through graded alcohols, cleared in xylene, and coverslipped with Cytoseal™ mounting medium (Fisher Scientific Canada, Ottawa, ON, Canada).

3.2.4 Thionin staining

Thionin acetate powder was obtained from Sigma-Aldrich. All solutions required for thionin staining were prepared the day before use. To make approximately 250 mL of the thionin staining solution, 0.125 g thionin was added to 110 mL deionized water and boiled for 5 minutes. After cooling to room temperature, 32.5 mL 1 N hydrochloric acid, 110 mL *tert*-butanol, and 2.175 g sodium bisulphite were added. The mixture was stirred for one hour, allowed to stand overnight, and filtered immediately prior to use.

All steps were performed at 23-24°C inside a temperature-controlled water bath. All steps were separated by thorough deionized water washes. The slides were post-fixed in Böhm-Sprenger fixative (methanol, formalin, and acetic acid, in a 16:3:1 volume ratio) for one hour, hydrolyzed for one hour in 5 N hydrochloric acid, immersed in the thionin staining solution for one hour, and rinsed thrice in a bisulphite rinse solution (0.5% sodium bisulphite (w/v) in 0.05 N hydrochloric acid), each separated by water rinses. After a final thorough wash, the slides were dehydrated through three changes of ethanol, 30 seconds each, cleared in xylene, and coverslipped before imaging. The hydrolysis period was varied in some experiments, so when multiple slides were stained on the same run with different hydrolysis times, the slides with the longest hydrolysis time were started first, with the other slides joining in such a manner that the hydrolysis period for all slides ended together. All thionin staining runs included at least one HL-60 slide that was to be stained with only thionin and hydrolyzed for 60 minutes to act as a run control.

3.2.5 Double staining with thionin and immunocytochemistry

When immunocytochemistry (ICC) was performed first, the procedure for immunostaining was followed as above and the slides were left overnight fully coverslipped. The following day, the slides were decoverslipped in xylene and rehydrated through graded alcohols, ending in several water washes before proceeding with the thionin staining.

When thionin staining was performed before ICC, the thionin procedure was followed as described up to the final rinse before dehydration with ethanol. The slides were then placed in PBS overnight. The slides were rinsed briefly with deionized water before performing immunocytochemistry.

The majority of the double staining of patient samples was done over 3 batches, each comprising a mix of dysplastic grades and control slides.

3.2.6 Imaging and analysis

Thionin-stained cells were imaged and analyzed using the automated Cyto-Savant™ image cytometer (Oncometrics Inc, Vancouver, Canada) (186, 192). The system was programmed to collect a random sampling of about 7000-10000 cells. All objects were subjected to a classification tree to sort objects into different classes; the only class of objects used in our analysis are the epithelial cells (243). As the Feulgen-thionin stain is stoichiometric for DNA, DNA content is proportional to the integrated optical density (IOD) of the cell. Each cell's ploidy status was assessed by normalizing the cell's IOD against the mean IOD of the sample's diploid cell population, as determined from a frequency histogram of the nuclear IODs (241). Diploid cells were assigned a DNA index of 1.0, alternatively denoted 2c.

For double-stained slides, imaging was performed after thionin staining. As the imaging system was monochrome, there was no way of determining a cell's immunostaining status directly from the system's output. However, cells could be manually revisited under the microscope by selecting them from the image gallery of cells automatically collected, allowing a human operator to manually assess each cell's immunocytochemical staining status. Hence, Ki-67-positive cells must be captured by the cytometer in order for them to be counted in our analysis.

3.3 Results

As this work used residual specimens, summary data on the study population have previously been published (152). Moreover, the results from previously performed ploidy analyses, the cytological and histological diagnoses, and the HPV test results were all available. A subset of 49 specimens was selected from this set (Table 3.1). Using moderate dysplasia and worse as the threshold for defining a high-grade lesion, the study samples were from 29 low-grade (LGSIL) or negative and 20 high-grade dysplasia cases (HGSIL), as determined previously by histopathology (152). HPV status had previously been determined by the Hybrid Capture II (HC II) test (152) and those positive for both low-risk and high-risk strains were counted as high-risk for the purpose of this study.

Cytology	Histology				Total
	Negative	Atypia/HPV	LGSIL/CIN1	HGSIL/CIN2+	
Negative/atypia	7	11	1	5	24
LGSIL	0	0	2	6	8
HGSIL	0	0	5	8	13
No diagnosis	0	3	0	1	4
Total	7	14	8	20	49

Table 3.1: Patient specimens classified by conventional cytology and histopathology. Highlighting and italicized font denote those cases classified as positive for high-grade cervical dysplasia. All others were treated as negative.

3.3.1 Destaining of thionin by antigen retrieval

It was quickly discovered that heat-mediated antigen retrieval would destain thionin stained samples. Alternatively, the images and coordinates of the thionin stained cells could be stored on a computer before performing ICC. The ICC staining could then be matched up with the stored thionin data. However, even 15 minutes of Feulgen hydrolysis was sufficient to render the Ki-67 antigen undetectable by ICC. Hence, it was determined that thionin staining must follow immunocytochemical staining.

3.3.2 Reduction of thionin staining intensity after immunocytochemical staining due to antigen retrieval

In our hands, the MIB-1 monoclonal antibody used to detect Ki-67 required antigen retrieval. Omitting this step consistently resulted in a complete abrogation of staining. Initial tests of thionin staining following immunocytochemistry revealed that while double staining was attainable (Figure 3.1), there was a significant reduction in thionin staining intensity compared to thionin staining alone, as reflected in the mean IOD of the diploid histogram peak.

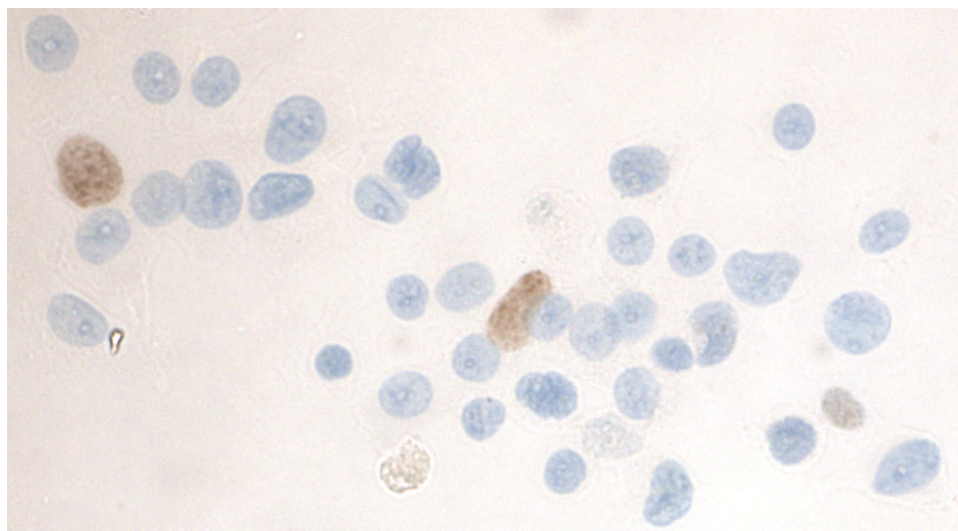


Figure 3.1: Successful double staining of HL-60 cells. Staining for anti-Ki-67 immunocytochemistry is brown and Feulgen-thionin for DNA is blue. Compared with regular thionin-only staining, however, thionin intensity was noticeably weaker and reaction conditions needed to be re-optimized. Antibodies were diluted 1:100, antigen retrieval conditions were 10.5 minutes in pH 6 citrate buffer, and the microscope image was obtained under 40X objective.

To determine whether this was due to the antigen retrieval step, a series of slides was tested in which immunocytochemistry was stopped at various steps in the protocol before thionin staining. Figure 3.2 shows that the greatest reduction in thionin staining intensity occurred after antigen retrieval and that subsequent immunocytochemical steps did not significantly alter the intensity of thionin staining beyond this initial reduction. We have also observed qualitatively weaker nuclear staining/fluorescence with hematoxylin or 4',6-diamidino-2-phenylindole (DAPI), both of which bind to DNA, after any procedures involving antigen retrieval.

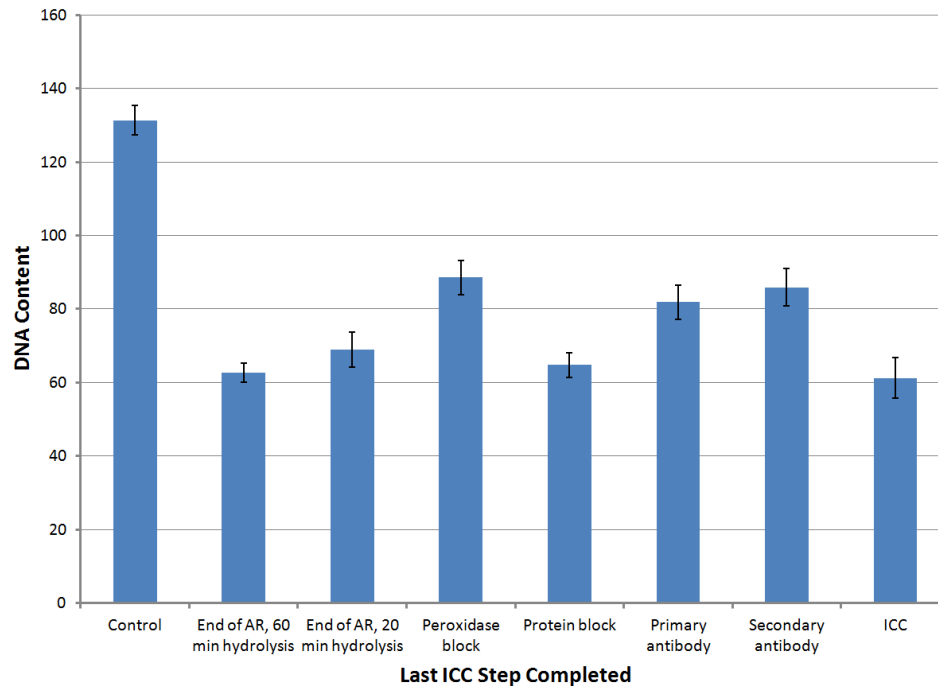


Figure 3.2: Aborted double staining test. HL-60 slides were treated first with an aborted ICC protocol, then placed into PBS after the step indicated. Slides were then subjected to Feulgen-thionin staining the next day. Control slide had no ICC steps performed whatsoever. Two slides were stopped after antigen retrieval, but were hydrolyzed for different durations during Feulgen staining, as indicated. All other Feulgen-stained slides were hydrolyzed for 20 minutes. Antibodies were diluted 1:100 and antigen retrieval conditions were 10.5 minutes in pH6 citrate buffer. In cultured HL-60 slides, the Ki-67 positivity rate is too low to significantly alter IOD means and coefficients of variation. Error bars show standard deviations.

3.3.3 Optimization of hydrolysis time

As antigen retrieval is required for MIB-1 staining, attempts were made to minimize its impact on thionin stain intensity by altering the hydrolysis time in the Feulgen-thionin staining procedure. A series of slides was treated with a mock immunocytochemical stain followed by Feulgen-thionin with various hydrolysis times, from 0 to 80 minutes, in 10-minute intervals. The primary antibody was replaced with just the diluent, ensuring that no immunostained cells would confound the automated imaging analysis. The DNA histograms were plotted and we sought to maximize the mean DNA amount of the diploid peak while minimizing the corresponding coefficient of variation (CV). There was some variability between runs, with optimal hydrolysis

times ranging between 20 and 40 minutes. For patient samples, 40 minutes was used as it was found to be the most consistent.

3.3.4 Image cytometry

DNA histograms of thionin-only HL-60 slides typically had a diploid peak around 130-150 units with a CV of 2-4%. Histogram bins were 5 units wide. Diploid and tetraploid peaks were considered to consist of all bins within approximately 2.5 standard deviations of the corresponding mean.

Even after optimization, the reduction in thionin staining intensity due to ICC persisted. Double-stained cells in the patient samples had diploid nuclear IODs averaging 56% of the corresponding HL-60 thionin-only staining control, while the thionin-only patient slides averaged 89%. In other words, after using thionin-only HL-60 slides to adjust for batch-to-batch variations in staining intensity, the double-stained slides averaged IODs of only 63% of those seen in thionin-only patient slides. As well, CVs were wider, as double-stained slides with more than 50 imaged cells showed a median CV of 11.8% (range 7.3%-25%), compared to a median CV of 4.3% (range 2.8%-11.7%) for thionin-only slides. For comparison, the corresponding ThinPrep slides prepared from these samples for the previous study had a median CV of 4.3% (range 2.7%-12.2%). Figure 3.3 shows a typical DNA histogram of Ki-67-negative cells from a patient who was negative for dysplasia.

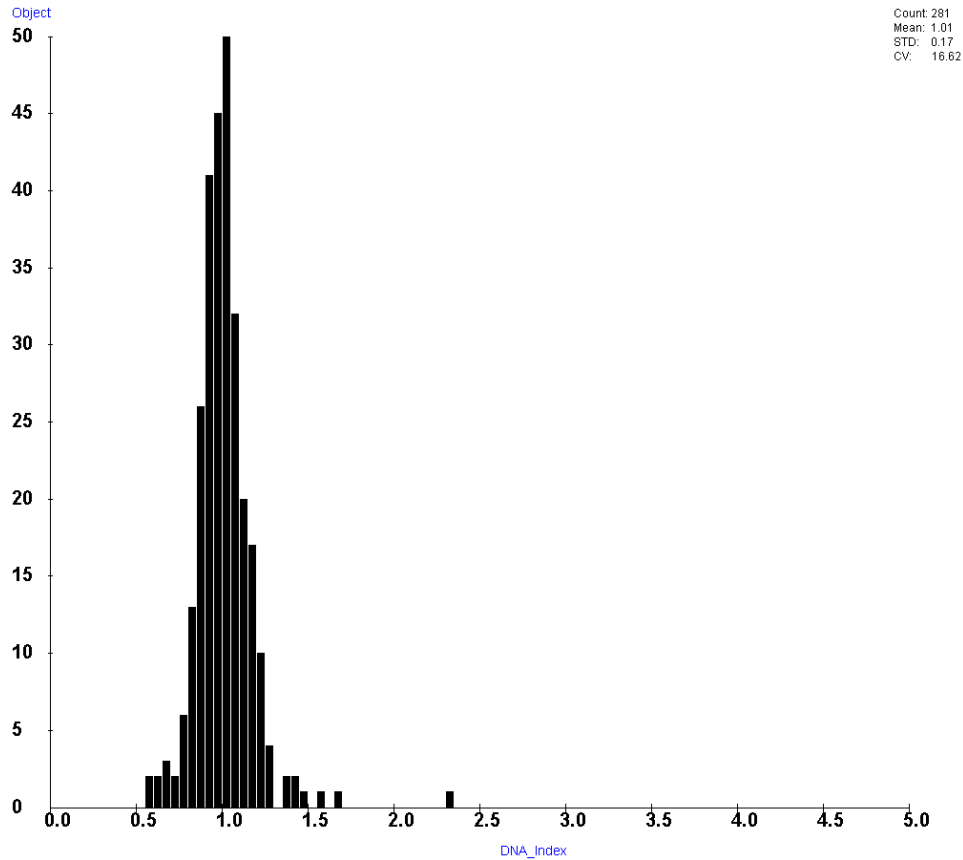


Figure 3.3: A typical DNA histogram of Ki-67-negative cells taken from a patient who was negative for dysplasia. Double staining was used to identify and remove Ki-67-positive cells. Cells were binned according to DNA index, where a value of 1.0 corresponds to the mean of the diploid peak.

3.3.5 Cervical cytology samples

Due to the weaker thionin staining after ICC, a lot fewer imaged cells were kept as they were too faint to be recognized as cells. In order to preserve statistical significance, a minimum cell count threshold of 50 was set for all patient slides. This resulted in about 12% of double-stained slides being excluded, while none of the thionin-only slides were excluded.

To assess the ability of double staining to detect the HGSIL, patients were classified based on diagnostic data from their prior study involvement. In that study, patients had conventional cytology and histopathological diagnoses of colposcopically directed biopsy. Based on these two diagnoses, presence of moderate dysplasia or worse in either defined the patient as

positive for high-grade dysplasia (HGSIL). This is the same criterion for treatment at the BCCA. Patients with other observed dysplasias were considered LGSIL, while those without dysplasia were classified as normal.

A critical test of the double staining method was whether the proportion of non-diploid cells became a better indicator of high-grade dysplasia once double staining was used to remove proliferating cells. In double-stained samples for the analyses described in this section, only Ki-67-negative cells were considered. The diploid-exceeding rate is calculated by dividing the number of cells with greater than diploid DNA content (i.e., >2.5 standard deviations above the diploid mean) by the number of all Ki-67-negative cells. In the thionin-only samples, the diploid-exceeding rate was determined from all imaged cells. Using the definition of high-grade dysplasia given above, receiver operating characteristic (ROC) curves for both these analyses could be constructed. Figure 3.4 shows that the two analyses gave similar results, with areas under the curve (AUC) of 0.73 and 0.74 (approximated by trapezoidal rule) for double-stained and thionin-only, respectively.

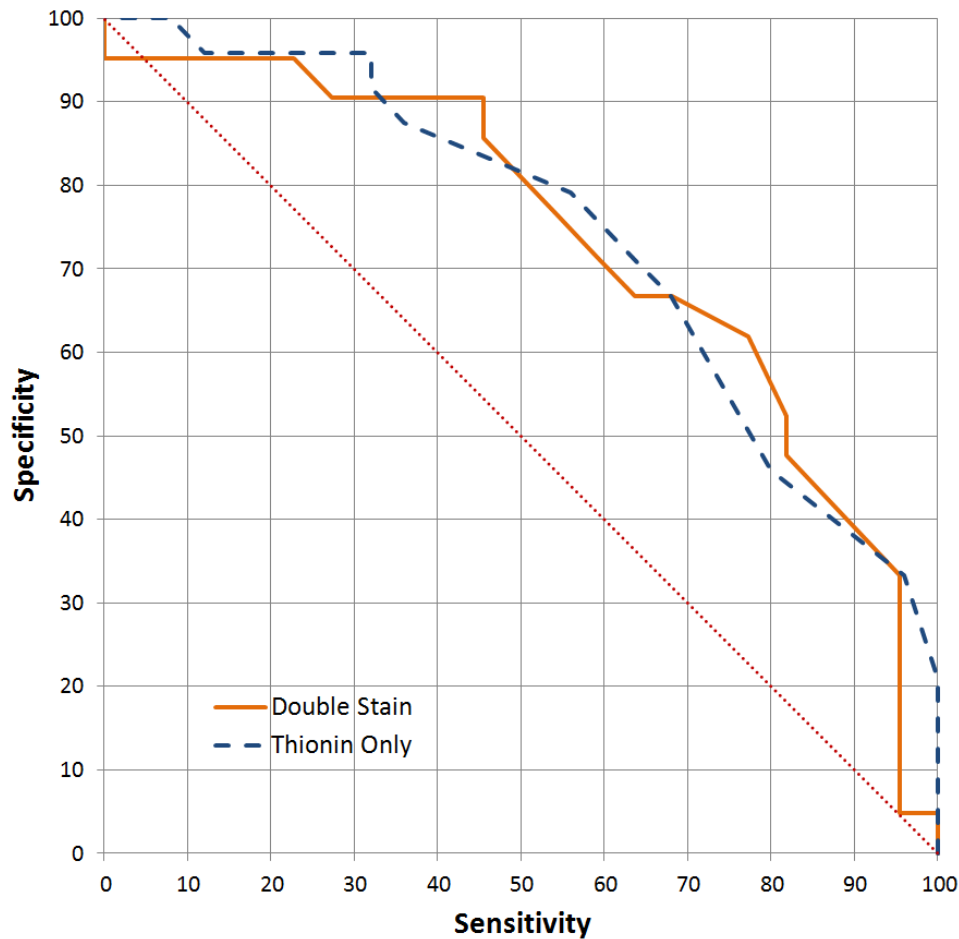


Figure 3.4: ROC curve of the performance of the diploid-exceeding rate in detecting high-grade cervical dysplasias (as defined in the text). AUCs were 0.73 and 0.74 (trapezoidal rule) for double-stained and thionin-only cytospins, respectively.

In addition to comparing double staining with thionin staining alone, one can also consider using Ki-67 alone as a potential marker for high-grade dysplasia. The Ki-67-positivity rate for the sample was determined by dividing the number of Ki-67 cells imaged by the total of all cells imaged (Ki-67 positive and negative) in a slide. Figure 3.5 shows that using Ki-67-positivity rate alone performs slightly worse than double staining at detecting high-grade dysplasias, with an AUC of 0.71.

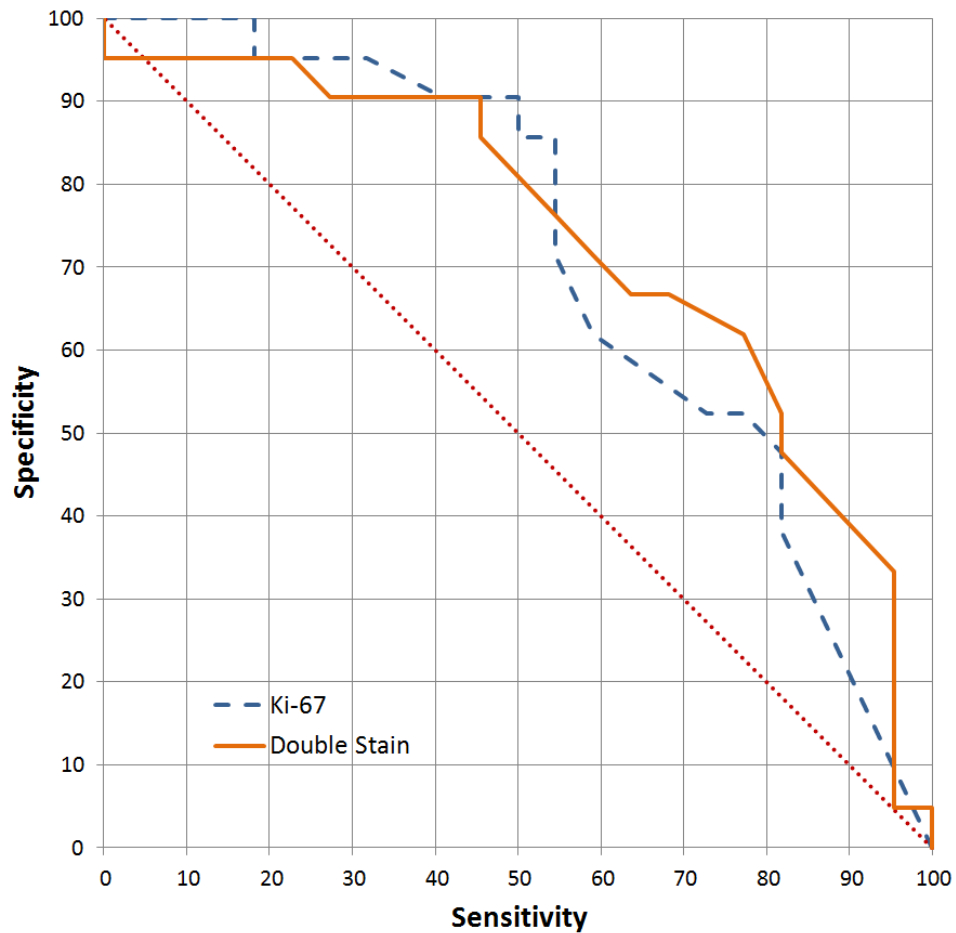


Figure 3.5: ROC curve comparing the use of Ki-67-positivity rate and double staining for detecting high-grade cervical dysplasias. Double staining performed better, with an AUC of 0.73, compared to 0.67 for Ki-67 alone.

A common approach to using ploidy measurements for cancer detection is to count the number of 5c exceeding cells. Due to the significant variation in the number of cells on each slide, the percentage of imaged cells with DNA content exceeding 5c (5cER) was used instead. In an analysis of the subset of our previously published results (152) corresponding to the samples in the present study, at a 5cER cutoff of 0.2%, sensitivity and specificity were 52% and 92%, roughly in line with our previously reported results where at least five 5c exceeding cells was used as a threshold (Table 4 in Guillaud *et al* (152)). The thionin-only cytopspins only had about 60% as many imaged cells per slide (median of 2032 versus 3373 for ThinPrep), resulting in a

reduced sensitivity and specificity of 56% and 83%, respectively, when using either a 5cER cutoff of 0.02% or at least one 5c exceeding cell to be considered positive. For the double-stained slides, however, a threshold of at least one Ki-67-negative 5c exceeding cell produced a very low sensitivity of 23% (95% specificity), likely because of the low cell count (median of 254 Ki-67-negative imaged cells, ranging from 1-2440). With 250 imaged cells per slide, a 5cER of 0.2% (i.e., the ThinPrep threshold) would be equivalent to half a 5c exceeding cell per slide.

Another approach to analyzing the ploidy data is to assess the discriminating ability of the frequency of cells falling within a series of DNA index ranges. The ranges considered were 1.3-1.6, 1.6-1.85, and 1.85-2.15. In all cases, ROC curves for the thionin-only cytopins had AUCs between 0.7-0.8, while double staining performed noticeably worse. Double staining performed best in the near-tetraploid 1.85-2.15 range, with an AUC of 0.65. Double-stained slides typically had very few cells in this range (maximum 12). Raising the minimum imaged Ki-67-negative cells per slide threshold for inclusion in this analysis to 150 improves the AUC for the 1.85-2.15 DNA index range to 0.79, virtually identical to thionin-only staining, but at a cost of excluding over 30% of the samples due to inadequate cell count.

3.4 Discussion

Automated image cytometry has demonstrated utility for early detection of various cancers (136, 148, 151, 153, 237, 269, 270). Our group had previously shown that ploidy analysis using Feulgen-thionin staining performs comparably with conventional cytology and HPV testing for detecting cervical high-grade lesions (152). However, as the ever-changing amount of DNA present within the nucleus of a normal cycling cell might confound a ploidy-based analysis, we

sought to determine if double staining with an additional proliferation marker might improve the use of ploidy in detecting high-grade cervical dysplasias.

A previous attempt at Feulgen/Ki-67 double staining used the Feulgen stain as a low-background nuclear counterstain to quantify Ki-67 labelling (271). However, the present study attempts to exploit the quantitative nature of the Feulgen reaction. This was the approach of Oud *et al*, who used an alkaline phosphatase detection of Ki-67 with the proprietary chromogen CAS Red (272). Their analysis was restricted to cell lines, while we applied this technique to patient samples to see if it would improve the clinical utility of ploidy analysis. Fleskens *et al* applied double staining to paraffin sections of oral dysplasias (273), although not in a screening or early detection setting. However, as even the authors themselves point out, direct ploidy analysis of tissue sections remains highly controversial, with studies arguing both for and against it, as it must contend with nuclear truncation and overlap (273).

3.4.1 A procedure for optimizing double staining conditions

Our results show that any attempt to double stain with Feulgen-thionin and an immunocytochemical marker must start with the ICC. The harsh conditions of microwave-induced antigen retrieval destained thionin and even a short period of acid hydrolysis as part of Feulgen-thionin staining rendered the Ki-67 antigen undetectable by ICC.

Contrary to the observations of Oud *et al* (272), we found that combining immunocytochemical staining with Feulgen-thionin had a significant impact on the intensity of thionin staining. Evidence suggests this is due to our use of antigen retrieval and we further observe that Oud and colleagues did not mention any use of antigen retrieval in their report. (While both Kolles *et al* (271) and Fleskens *et al* (273) reported using antigen retrieval, neither group attempted to compare Feulgen stain intensity with and without ICC.) As Feulgen staining

involves exposing a cell sample to an acid to hydrolyze off purine bases of DNA to generate reactive sites for the thionin stain (143, 144), Feulgen staining intensity might be impacted by any reactions that might also result in hydrolysis of DNA. Antigen retrieval, especially when mediated by heat, is hypothesized to work at least partly through hydrolysis (224). Optimization of thionin staining involves balancing the creation of more abasic reactive sites against the destruction of the DNA backbone (where the shorter segments can be lost to diffusion) using longer and more potent hydrolysis reactions (143, 144). It appears antigen retrieval both disrupts this balance and permanently reduces the quantity of DNA available for Feulgen reaction by destroying some of the DNA backbone.

As many antibodies available today require some form of antigen retrieval, our experiences with double staining with MIB-1 might enable countless scientific questions to be answered by simultaneously staining with Feulgen-thionin and any immunostain requiring antigen retrieval. Future investigations may consider, for example, the biological mechanisms by which premalignant cells become aneuploid and the consequences of such transformations. By understanding such mechanisms, biomarkers and interventions may yet be developed that target such abnormalities to help detect or even treat precancerous changes even earlier.

In order to successfully double stain a cytological sample, one must optimize a series of parameters in a specific order. First, immunocytochemical staining conditions must be optimized, with special attention given to antigen retrieval. The mildest form of antigen retrieval required to get an acceptable level of staining should be chosen. Second, Feulgen hydrolysis conditions must be optimized to account for the effects of the antigen retrieval method selected in the first step. This is best achieved by subjecting a series of slides to double staining with mock ICC and various Feulgen acid hydrolysis conditions. Antigen retrieval will generally lead to shorter optimal hydrolysis times and reduced overall Feulgen staining intensity. Finally, imaging and

analysis protocols may need to be optimized to handle the lower intensities expected with double-stained samples.

3.4.2 Ability to discriminate HGSIL and LGSIL samples

Double staining was found to match thionin staining alone in its ability to discriminate between high- and low-grade cervical dysplasias (Figure 3.4). In the high-specificity operating range, where HPV testing has typically not fared as well, double staining performs better than thionin alone on cytospins, but Fisher's exact test showed the sensitivities were not statistically significant at 90% specificity (one-tailed, $P = 0.26$). At this specificity, for the samples available, a test with twice the sensitivity of thionin alone (64% versus 32%) would be statistically distinguishable ($P = 0.03$). Overall, though, double staining failed to show any significant improvement. This is likely due to the low percentages of both Ki-67-positive cells and non-diploid cells in our samples. The median Ki-67-positivity rate across the sample set was 0.4%, while the median rate of cells with greater than diploid DNA content was 3.2%. This meant that even if the Ki-67 staining was contributing information that would allow us to better discriminate between high- and low-grade dysplasias, the effect was so small that it could not be detected with the sample sizes utilized.

A low Ki-67-positivity rate is expected for many low-grade dysplasias (261, 274, 275). Moreover, cytological specimens are preferentially sampled from the uppermost layers of the epithelium, where proliferation rates are generally substantially lower than in more basal cell layers, except for cancers and the most severe dysplasias (274). Sahebali *et al*, using Ki-67 immunostaining on cervical cytology, found only about 0.35% average Ki-67 positivity in high-grade squamous intraepithelial lesions (261). This is an even lower rate than we observed, underscoring one of the challenges of using these rare cells to improve cancer detection.

Double staining outperformed Ki-67 staining alone in detecting high-grade dysplasias. However, this analysis of Ki-67 staining is far from perfect, hampered by the inherent low rate of Ki-67 staining and our system's inability to image every individual cell. While these results show promise that double staining is an improvement over Ki-67 staining alone, more work and perhaps an improved imaging system will be needed to show this conclusively.

Any cervical cancer screening technologies are invariably compared to HPV testing. Within our sample set, using the presence of high-risk strains of HPV as the criterion for positivity, HC II testing was found to have a 92% sensitivity and 79% specificity, which is similar to the result previously reported for the full sample population from which our set was derived (152). This is considerably better than double staining (Figure 3.4) or thionin-only among this sample set.

Although double staining with thionin and anti-Ki-67 immunocytochemistry does not appear to be an improvement over regular thionin staining for the identification of high-grade dysplasias, our results suggest that double staining is a feasible assay that could be extended to other immunocytochemical stains that might demonstrate a greater improvement when paired with thionin. Perhaps some or most of the dysplastic cells are stuck at check points within the non-resting phases of the cell cycle (i.e., not in G_0) and are therefore seen as cycling cells. A more specific marker for S phase cells (e.g., proliferating cell nuclear antigen or cyclin A) or mitotic cells (e.g., phosphohistone-H3 (276)) might be better in this case, although the increased specificity of the immunostain would require a higher sample cellularity in order to observe any improvement in the ability to discriminate between high- and low-grade dysplasias. Alternatively, ploidy analysis has been shown to complement high-risk HPV testing done in parallel (277), but a double stain approach might prove even more beneficial. Another potential application of double staining could be to restore some of the tissue architectural information that is lost when

collecting cytology specimens. An immunostain could be used, for example, to label only basal epithelial cells, enabling a ploidy analysis on a defined subset of the epithelial cells.

3.4.3 Limitations of double staining

While adding anti-Ki-67 ICC does not appear to improve ploidy's ability to separate high- and low-grade dysplasias over thionin staining alone in cervical dysplasia cytology specimens, double staining remains an intriguing approach to improving ploidy analysis as a screening technique. Double staining with Feulgen-thionin and ICC offers a new approach to studying mechanisms of aneuploidy and possibly novel biomarkers for precancerous changes, but our investigations have revealed several important caveats.

The chromogen we used for ICC was DAB, which is known for its insolubility and general lack of chemical reactivity. Its ubiquity and the ability to perform Feulgen-thionin staining on slides previously stained with hematoxylin and eosin (278) or Pap stain (162) raise the possibility that previously immunostained specimens might be retrospectively stained with Feulgen-thionin. However, care must be taken when analyzing the results as the Feulgen-thionin stain reacts with deposited DAB stain. To demonstrate this, DAB substrate solution was reacted with HRP-conjugated secondary antibody and allowed to settle and air dry completely, creating a sample of DAB chromogen free of cellular material and any other compounds that might cross-react with Feulgen-thionin staining. After going through the Feulgen-thionin staining procedure, the patch of DAB chromogen demonstrated the colour change characteristic of a reaction with thionin. This was further confirmed via microspectrophotometry, a technique that enables the measurement of absorbance spectra of localized areas of microscopic samples (279), as shown in Figure 3.6.

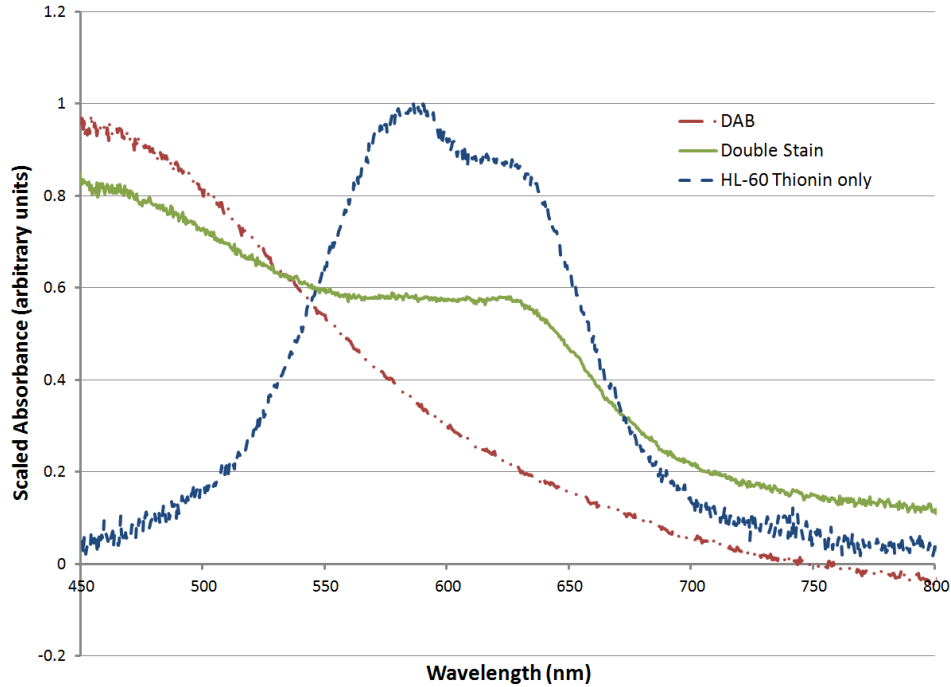


Figure 3.6: Microspectrophotometry data measured from spots of oxidized DAB chromogen deposited on a slide, before and after thionin staining, compared with thionin-stained HL-60 nuclei. Even in the absence of cells, deposited DAB chromogen reacts with thionin, giving a characteristic shoulder in the microspectra.

Analysis of double-stained specimens with DAB as the ICC chromogen should therefore be limited to DAB-negative cells. The present analysis fits this requirement, but with the expanding use of colour image analysis, quantitation of the double-stained cells may in future become a technical possibility. If analysis of ploidy of immunostain-positive cells is desired, an alternative chromogen should be sought. A washable chromogen could be used as per Fleskens *et al* (273), but that would require ensuring that every immunostain-positive cell is imaged. Although considered beyond the capabilities of our present system, this type of analysis should be possible with whole slide scanners becoming available for clinical digital pathology (280-282).

Double staining also comes with a significant cost. In addition to the time and reagents required to process the samples, the antigen retrieval of ICC significantly widens histogram peaks while reducing the intensity of the thionin stain, making it technically more challenging to ensure

accurate and reproducible results in quantitative imaging. The diploid peaks in double staining histograms had larger CVs than typical DNA measurements made by flow or standard image cytometry, but are superior to DNA cytometry measurements of tissue sections (273). Even with 10% CVs, cells with DNA indices greater than 1.25 can still be classified as non-diploid with some confidence, an observation that we had hoped would improve the sensitivity and specificity of ploidy-based detection of high-grade dysplasias. Chromosomal mutations may have given rise to an aneuploid stemline, for example, which manifests itself as a distinct population of mostly Ki-67-negative cells with DNA indices between diploid and tetraploid. Unfortunately, the observed scarcity of cells with DNA indices between 1.25 and 2.5, combined with the reduction in imaged cells per sample as a result of double staining, meant that even if double staining were an improvement, it could not be observed under the present conditions. It might be possible to improve this through the use of more cellular fresh samples.

The fainter Feulgen staining resulting from double staining could be addressed by adjusting the imaging settings. Nevertheless, sampling will undoubtedly be biased in favour of cells with more intense staining, so care must be taken in interpreting and comparing the results. By studying whether the entire double staining analysis would be superior to Feulgen-thionin staining alone, this bias becomes embedded into the reported sensitivities and specificities and does not need to be separately controlled. However, one must be careful not to assume that the Ki-67 positivity “rates” or proportions of aneuploid cells, for example, are absolute and comparable between single stain and double stain procedures. Due to the imaging bias, this is not necessarily the case, so the “rates” are more like scores that correlate with the underlying real rates.

Moreover, in a cancer screening setting, the diagnostic dysplastic cells are quite rare (typically only a few, if any, are observed per sample) (147, 151, 154). Cytological methods are

attractive for cancer screening because they do not rely on the physician knowing the precise location of a suspected lesion, thereby theoretically allowing greater sensitivity over biopsy or direct visualization-based methods. However, as cytological specimens represent an averaging of both diseased and normal cells, detecting the rare dysplastic cells can be quite difficult, a situation that is compounded by the weaker double stain that might lead many cells to be missed by the imaging system.

3.5 Conclusion

Future investigations into other ICC markers to be paired with Feulgen-thionin staining for screening or diagnostic purposes will need to show statistically significant improvement at distinguishing cases of different severities to justify the added costs of double staining. While combining Ki-67 with ploidy analysis did not show a statistically significant benefit, an improved sensitivity trend was seen in the high-specificity range of the ROC result. Meanwhile, a protocol has been demonstrated that can be used for countless other ICC markers. Unlike prior attempts at combining ICC with Feulgen staining, we have considered the effects of antigen retrieval, thus expanding the universe of ICC markers that might be suitable for combination with Feulgen-thionin staining. As long as a suitable ICC marker is chosen and proper care is taken in the analysis of double staining data, it seems that combining ICC with Feulgen-thionin staining could prove advantageous.

4 Microarray Analysis of Microdissected Molecular Fixative Cervical Dysplasias: Technical Aspects

While double staining with thionin and anti-Ki-67 immunocytochemistry failed to show significant clinical improvement over thionin staining alone, double staining as a technique appears to be feasible and potentially applicable to a wide range of markers. Ploidy analysis is already being used as a cervical cancer screening test in China and double staining using another marker in place of Ki-67 might one day prove to be a clinically significant improvement. To achieve this goal, however, novel biomarkers for cervical cancer will likely be needed.

Normal human cervical squamous epithelium consists of a differentiating continuum of cell layers. It is hypothesized that the basal layer consists of stem cells and that as cells mature and differentiate they migrate towards the surface. Hence, cells in different layers of the epithelium are expected to express different genes. Carcinogenesis is a long, multi-step process that upsets this regulated program of cell maturation. By studying differences in expression between cervical epithelial layers across various grades of cervical intraepithelial neoplasia (CIN), we seek to explore the molecular basis of the carcinogenic process.

One approach to studying genome-wide expression is to use oligonucleotide microarrays (205). While such analyses have been applied to cervical cancer in the past, such studies have typically compared invasive cancer to normal controls (283-288), ignoring any changes that might be occurring during the carcinogenic process through the various grades of dysplasia. Even when CIN is studied, such studies tend to treat the entire epithelium as one homogeneous whole (289-293), ignoring subtle differences in expression between epithelial layers that might be playing a crucial role in carcinogenesis and that do play a significant role in the pathological identification of the different grades of CIN. Microarray analysis of microdissected samples of

CIN layers, then, might offer a novel approach to understanding the early genetic changes underpinning carcinogenesis in cervical squamous epithelium.

An improved understanding of the biology will hopefully lead to better biomarkers for detecting CIN at highest risk of progression and perhaps even identify targets for early intervention. Furthermore, cytology specimens are preferentially sampled from the upper half of the epithelium, so targets/alterations that reside in the upper half are likely to be of particular interest from a screening perspective.

4.1 Molecular fixative

Advances in molecular biology have greatly improved our understanding of biological systems. Despite the wealth of cell lines, animal models, and other model systems, there remains no true substitute for clinical specimens to probe the molecular mechanisms underpinning human health and disease.

Once collected, specimens must be fixed to preserve them in a state as close to their native state as possible and to prevent tissue response to removal/wounding and further decay. Two commonly used fixation techniques in clinical settings are immersion in formalin solution and freezing. Formalin fixation is considered the gold standard for clinical diagnosis. Formalin-fixed samples are embedded in paraffin (FFPE) and the resulting block is sectioned on to glass slides. These are then stained with hematoxylin and eosin and interpreted by a pathologist. Formalin fixation preserves the tissue and cellular morphology that pathologists rely upon to make their diagnoses, while maintaining the immunoreactivity of many antigens. Consequently, there is increasing interest in extracting biomolecules from archival FFPE samples in clinical laboratories around the world for molecular studies. However, formalin fixation can be quite

damaging to many of the biomolecules of interest to molecular biologists. Damage to proteins, DNA, and especially RNA caused by formalin has been well-documented (220, 294-296), including chemical cross-linking of proteins and fragmentation and covalent modifications of nucleic acids (297). In contrast, freezing specimens preserves biomolecules better, but affects the morphology, limiting the use of frozen specimens for diagnostic purposes (298). Moreover, freezing must be done immediately, as any delay will result in biochemical changes in the tissue. In order to align molecular analyses with clinical diagnoses, researchers have typically collected adjacent specimens, fixing one in formalin and freezing the other.

The observation that a high-concentration aqueous sulphate solution precipitated out RNases at room temperature led to the development of *RNAlater* (originally marketed by Ambion, now part of Life Technologies, Carlsbad, CA, USA) (298, 299). *RNAlater* was found to effectively preserve RNA, allowing samples to be processed at a different place and time from collection (299, 300). Unfortunately, *RNAlater* alone resulted in uneven immunohistochemical staining and preserved noticeably less of the finer structural details compared to formalin (301). This could be improved by post-fixing with formalin, but formalin is known to damage RNA (297, 302).

More recently, alcohol-based molecular fixatives have been introduced that aim to combine the best attributes of formalin and freezing (298, 303-305). Samples are processed in a manner similar to formalin, including embedding in paraffin and subsequent sectioning to glass slides. However, preservation of biomolecules is decidedly superior to that of formalin (306, 307). Moreover, clinical diagnosis and molecular analysis can now be performed from the same sample block. While molecular fixatives have been tested on a number of human tissue types, little is known about their effect on cervical tissue. Successes reported for other tissue types suggest that molecular fixative will preserve biomolecules while maintaining morphological

features necessary for clinical diagnosis in cervical specimens. The improved preservation of biomolecules should enable the use of molecular fixative preserved paraffin-embedded (MFPE) cervical samples for microdissection and subsequent microarray analysis.

4.2 Materials and methods

4.2.1 Samples

Thirty cervical biopsies were fixed and frozen for long-term storage at -80°C in RNAlater (Life Technologies, Burlington, ON, Canada) as part of a previous study (308, 309), as previously described. The samples were collected from patients with various grades of cervical dysplasia. The majority were CIN II or III, with 3 carcinomas *in situ*, 3 metaplasias, one reactive atypia, and one negative for dysplasia.

Seven recently collected biopsies from patients about to undergo loop electrosurgical excision procedure (LEEP) were rapidly (all within 15 minutes, typically within 5 minutes) fixed in Tissue-Tek® Xpress® Molecular Fixative (Sakura Finetek, Torrance, CA, USA) and embedded in paraffin (MFPE). Except where otherwise noted, all references to molecular fixative and MFPE in this thesis refer to this one from Sakura Finetek. The samples were collected primarily from patients with CIN II or CIN III, but the actual regions collected had a range of histopathological grades ranging from normal to CIN III. In this study, all regions with histopathological grades of CIN II or worse will be considered high-grade and all others will be low-grade. Approval was granted by the Research Ethics Boards of the BC Cancer Agency and The University of British Columbia. Written informed consent was obtained from all participants.

4.2.2 Sample preparation

Frozen samples were rinsed with cold phosphate-buffered saline to remove *RNAlater* before embedding in ice-cold O.C.T. embedding medium (Sakura Finetek). Sectioning was performed on a -20°C cryostat. For laser microdissection, 40 sections were cut at 6 µm each to membrane slides designed for use with the microdissection machine. Slides were fixed overnight in 100% ethanol at -20°C, removed from ethanol, and then stored at -80°C with a thin film of residual alcohol. For manual microdissection, 80 sections of 10 µm each were cut on to glass slides. Slides were treated with *RNAlater*, dehydrated through graded alcohols, then air dried before storage at -80°C. In both cases, for about every ten slides generated, one section was mounted on to a regular glass slide, fixed with formalin, and stained with hematoxylin and eosin for reference.

Molecular fixative samples were handled like routine formalin-fixed paraffin-embedded blocks. They were cut at 8 µm, 4 sections per slide, and deparaffinized before use. 100 sections per block were cut, with reference slides cut about every 20 sections and stained with hematoxylin and eosin.

4.2.3 Microdissection

The hematoxylin and eosin-stained reference slides were scanned using a whole slide imager (Pannoramic MIDI, 3DHistech, Budapest, Hungary). Using the digital images, the study pathologist (Dr. Dirk van Niekerk) graded and circled all the regions of abnormality. This information was used to guide the microdissection of the adjacent unstained sections.

Laser-assisted microdissection involved lightly staining the sample with hematoxylin before mounting it on the microdissection machine. The Molecular Machines and Industries CellCut (Haslett, MI, USA) was equipped with an ultraviolet laser that cut the tissue and its

associated membrane around the region of interest. Once the desired region was dislodged, it was collected into a microcentrifuge tube, where it was vortexed with TRIzol reagent (Life Technologies). The epithelium was cut into three roughly equal thickness layers: outer, intermediate, and basal. In addition to the epithelial layers, the leftover tissue (consisting primarily of stromal tissue) was collected in a separate set of tubes. All the tubes from each layer were then pooled together to yield one tube per layer. Laser microdissection was only tested on frozen samples because, as discussed below, manual microdissection was found to be more promising.

Manual microdissection was performed at room temperature using a needle and a dissecting microscope. The dewaxed slides were kept on dry ice until just before microdissection. The collected sample was transferred to a tube of TRIzol (for frozen samples) or Buffer PKD (for MFPE samples, this is a component of the QIAGEN RNeasy FFPE Kit used for RNA purification, see below). As the manual method is considerably less precise than the laser method, we only collected two layers: top (superficial) half and bottom (basal) half. For frozen samples, the remaining tissue was scraped off with a razor blade and processed as “stroma.” For MFPE samples, the actual stroma underlying the collected epithelium was selectively scraped off as a separate layer. Separate needles were used for each layer of each sample.

4.2.4 RNA extraction and purification

RNA from the frozen samples was extracted using the TRIzol method. RNA was then purified by DNase I treatment, followed by a phenol-chloroform extraction. RNA from molecular fixative samples was purified using the QIAGEN RNeasy FFPE Kit (Toronto, ON, Canada). RNA amount and purity were assessed using NanoDrop (Thermo Scientific, Wilmington, DE, USA), while degradation was assessed by 1% agarose gel electrophoresis or Agilent Bioanalyzer (Mississauga, ON, Canada) assay of the sample with the most RNA in a batch.

4.2.5 Microarrays

RNA was amplified and labelled using Agilent Quick Amp Labeling kit (laser microdissected) or Agilent Low Input Quick Amp Labeling kit (manually microdissected). Labelled cRNA yield and quality was assessed according to the instructions provided by Agilent. Expression analysis on successfully labelled samples was performed using the Agilent Whole Human Gene Expression Microarray Kit, 4×44K, following Agilent's recommended protocol. These arrays assay over 41000 unique probes spanning the human genome. Each array slide allows up to 4 samples to be assayed simultaneously. In some instances, the same sample was run twice (e.g., with different RNA input amounts into the labelling reaction) to serve as an indicator of the reproducibility of the data. The hybridized microarray slides were scanned using a GenePix 4000B Microarray Scanner (Molecular Devices, Sunnyvale, CA, USA) running GenePix Pro version 6.1 software.

4.2.6 Data analysis

Data manipulation was performed in Microsoft Office Excel (Redmond, WA, USA) for spreadsheet functions and STATISTICA (StatSoft Inc., Tulsa, OK, USA) for statistical analysis.

The microarray data (background-subtracted intensity values) were normalized for each array by dividing by the median intensity value of the spots on the array. Zero and negative values were deleted and data from probes with multiple spots on the array were consolidated by removing the highest and lowest intensity values and averaging. From this, a set of normalized intensity values for each unique probe was obtained. Some genes were represented by multiple probes, but these data were not averaged. The resulting data set consisted of 41000 intensity values per sample, one for each unique probe on the microarray.

For each sample, modified M-A plots (210, 211) were generated by plotting $M = \log_2(T/B)$ against $A = \log_2(T \cdot B)$, where T and B are the array-median normalized intensity values for the top and bottom layers, respectively. Assuming that the majority of probes are expressed at similar levels between the layers, a further adjustment to the data can be made to set the central log-ratio M of each plot to be zero. Using all data with $A > 4$ to avoid fitting to data that is excessively noisy due to weak signal, a linear regression for each M-A plot was calculated and subtracted from all data in that plot.

Samples that were run in duplicate were used to determine technical scatter. Instead of comparing top against bottom layers, M-A plots were constructed by comparing the duplicate samples. After linear adjustment as above, the regions of technical scatter generated by M-A plots of duplicate samples can be used to set thresholds of significance with which to analyze the top versus bottom data of the other samples. Double exponential functions were fit manually in an attempt to replicate the envelopes traced by the duplicate data M-A plots. These fits will be overlaid on the top versus bottom data in subsequent analyses to define the range of expected technical variability and thus identify candidate targets outside this range. As the range and scaling of microarray data is somewhat arbitrary, the fits were translated so that the maximum (saturation) A values of the duplicate data underlying the fits and the test data were aligned. Test data points lying outside these fits could then be considered potential targets. To estimate the false discovery rate of using these fits, this method is applied to the duplicate data, where the biological variation between the samples is known to be zero.

To further evaluate the quality of the microarray data, unsupervised hierarchical cluster analysis was performed on all log-transformed data collected from arrays run with at least 50 ng RNA to start the labelling reaction. All saturated and low-intensity (<16 or $\log_2 < 4$) data were removed from this analysis. Complete linkage was required between grouped clusters and the

Pearson distance metric, which is more likely to capture overall correlations within the data and is less sensitive to imperfect between-array normalization than metrics like Euclidean distance (310), was used.

4.3 Results

Six frozen and seven molecular fixative cases were used in this study. These are summarized in Table 4.1. Multiple regions were microdissected from some cases and are therefore listed twice. Upon examining a couple cases in which adjacent pieces of LEEP tissue were fixed in formalin or molecular fixative in an alternating pattern, our study pathologist felt that the molecular fixative samples looked acceptable (311).

Case	Fixation	Microdissected grade	Highest grade in patient	Grade: High or Low
5034	Frozen	N/A	CIN II	High
5065	Frozen	N/A	CIN II	High
5054	Frozen	N/A	CIN II	High
5079	Frozen	N/A	Imm Sq Meta	Low
5047	Frozen	N/A	CIN III	High
5067	Frozen	N/A	CIN III	High
0027	Molecular	CIN I	CIN II	Low
0028	Molecular	CIN III	CIN III	High
0030	Molecular	CIN I	CIN III	Low
0033A	Molecular	CIN III	CIN III	High
0033B	Molecular	CIN II	CIN III	High
0043	Molecular	Normal	CIN II	Low
0044	Molecular	CIN I	CIN III	Low
0053A	Molecular	CIN II	CIN III	High
0053B	Molecular	CIN III	CIN III	High

Table 4.1: Summary of all cases used in this study. RNA was purified and analyzed from all the molecular fixative samples listed, plus frozen case 5065. Microdissected regions of frozen cases could not be graded directly and were therefore classified according to the highest grade in the patient. Regions with histopathological grades of CIN II or worse were classified as high-grade while the rest were low-grade. This is denoted in the final column. Imm Sq Meta = Immature Squamous Metaplasia.

4.3.1 Laser microdissection

Two frozen samples were laser microdissected, but both yielded unsatisfactory results. This appears to be due to the cervical tissue being very tough to cut with the laser system, taking twice as long (2 hours per section) versus typical samples of other tissue types. Since laser microdissection was performed at room temperature, this was likely to have affected the integrity of the RNA we collected and analyzed.

The first frozen sample attempted was case 5034. About 810 ng RNA was purified from the outer layer. Inadequate material was isolated from the basal layer, so it was pooled together with the intermediate layer to yield 1034 ng. All the purified RNA was used for microarray analysis, performed by the laboratory of Dr. Cathie Garnis. The outer layer failed due to low labelling yield (0.762 μ g compared to about 10 μ g for Garnis Lab's concurrently run samples),

while the combined intermediate/basal layer had a low specific activity after labelling (0.94 versus about 3 pmol Cy3 per μg cRNA).

The second sample attempted was case 5065. The Bioanalyzer reported that RNA from whole sections had an RNA Integrity Number of 7.8. However, after laser microdissection and extraction, the 260/280 ratio remained low (1.5-1.65) and would likely require further cleanup before microarray analysis. With some initial successes from the following manual microdissection of molecular fixative samples, these frozen samples (and laser microdissection) were abandoned.

4.3.2 Manual microdissection

Manual microdissection was tedious, but largely successful and feasible. Upon moving slides from dry ice to the room temperature dissecting microscope, a small amount of moisture condensed on sections. This was found to aid in microdissection by binding the collected tissue to the needle, preventing it from “flying away” as the tissue was scraped. Moreover, the slightly moist epithelium was found to hold together well and separate from the stroma cleanly at the basement membrane when gently pulled by the needle. Separating the top and bottom layers of the epithelium was best done by scoring completely dry sections with the needle point. Working on multiple slides at once allowed microdissection to proceed more quickly, allowing 12 sections to be completed in about 30 minutes. An example of one sample undergoing the various stages of microdissection is shown in Figure 4.1. Photos documenting microdissection of all the samples can be found in the Appendix.

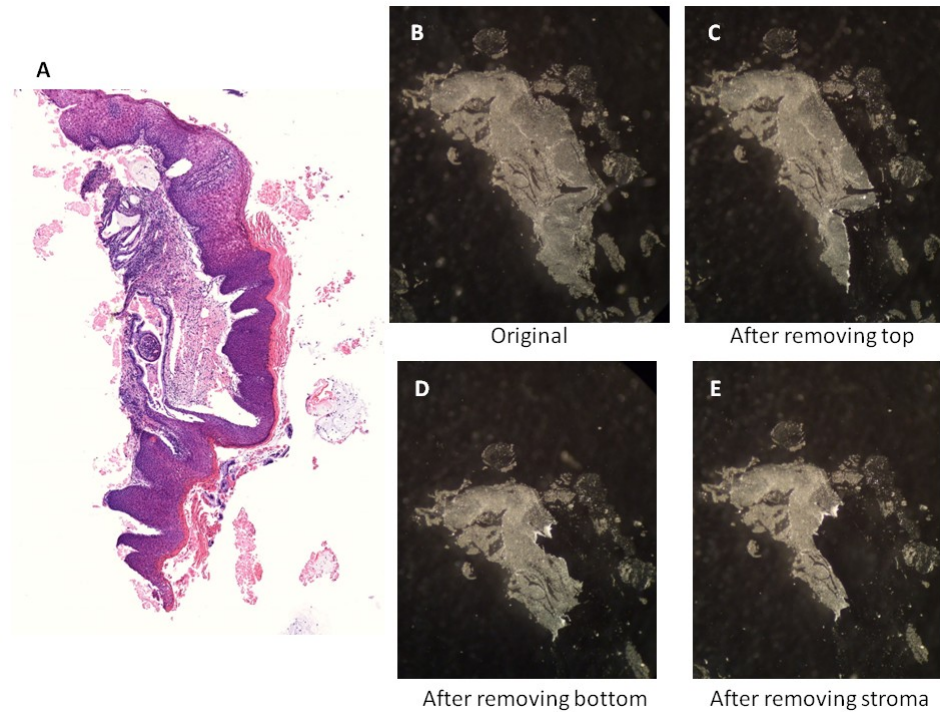


Figure 4.1: Example composite figure documenting microdissection of case 0028. A reference H&E section (**A**) is shown alongside photographs of an adjacent section before microdissection (**B**), after removing the top layer (**C**), after removing the bottom layer (**D**), and after removing the stroma (**E**). Microdissection was always performed in this order. See appendix for all other cases.

4.3.3 Frozen samples

A total of six frozen samples were processed. Two have already been detailed in the laser microdissection results (Section 4.3.1). Case 5065 was also manually microdissected and successfully generated microarray data. The stroma was run in duplicate, with 100 ng and 200 ng RNA input into the labelling reaction. When plotted against each other, most probes showed a linear relationship between the replicates, indicating that the assay is quantitative and reproducible over the RNA input range of 100-200 ng (Figure 4.2). The top and bottom layers were also run on arrays, using all available purified RNA for each layer as starting material for the labelling reactions.

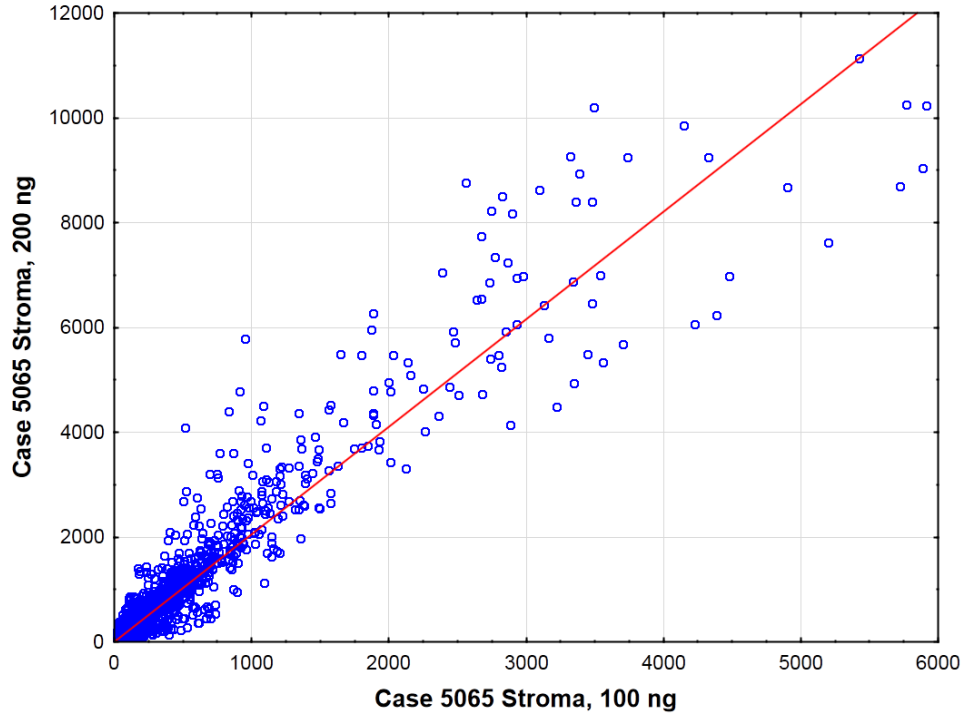


Figure 4.2: Plot of expression data from two aliquots of frozen sample case 5065 stroma. One aliquot had 200 ng RNA starting material in the labelling reaction (ordinate) while the other used only 100 ng (abscissa). Most data points lie along the linear regression line ($R^2 = 0.91$). Data from saturated spots have been removed (i.e., 200 ng data > 12000).

Another sample attempted was case 5054. This sample was initially difficult to section. This was believed to be attributable to excess residual RNA_{later} in the sample, causing it to set improperly in the cryostat. The sample was thawed and rinsed with PBS before being re-embedded. After successful sectioning, a few complete sections were used to test RNA quality. Both gel electrophoresis and Agilent Bioanalyzer confirmed that the RNA was degraded even prior to microdissection.

The fourth frozen sample processed (case 5079) was found to contain no extractable RNA, even after repeated attempts, including attempts to extract RNA from solutions that would normally be discarded during the normal extraction protocol. The fifth frozen sample (case 5047) was found to contain no epithelium. The sixth frozen sample (case 5067) was found to contain degraded RNA, even without microdissection. This was independently verified by testing several

whole sections in the laboratory of Dr. Wan Lam. Despite the initial success with case 5065, attempts to replicate the results using the remainder of the frozen samples were largely unsuccessful. Consequently, this effort was abandoned to focus on molecular fixative samples.

4.3.4 Molecular fixative samples

RNA was successfully purified from manually microdissected MFPE blocks from 7 patients. Nine regions were microdissected, including 1 normal, 3 CIN I, 3 CIN II, and 2 CIN III (Table 4.1). Most of these were successfully assayed on the gene expression microarrays. Two patients had regions of both CIN II and CIN III, which were collected separately and labelled as A and B. For one CIN I case (case 0027), the stroma was microdissected but RNA has not yet been purified. For a CIN II case (0053A), the stroma RNA was purified, but yield and quality were insufficient and the labelling reaction failed. Generally, purified RNA from the MFPE samples was superior in yield and purity compared to the frozen samples, as determined by NanoDrop. This is likely due to the difference in purification methods. RNA also appeared mostly intact, with the 18S band appearing distinctly, while the 28S band was less well-defined (Figure 4.3).

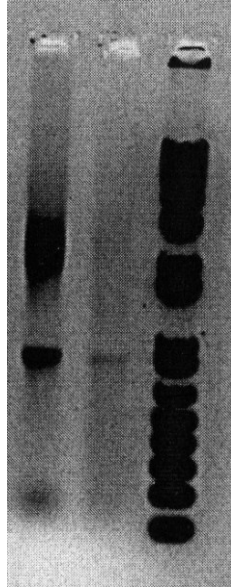


Figure 4.3: 200 ng of basal layer from case 0028 in the middle lane of a 1% agarose gel. The sample may have been over-diluted, causing it to appear faint. The TrackIt 1Kb Plus DNA Ladder (Life Technologies) is on the right. An oral cancer sample from the laboratory of Dr. Cathie Garnis is on the left as a comparison. The 18S band appears distinctly, while the 28S band is less well-defined. Gel and image were done by the lab of Dr. Cathie Garnis.

Among the molecular fixative samples, three were assayed on the microarrays in duplicate: A normal (case 0043) bottom layer run with 200 ng and 50 ng input into the labelling reaction, a CIN I (0044) top layer run with two 200 ng aliquots input, and a CIN III (0028) bottom layer run with 200 ng and 25 ng input. The first two sets of duplicates showed an excellent linear correlation (e.g., Figure 4.4), but the third set exhibited a significant side branch (Figure 4.5).

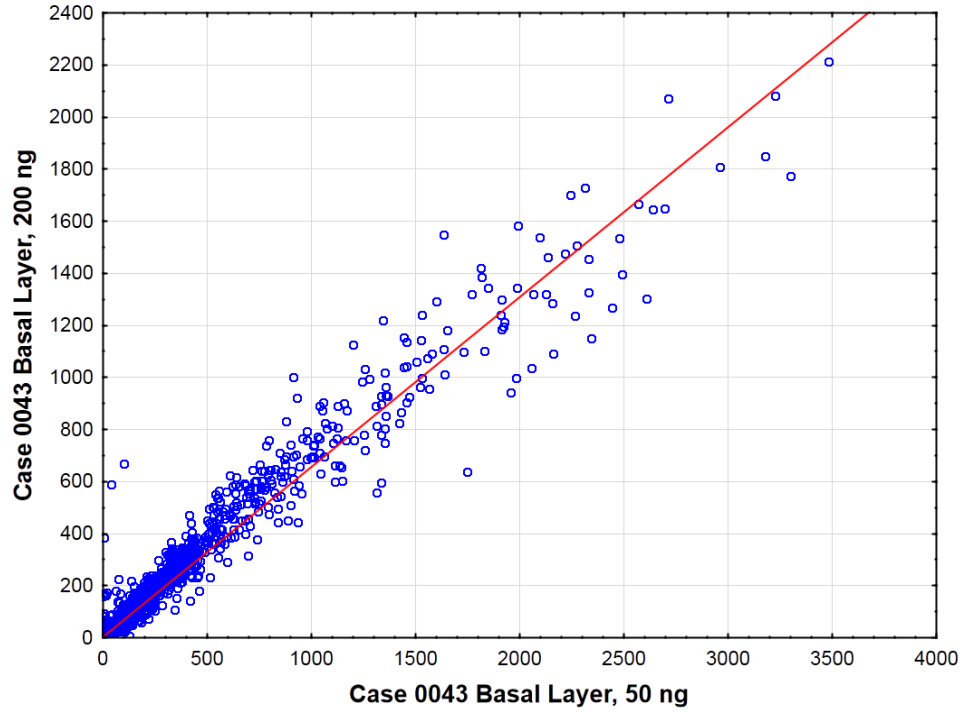


Figure 4.4: Plot of expression data from two aliquots of case 0043 basal layer. One aliquot had 200 ng RNA starting material in the labelling reaction (ordinate) while the other used only 50 ng (abscissa). Most data points lie along the linear regression line ($R^2 = 0.96$). Data of saturated spots have been removed (e.g., 200 ng > 2400).

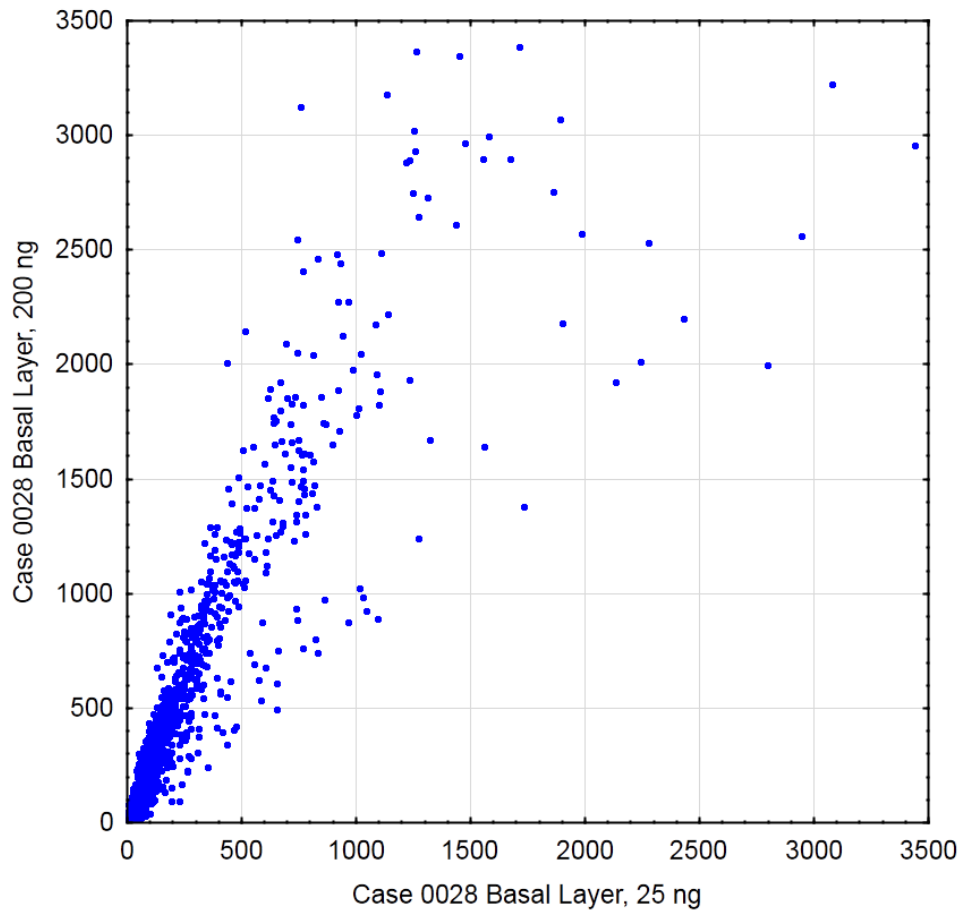


Figure 4.5: Plot of expression data from two aliquots of case 0028 basal layer. One aliquot had 200 ng RNA starting material in the labelling reaction (ordinate) while the other used only 25 ng (abscissa). A distinct side branch in the data is visible, suggesting 25 ng might be too little starting material. Data of saturated spots have been removed.

Hence, good microarray data can be obtained using 50-200 ng RNA in the labelling reaction. The scatter in the molecular fixative replicate data was also less than in the frozen sample (Figure 4.6). The normal data produced slightly more scatter than the CIN I data, so their fits will be referred to as the wide and narrow thresholds, respectively.

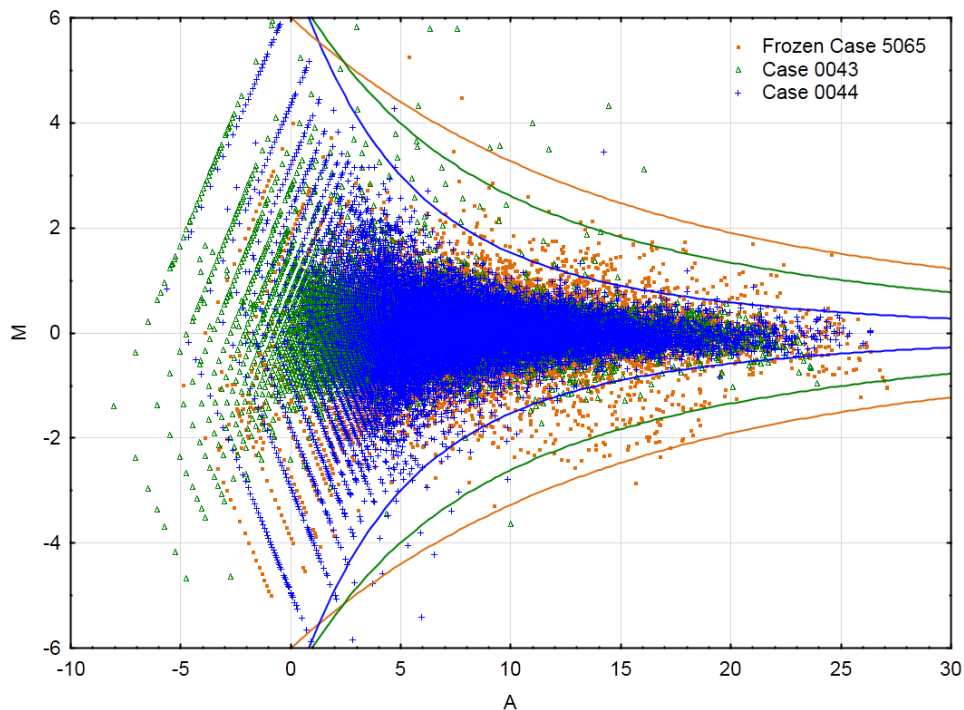


Figure 4.6: Overlaid M-A plots of duplicate data for the frozen and two MFPE samples. Data have been adjusted so that the mean M is zero. The traces are manually fit double exponentials that can be used as thresholds separating data that is indistinguishable from technical scatter and likely differentially expressed targets.

4.3.5 False discovery rates

The accuracy of the fitting can be assessed by counting the number of false positives that the fit would “discover” in the duplicate data from which it was derived. Each set of thresholds was overlaid over each set of duplicate data and the number of data points outside the thresholds was determined. Additionally, as a narrow threshold would result in a high false discovery rate on its own, additional criteria of a greater than two-fold change ($|M| > 1$) and a sufficiently high intensity ($A > (\max A) - 20$) were applied. The results are shown in Table 2.

Threshold bands	Number of false discoveries in duplicate samples					
	Frozen		Normal (Wide)		CIN I (Narrow)	
Frozen	4	(0.0098%)	27	(0.0659%)	18	(0.0439%)
Wide	15	(0.0366%)	16	(0.0390%)	4	(0.0098%)
Narrow	315	(0.7683%)	120	(0.2927%)	43	(0.1049%)
Narrow with extra criteria	247	(0.6024%)	63	(0.1537%)	20	(0.0488%)

Table 4.2: False discovery rates when differential expression criteria based on the scatter of duplicate data are applied to the three sets of duplicate data. For each condition, the counts and percentages of data points outside the indicated threshold bands are given.

4.3.6 Cluster analysis

Cluster analysis revealed that replicate samples and epithelial samples from the same patient tended to cluster together (Figure 4.7). The two data sets corresponding to the same hybridized array scanned twice with different gain settings (0044 TopB) were the most similar, followed by a replicate array of the same sample (0044 TopA). The frozen (5065) and stroma samples formed their own clusters. For case 0053, the CIN II and III regions were identified in essentially the same region on different H&E slides. Hence, the A and B samples constitute adjacent regions that differ only by their depth in the tissue block. As high-grade lesions, the top and bottom layers also appear morphologically similar. We observe, then, that all four epithelial samples from case 0053 cluster closely together. A similar situation is observed with case 0033 and again, its four epithelial samples cluster together. Case 0028 was somewhat anomalous as it clustered closer to the frozen sample and not the other MFPE ones.

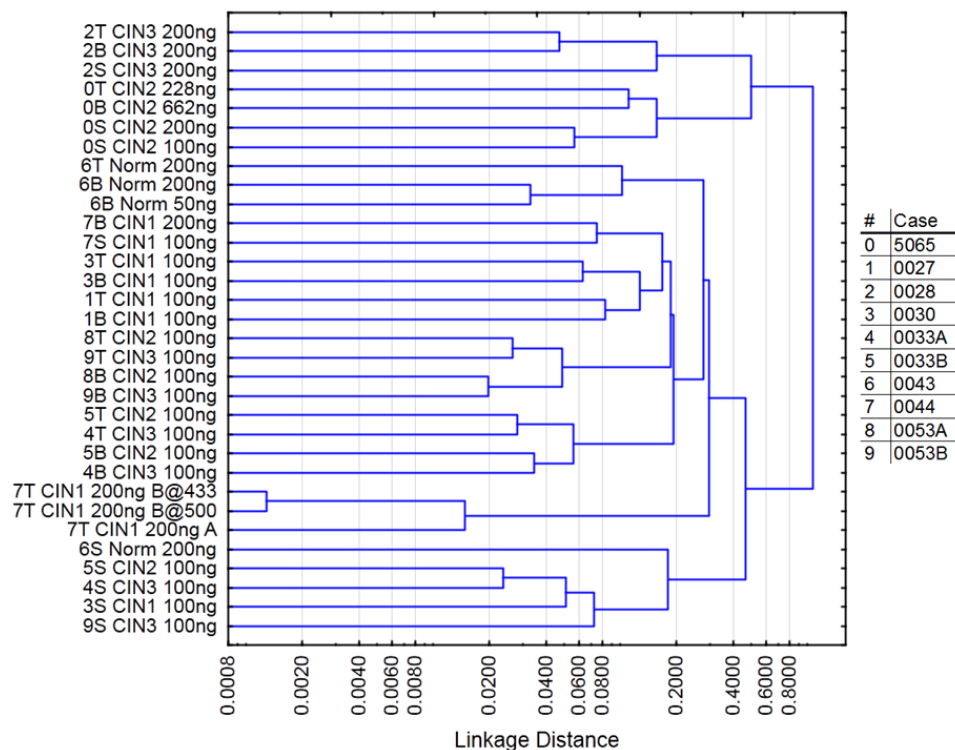


Figure 4.7: Cluster analysis of all log-transformed microarray data. Pearson distances and complete linkage were used. Each array is labelled by case number (abbreviated as per the table at right), layer (T = Top, B = Bottom, S = Stroma), grade (Norm = Normal), and the amount of RNA used in the labelling reaction. Where given, the number after the @ symbol denotes the detector gain setting on the microarray scanner. In all other cases (plus the one where gain = 433), the detector gain was set automatically by the scanning software. Two samples that differ only by detector gain were the same hybridized array imaged twice with different settings. All other samples represent distinct arrays. Two 200 ng aliquots of top layer RNA were assayed from case 0044 (case 7 in this figure), so these are denoted A and B. Note that in this figure, cases 4 and 5 are from the same patient, as are cases 8 and 9.

4.4 Discussion

Formalin fixation has routinely been used in laboratories all around the world for processing human specimens for clinical use. However, the deleterious effects of formalin fixation on biomolecules have been well documented and present a significant challenge to attempts to use such material for molecular biological studies (220, 294-296). Frozen samples, on the other hand, produce better quality biomolecules (306, 312), but at the cost of being unreliable

for clinical use. Molecular fixative has been proposed as a fixative that will allow clinical diagnosis and molecular biology to be performed on the same sample (298). Molecular fixative in some tissues has already been shown to preserve morphology of tissues nearly as well as formalin (298, 306). Meanwhile, RCL2, a commercially available alcohol-based fixative that competes with the Tissue-Tek® Xpress® Molecular Fixative used here, very recently was shown to perform well on a number of tissue types, including cervical specimens (313) supporting our efforts to evaluate the technical feasibility of using molecular fixative samples for microdissection and subsequent gene expression microarray analysis.

Investigations into microdissection techniques showed that laser microdissection was unsatisfactory when applied to the small sections from cervical biopsies. It was slow and produced unusable RNA, although subsequent experiments suggested that the latter may have been a result of the suboptimal samples being used. While laser microdissection has helped shape our understanding of many aspects of pathology and cancer biology, it is not often applied to cervical neoplasms (314). Laser microdissection has been used on 5 μ m sections of cervical tissue (314, 315), not much thinner than what we used, but these reports do not indicate how long this took. Wilting *et al* used laser microdissection to selectively increase the proportion of epithelial cells on some of their 8 μ m sections, but again did not indicate how long this took (283). While the laser microdissection system used here relies on cutting through the tissue, there are other systems that operate by using an infrared laser to melt a transfer membrane, fusing with the sample (316). The membrane is then lifted off, taking the fused sample along with it. Gius *et al* reported using this type of system on 10 μ m LEEP sections, taking about 5 hours to microdissect about 20 slides (289). This is much slower than our manual microdissection. Moreover, such systems are better suited for collecting smaller regions as only cells in the vicinity of the laser pulse are collected, making such systems a poor choice for our study. In our experience, it is clear that laser microdissection takes significantly longer than manual

microdissection. Consequently, laser microdissection was deemed to not be worth the effort and abandoned.

Most of the frozen samples were found to have unusable RNA. Many attempts were made to tweak the protocol, including re-extracting from what would normally be waste solutions, but all to no avail. In many cases, degraded RNA was found in whole sections, meaning microdissected samples would yield even more degraded material. For the one frozen sample that was usable, the RNA was less pure and generated microarray data with more scatter than any of the MFPE samples. The poor results are likely due to a combination of the age of the samples, the fixative, and the RNA purification method, although we are unable to tease apart the contribution from each. A direct comparison between frozen and MFPE samples is not possible in this study.

Cluster analysis was used to get an impression of the reliability of the data by looking at the general relationships between the samples. Many studies employ single linkage clustering, in which the distance between two clusters at each step is defined by the minimum distance between an element from one cluster and an element from the other. In some cases, two clusters in which the majority of elements are far apart may be joined by single linkage if one element from each is close to each other. Complete linkage, on the other hand, avoids this problem by using the maximum distance between clusters, generally resulting in more compact clusters. We also chose to use the Pearson distance metric as it captures the correlations between genes via the Pearson correlation coefficient (310). Relative expression levels for the majority of genes are expected to be similar in all the samples. The Euclidean distance, another common metric for cluster analysis, depends on the absolute expression differences between samples, leaving it sensitive to imperfections in between-array normalization. Cluster analysis was found to be sensitive to saturated data, as saturation of a gene signal in only one sample resulted in a significant deviation

from the regression line. On the other hand, cluster analysis was sensitive to low-intensity data only when performed on log-transformed data. At low intensities, small deviations are magnified once the logarithm is taken. For these reasons, both saturated and low-intensity data were removed from the cluster analysis. There was little difference between the clustering results performed on log-transformed and untransformed data.

We would expect replicate samples to cluster most tightly together. Samples taken from the same patient would also be expected to cluster together as would perhaps samples of the same CIN grade. On the other hand, we would expect epithelial samples to cluster separately from stromal samples and, ideally, high-grade samples would cluster separately from normal and low-grade samples. In our data, we found that replicate and adjacent samples tended to cluster together and stroma generally clustered separately from epithelium. Aside from case 0028, frozen samples behaved differently from the MFPE ones. As noted, the frozen samples were collected and handled differently, which likely accounted for much of the difference. The behaviour of case 0028 is a little more difficult to explain. However, it was the first MFPE sample processed and my relative inexperience with the procedures at the time may have played a role. Besides these two cases, most of the data clustered as expected and give us additional confidence in our data. Of note, samples from the same patient tended to cluster together, suggesting that there might be significant biological variation between patients, even those harbouring lesions of the same CIN grade.

All pair-wise comparisons of microarray data showed similar levels of expression between samples for the majority of probes. In MFPE samples, this was valid down to 50 ng of purified total RNA. Duplicate samples, in particular, allow us to quantify the level of technical scatter in the data. The frozen sample exhibited the widest scatter, while the MFPE samples

varied somewhat in the amount of scatter present, showing that there is some variability between different runs or samples in terms of the level of scatter one can expect.

With an eye towards using the technical scatter to analyze the epithelial layer comparison data, thresholds were set to approximate the scatter inherent in the duplicate samples, separating data that is indistinguishable from noise from differentially expressed potential targets. Applying these thresholds back on to the duplicate data allowed the false discovery rate to be estimated (Table 4.2) for these thresholds. The results trend mostly as expected, although the frozen data threshold is somewhat anomalous when applied to the molecular fixative data because that threshold is relatively flat and picks up a lot of low-intensity false positives. If the intensity criterion were also applied to the frozen data threshold, the false discovery rate for the molecular fixative samples would be greatly reduced. Focussing in on the MFPE sample data, the wide thresholds will certainly be specific, but might not be sensitive enough to detect many valid targets. On the other hand, the narrow thresholds might result in too many hits that would then need to be whittled down to a more manageable number. Adding the extra intensity and fold-change criteria helps reduce the number of false positive hits. Depending on the scatter in the specific data set, this might be sufficient. Since we have 5 high-grade (CIN II or worse) and 4 low-grade dysplasia samples (CIN I or normal), we could require that changes detected occur in at least 3 samples of the same group (high- or low-grade) and must agree before the change is accepted as significant. Taking the most permissive scenario, narrow thresholds on widely scattering data, and accounting for all the possible combinations, this yields a probability of $2.5 \times 10^{-5} \%$ that a probe would register as a false positive for at least 3 out of 5 samples, or about 0.01 probes in a microarray of 41000. This number drops further if we add the intensity and fold-change criteria. Hence, we can be confident that even the narrow thresholds will yield a small number of false positives as long as at least three samples are required to agree.

From a technical perspective, then, gene expression microarray analysis of cervical squamous epithelium samples in which the layers of the epithelium are microdissected appears very promising. As microdissection entails additional sample handling and smaller volumes of tissue being processed for analysis, RNA quality and yield would be expected to be worse than comparable analyses of whole sections or biopsies. Nevertheless, replicate samples show good agreement with a quantifiable amount of technical scatter that can be exploited to identify potential targets of interest, even when using as little as 50 ng total RNA for the labelling reaction. Molecular fixative appears to generate results as good as or even better than frozen samples while preserving morphology. A very recent publication not available at the start of this work confirmed the suitability of alcohol-based molecular fixatives in studying cervical tissues for morphology and molecular analyses (313). However, as with any other genome-wide expression analysis, the ultimate test will come from biological validation of the identified targets. This will be the focus of the following chapter.

5 Microarray Analysis of Epithelial Layers in Cervical Dysplasia: Biological Validation

5.1 Introduction

We have demonstrated that gene expression microarray analysis of microdissected molecular fixative preserved paraffin-embedded (MFPE) cervical epithelial samples generates data of adequate quality for analysis. Comparisons of replicate data show high degrees of correlation and we have used these analyses to define thresholds that could be used to identify differentially expressed genes. The true test of any expression microarray analysis, however, is biological validation: Are the results biologically relevant and meaningful? In this chapter, I tackle this question by identifying candidate biomarkers from the gene expression profiling data and then attempting to validate these results at the protein level using immunohistochemistry (IHC). Hopefully, the improved biological understanding provided by microarray analysis of epithelial layers in cervical dysplasia will lead to the discovery of novel biomarkers of CIN at highest risk of progression and that can be assayed on cervical cytology specimens.

5.2 Materials and methods

5.2.1 Data analysis

Gene expression microarray data from the previous chapter was analyzed. Instead of focussing on the duplicate samples, all data from the MFPE epithelial layers was incorporated into the analysis. The frozen sample data will not be considered here as there was only one sample and it had considerably wider scatter than the molecular fixative data. Data manipulation

was performed in Microsoft Office Excel (Redmond, WA, USA) for spreadsheet functions and STATISTICA (StatSoft Inc., Tulsa, OK, USA) for statistical analysis.

For each microdissected region, the top and bottom epithelial layers were compared by constructing M-A plots. The plots were normalized by linear regression according to the protocol described in Section 4.2.6. The two thresholds defined in Section 4.3.4 for molecular fixative data were then overlaid on each plot. Data points lying outside the thresholds were considered to represent probes for differentially expressed genes. For each probe, the number of high-grade (CIN II or worse) and low-grade (normal or CIN I) samples for which the gene was differentially expressed was counted. From these counts, different filters could be set up to generate gene lists matching defined criteria.

Using each of the two thresholds, probes which were overexpressed in the top layers of at least half (three) of the high-grade samples and less than half (i.e., at most two) of the low-grades were identified. As one of the ultimate goals of this analysis is to discover novel biomarkers for cervical intraepithelial neoplasia, this comparison might yield markers that are overexpressed in the upper layer of high-grade lesions. Cytology specimens are preferentially sampled from the upper half of the epithelium, so targets in the upper half are of particular interest.

The layers of normal cervical epithelium have distinct morphologies and these differences tend to disappear as a lesion progresses toward malignancy. To study basal layer markers that might also display this pattern, probes were identified that showed overexpression in the bottom layers of at least half of the low-grades but less than half of the high-grades. This would include markers that are no longer differentially expressed in high-grades because they are highly expressed in both top and bottom layers.

One can also look for negative markers of high-grade dysplasia that show decreased expression in the upper layer in high-grade lesions. To find these, probes with overexpression in the top layers of at least half of low-grades but not more than half of high-grades were selected.

An alternative approach to analyzing the data is to compare the data from each layer and grade in aggregate. This method sacrifices some of the advantages of the pair-wise comparisons of top and bottom samples from the same lesion with respect to insensitivity to patient variability, but gains the ability to detect trends across the data that might not be apparent in the pair-wise analysis. To directly compare between the upper layers of high- and low-grade lesions, a t-test was performed for each probe, using M-A normalized data. The top and bottom layer data implied by the M-A normalized data can be calculated by undoing the log transforms with the following formulae: $T_{norm} = 2^{(A+M_{norm})/2}$ and $B_{norm} = 2^{(A-M_{norm})/2}$. The probes were then ranked by increasing P-value and sorted according to which layer showed higher expression.

5.2.2 Human Protein Atlas

The most promising targets will be validated by immunohistochemistry. However, each of the screens described in the previous section returns quite a few candidates, which would translate into a lot of antibodies and a lot of validation testing. One way to both narrow down the target list and to do a preliminary round of validation is to use the Human Protein Atlas (HPA, www.proteinatlas.org), an online compilation of IHC staining of thousands of antibodies across various types of human tissue (317, 318). The present analysis uses Version 9.0 of HPA, released on November 11, 2011. It contains data and images of staining for 15598 antibodies targeting 12238 genes. Unfortunately for the question being examined here, HPA only includes IHC results for normal and cancer samples from the cervix on only a few cases (typically 1-3 for normal and a few more for cancers) for each of these genes. However, it provides a very valuable independent check going forward for the differentially expressed genes found in this study.

A search was performed on HPA for each of the genes identified in the screens and the images of the stained normal and cancerous cervical tissue cores were viewed. Those for which the HPA images agreed with the microarray data and appeared to show the greatest contrast in top layer staining between normal and cancer (i.e., those most likely to be suitable as a cytology biomarker) were flagged for further validation. Antibodies for a subset of these were then purchased to complete IHC validation.

5.2.3 Validation samples

Tissue excised from patients having undergone loop electrosurgical excision procedures (LEEP) were fixed in formalin and embedded in paraffin according to standard clinical protocols. Each LEEP specimen was a ring of tissue too large for a single paraffin block, so it was cut open and divided sequentially over multiple blocks, typically 6 or 7. Each block generally contained two adjacent tissue pieces and was assessed by a licensed surgical pathologist (Dr. Dirk van Niekerk) by examining representative sections that had been stained with hematoxylin and eosin (H&E). These slides were digitized on the Pannoramic MIDI system (3DHistech, Budapest, Hungary) so the pathologist could circle precisely the regions of abnormality. Sections adjacent to the diagnostic H&E sections were cut at 4 μ m on to glass slides and left unstained for use in the subsequent validation experiments. Table 5.1 presents a summary of the LEEP specimens used in this validation analysis. The samples were collected from patients with various grades of cervical dysplasia and many LEEP samples had regions of different histopathological diagnoses identified, allowing staining of different pathological grades to be compared against a more consistent genetic background.

Case	H&E slide	CIN grades present	Highest grade in patient	IHC staining
0013	A2-1	III	CIN III	CRNN
0018	A1-1	II/III	CIN III	CLDN1, CRNN
0018	A1-2	II/III	CIN III	IFITM3, KLK7
0018	A3-1	I	CIN III	CLDN1, CRNN
0018	A4-1	I	CIN III	IFITM3, KLK7
0019	A2-1	I/II/III	CIN III	CLDN1, IFITM3
0019	A7-1	I/III	CIN III	CRNN, KLK7
0028	A1-1	I/II	CIN III	CRNN
0028	A1-2	I/II	CIN III	IFITM3
0028	A2-3	II	CIN III	CRNN, KLK7
0030	A3-1	II/III	CIN III	CLDN1, CRNN
0030	A3-3	II/III	CIN III	IFITM3, KLK7
0037	A4-1	III	CIN III	CLDN1, IFITM3
0037	A5-1	II	CIN III	CLDN1, IFITM3
0037	A6-1	II/III	CIN III	CRNN, KLK7
0047	A2-2	I	CIN III	CLDN1, CRNN
0047	A2-3	I	CIN III	IFITM3, KLK7
0047	A5-1	I/III	CIN III	CLDN1, CRNN
0047	A5-2	I/III	CIN III	IFITM3, KLK7
0050	A2-2	I	CIN I	IFITM3
0052	A1-1	II	CIN II	CLDN1, CRNN
0052	A1-2	II	CIN II	IFITM3, KLK7

Table 5.1: Summary table of all the LEEP specimens used in the validation of microarray targets by IHC. All grades refer to CIN grades. Each block may have multiple regions present with different grades, as indicated. All blocks have regions of normal pathology.

In addition to the LEEP specimens, cervical biopsies from a previous study (185, 194, 241) and human tissue and tissue microarray slides from US Biomax (Rockville, MD, USA) and Pantomics (Richmond, CA, USA) were used to titrate antibodies.

5.2.4 Immunohistochemistry

Immunohistochemistry was performed according to a standard protocol similar to the one in Chapter 3, with modifications to account for the different nature of the paraffin embedded

material used here. For each antibody, various antigen retrieval methods and antibody dilutions were first tested on control tissues for optimization.

Concentrate buffer solutions for antigen retrieval (pH 6 and 9) were obtained from Vector Laboratories (Burlingame, CA, USA) and used at 1:100 dilution. Bovine serum albumin (BSA) and SIGMAFAST™ 3,3'-diaminobenzidine (DAB) chromogen tablets were obtained from Sigma-Aldrich Canada (Oakville, ON, Canada). Serum-free protein block and EnVision+ HRP-labelled polymer were purchased from Dako Canada (Mississauga, ON, Canada). Both anti-mouse and anti-rabbit versions of the EnVision+ reagent were used, depending on the primary antibody being tested. All antibodies were diluted in 1% BSA in phosphate-buffered saline (PBS) just prior to use.

Mouse anti-connexin 26 (CX-1E8), rabbit anti-connexin 26 (UM214), and rabbit anti-claudin-1 (JAY.8) antibodies were obtained from Life Technologies (Burlington, ON, Canada). Rabbit anti-kallikrein 7 antibody (ab40953) was obtained from Abcam (Cambridge, MA, USA). Rabbit anti-KLK7 (HPA018994) and rabbit anti-COL16A1 (HPA027235) antibodies were obtained from Sigma-Aldrich. Rabbit anti-TMEM45B (NBP1-88686) antibody was obtained from Novus Canada (Oakville, ON, Canada). Mouse anti-IFITM3 (4C8-1B10) and rabbit anti-cornulin (SZ1229) antibodies were obtained from Cedarlane (Burlington, ON, Canada). Mouse anti-stathmin (sc-48362) and rabbit anti-stathmin (sc-20796) antibodies were obtained from Santa Cruz Biotechnology (Santa Cruz, CA, USA). Rabbit anti-stathmin (3352S) antibody was obtained from New England Biolabs (Pickering, ON, Canada).

Slides were baked on a slide warmer (GCA/Precision Scientific, Chicago, IL, USA) held at around 55°C (so that the paraffin just melted) for at least 30 minutes before deparaffinization. The slides were then immersed in 3 changes of xylenes, at least 10 minutes each, followed by rehydration through graded alcohols and finally deionized water. Antigen retrieval was performed

using the microwave method, followed by cooling for 20 minutes. For each antibody, different retrieval buffers and times were tried (in order of increasing antigen retrieval strength): no retrieval (immersion in PBS for at least 30 minutes), 10.5 minutes in pH 6 citrate buffer, 12 minutes in pH 9 buffer, or 22.5 minutes in pH 9 buffer. Blocking steps were 20 minutes with 3% v/v H₂O₂ in methanol for endogenous peroxidase, 5 minutes with 1% BSA and 0.1% Triton® X-100 in PBS for permeabilization, and 60 minutes with protein block for non-specific binding. Slides were incubated with primary antibody for one hour at room temperature. A 30-minute incubation with the secondary antibody (polymer linker) followed, which was visualized via a 7-minute incubation with the DAB chromogen solution. After a thorough rinse, slides were dehydrated through graded alcohols, cleared in xylene, and coverslipped with Cytoseal™ mounting medium (Fisher Scientific Canada, Ottawa, ON, Canada).

All stained LEEP slides were digitized by scanning them into the Pannoramic MIDI system.

5.2.5 Data analysis of immunohistochemistry

Digital images of the IHC stained slides were compared with the corresponding H&E reference section. The pathologist's annotations were then transcribed on to the IHC images. All regions of abnormality were assessed separately as well as at least one representative region of normal epithelium per tissue piece. The top and bottom halves of the epithelium were scored separately, except where noted below.

Scoring of IHC staining was performed semi-quantitatively on a 0-3 scale, with 0 indicating no staining and 3 being intense staining. Each region was assessed visually by estimating both the proportion of stained cells and the intensity of the stain. In most cases where staining was present, most of the cells in the region were stained, but staining intensity varied

from one region to the next and an estimate of this intensity was the deciding factor in assigning a staining score.

In some cases, different parts of the same tissue piece exhibited different staining patterns even though they were assigned the same histopathological grade by the pathologist. In these cases, each region was assessed separately and then all the IHC scores from the same histopathological grade in the tissue piece were averaged together. Hence, for each slide, there can be up to two different scores per layer per histopathological grade (i.e., one per tissue piece). A minimum of 5 regions of each grade were assessed for each antibody.

A non-parametric test (Kruskal-Wallis), which is only dependent on the ordering of the scores and not on the linearity of the scoring system used, was used to evaluate the IHC results. Kruskal-Wallis tests were performed to find differences in staining between high-grades (CIN II and CIN III) and low-grades (normal and CIN I).

5.3 Results

Comparing the top and bottom layers of each case, most genes were not differentially expressed. There were, however, some general trends in the number of differentially expressed genes found when performing such comparisons. Figure 5.1 shows representative plots of a CIN I case and a CIN III case. There are far more differentially expressed genes in the CIN I. Generally, there were also more genes overexpressed in the top than overexpressed in the bottom. We can quantify these trends by counting the number of probes that were differentially expressed in at least half (three) of the high-grade or low-grade samples. Using the narrow thresholds combined with the minimum ratio and intensity criteria described in Section 4.3.5, differential expression was found for 166 probes in the low-grades compared to only 142 in the high-grades. Keeping in

mind that these are counts and that there was one fewer low-grade sample, this difference would have been even greater if there were the same number of high- and low-grade samples. In the low-grades, 130 probes were overexpressed in the top layer and 36 were overexpressed in the bottom layer. This trend is mirrored in the high-grades, with 105 overexpressed in the top and 37 overexpressed in the bottom. Using the wide thresholds, there are far fewer differentially expressed genes, but the difference between numbers of overexpressed genes in the top and bottom layers is still apparent.

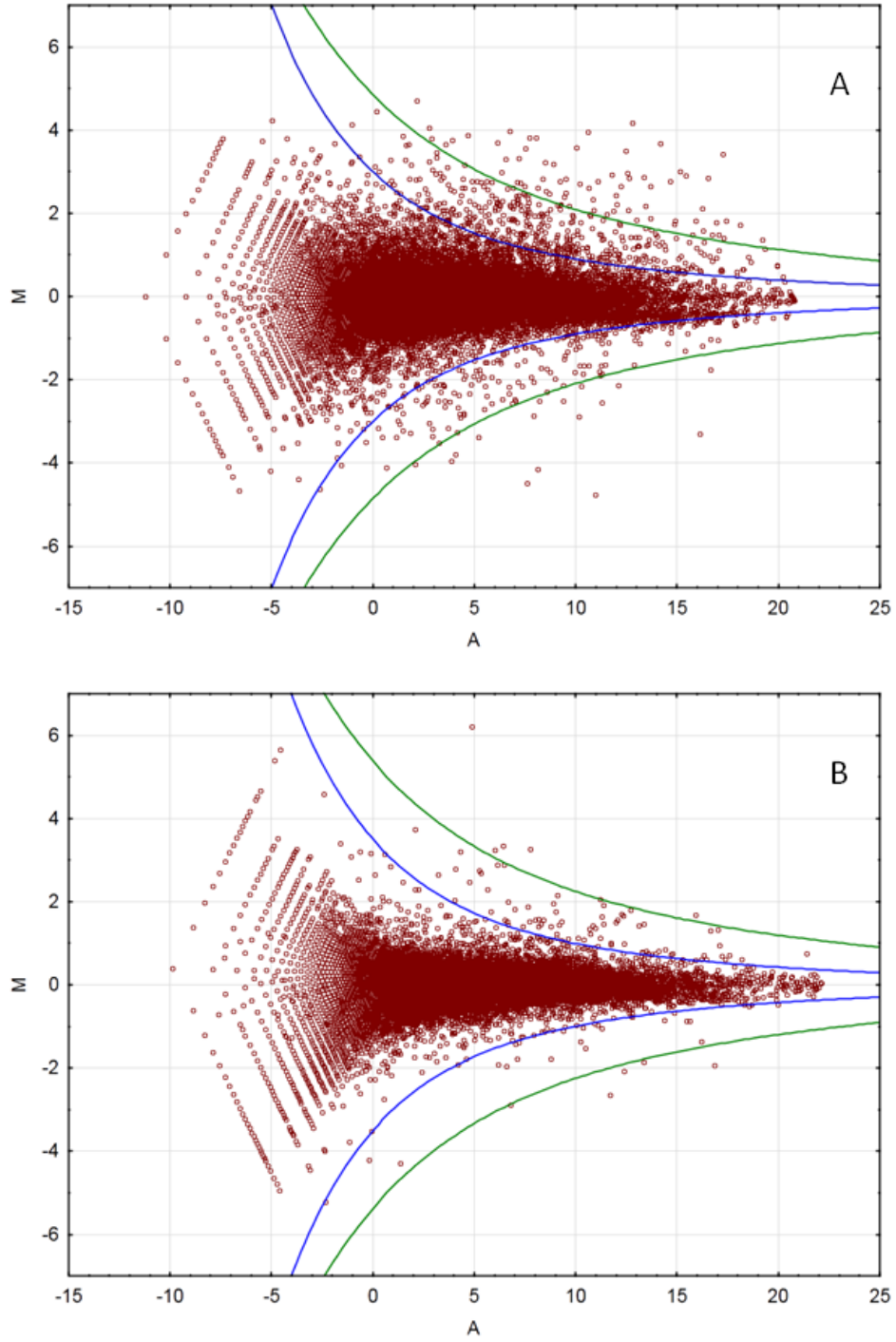


Figure 5.1: M-A plots of representative CIN I (0030) (**A**) and CIN III (0053B) (**B**) cases. Overlaid on each are the narrow (blue) and wide (green) thresholds from Section 4.3.4.

5.3.1 Target lists

The different filtering criteria described in the Materials and methods section each yielded different lists of candidate targets. 35 probes were overexpressed in the top layers of at least half (three) of the high-grade samples and less than half (i.e., at most two) of the low-grades, when using the narrow thresholds with the minimum ratio and intensity criteria. When the wide thresholds are used, only 9 probes were found. These results are summarized in Table 5.2. In all of the target lists, the highlighted rows indicate targets for which antibodies were purchased (Section 5.2.4) for the purpose of validation by IHC on LEEPs.

Table 5.2A: Narrow thresholds

Gene Symbol	Agilent Probe ID	Top > Bottom		Top < Bottom		Description	Human Protein Atlas	
		CIN2/3	Normal/ CIN1	CIN2/3	Normal/ CIN1		Normal	Cancer
GJB2	A_23_P204941	5	0	0	0	Gap junction protein, beta 2	Confined to basal layer	Strong staining in many cancer cases
GJB2	A_23_P204947	5	1	0	0	Gap junction protein, beta 2	Confined to basal layer	Strong staining in many cancer cases
SCGB1D2	A_23_P150555	4	0	0	0	Secretoglobin, family 1D, member 2 (SCGB1D2)	Weak staining in lower 2/3	Weak staining across cancer cases
TFF3	A_23_P257296	4	0	0	1	Trefoil factor 3 (intestinal) (TFF3)	Negative	Negative
TFF3	A_23_P393099	4	1	0	1	Trefoil factor 3 (intestinal) (TFF3)	Negative	Negative
CDH26	A_23_P502957	3	0	0	0	Cadherin-like 26 (CDH26)	Cadherin 26: Weak staining confined to basal layer and stroma	Cadherin 26: Strong staining in some cancers
CXCL1	A_23_P7144	3	0	0	0	Chemokine (C-X-C motif) ligand 1 (CXCL1)	No data	No data
SCGB1D1	A_23_P127781	3	0	0	0	Secretoglobin, family 1D, member 1 (SCGB1D1)	No data	No data
TFF1	A_23_P68759	3	0	0	0	Trefoil factor 1 (TFF1)	One antibody stains top layer, the other doesn't stain	Weak if any staining
KRT24	A_23_P4387	4	2	0	0	Keratin 24 (KRT24)	Staining across all layers in normal	Staining across all layers in cancer
PIM1	A_23_P345118	4	2	0	0	Pim-1 oncogene (PIM1)	Weak diffuse staining, primarily in lower layers	Extensive weak staining, but mainly in stroma
PLAC8	A_23_P81219	4	2	0	0	Placenta-specific 8 (PLAC8)	Confined to glands and some stromal cells	Some cancer cases staining throughout, but in others, confined to rare infiltrating cells
SLPI	A_23_P91230	4	2	0	0	Secretory leukocyte peptidase inhibitor (SLPI)	Weak staining in middle third and strong staining in glands	Pockets of staining in some cancers
SLPI	A_24_P190472	4	2	0	0	Secretory leukocyte peptidase inhibitor (SLPI)	Weak staining in middle third and strong staining in glands	Pockets of staining in some cancers
SPRR3	A_23_P62709	4	2	0	0	Small proline-rich protein 3 (SPRR3)	No data	No data
APOL1	A_24_P87931	3	1	0	0	Apolipoprotein L, 1 (APOL1)	Apolipoprotein L, 1: Some weak basal layer staining	Apolipoprotein L, 1: Some staining
ENST00000304963	A_23_P418031	3	1	0	0	Novel protein (Fragment)	IFFO2: Diffuse staining throughout lower 2/3 and stroma	IFFO2: Extensive staining
HBA1	A_23_P37856	3	1	0	0	Hemoglobin, alpha 1 (HBA1)	Blood marker: Staining confined to topmost layer; Same antibodies as HBA2	Some cancers show extensive staining but most negative; Same antibodies as HBA2

Table 5.2A continued

Gene Symbol	Agilent Probe ID	Top > Bottom		Top < Bottom		Description	Human Protein Atlas	
		CIN2/3	Normal/ CIN1	CIN2/3	Normal/ CIN1		Normal	Cancer
HBA2	A_23_P26457	3	1	0	0	Hemoglobin, alpha 2 (HBA2)	Blood marker: Staining confined to topmost layer; Same antibodies as HBA1	Some cancers show extensive staining but most negative; Same antibodies as HBA1
HS3ST1	A_23_P121657	3	1	0	0	Heparan sulfate (glucosamine) 3-O-sulfotransferase 1 (HS3ST1)	No data on normal squamous	No staining
PIGR	A_24_P844984	3	1	0	0	Polymeric immunoglobulin receptor (PIGR)	Very weak diffuse staining across all layers	Negative in squamous cell carcinoma, except for rare infiltrating positive cells
UNQ467	A_23_P90453	3	1	0	0	KIPV467 (UNQ467)	KRTDAP: No data	KRTDAP: No data
VEGFA	A_23_P70398	3	1	0	0	Vascular endothelial growth factor A (VEGFA)	Strong staining across all layers	Strong staining across all layers
A_32_P218707	A_32_P218707	3	2	0	0	Unknown	Not listed	Not listed
CD55	A_24_P188377	3	2	0	0	Decay-accelerating factor 4ab	1/3 antibodies stained diffusely, others negative with some stromal staining	Moderate-strong staining in some cancers
FLJ22662	A_23_P87709	3	2	0	0	Hypothetical protein FLJ22662 (FLJ22662)	PLBD1: Nuclear staining in lower 2/3 and stroma	PLBD1: Extensive moderate-level staining
HBB	A_23_P203558	3	2	0	0	Hemoglobin, beta (HBB)	Blood marker: Staining confined to topmost layer	Not much staining of tumour cells
HK2	A_32_P175739	3	2	0	0	Hexokinase 2 (HK2)	Weak diffuse staining across all layers	Diffuse staining
HMOX1	A_23_P120883	3	2	0	0	Heme oxygenase (decycling) 1 (HMOX1)	Weak staining confined to thin layer near top	Sporadic staining (maybe infiltrating immune cells)
KLK7	A_23_P39056	3	2	0	0	Kallikrein-related peptidase 7 (KLK7)	Mostly negative, with occasional positive cell	One antibody showed extensive strong staining while the other showed none
KRT4	A_23_P2674	3	2	0	0	Keratin 4 (KRT4)	Staining across all layers in normal	Staining across all layers in cancer
MALL	A_24_P80204	3	2	0	0	Mal, T-cell differentiation protein-like (MALL)	No data	No data
SAMD9	A_23_P355244	3	2	0	0	Sterile alpha motif domain containing 9 (SAMD9)	Staining across all layers in normal	Staining across all layers in cancer
TGM1	A_23_P65618	3	2	0	0	Transglutaminase 1 (TGM1)	Staining top 2/3	Positive in some, but negative in most
TMEM45B	A_23_P1682	3	2	0	0	Transmembrane protein 45B (TMEM45B)	Weak staining in lower 2/3	Extensive staining

Table 5.2B: Wide thresholds

Gene Symbol	Agilent Probe ID	Top > Bottom		Top < Bottom		Description	Human Protein Atlas	
		CIN2/3	Normal/ CIN1	CIN2/3	Normal/ CIN1		Normal	Cancer
CSTA	A_23_P170 233	4	1	0	0	Cystatin A (stefin A) (CSTA)	Staining throughout epithelium but weaker in basal-most layer for 1/3 antibodies, no data for other antibodies	Extensive staining for 2/3 antibodies, but for antibody with data for normal, see only staining in some sections and spotty staining in others
TFF3	A_23_P393 099	3	0	0	1	Trefoil factor 3 (intestinal) (TFF3)	Negative	Negative
DUOX2	A_23_P151 851	3	1	0	0	Dual oxidase 2 (DUOX2)	No data	No data
HBB	A_23_P203 558	3	1	0	0	Hemoglobin, beta (HBB)	Blood marker: Staining confined to topmost layer	Not much staining of tumour cells
CRCT1	A_23_P121 55	3	2	0	0	Cysteine-rich C-terminal 1 (CRCT1)	No data	No data
RHCG	A_23_P151 975	3	2	0	0	Rh family, C glycoprotein (RHCG)	No data	No data
SLPI	A_23_P912 30	3	2	0	0	Secretory leukocyte peptidase inhibitor (SLPI)	Weak staining in middle third	Pockets of staining in some cancers
SPRR2D	A_23_P116 44	3	2	0	0	Small proline-rich protein 2D (SPRR2D)	No data	No data
SPRR3	A_23_P627 09	3	2	0	0	Small proline-rich protein 3 (SPRR3)	No data	No data

Table 5.2: Target list found by selecting only those probes that were overexpressed in the top layers of at least half of the high-grade squamous intraepithelial dysplasia samples and less than half of the low-grade squamous intraepithelial dysplasias, when using the narrow thresholds combined with the minimum ratio and intensity criteria (**A**) or the wide thresholds (**B**). The counts in the middle columns indicate the number of high- or low-grade samples for which each probe was differentially expressed. The HPA images of cervical tissue for each probe are briefly summarized in the last column. Not listed means the gene was not found in the HPA database at all. No data means that the gene was listed, but the indicated data (image(s) of staining in normal or cancerous cervical squamous epithelium) was missing. The highlighted probes are the ones for which IHC validation was attempted.

Looking for genes that might be responsible for a basal-like phenotype, 20 probes were identified that showed overexpression in the bottom layers of at least half of the low-grades but less than half of the high-grades, when using the narrow thresholds with additional ratio and intensity criteria. Using the wide thresholds, only two probes were found. Table 5.3 summarizes these findings.

Table 5.3A: Narrow thresholds

Gene Symbol	Agilent Probe ID	Top > Bottom		Top < Bottom		Description	Human Protein Atlas	
		CIN2/3	Normal/ CIN1	CIN2/3	Normal/ CIN1		Normal	Cancer
ENST00000339867	A_32_P167592	0	0	1	4	Similar to Interferon-induced transmembrane protein 3 (Interferon-inducible protein 1-8U) (LOC650205)	Not listed	Not listed
IFITM3	A_23_P87545	0	0	1	4	Interferon induced transmembrane protein 3 (1-8U) (IFITM3)	Weak staining in basal-most layer and some intermediate layer; Same antibodies as IFITM2	Some strong staining, but some cases negative; Same antibodies as IFITM2
KRT14	A_23_P4335	0	0	0	3	Keratin 14 (epidermolysis bullosa simplex, Dowling-Meara, Koebner) (KRT14)	Staining across all layers	Staining across all layers
SEPP1	A_23_P121926	0	0	0	3	Selenoprotein P, plasma, 1 (SEPP1)	Staining throughout epithelium, primarily lower 2/3 and stroma	Moderate staining across many cancer cases, with staining of stroma
MT2A	A_23_P252413	0	0	2	4	Metallothionein 2A (MT2A)	No data	No data
PTRF	A_23_P394064	0	0	2	4	Polymerase I and transcript release factor (PTRF)	No data	No data
ATP1B3	A_23_P68007	0	0	1	3	ATPase, Na ⁺ /K ⁺ transporting, beta 3 polypeptide (ATP1B3)	Staining throughout, but strongest in basal layer	Staining throughout
COL27A1	A_23_P158096	0	0	1	3	cDNA FLJ11895 fis, clone HEMBA1007301, weakly similar to COLLAGEN ALPHA 1(III) CHAIN	No data	No data
IFITM2	A_24_P287043	0	0	1	3	Interferon induced transmembrane protein 2 (1-8D) (IFITM2)	Weak staining in basal-most layer and some intermediate layer; Same antibodies as IFITM3	Some strong staining, but some cases negative; Same antibodies as IFITM3
TPM2	A_23_P216501	0	0	1	3	Tropomyosin 2 (beta) (TPM2)	No data	No data
COL16A1	A_23_P160318	0	0	2	3	Collagen, type XVI, alpha 1 (COL16A1)	One antibody showed no staining, the other primarily lower 2/3 and stroma	Staining primarily labelling tumour-infiltrating cells
CXCL14	A_23_P213745	0	0	2	3	Chemokine (C-X-C motif) ligand 14 (CXCL14)	Staining middle third	Weak and diffuse staining
ENST00000285605	A_23_P138725	0	0	2	3	MARVEL domain-containing protein 1	MARVELD1: No data	MARVELD1: No data
FBLN1	A_23_P211631	0	0	2	3	Fibulin 1 (FBLN1)	Staining in lower half and stroma	Staining present, but primarily in stromal cells
HTRA1	A_23_P97990	0	0	2	3	HtrA serine peptidase 1 (HTRA1)	No data on normal squamous, but glands positive	Moderate staining

Table 5.3A continued

Gene Symbol	Agilent Probe ID	Top > Bottom		Top < Bottom		Description	Human Protein Atlas	
		CIN2/3	Normal/ CIN1	CIN2/3	Normal/ CIN1		Normal	Cancer
LAMB2	A_23_P21 382	0	0	2	3	Laminin, beta 2 (laminin S) (LAMB2)	Stains basement membrane and stroma, one antibody additionally stains top layer	Weak staining; Might be marker of adenocarcinoma as normal glands negative
MT1H	A_23_P41 4343	0	0	2	3	Metallothionein 1H (MT1H)	Staining throughout epithelium, primarily lower 2/3 and stroma	Weak to moderate staining throughout
MT1X	A_23_P30 3242	0	0	2	3	Metallothionein 1X (MT1X)	No data	No data
MT2A	A_23_P10 6844	0	0	2	3	Metallothionein 2A (MT2A)	No data	No data
SPON2	A_23_P12 1533	0	0	2	3	Spondin 2, extracellular matrix protein (SPON2)	No data	No data

Table 5.3B: Wide thresholds

Gene Symbol	Agilent Probe ID	Top > Bottom		Top < Bottom		Description	Human Protein Atlas	
		CIN2/3	Normal/ CIN1	CIN2/3	Normal/ CIN1		Normal	Cancer
CXCL14	A_23_P21 3745	0	0	1	3	Chemokine (C-X-C motif) ligand 14 (CXCL14)	Staining middle third	Weak and diffuse staining
ENST00000390539	A_24_P16 9873	0	0	2	3	Immunoglobulin heavy chain C gene segment	Not listed	Not listed

Table 5.3: Target list found by selecting only those probes that were overexpressed in the bottom layers of at least half of the low-grade squamous intraepithelial dysplasias but less than half of the high-grade squamous intraepithelial dysplasias, when using the narrow thresholds with additional ratio and intensity criteria (A) or the wide thresholds (B). The counts in the middle columns indicate the number of high- or low-grade samples for which each probe was differentially expressed. The HPA images of cervical tissue for each probe are briefly summarized in the last column. Not listed means the gene was not found in the HPA database at all. No data means that the gene was listed, but the indicated data (image(s) of staining in normal or cancerous cervical squamous epithelium) was missing. The highlighted probes are the ones for which IHC validation was attempted.

An attempt was also made to find a negative marker for high-grade lesions (i.e., one that is expressed in the upper layers of low-grade but not high-grade regions). Using the narrow thresholds with the extra ratio and intensity criteria, 60 probes were identified showing overexpression in the top layers of at least half of low-grades but not more than half of high-grades. 9 probes were found using the wide thresholds. Table 5.4 presents a summary of this analysis.

Table 5.4A: Narrow thresholds

Gene Symbol	Agilent Probe ID	Top > Bottom		Top < Bottom		Description	Human Protein Atlas	
		CIN2/3	Normal/ CIN1	CIN2/3	Normal/ CIN1		Normal	Cancer
A_32_P113646	A_32_P113646	1	4	0	0	Unknown	Not listed	Not listed
BC037919	A_32_P47538	0	3	0	0	cDNA clone IMAGE:5278089	Not listed	Not listed
C8orf73	A_23_P369634	0	3	0	0	Clone pp7882 unknown mRNA	No data	No data
C9orf58	A_23_P392384	0	3	0	0	Chromosome 9 open reading frame 58 (C9orf58), transcript variant 1	AIF1L: No data	AIF1L: No data
CTNNA1	A_24_P80633	0	3	0	0	Catenin (cadherin-associated protein), alpha 1	Weak staining in epithelium, mainly in top half	Stains some cancers, primarily adenocarcinomas
ERO1L	A_24_P407311	0	3	0	0	ERO1-like (S. cerevisiae) (ERO1L)	Staining top half, with antibody-dependent specificity and intensity	Staining in many cancers
FTH1	A_32_P342064	0	3	1	0	Ferritin, heavy polypeptide 1 (FTH1)	Staining confined to basal-most layer	Extensive staining in many cancers
FTH1	A_32_P820503	0	3	1	0	Ferritin, heavy polypeptide 1 (FTH1)	Staining confined to basal-most layer	Extensive staining in many cancers
GNE	A_23_P216489	0	3	0	0	Glucosamine (UDP-N-acetyl)-2-epimerase/N-acetylmannosamine kinase (GNE)	Weak staining throughout, but some cases show preferential staining in upper third	Weak or faint staining throughout
GPRC5D	A_23_P105691	0	3	0	0	G protein-coupled receptor, family C, group 5, member D (GPRC5D)	No data	No data
LIPH	A_23_P84219	0	3	0	0	Lipase, member H (LIPH)	No data	No data
NAPRT1	A_23_P258312	0	3	0	0	Nicotinate phosphoribosyltransferase domain containing 1 (NAPRT1)	Strongest staining in basal layer, but one case completely negative	Weak to no staining in many cancers, but strong staining in some
ORMDL2	A_23_P87500	0	3	0	0	ORM1-like 2 (S. cerevisiae) (ORMDL2)	No data	No data
RIPK4	A_24_P125871	0	3	0	0	Receptor-interacting serine-threonine kinase 4 (RIPK4)	Negative	Mostly negative
TPM4	A_32_P28284	0	3	0	0	Tropomyosin 4 (TPM4)	Weak diffuse staining in lower 2/3 and stroma	Weak diffuse staining
ZNF12	A_24_P337774	0	3	0	0	Zinc finger protein 12 (ZNF12)	No data	No data
CST6	A_23_P146946	2	4	0	0	Cystatin E/M (CST6)	Weak staining throughout epithelium, mainly lower 2/3 and stroma	Weak staining in many cancers

Table 5.4A continued

Gene Symbol	Agilent Probe ID	Top > Bottom		Top < Bottom		Description	Human Protein Atlas	
		CIN2/3	Normal/ CIN1	CIN2/3	Normal/ CIN1		Normal	Cancer
EMP1	A_23_P76 488	2	4	0	0	Epithelial membrane protein 1	No data	No data
ANXA9	A_23_P10 3617	1	3	0	0	Annexin A9 (ANXA9)	Staining through most of epithelium, with one antibody staining strongest in bottom third while another strongest in middle third	Weak staining in many cancers
C2orf54	A_23_P60 990	1	3	0	0	Hypothetical protein FLJ22671 (FLJ22671)	No data	No data
C9orf58	A_23_P94 380	1	3	0	0	Chromosome 9 open reading frame 58 (C9orf58), transcript variant 1	AIF1L: No data	AIF1L: No data
CD55	A_23_P37 4862	1	3	0	0	Decay-accelerating factor 4ab	1/3 antibodies stained diffusely, others negative with some stromal staining	Moderate-strong staining in some cancers
CEACAM1	A_23_P55 738	1	3	0	0	Carcinoembryonic antigen-related cell adhesion molecule 1 (biliary glycoprotein) (CEACAM1)	Negative; One antibody is pan-CEACAM	Negative with pan-CEACAM antibody, a few positive cases with other antibody
CEACAM7	A_24_P22 8302	1	3	0	0	Carcinoembryonic antigen-related cell adhesion molecule 7 (CEACAM7)	No data	No data
DHRS9	A_23_P56 559	1	3	0	0	Dehydrogenase/reductase (SDR family) member 9 (DHRS9)	Weak staining throughout epithelium, mainly lower 2/3 in one antibody and basal-most layer in the other	Weak staining in some cancers
FAM129B	A_23_P41 7404	1	3	0	0	MEG3 (MEG3)	Weak staining throughout	Mostly negative for 2/4 antibodies, weak to moderate staining for other antibodies
FTH1	A_32_P11 1565	1	3	1	0	Ferritin, heavy polypeptide 1 (FTH1)	Staining confined to basal-most layer	Extensive staining in many cancers
GLTP	A_23_P25 336	1	3	0	0	Glycolipid transfer protein (GLTP)	No data	No data
HPGD	A_23_P21 3050	1	3	0	0	Hydroxyprostaglandin dehydrogenase 15-(NAD) (HPGD)	No staining	Mostly negative with some positive cells in a few cases
KLK10	A_24_P39 9490	1	3	0	0	Kallikrein-related peptidase 10 (KLK10)	Top half preferentially stained, but one case shows staining of basal-most layer	Weak or no staining, especially with antibody that stains normal basal layer
PRSS27	A_23_P10 6806	1	3	0	0	Protease, serine 27 (PRSS27)	Staining mainly upper third	Some staining in many cancers

Table 5.4A continued

Gene Symbol	Agilent Probe ID	Top > Bottom		Top < Bottom		Description	Human Protein Atlas	
		CIN2/3	Normal/ CIN1	CIN2/3	Normal/ CIN1		Normal	Cancer
RAB11FIP1	A_23_P39 1198	1	3	0	0	RAB11 family interacting protein 1 (class I) (RAB11FIP1)	Basal half staining in normal for 2/4 antibodies, no staining throughout for 1/4 antibodies, no normal squamous data for last antibody	Staining throughout
SH3GL1	A_24_P13 9094	1	3	0	0	SH3-domain GRB2-like 1 (SH3GL1)	Strong staining except for top layer	Staining throughout, but intensities not consistent
THC2682885	A_24_P69 1826	1	3	0	0	Q6BEA3_RAT (Q6BEA3)	Not listed	Not listed
THC2714090	A_32_P12 1140	1	3	0	0	WDNM1 homolog, partial (33%)	Not listed	Not listed
TMPRSS11D	A_23_P14 4417	1	3	0	0	Unknown	Not listed	Not listed
TPM4	A_32_P21 993	1	3	0	0	Transmembrane protease, serine 11D (TMPRSS11D)	No data	No data
VPS25	A_23_P66 599	1	3	0	0	Tropomyosin 4 (TPM4)	Weak diffuse staining in lower 2/3 and stroma	Weak diffuse staining
AA593970	A_32_P30 898	2	3	0	0	Vacuolar protein sorting 25 homolog (S. cerevisiae) (VPS25)	No data	No data
ANXA1	A_23_P94 501	2	3	0	0	AA593970 nn01c05.s1 NCI_CGAP_Co9 cDNA clone IMAGE:1076456 3'	Not listed	Not listed
BU943730	A_24_P98 948	2	3	0	0	Annexin A1 (ANXA1)	Staining across all layers	Staining across all layers
CEACAM1	A_23_P43 4118	2	3	0	0	AGENCOURT_10544326 NIH_MGC_126 cDNA clone IMAGE:6723988 5'	Not listed	Not listed
IL1RN	A_23_P20 9995	2	3	0	0	Carcinoembryonic antigen-related cell adhesion molecule 1 (biliary glycoprotein) (CEACAM1)	Negative; One antibody is pan-CEACAM	Negative with pan-CEACAM antibody, a few positive cases with other antibody
ITPKC	A_24_P20 2567	2	3	0	0	Interleukin 1 receptor antagonist	Staining throughout, but 1/2 antibodies doesn't stain basal-most layer	Staining in most cancers, especially with antibody that stains normal basal layer
KLK12	A_23_P50 0010	2	3	0	0	Inositol 1,4,5-trisphosphate 3-kinase C (ITPKC)	No data	No data
KRTAP8-1	A_23_P31 2932	2	3	0	0	Kallikrein-related peptidase 12 (KLK12)	No data on normal squamous	Negative
LOC146439	A_24_P27 3647	2	3	0	0	Keratin associated protein 8-1 (KRTAP8-1)	Staining mainly lower 2/3 and stroma	Staining throughout
						mRNA; cDNA DKFZp666L166 (from clone DKFZp666L166)	Not listed	Not listed

Table 5.4A continued

Gene Symbol	Agilent Probe ID	Top > Bottom		Top < Bottom		Description	Human Protein Atlas	
		CIN2/3	Normal/ CIN1	CIN2/3	Normal/ CIN1		Normal	Cancer
LOC147645	A_23_P10 1246	2	3	0	0	Clone IMAGE:4401841	Not listed	Not listed
NDRG2	A_23_P37 205	2	3	0	0	NDRG family member 2 (NDRG2)	Staining throughout	Weak or no staining throughout
PBEF1	A_32_P79 396	2	3	0	0	Pre-B-cell colony enhancing factor 1 (PBEF1)	NAMPT: No data for normal squamous, but staining in glands	NAMPT: Staining throughout
RAB11FIP1	A_23_P31 873	2	3	0	0	RAB11 family interacting protein 1 (class I) (RAB11FIP1)	Basal half staining in normal for 2/4 antibodies, no staining throughout for 1/4 antibodies, no normal squamous data for last antibody	Staining throughout
SERPINB1	A_23_P21 4330	2	3	0	0	Serpin peptidase inhibitor, clade B (ovalbumin), member 1 (SERPINB1)	Weak staining mainly in upper 2/3	Weak staining
SERPINB2	A_23_P15 3185	2	3	0	0	Serpin peptidase inhibitor, clade B (ovalbumin), member 2 (SERPINB2)	Negative	Negative with weak staining in parts, except a few more strongly staining cases
SLC16A3	A_23_P15 8725	2	3	0	0	Solute carrier family 16, member 3 (monocarboxylic acid transporter 4) (SLC16A3)	No staining for one antibody, staining throughout but primarily lower 2/3 for the other antibody	Weak staining with one antibody, extensive strong staining with the other
SLC47A1	A_23_P20 7221	2	3	0	0	Hypothetical protein FLJ10847 (FLJ10847)	Negative	Negative, but with positive stromal cells in some cases
SPRR1A	A_23_P34 8208	2	3	0	0	Small proline-rich protein 1A (SPRR1A)	Strong staining throughout epithelium, but a bit weaker in basal layer	Extensive strong staining
SPRR1B	A_23_P15 9406	2	3	0	0	Small proline-rich protein 1B (cornifin) (SPRR1B)	No data	No data
TMPRSS11B	A_23_P81 190	2	3	0	0	Transmembrane protease, serine 11B (TMPRSS11B)	Incomplete normal squamous	Staining in most cancers
TMPRSS11E	A_23_P18 751	2	3	0	0	Transmembrane protease, serine 11E (TMPRSS11E)	No data	No data
TTC9	A_23_P14 508	2	3	0	0	Tetratricopeptide repeat protein 9 (TPR repeat protein 9)	No data	No data

Table 5.4B: Wide thresholds

Gene Symbol	Agilent Probe ID	Top > Bottom		Top < Bottom		Description	Human Protein Atlas	
		CIN2/3	Normal/ CIN1	CIN2/3	Normal/ CIN1		Normal	Cancer
C15orf48	A_23_P26024	1	4	0	0	Chromosome 15 open reading frame 48 (C15orf48)	NMES1: No data	NMES1: No data
CEACAM5	A_23_P153301	0	3	0	0	Carcinoembryonic antigen-related cell adhesion molecule 5 (CEACAM5)	Some staining for 2/4 antibodies above bottom third, pan-CEACAM antibody is negative, no data for last antibody	Staining in many cancers for 3/4 antibodies, pan-CEACAM antibody negative
CEACAM7	A_23_P130573	1	3	0	0	Carcinoembryonic antigen-related cell adhesion molecule 7 (CEACAM7)	No data	No data
KLK12	A_23_P500010	1	3	0	0	Kallikrein-related peptidase 12 (KLK12)	No data on normal squamous	Negative
CEACAM6	A_23_P421483	2	3	0	0	Carcinoembryonic antigen-related cell adhesion molecule 6 (non-specific cross reacting antigen) (CEACAM6)	Data only on 1/2 antibodies, showing only rare stromal cell staining and no epithelial staining, this antibody is pan-CEACAM	Weak staining in some cancer; This antibody is pan-CEACAM
CNFN	A_23_P27473	2	3	0	0	Cornifelin (CNFN)	No data	No data
CRNN	A_23_P115202	2	3	0	0	Cornulin (CRNN)	Strong staining localized to upper 2/3	Mostly negative but with some pockets of positive staining
S100P	A_23_P58266	2	3	0	0	S100 calcium binding protein P (S100P)	Negative	Staining in some of the cancers, especially for 1/2 of the antibodies
SPRR1A	A_23_P74012	2	3	0	0	Small proline-rich protein 1A (SPRR1A)	Strong staining throughout epithelium, but a bit weaker in basal layer	Extensive strong staining

Table 5.4: Target list of negative markers for high-grade squamous intraepithelial dysplasia, showing probes that are overexpressed in the upper layers of a majority of low-grade but not high-grade squamous intraepithelial dysplasia regions, using narrow thresholds with extra ratio and intensity criteria (**A**) or wide thresholds (**B**). The counts in the middle columns indicate the number of high- or low-grade samples for which each probe was differentially expressed. The HPA images of cervical tissue for each probe are briefly summarized in the last column. Not listed means the gene was not found in the HPA database at all. No data means that the gene was listed, but the indicated data (image(s) of staining in normal or cancerous cervical squamous epithelium) was missing. The highlighted probes are the ones for which IHC validation was attempted.

T-test analysis of the top layer data generated many targets using the standard statistical cutoff of $P < 0.05$. To make this list a more manageable size, probes for which there was unusable data from any of the 9 samples were discarded. A zero or negative normalized signal intensity, such as those removed from the data set in Section 4.2.6, would qualify as an unusable data point. Additionally, as the goal is to find a novel and detectable biomarker, a minimum intensity criterion can be applied. If the mean top layer signal is required to be at least 10 for the high-grades, 27 probes show significantly higher expression in the top layers of high-grades compared to the low-grades. The reverse analysis with a minimum mean low-grade signal of 150 finds 35 probes showing significantly higher expression in the top layers of low-grades versus high-grade regions. These analyses are summarized in Table 5.5.

Table 5.5A: High-grade > Low-grade

Gene Symbol	Agilent Probe ID	High-Grade	Low-Grade	t-value	P-value	Description	Human Protein Atlas	
							Normal	Cancer
IFI16	A_23_P217866	29.723	9.864	3.6223	0.008484	Interferon, gamma-inducible protein 16 (IFI16)	Weak nuclear stain near basal layer	Weak to moderate staining in many cancers
HRB	A_23_P169529	13.838	6.132	3.4511	0.010675	HIV-1 Rev binding protein (HRB)	AGFG1: Very weak staining throughout, with slight bias toward bottom half	Weak to moderate staining
CX3CL1	A_23_P37727	10.884	3.301	3.3559	0.012152	Chemokine (C-X3-C motif) ligand 1 (CX3CL1)	Staining throughout epithelium, with slight bias toward bottom half in one antibody	Staining throughout
SOD2	A_23_P134176	16.585	8.623	3.1810	0.015470	Superoxide dismutase 2, mitochondrial	Speckled staining pattern in bottom half for 1/2 antibodies, no squamous for other antibody	Extensive staining throughout
LOC399744	A_32_P11230	28.542	21.201	3.1743	0.015613	cDNA FLJ44672 fis, clone BRACE3006553	Not listed	Not listed
INDO	A_23_P112026	10.198	1.406	3.0114	0.019622	Indoleamine-pyrrole 2,3 dioxygenase (INDO)	IDO1: 1/3 antibodies shows faint staining throughout, with strongest staining in basal layer, 1/3 antibodies shows no staining, 1/3 antibodies shows faint staining with strongest staining at top edge (probably edge effect)	Staining of some cancers (spotty at times), negative in others
ENST00000270031	A_24_P254933	26.721	16.874	2.9759	0.020633	Interferon induced transmembrane protein 3 (1-8U) (IFITM3)	IFITM3: Weak staining in basal-most layer and some intermediate layer of normal; Same antibodies as IFITM2	Some strong staining, but some cases negative; Same antibodies as IFITM2
DUT	A_24_P160874	13.651	7.265	2.9503	0.021397	dUTP pyrophosphatase (DUT), nuclear gene encoding mitochondrial protein	No data	No data
NUP50	A_32_P121303	14.527	9.258	2.9428	0.021628	Nucleoporin 50kDa (NUP50)	No data	No data
CLDN1	A_24_P165949	21.827	6.129	2.9165	0.022451	Claudin 1 (CLDN1)	Staining in bottom half	Staining throughout

Table 5.5A continued

Gene Symbol	Agilent Probe ID	High-Grade	Low-Grade	t-value	P-value	Description	Human Protein Atlas	
							Normal	Cancer
KPNA4	A_24_P11791	15.823	7.782	2.9093	0.022683	Karyopherin alpha 4 (importin alpha 3) (KPNA4)	No data	No data
IFITM1	A_23_P72737	35.676	11.490	2.7910	0.026871	Interferon induced transmembrane protein 1 (9-27) (IFITM1)	Staining primarily basal-most layer	Extensive staining in many cancers
CKS2	A_23_P71727	21.784	10.146	2.7548	0.028308	CDC28 protein kinase regulatory subunit 2 (CKS2)	Nuclear staining primarily in bottom 2/3	Staining throughout
HLA-F	A_23_P145264	20.584	10.725	2.6910	0.031042	Major histocompatibility complex, class I, F (HLA-F)	No data	No data
KIAA0430	A_23_P26674	15.976	7.739	2.6906	0.031060	KIAA0430 (KIAA0430)	Staining throughout	Staining throughout
NCOA7	A_24_P12435	11.737	4.017	2.6830	0.031404	Nuclear receptor coactivator 7 (NCOA7)	Staining favours bottom half, with intensity varying with antibody, ranging from no staining to extensive	Weak staining in some cancers, correlating with sensitivity of antibody
NUP62	A_24_P322444	32.609	13.011	2.6736	0.031833	Nucleoporin 62kDa (NUP62)	1/2 antibodies staining perinuclear in bottom third, other antibody more diffuse in bottom half	Perinuclear antibody showing strong staining throughout, other antibody showing weak staining in some cancers
THC2572108	A_24_P707543	160.064	67.066	2.6680	0.032091	C15orf21 protein, partial (75%)	Not listed	Not listed
CLDN1	A_23_P57784	18.896	3.442	2.6252	0.034151	Claudin 1 (CLDN1)	Staining in bottom half	Staining throughout
SMC4	A_23_P91900	11.268	5.948	2.6111	0.034857	cDNA FLJ11338 fis, clone PLACE1010720, highly similar to mRNA for chromosome-associated polypeptide-C	Staining in bottom 2/3, particularly for 1/2 antibodies	Staining in about half of cancers
GMPS	A_23_P21033	17.251	6.474	2.5330	0.039064	Guanine monophosphate synthetase (GMPS)	No data	No data
THC2583762	A_32_P71437	42.076	20.682	2.5160	0.040042	Brain cDNA, clone: QccE-17725, partial (28%)	Not listed	Not listed
SGK	A_23_P19673	156.864	54.381	2.5068	0.040587	Serum/glucocorticoid regulated kinase (SGK)	SGK1: Mainly lower half staining	Moderate staining

Table 5.5A continued

Gene Symbol	Agilent Probe ID	High-Grade	Low-Grade	t-value	P-value	Description	Human Protein Atlas	
							Normal	Cancer
STMN1	A_23_P200866	161.243	23.867	2.4859	0.041847	Stathmin 1/oncoprotein 18 (STMN1)	Stathmin 1: Confined to parabasal layer	Strong staining throughout
AJ227863	A_24_P925361	14.156	6.692	2.4804	0.042182	partial mRNA; ID YG39-2B	Not listed	Not listed
PTMA	A_24_P34632	262.714	118.788	2.4333	0.045203	Prothymosin, alpha (gene sequence 28) (PTMA)	Prothymosin alpha: No data	Prothymosin alpha: No data
ABCC5	A_23_P258221	17.826	6.575	2.3918	0.048041	ATP-binding cassette, sub-family C (CFTR/MRP), member 5 (ABCC5)	No data	No data

Table 5.5B Low-grade > High-grade

Gene Symbol	Agilent Probe ID	High-Grade	Low-Grade	t-value	P-value	Description	Human Protein Atlas	
							Normal	Cancer
SCC-112	A_24_P52004	117.592	372.675	-4.7679	0.002041	SCC-112 protein	PDS5A: Nuclear staining across all layers	PDS5A: Nuclear staining across all layers
AK092577	A_24_P170717	144.028	474.853	-4.5739	0.002562	cDNA FLJ35258 fis, clone PROST2004146	Not listed	Not listed
BQ365891	A_24_P926484	100.159	213.994	-4.1657	0.004212	BQ365891 CM0-GN0111-230900-569-g06 GN0111 cDNA	Not listed	Not listed
A_32_P234913	A_32_P234913	178.694	264.132	-3.6265	0.008437	Unknown	Not listed	Not listed
GIPR	A_23_P119395	265.124	681.446	-3.5482	0.009366	Gastric inhibitory polypeptide receptor (GIPR)	No data	No data
LOC647580	A_24_P625683	124.986	168.794	-3.5437	0.009422	PREDICTED: hypothetical LOC647580	Not listed	Not listed
A_24_P853410	A_24_P853410	107.731	220.461	-3.4338	0.010929	Unknown	Not listed	Not listed
BX537551	A_32_P52076	113.018	273.659	-3.2326	0.014400	mRNA; cDNA DKFZp686M0346	Not listed	Not listed
ART4	A_23_P116902	114.255	303.948	-3.1547	0.016047	ADP-ribosyltransferase 4 (Dombrock blood group) (ART4)	Weak staining, mainly middle third	Some thorough staining

Table 5.5B continued

Gene Symbol	Agilent Probe ID	High-Grade	Low-Grade	t-value	P-value	Description	Human Protein Atlas	
							Normal	Cancer
SNTA1	A_23_P57227	79.411	177.014	-3.1069	0.017157	Syntrophin, alpha 1 (dystrophin-associated protein A1, 59kDa, acidic component) (SNTA1)	No data	No data
PABPC3	A_23_P48307	163.066	363.750	-3.0341	0.019004	Poly(A) binding protein, cytoplasmic 3 (PABPC3)	No data	No data
CAST	A_23_P434352	118.272	204.333	-2.9981	0.019996	Calpastatin (CAST)	Moderate staining throughout, but mostly lower 2/3	Extensive staining
EFNB3	A_23_P433588	131.196	344.515	-2.9440	0.021589	Ephrin-B3 (EFNB3)	No data on normal squamous, although negative in glands	Negative
SSTR3	A_23_P68910	226.841	558.467	-2.9430	0.021622	Somatostatin receptor 3 (SSTR3)	No data on normal squamous, although positive in glands	Weakly staining
BX106126	A_32_P216041	71.546	412.789	-2.9425	0.021635	BX106126 Soares_testis_NHT cDNA clone IMAGp998M074160	Not listed	Not listed
GPR78	A_23_P69652	69.437	165.520	-2.9035	0.022871	G protein-coupled receptor 78 (GPR78)	No staining; Staining appears to be on endothelial cells	No staining; Staining appears to be on endothelial cells
A_24_P289884	A_24_P289884	75.415	150.526	-2.8634	0.024219	Unknown	Not listed	Not listed
NRN1	A_23_P82088	85.337	179.673	-2.8515	0.024635	Neuritin 1 (NRN1)	No data	No data
BC028232	A_24_P102456	88.184	324.486	-2.8364	0.025173	Clone IMAGE:5221276	Not listed	Not listed
LOC341412	A_24_P315326	98.032	210.433	-2.8174	0.025870	AGENCOURT_10640955 NIH_MGC_126 cDNA clone IMAGE:6723568 5'	Not listed	Not listed
CEACAM6	A_23_P421483	89.056	335.953	-2.7729	0.027581	Carcinoembryonic antigen-related cell adhesion molecule 6 (non-specific cross reacting antigen) (CEACAM6)	Data only on 1/2 antibodies for squamous, showing only rare stromal cell staining and no epithelial staining, this antibody is pan-CEACAM	Weak staining in some cancer
LOC391701	A_24_P118281	60.096	175.617	-2.7442	0.028744	PREDICTED: similar to ribosomal protein S23 (LOC391701)	Not listed	Not listed
ATP1B1	A_23_P62932	119.372	180.163	-2.6962	0.030807	ATPase, Na ⁺ /K ⁺ transporting, beta 1 polypeptide (ATP1B1)	Staining confined to basal-most layer	Extensive staining

Table 5.5B continued

Gene Symbol	Agilent Probe ID	High-Grade	Low-Grade	t-value	P-value	Description	Human Protein Atlas	
							Normal	Cancer
A_24_P752279	A_24_P752279	89.107	195.464	-2.6742	0.031806	Unknown	Not listed	Not listed
MAL	A_23_P17134	144.118	1629.454	-2.6228	0.034269	Mal, T-cell differentiation protein (MAL)	No data	No data
LCE2D	A_23_P371284	111.327	177.699	-2.5884	0.036028	Late cornified envelope 2D (LCE2D)	No data	No data
RP6-166C19.11	A_32_P207124	101.736	189.957	-2.5308	0.039188	Cancer/testis CT47 family, member 11 (CT47.11)	Not listed	Not listed
SERPINB1	A_23_P214330	74.058	339.280	-2.5140	0.040163	Serpin peptidase inhibitor, clade B (ovalbumin), member 1 (SERPINB1)	Weak staining mainly in upper 2/3	Weak staining
FAM127B	A_23_P62429	156.340	221.300	-2.4989	0.041060	Family with sequence similarity 127, member B (FAM127B)	No data	No data
EMP1	A_23_P76488	159.840	651.025	-2.4953	0.041273	Epithelial membrane protein 1	No data	No data
SPATA2L	A_23_P118086	100.203	161.677	-2.4820	0.042088	Spermatogenesis associated 2-like (SPATA2L)	No data	No data
PSCA	A_23_P71379	35.741	310.112	-2.4752	0.042508	Prostate stem cell antigen (PSCA)	No data	No data
SMARCC2	A_24_P129813	266.632	459.095	-2.4703	0.042813	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c, member 2 (SMARCC2)	Nuclear staining across all layers	Nuclear staining across all layers
THC2583971	A_32_P217523	63.412	186.532	-2.4013	0.047377	Q5PR09_MOUSE (Q5PR09) Ribosomal protein L32, partial (87%)	Not listed	Not listed
ENST00000324745	A_24_P255609	256.119	454.254	-2.3949	0.047820	mRNA for FLJ00388 protein	Not listed	Not listed

Table 5.5: Target list from t-test analysis comparing only the top layers of high- and low-grade squamous intraepithelial dysplasias. **(A)** shows probes that are overexpressed in high-grades and **(B)** shows those that are overexpressed in low-grades. Mean signal intensities are listed for high- and low-grade top layers. Probes are listed in order of increasing P-value and only those for which all samples produced non-zero, non-negative data are included. The HPA images of cervical tissue for each probe are briefly summarized in the last column. The highlighted probes are the ones for which IHC validation was attempted.

A number of the targets listed in the above tables have been found in previous studies on cervical cancer or CIN. To assist in sorting through the published data, the gene lists were queried in the online Cervical Cancer Gene Database (<http://crdd.osdd.net/raghava/ccdb/>) (319), supplemented with additional literature search on select targets. Some of these are listed in Table 5.6, divided into expression microarray studies and other studies.

Table 5.6A: Microarray studies

Study with Reference	Gene	Change in Malignant Sample (Reference)	Malignant Grade (Reference)	Layer with Overexpression (Our Data)	Agrees?
Wilting <i>et al</i> 2008 (283)	CLDN1 (SEMP1)	Increased	Cancer	T-test HG Top	Yes
	APOL1	Increased	Cancer	HG Top	Yes
Gius <i>et al</i> 2007 (289)	EMP1	Decreased	Cancer	LG Top	Yes
	IL1RN	Decreased	Cancer	LG Top	Yes
	EMP1	Decreased	Cancer	T-test LG Top	Yes
Manavi <i>et al</i> 2007 (284)	VEGFA (VEGF)	Increased	SCC	HG Top	Yes
Wong <i>et al</i> 2006 (286)	SPRR3	Decreased	Cancer	HG Top	No
	HBA2	Decreased	Cancer	HG Top	No
	VEGFA (VEGF)	Increased	Cancer	HG Top	Yes
	KLK7	Decreased	Cancer	HG Top	No
	KRT4	Decreased	Cancer	HG Top	No
	CRCT1 (c1orf42)	Decreased	Cancer	HG Top	No
	RHCG	Decreased	Cancer	HG Top	No
	EMP1	Decreased	Cancer	LG Top	Yes
	GLTP	Decreased	Cancer	LG Top	Yes
	HPGD	Decreased	Cancer	LG Top	Yes
	KLK10	Decreased	Cancer	LG Top	Yes
	TMPRSS11D (HAT)	Decreased	Cancer	LG Top	Yes
	IL1RN	Decreased	Cancer	LG Top	Yes
	KLK12	Decreased	Cancer	LG Top	Yes
	SPRR1A	Decreased	Cancer	LG Top	Yes
	TMPRSS11E (DESC1)	Decreased	Cancer	LG Top	Yes
	CRNN (c1orf10)	Decreased	Cancer	LG Top	Yes
	IFI16	Increased	Cancer	T-test HG Top	Yes
	CKS2	Increased	Cancer	T-test HG Top	Yes
	EMP1	Decreased	Cancer	T-test LG Top	Yes
Chen <i>et al</i> 2003 (293)	CLDN1	Increased	HGSIL/Cancer	T-test HG Top	Yes
	APOL1	Increased	HGSIL/Cancer	HG Top	Yes
Ahn <i>et al</i> 2004 (287)	SEPP1	Increased	Cancer	LG Bottom	Yes

Table 5.6B: Other studies

Study with Reference	Gene	Change in Malignant Sample (Reference)	Malignant Grade (Reference)	Layer with Overexpression (Our Data)	Agrees?
Hammes <i>et al</i> 2008 (320)	VEGFA (VEGF)	Increased	HGSIL	HG Top	Yes
Shadeo <i>et al</i> 2008 (308)	SPRR3	Decreased	CIN III	HG Top	No
	SPRR1A	Decreased	CIN III	LG Top	Yes
	TMPRSS11B	Decreased	CIN III	LG Top	Yes
Shadeo <i>et al</i> 2007 (309)	GJB2	Decreased	CIN III	HG Top	No
	FTH1	Increased	CIN III	LG Top	No
	KLK12	Increased	CIN III	LG Top	No
	SERPINB2	Increased	CIN III	LG Top	No
Kneller <i>et al</i> 2007 (321)	CEACAM5	Increased	CIN III	LG Top	No
	GJB2	Decreased	CIN III	HG Top	No
Narayan <i>et al</i> 2007 (322)	COL16A1	Amplified (by CGH)	Cancer	LG Bottom	Yes
Sova <i>et al</i> 2007 (323)	c15orf48 (NMES1)	Hypermethylated	Cancer	LG Top	Yes
Soufla <i>et al</i> 2005 (324)	VEGFA (VEGF)	Increased	CIN	HG Top	Yes
Xu <i>et al</i> 1999 (325)	SPRR1B (Cornifin)	Decreased	CIN	LG Top	Yes

Table 5.6: List of genes found in the present analysis that have previously been associated with cervical cancer or CIN. Published studies have been divided into (A) gene expression microarray studies and (B) other studies. Results from this study are compared with the direction of change of the malignant case relative to normal reported in the published studies. Hits in each cited study are listed in the order in which they appear in the gene lists of the present study. The second-last column references the gene list from this study in which the target was found: HG Top for Table 5.2, LG Bottom for Table 5.3, LG Top for Table 5.4, T-test for Table 5.5. In the present study, only Table 5.4 and Table 5.5B listed markers expected to show reduced expression in high-grade lesions or cancer. It is assumed that copy number gains result in increased expression while hypermethylation leads to decreased expression. HG = High-grade, LG = Low-grade, SCC = Squamous cell carcinoma, CGH = Comparative genomic hybridization.

5.3.2 Human Protein Atlas

Searches on HPA were performed for each of the probes listed in the previous section.

All the images of staining performed on normal and cancerous cervical tissue were examined visually and a brief description of the staining observed is included in Table 5.2 through Table 5.5.

5.3.3 Immunohistochemistry

Despite the age of the biopsy specimens (most were collected 8-10 years ago), they were found to be suitable for immunohistochemical analysis. All antibodies that worked on the fresher

LEEP specimens also worked at a similar dilution on the older biopsies. Antibodies that failed on the older biopsies were generally tested on at least one LEEP, which also failed. The summaries below are presented in the order in which they appear in the target list tables of Section 5.3.1.

Gap junction protein, beta 2 (GJB2), also known as connexin 26 and overexpressed in the top layers of high-grades (CIN II or CIN III) but not low-grades (normal or CIN I) (Table 5.2A), was tested using 2 different antibodies, chosen based on references in the literature (CX-1E8) (326) and HPA (UM214). However, staining patterns observed in HPA could not be reproduced. Using CX-1E8 at 1:50 dilution and UM214 at 1:10, staining was faint and primarily in the bottom half of normal epithelium. Staining was not confined to the basal-most layer as seen in HPA, but rather diffuse throughout the lower half. Few of the basal cells were actually stained. There was little difference in staining between different antigen retrieval conditions. This target was not pursued further.

Kallikrein-related peptidase 7 (KLK7), previously known as stratum corneum chymotryptic enzyme and overexpressed in the top layers of high-grades but not low-grades (Table 5.2A), was tested using two antibodies, including the one used by HPA (HPA018994). The HPA antibody failed using all antigen retrieval methods on bone marrow and cervical cancer. HPA images suggested that bone marrow should be a good positive control for KLK7. The other antibody, ab40953 from Abcam, fared better. Staining was observed in bone marrow and cervical cancer, regardless of antigen retrieval conditions. For the LEEPs, the antibody was diluted 1:35 and citrate antigen retrieval was used.

Kallikrein-related peptidase 7 is a serine protease normally found in the skin. KLK7 staining of LEEPs was generally stronger in the top layer. Staining was generally weak and pervasive, with only a weak upward trend in staining of the top layer with increasing grade (Figure 5.2) ($P = 0.01$ for the top layer, $P = 0.70$ for the bottom layer, by Kruskal-Wallis

comparing low-grades versus high-grades). Among individual cases, some LEEPs showed increased staining in higher grade regions (Figure 5.3A), but others showed little difference between even CIN III and normal (Figure 5.3C).

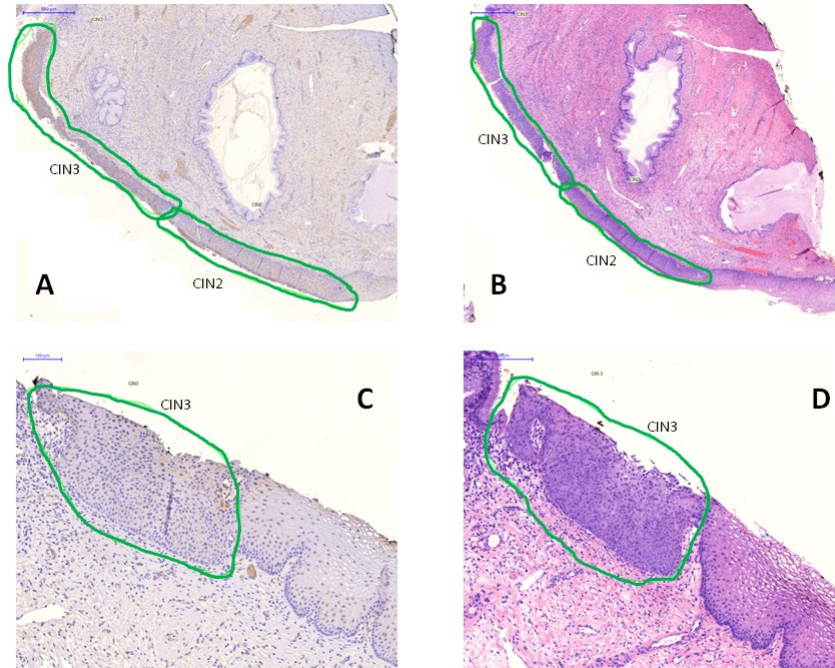


Figure 5.3: Sample LEEP sections (cases 0018 (A, B) and 0030 (C, D)) stained with KLK7 IHC (A, C), along with nearby reference H&E sections (B, D). High-grade squamous intraepithelial neoplasia regions have been circled with green, while normal epithelium is uncircled. Scale bars in the top-left of each panel are 500 μ m in length for (A) and (B), 100 μ m for (C), and 200 μ m for (D).

Transmembrane protein 45B (TMEM45B), overexpressed in the top layers of high-grades but not low-grades (Table 5.2A), was tested using only one antibody. There was very little literature to go on for this target and HPA antibodies seem to have a poor success rate. Unfortunately, the antibody we used from Novus fared no better, showing no staining in normal or severe dysplasias at concentrations up to 1:15 dilution. None of the antigen retrieval methods tested helped. Hence, we didn't pursue this target any further.

Interferon induced transmembrane proteins 2 and 3 (IFITM2 and IFITM3), overexpressed in the bottom layers of low-grades but not high-grades (Table 5.3A), were tested with one antibody. 4C8-1B10 was raised against IFITM3, but HPA used only one antibody to detect both these proteins, suggesting a high degree of similarity between these two related proteins. No attempt was made to quantify the degree of cross-reactivity of this antibody towards

IFITM2. This antibody was used at 1:250 dilution with citrate antigen retrieval on the LEEP specimens. Much of the stroma was stained in all sections.

Interferon induced transmembrane protein 3 plays a role in the immune system's defence against a number of viral infections. IFITM3 staining showed an upward trend in both top and bottom layers going from normal to CIN III (Figure 5.4). However, there was a lot of variability between individual cases, so the trend is quite weak and, in the case of the bottom layer, not significant ($P = 0.003$ and 0.13 for top and bottom layers, respectively). Within individual LEEPs, some cases show primarily increasing basal layer staining with increasing histopathological grade (Figure 5.5A) while others show the trend in the upper layer (Figure 5.5C).

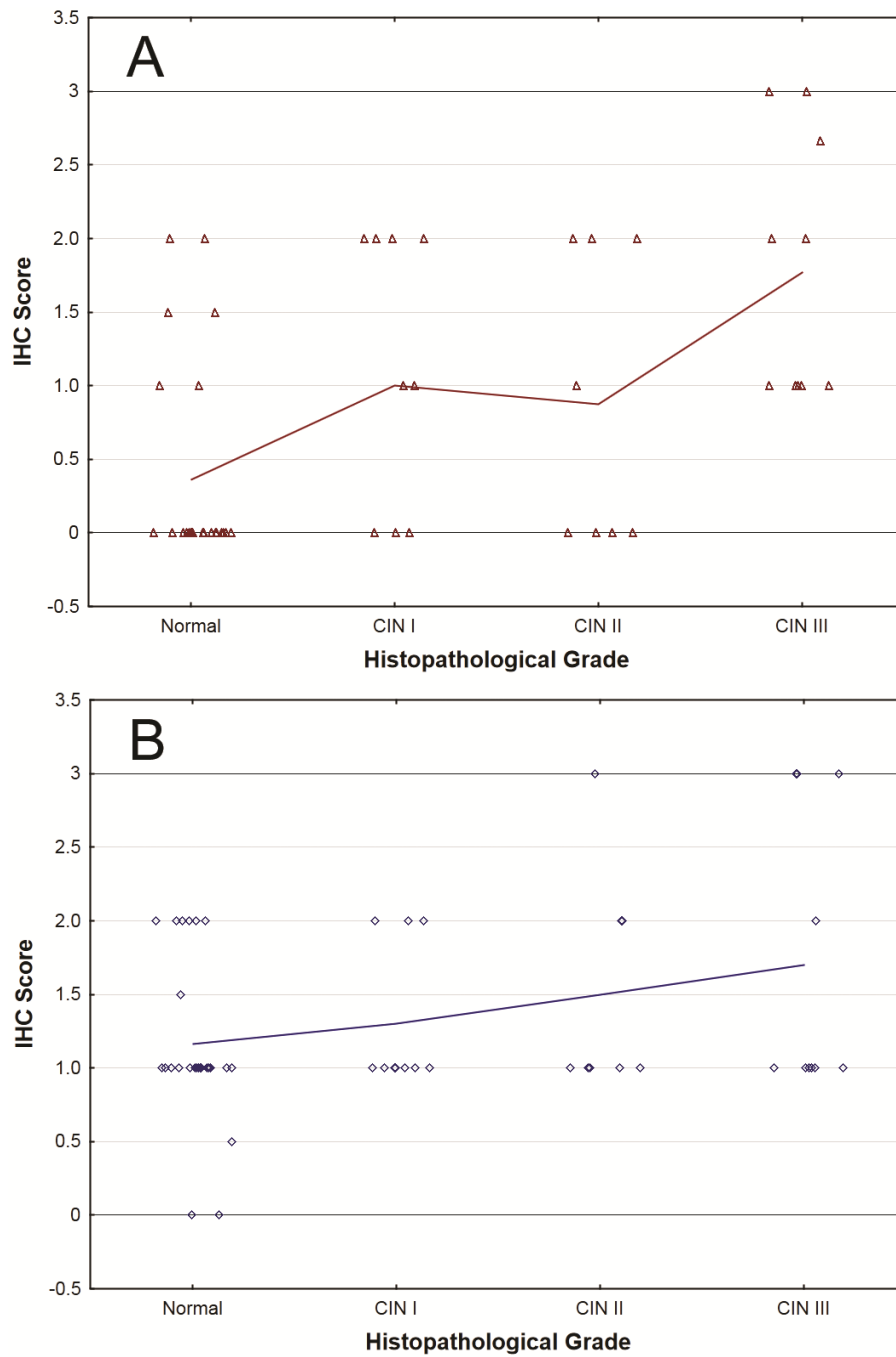


Figure 5.4: Summary of IFITM3 staining in LEEP specimens, separated into top (**A**) and bottom (**B**) halves of the epithelium. Scoring was performed as described in the text and each data point shown represents the average scores for all regions of the corresponding histopathological grade in one piece of tissue. A line connecting the mean scores for each grade is overlaid on each plot.

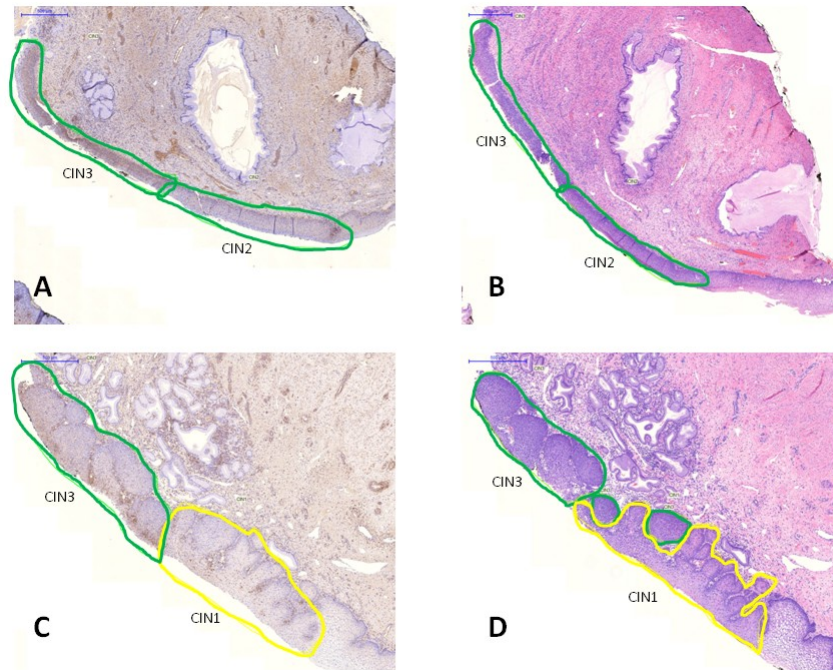


Figure 5.5: Sample LEEP sections (cases 0018 (**A, B**) and 0047 (**C, D**)) stained with IFITM3 IHC (**A, C**), along with adjacent reference H&E sections (**B, D**). High-grade squamous intraepithelial neoplasia regions have been circled with green, CIN I with yellow, and normal epithelium is uncircled. Scale bars in the top-left of each panel are 500 μ m in length.

Collagen, type XVI, alpha 1 (COL16A1), overexpressed in the bottom layers of low-grades but not high-grades (Table 5.3A), was tested with the HPA antibody. A LEEP sample containing CIN III was negative even at 1:20 dilution. Very weak staining was observed in a placenta tissue core, which should have been strongly stained according to HPA. Different antigen retrieval conditions made no difference. Consequently, this target was not pursued any further.

Cornulin (CRNN), a marker of epidermal differentiation, was the only negative marker of CIN tested, with the microarrays showing overexpression in the upper layers of low-grades but not high-grades (Table 5.4B). The antibody was used at 1:50 and antigen retrieval was found to be optional, although citrate retrieval was used for all the LEEPs. As seen in HPA, normal

epithelium showed strong staining concentrated in the upper half. Staining was generally weaker in higher grade lesions, but this was not always the case.

Cornulin staining was assessed slightly differently from the others. As the staining pattern in normal tissue typically includes intense staining of the top two-thirds with minimal to no staining in the basal third, staining was assessed between the top two-thirds and bottom third instead of top and bottom halves. The data shows a downward trend in staining of the upper layer with increasing grade ($P = 0.0001$), primarily due to the consistency of the intense staining in normal regions (Figure 5.6). The trend is less clear when considering just the dysplasias and there is no discernible trend in the basal layer ($P = 0.99$). Even in CIN III, upper layer staining was often observed, although it was usually quite weak. In individual sections, CIN regions are often clearly distinguishable from normal regions (Figure 5.7).

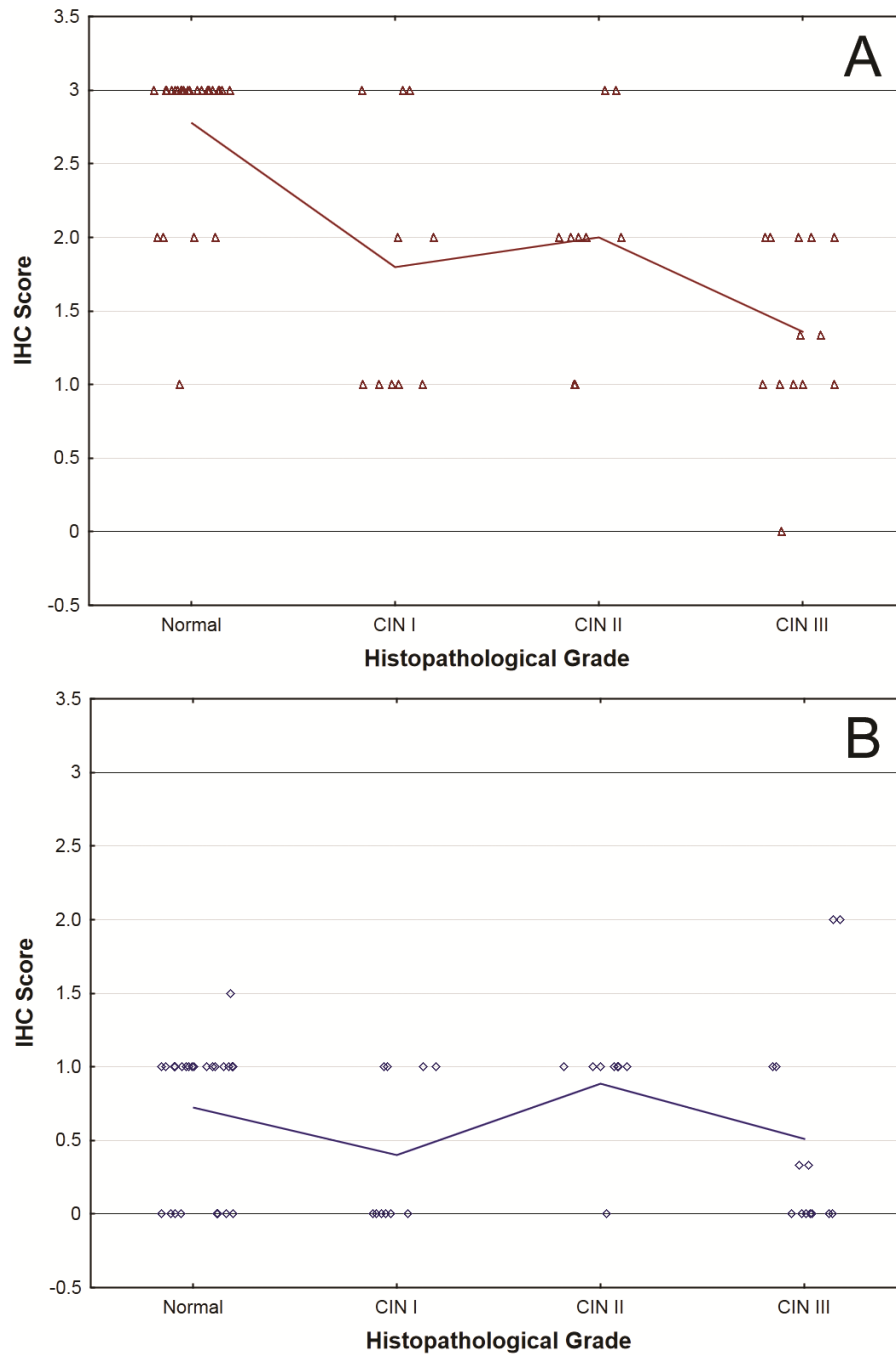


Figure 5.6: Summary of CRNN staining in LEEP specimens, separated into top two-thirds (A) and bottom third (B) of the epithelium. Scoring was performed as described in the text and each data point shown represents the average scores for all regions of the corresponding histopathological grade in one piece of tissue. A line connecting the mean scores for each grade is overlaid on each plot.

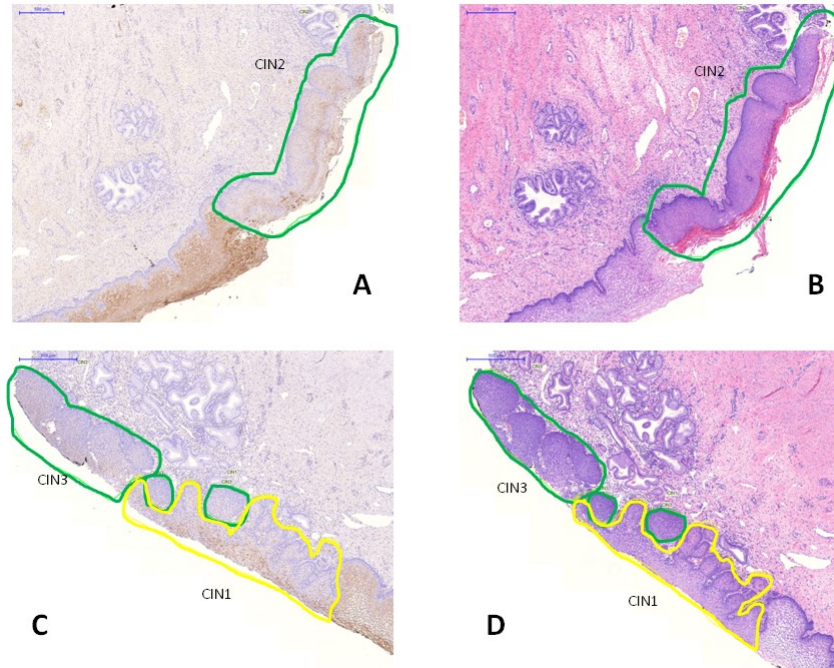


Figure 5.7: Sample LEEP sections (cases 0028 (A, B) and 0047 (C, D)) stained with CRNN IHC (A, C), along with adjacent reference H&E sections (B, D). High-grade squamous intraepithelial neoplasia regions have been circled with green, CIN I with yellow, and normal epithelium is uncircled. Scale bars in the top-left of each panel are 500 μ m in length.

Claudin 1 (CLDN1) was expressed more strongly in the top layers of high-grade regions than low-grades (Table 5.5A), based on microarray data. Two different probes for CLDN1 appeared on this list. Used at 1:15 dilution with pH 9 antigen retrieval for 22.5 minutes, staining was localized to cell membranes of the bottom half as seen in HPA. Staining was generally stronger and spanned more of the epithelium in higher grade LEEPs.

Claudin 1, a tight junction protein, had the most promising LEEP results. Staining in the upper layers displayed a strong upward trend with increasing dysplastic grade ($P < 0.0001$), although the trend in basal layer staining was much weaker ($P = 0.09$) (Figure 5.8). Looking at individual LEEPs, it was often easy to identify regions of high-grade dysplasia as they often stained much more intensely than normal tissue, especially in the superficial layer (Figure 5.9).

However, not every case fit the pattern exactly, including case 0018 in Figure 5.9A, where CIN II appears only weakly stained, similar to the adjacent normal tissue.

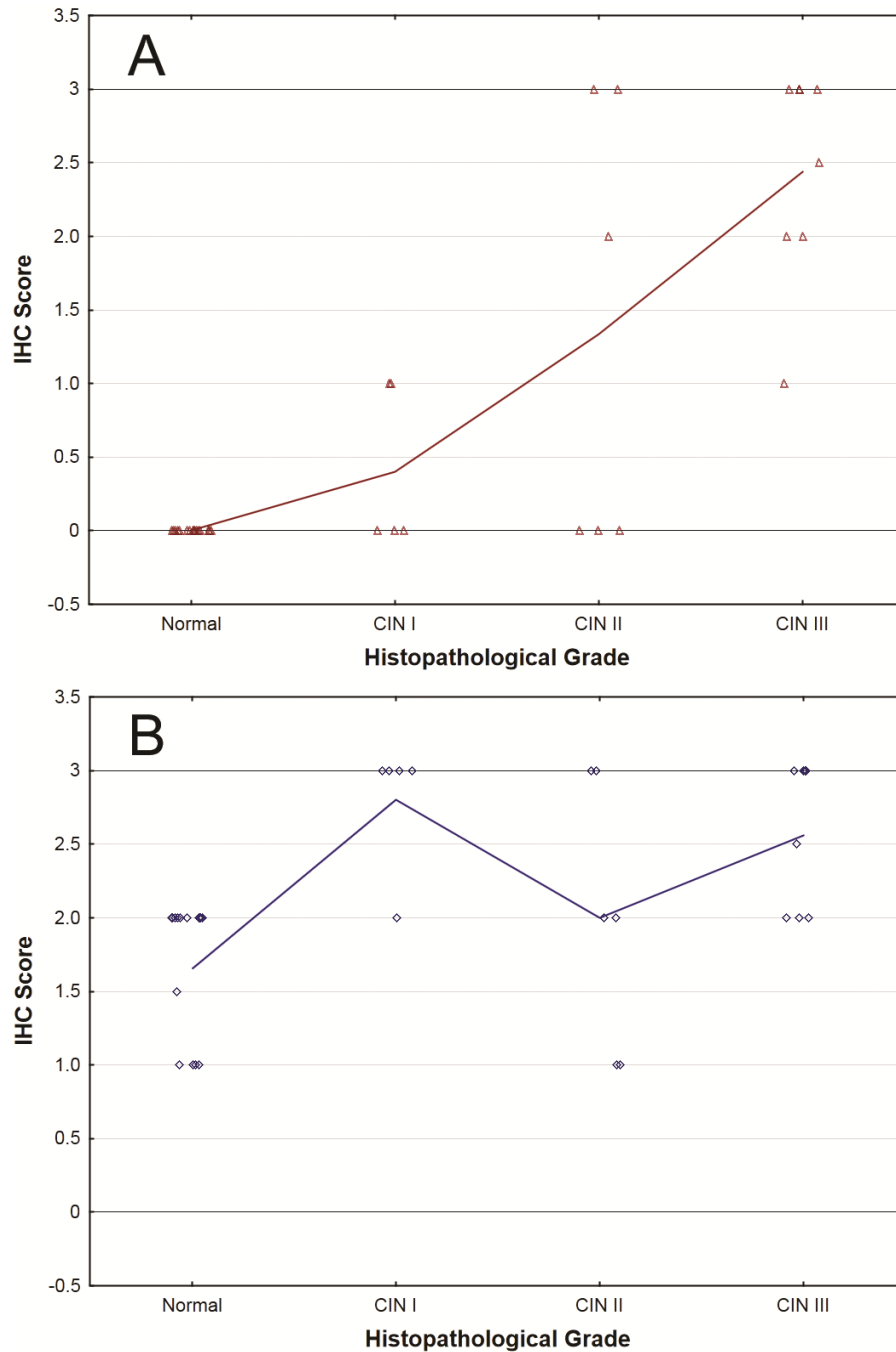


Figure 5.8: Summary of CLDN1 staining in LEEP specimens, separated into top (A) and bottom (B) halves of the epithelium. Scoring was performed as described in the text and each data point shown represents the average scores for all regions of the corresponding histopathological grade in one piece of tissue. A line connecting the mean scores for each grade is overlaid on each plot.

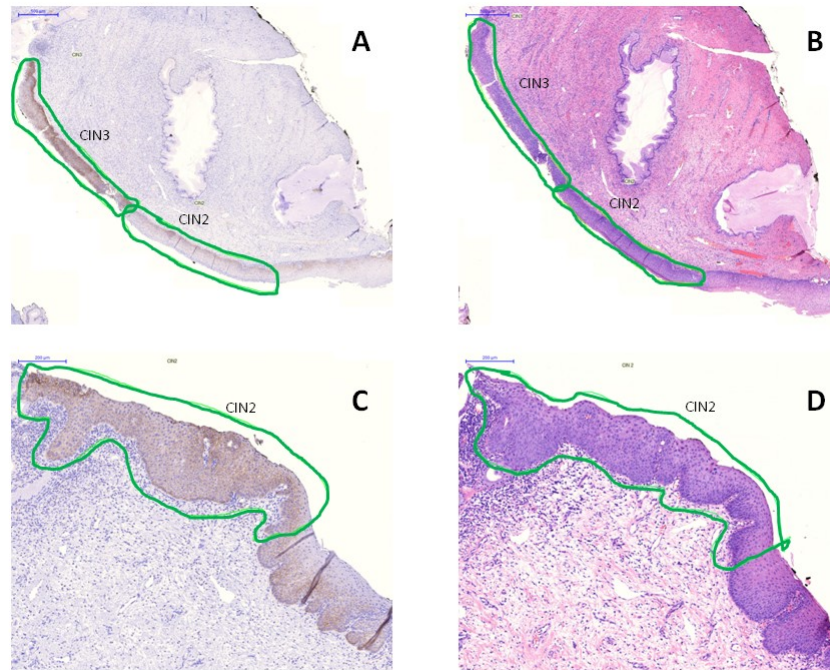


Figure 5.9: Sample LEEP sections (cases 0018 (A, B) and 0030 (C, D)) stained with CLDN1 IHC (A, C), along with adjacent reference H&E sections (B, D). High-grade squamous intraepithelial neoplasia regions have been circled with green and normal epithelium is uncircled. The lengths of the scale bars in the top-left of each panel are 500 μm for (A) and (B) and 200 μm for (C) and (D).

Stathmin 1 (STMN1) was also found to be expressed more strongly in the top layers of high-grade regions than low-grades from the microarray experiments (Table 5.5A). The first antibody to be tried was the same one that HPA used (sc-48362). No staining was observed in cervical epithelium with all antigen retrieval conditions, at concentrations up to 1:25. There was diffuse staining in the bony part of a marrow core, but likely attributable to non-specific staining. The vendor offered a replacement product, sc-20796, which fared no better. At 1:25 dilution and high pH antigen retrieval, staining was finally coaxed out of this antibody, but it was faint and diffuse throughout the normal epithelium, unlike the well-defined basal layer staining seen in HPA. The third antibody tried, 3352S, failed as well despite reported success in the literature (327). As with the previous antibody, 1:25 dilution yielded very faint and diffuse staining throughout the normal epithelium. Hence, this target was not pursued further.

5.4 Discussion

Advances in molecular biology have directed much attention towards understanding the biochemical mechanisms of disease, both to potentially improve treatment and to develop better strategies for early detection and screening. Previous molecular studies on cervical cancer have tended to focus on comparing invasive cancer with normal controls and have tended to treat the full thickness of the epithelium as one homogeneous unit. The present study is the first to our knowledge to use gene expression microarray analysis of molecular fixative preserved cervical tissue samples to consider differences in expression between epithelial layers and how this profile evolves with progression through the multistep carcinogenesis process.

Having established the technical validity of using microdissected molecular fixative paraffin-embedded samples for gene expression microarray analysis, we turned our attention to identifying and validating potential biomarkers in the collected data. Ideally, a positive biomarker for high-grade dysplasia would be present only in CIN II or worse, with higher expression in more severe lesions. The upper layer of epithelium is of particular interest because this layer is preferentially sampled in cytology. The present analysis, which treats the upper and lower halves of the epithelium separately, should be more sensitive to biomarkers which may be present in the basal layer of normal epithelium, but present in the top or both layers in high-grade lesions.

Morphologically, the top and bottom layers appear more similar in high-grade lesions than in low-grade ones. Disruptions in the cellular differentiation program of the epithelial cells in high-grade lesions may lead to a basal-like phenotype even for cells that are higher up in the epithelium. This is reflected in our observation that more probes are differentially expressed between the layers in low-grade lesions than in high-grades in our data. Moreover, cells that are

higher up in the epithelium are generally further along in their differentiation, turning on genes specific to their specialized function. Consequently, we observed that most differentially expressed probes are overexpressed in the top layer. Both these trends are evident in our microarray data because we treat the top and bottom layers separately in this study. These trends would be obscured in studies where the entire epithelium is treated as one unit.

To quantify the rate at which the microarray data validates, screens were performed to find the probes for which differential expression between the layers was observed in the low-grade regions. These were compared with the HPA images for normal cervical epithelium. 36 probes were found to be overexpressed in the majority of low-grade basal layer samples, based on the narrow thresholds with high ratio and high intensity criteria, of which 12 matched the microarray data. Subtracting out the 17 for which suitable images couldn't be found on HPA, this means 63% of the probes passed HPA validation. The opposite test, for probes overexpressed in the top layer, found 130 probes, for which 55 were missing data and only 22 validated. This is a success rate of only 29% and a missing data rate of 42%. Using the wide thresholds for top layer overexpression produces a shorter list of only 14 probes, of which 6 were missing HPA images and 4 validated, for a success rate of 50% and a missing data rate of 43%. Although HPA is a quick way of narrowing down the list of potential targets for validation, the high rate of missing data in HPA means that there is a chance that a legitimate hit was found in the microarray data that would not be followed up in the present analysis.

The disparity in validation success between top layer overexpression and bottom layer overexpression might be due to one of the assumptions inherent in the present analysis. All the data used was normalized on the assumption that the majority of genes are expressed at similar levels between the top and bottom layers. If basal layer cells inherently express all genes to a higher level, some genes in the top layer may be expressed at a higher level than the bulk of

genes in the top layer, but still be expressed at a lower absolute level than in the basal layer where it may be expressed at or below the level of the majority of bottom layer genes. Cells in the upper layers may be shutting off transcription as they prepare to be sloughed off as part of the body's natural renewal process. Looking back at the RNA purification data, more bottom layer RNA was collected in all cases except 0044 and 0053A. In case 0033A, similar amounts of RNA were collected between the layers. Normalizing on the assumption of similar expression levels between layers might highlight markers of interest in understanding the underlying biology of the neoplastic process. However, this assumption might not be as valid or useful in the context of biomarker discovery. A future analysis of this data might consider adjusting the normalization to account for the difference in overall expression levels between the layers. However, no attempt was made to standardize the volume of tissue microdissected from each layer, so it's unclear if the proper ratio can be determined for the present data set.

A number of targets on the shortlists of this analysis have also been found in microarray or similar analyses reported in the literature, but the majority of the candidate targets have not previously been found in a genome- or expressome-wide analysis. Among the targets that were found in previous studies, our results match the direction of expression change in the literature for most of them, giving us some measure of confidence in our data. The only exceptions come from the study by Wong *et al* (286) or from the serial analysis of gene expression (SAGE) studies by Shadeo *et al* and Kneller *et al* (308, 309, 321). These studies also happened to find the most targets in common with our study. While we focussed on CIN, Wong *et al* compared invasive cancers against normals (286). In addition, the SAGE and Wong *et al* studies did not separate the epithelium from the stroma. The SAGE studies used whole biopsies, likely including significant amounts of stroma, while Wong *et al* reported at least 80% malignant cells in their samples (286) without a means of standardizing this any further, meaning that some might have 80% while others might have closer to 100% malignant cells. Among the targets for which there was

disagreement between the present study and the published studies, only KLK12 did not show any staining in any combination of normal or cancerous epithelium or stroma, according to HPA. Hence, the discrepancies may be reflections of the different amounts of stroma in the normal and malignant samples used in the previous studies. Additionally, the observation by Wong *et al* of decreased expression of KLK7 in cancer cases (286) contrasts sharply with previous reports by Termini *et al* (328) and Santin *et al* (329), who studied KLK7 using IHC and, in the case of Santin *et al*, reverse transcription polymerase chain reaction (RT-PCR). Meanwhile, GJB2 is expressed in only the basal layer in normal and more pervasively in cancer, according to HPA, which is in closer agreement with the present study than the SAGE studies, where they were downregulated in CIN III compared to normal (309, 321). Hence, some of the observed discrepancies may be a result of false discoveries in those studies.

Looking at only the genes that passed HPA validation, a few were selected from across the various shortlists for further validation by IHC on LEEP specimens. A disappointingly low number of the antibodies ordered worked as advertised. Fewer than half of the antibodies, and none of the ones listed as being used to generate HPA images, worked despite trying different dilutions and antigen retrieval methods. In most cases, a general lack of staining was observed. In others, especially at higher antibody concentrations, staining was observed in a pattern inconsistent with HPA or the antibody information sheet, likely due to non-specific staining. This raises a more general question of the validity of using IHC as a validation method for expression studies. One could consider limiting IHC testing to those for which well-validated antibodies are available. However, these are typically the ones that have already drawn the most interest due to previously discovered disease links while truly novel potential biomarkers would be under-tested. Even for the four antibodies that produced staining consistent with HPA, more testing will be needed to draw any firm conclusions on the validity of any of these genes as biomarkers for high-grade dysplasia.

Among the targets tested with IHC, the most promising one appears to be claudin 1. It was found to be upregulated in high-grade top layers relative to low-grade regions (Table 5.5A). It is a component of tight junctions and previously found to correlate with grade in cervical dysplasia (330, 331). The claudin family of proteins have been investigated in many cancers (reviewed in (332) and (333)). Claudin 1 has been found to be downregulated in breast cancer (334) but upregulated in colon cancer (335, 336). In cervical cancer, previous microarrays studies demonstrated that overexpression of claudin 1 was particularly associated with squamous cell carcinoma, while association with adenocarcinoma was ambiguous (283, 293). The exact biological functions and the role of claudin 1 in carcinogenesis are still poorly understood (335), but involvement in the beta catenin pathway in the colon has been suggested (337). Our results show claudin 1 staining in the basal layer cells of normal cervical epithelium, but staining in upper layer cells is absent in all normal samples. The marked difference in expression of CLDN1 in the upper layer of cervical epithelium between high- and low-grade dysplasias make it a potential candidate as a cytological biomarker. The validation set in the current analysis, however, is quite small so further study will be required to draw more general conclusions about the applicability of claudin 1 as a biomarker.

The other two positive markers for high-grade dysplasia showed weaker but still positive trends in staining with increasing grade. *IFITM3* was found to be overexpressed in the basal layers of a majority of low-grade samples but not in a majority of the high-grades (Table 5.3A), making it potentially a marker for a basal-like phenotype. Additionally, *IFITM3* was found to be upregulated in the upper layers of high-grade versus low-grade dysplasias (Table 5.5A, as ENST00000270031), one of very few probes to be found in more than one screen. Interferon induced transmembrane protein 3 is involved in the body's response to viral infection. Although not well-studied in the context of cancer, altered expression of *IFITM3* has been found to play a role in colorectal carcinogenesis (338). Other members of the *IFITM* family have also been

studied in colorectal and gastric cancers (339-341). IHC staining of IFITM3 generally trended upward in both top and bottom layers with increasing CIN grade, but staining was sometimes inconsistent. The underlying stroma was also quite strongly and consistently stained regardless of the state of the overlying epithelium. Staining patterns in the epithelium often consisted of isolated positive cells infiltrating as if from the stroma. Perhaps IFITM3 could be a marker of inflammation associated with dysplasia, such as HPV associated changes. More investigation will be needed to clarify the exact role that IFITM3 might play in cervical carcinogenesis.

Of the positive markers tested with IHC, the one correlating the least with dysplastic grade was kallikrein-related peptidase 7, which was originally found to be overexpressed in the top layers of a majority of high-grade lesions and not a majority of low-grades (Table 5.2A). It is a serine protease that cleaves intercellular cohesive structures in a process required for desquamation (shedding of the outer layer of epidermis) in the skin. KLK7 has been reported to be overexpressed in CIN (328) and cervical cancer (329), although another microarray study found underexpression in cervical cancer (286). Among the LEEP specimens of the present study, however, there was little trend in staining in the bottom layer and only a weak positive correlation with grade in the upper layer. This might be due to the relatively small size of the present validation set, or this could be a case of protein levels not correlating with RNA levels. Additional tests, such as quantitative RT-PCR, might shed some light on whether this is an instance of the latter. The lack of a trend in our staining results might also reflect differences between the antibodies used. Most previous studies of kallikrein-related peptidase 7 used antibodies made in-house or from other researchers. The commercial antibody used on the LEEP specimens (from Abcam) had no references listed on its information sheet. Even on HPA, the two antibodies used there showed markedly different staining patterns in cervical tissue, with one of them not really staining malignant tissues at all. As suggested above, it is unclear to what extent little-used antibodies can be trusted without further validation. This is an important consideration

and one that will become even more relevant with the currently expanding use of high-throughput technologies to identify candidates that will undoubtedly require validation.

Among the candidate negative markers for cervical dysplasia, only one, cornulin, was tested with IHC. It was found to be overexpressed in the top layers of a majority of low-grade regions and not a majority of the high-grade lesions (Table 5.4B). Cornulin is a late marker of epidermal differentiation and loss of IHC staining has been observed in other cancers such as head and neck squamous cell carcinoma (342). In the cervix, a previous microarray study found lower expression of *CRNN* in cancers than normals (286), while IHC staining of cornulin has been reported to be strongest in normal tissues, weakening in high-grade dysplasias, and weakest in invasive cancers (343). The LEEP data shows this trend well in the top layer, although there is no clear trend in the bottom layer. Upper layer staining is strong in nearly all normal regions, but even in CIN III, there is some variability in the presence and intensity of staining. Furthermore, loss of staining in the upper layers of high-grade dysplasias is rarely absolute, further reducing its appeal as a potential negative biomarker for high-grade dysplasia.

The link between HPV and cervical cancer has been well established. Although it was not picked up on any of our screens, p16^{INK4A} has been used as a proxy for HPV activity (344) and generally correlates with CIN grade (262, 345, 346). In the microarray data, there were 2 probes for cyclin-dependent kinase inhibitor 2A (*CDKN2A*), the gene that encodes for p16. Both probes showed higher top layer expression in high-grades than low-grades, but both showed less than 2-fold change and P-values were not statistically significant ($P = 0.32$ and $P = 0.18$). Comparing top and bottom layers, only one probe (the one with the smaller P-value in the top layer t-test) showed differential expression and only in one case. Based on that probe, *CDKN2A* was overexpressed in the bottom layer of case 0044B (CIN I) using narrow thresholds. Where positive, literature reports on p16 show staining throughout the full thickness of the epithelium

(347), with a bias toward slightly stronger bottom layer staining. Our data generally agree with these trends, but the differences between layers and disease states were too small to be picked up in the target screens. The difference in staining between the layers may have been too small to be detected by even the narrow thresholds on the M-A plots. On the other hand, the t-tests may have simply needed a larger sample size to reach statistical significance.

The target lists generated in this analysis appear to be dominated by extracellular matrix components. Proliferation genes, which would be expected to play a crucial role in cancer, do not seem to factor prominently in any of the lists. This might be a result of our interest in biomarkers, as our screens favour genes that are highly expressed. Many biochemical signalling cascades amplify the signal from a small number of molecules to many copies of the downstream effector molecules. By focussing on abundantly expressed genes and their protein products, we may have been biasing our gene lists toward these downstream targets. This might give us acceptable biomarkers for early cancer detection, but it would be more difficult to use this approach to gain deeper insights into the molecular mechanisms of the carcinogenic process.

A way around this apparent bias towards extracellular matrix genes might be to relax the selection criteria. The important driver genes might not be greatly differentially expressed. More relaxed selection criteria would result in longer target lists, though, so more follow-up work would be necessary to validate the new targets. Additional strategies will need to be employed to reduce the lists to a more manageable size. Bioinformatics and pathway analysis tools might offer a potential solution to achieve this. Another side-effect of more permissive filtering criteria is the possibility of more false positives. The susceptibility of the t-test lists to false discoveries is a function of the P-value cutoff. Certainly, then, we expect a few of the hits in the t-test gene lists will be false positives. There is a wide variety of criteria for identifying differentially expressed genes in microarrays studies reported in the literature and, despite their simplicity, t-tests and

fold-change remain popular. A common method of estimating false discovery rates from such analyses is permutation analysis, in which the data for each gene is permuted between all patients in both malignant and normal groups to get an estimate of the null distribution. The significance analysis of microarrays algorithm, for example, uses a modified t-statistic and permutation analysis to calculate an estimated false discovery rate for each gene, which is then used as the threshold for differential expression (348). These methods, however, are all dependent on proper normalization of the data. Any normalization strategy makes fundamental assumptions about the behaviour of the data. A number of housekeeping genes known to be expressed at similar levels among all cells can be used to normalize the data. Alternatively, one can assume, as was done in the present analysis, that all samples should contain similar levels of RNA and that the majority of genes should be expressed at similar levels between samples. This is effectively using the bulk of the non-differentially expressed genes as housekeeping genes, an arguably more robust method that protects against the possibility that one of a few pre-selected housekeeping genes might not be as consistently expressed as previously believed. We have already seen, though, that even the assumption that the majority of genes are expressed at similar levels between samples might not be valid when comparing between layers. Meanwhile, Kendrick *et al* have reported higher overall RNA levels in CIN compared to normal (290). Future biomarker discovery analyses should consider the RNA content in the study samples.

An alternative strategy for identifying differentially expressed genes is to make per-case comparisons between the top and bottom layers, minimizing the potential impact of biological variation highlighted by our cluster analysis (Figure 4.7). Our calculations in Section 4.4 suggest that by requiring differential expression in at least three samples, we should not expect any false positives in the gene lists built from counting the numbers of samples showing differential expression. However, this assumes that the technical scatter observed in the duplicate sample experiments in Chapter 4 are representative of the technical component of the scatter observed in

the top versus bottom data studied here. Those estimates were made on the basis of only two sets of data, so it is quite possible that some of the top versus bottom data here displayed a wider degree of technical scatter than anticipated, resulting in a higher than expected false discovery rate. More testing on duplicate samples will be needed to get a clearer picture of the level of technical scatter that can be expected. Nevertheless, the expected false discovery rates are much lower than that implied by the HPA validation rate. As discussed above, only 29-63% of probes with suitable HPA data had expression data that matched. There are many possible reasons for this, including the possibility that for some genes, RNA expression doesn't necessarily correlate with protein expression. Moreover, for some genes, HPA data show different antibodies with different staining patterns. Even in our own experience, antibodies do not always work as advertised. Hence, when only one set of data is shown on HPA, it is unclear to what extent this data would be corroborated with other assays. For VEGFA, for example, HPA showed a single image of normal cervical epithelium stained throughout. However, other researchers have found that VEGF staining is weak and localized to the bottom half in normal cervix (320). Further experiments would be needed to explore these questions, such as using RT-PCR to distinguish between false discovery, RNA not translating into protein, and problems with IHC.

As alluded to previously, although this study has tested a few candidate biomarkers and one of them has even proven quite promising, this study is limited by a small sample size, both in the discovery and validation of potential markers. Consequently, much more testing will be needed to further validate the biological relevance of the results generated. Pathway analysis using tools such as IPA (Ingenuity Systems, www.ingenuity.com, Redwood City, CA, USA) might also be useful for confirming biological relevance. Nonetheless, this analysis is an important proof-of-principle that demonstrates the feasibility of microarray analysis on sub-epithelial layer microdissected MFPE cervical specimens. Even within the present data set, other promising markers that passed HPA validation have yet to be followed up with IHC. Much

remains to be learned from the stroma data which has been collected but not analyzed. While cancer research has typically focussed on the malignant epithelial cells, there has recently been increasing interest in studying the role of the underlying stroma in both the development and progression of cancer (349). This has led in turn to studies on stromal biomarkers of many types of cancer, including cervical cancer (289, 349-352). Future analyses might also try to account for differences in total RNA between the layers. The target screens applied in this analysis were quite rudimentary and only intended to demonstrate proof-of-concept. Hence, a future expanded data set with more microdissected samples could be analyzed with more rigorous bioinformatics approaches. As with our various target screens, different analysis approaches make different assumptions both about the data and about what we are looking for. We set out to look for novel biomarkers and consequently, our present analysis is biased toward looking for highly expressed genes. As such, it is possible that we missed some genes that are more weakly expressed but still important biochemically. Future work could include analyzing our data from the perspective of trying to understand the underlying biology. This might present an alternative route to uncovering biomarkers. Meanwhile, the most promising biomarkers arising from this analysis can be tested for their ability to discriminate high- and low-grade dysplasias. Even though a marker appears to correlate with histopathological grade based in IHC scoring, high degrees of variability among individual cases may cause such markers to be less attractive when their discriminating power is assessed. A larger validation set would assist in making such an analysis more statistically robust.

The use of MFPE specimens means that histopathological grading can be used to directly guide microdissection or compare with IHC. Without the use of molecular fixative, adjacent tissue pieces would have had to be fixed and processed differently under the assumption that the abnormal lesion was distributed across both pieces. Even within one tissue biopsy for microdissection, significant changes in tissue morphology were often witnessed going through the block. Slides from the last microdissected slide were sometimes hardly recognizable

compared to the first, which was sectioned about 800 μm apart in the block. Hence, it was important to have H&E reference slides along the way graded by the study pathologist.

This study also demonstrates the feasibility of studying the cervical epithelium not as one unit but rather as a series of cell layers that might have distinct biological differences. To our knowledge, this study is the first microarray analysis of microdissected cervical intraepithelial neoplasia to compare expression profiles between epithelial layers. By studying the top and bottom layer separately, trends in expression between the layers that might otherwise be obscured become apparent. High-grade lesions exhibit fewer differentially expressed genes than low-grades, while more differentially expressed genes are overexpressed in the top layer than the bottom (Figure 5.1). These trends are consistent with a model of cervical epithelium consisting of a continuum of cell layers where those at the top are further along the differentiation program that is disrupted during the carcinogenic process. Although the original plan was to microdissect the epithelium into three layers, technical difficulties dictated that only two layers would be feasible. The laser microdissection system could not cut the tissue within a reasonable length of time. However, improvements in technology might one day make it possible to microdissect more layers. Alternatively, future expression analyses may require less starting material, mitigating the impact of the slowness of the laser microdissection system. We believe the epithelium should be viewed as a continuum of layers, so separating the epithelium into a small number of distinct layers is by its nature a somewhat arbitrary exercise. It is one, though, that simplifies the system enough to make it comprehensible yet maintains enough of the original complexity to allow a more complete picture of the biochemical underpinnings of cervical epithelial differentiation and maturation to be put together, while allowing us to study how this process may be disrupted by the carcinogenic process.

The quest for novel biomarkers for high-grade cervical dysplasia has led us to gene expression array analysis of microdissected MFPE samples. While both the discovery and validation sample sizes were small in this pilot study, the data demonstrate that this approach is feasible and has already produced some promising biomarker candidates. Staining of claudin 1 in the upper layer of the epithelium, for example, appears to correlate with CIN grade. However, more studies will be needed to confirm these preliminary findings and to validate the potential use of claudin 1 as a cytological biomarker. Meanwhile, the use of MFPE samples enables the direct use of H&E slides to accurately guide microdissection and IHC interpretation, while the analysis of stratified squamous epithelium as distinct layers as opposed to a homogeneous unit unveils a new avenue for understanding the basic biology of cervical epithelium and carcinogenesis.

6 Conclusion

Much work has gone into developing novel biomarkers for the screening and early detection of cancer. Image cytometric measurement of DNA ploidy is one such technique that has shown great potential in many cancers (56) and has even been used for large-scale screening of cervical cancer in China. However, there is still much room for improvement and this thesis sought to build upon the successes of cytometric ploidy analysis by layering on additional biological information.

6.1 Recapitulation of aims

In lung cancer, previous studies on conventional sputum cytology have yielded varying results (237), with many of them focussing on trying to distinguish between invasive cancer and normal cases. By combining ploidy measurements with malignancy associated changes (MAC) features, a novel biomarker, the Combined Score (CS), was created. CS correlated with a number of known lung cancer risk factors, including histopathological grade, age, smoking status, quantitative morphometry, and p53 and Ki-67 staining. As a biomarker to identify the patients at highest risk to progress to invasive disease, CS performed comparably with another similar biomarker, LungSign (191), despite the fact that LungSign sought to distinguish between cancerous and non-cancerous samples, while CS was able to separate high-grade dysplasias from normals. By employing MAC features, CS could use subtle changes present in the far more numerous non-malignant cells to indicate the presence of disease, rather than trying to detect the rare malignant cells that other cytological tests rely on.

In cervical cancer, ploidy has previously been shown to be an effective cancer screening tool (151, 152) and its adoption into clinical practice in China continues to demonstrate this. We

attempted to use double staining with anti-Ki-67 immunocytochemistry and Feulgen-thionin to improve upon ploidy analysis but found little improvement. Double staining was technically successful, but perhaps Ki-67 was a poor choice of ICC marker. Motivated by this and the overall pursuit of novel biomarkers for early detection and screening of cervical cancer, we turned our attention to a gene expression microarray analysis of cervical dysplasia. By isolating different layers of the epithelium and comparing them against one another and across different histopathological grades, an improved understanding of cervical carcinogenesis could be sought and, with it, novel candidate biomarkers that might be practical when applied to cytology. A number of promising targets were identified, but as this was only intended to be a proof-of-principle, the present analysis used a small discovery and validation data set and further validation will be necessary. Nevertheless, it is clear that the approach holds promise for future biomarker discovery.

6.2 Impact and final remarks

We initially set out to improve upon ploidy-based image cytometry as a screening tool for preinvasive neoplastic lesions. Combining MAC features with ploidy in lung sputum generates a promising novel biomarker for risk assessment. However, combining anti-Ki-67 immunocytochemistry with ploidy in cervical cancer did not result in a clear improvement and it remains unclear to what extent ploidy would benefit from double staining with ICC using another marker, in particular one that would generate new information not associated with changes in cell cycling. The investigations into cervical biomarkers, however, resulted in a number of technical advances that might be useful for future work in this field. Double staining with Feulgen-thionin and immunocytochemistry was shown to be feasible even in cases where antigen retrieval is required. A protocol for optimizing the double staining was developed and potential pitfalls of the

process were studied. If the right immunocytochemical marker is used, double staining might not only improve the screening performance of ploidy, but also answer important questions about the role of ploidy in carcinogenesis. Although it is evident that aneuploidy manifests in many neoplastic lesions and can be a prognostic indicator, it is not clear what the biochemical mechanisms underpinning these correlations are. Is aneuploidy a driver of carcinogenesis or just a symptom of an increasingly dysregulated genome? Is it a little of both? Double staining might be one tool to help shed some light on this question.

Another technical advance in this thesis is the use of molecular fixative preserved paraffin-embedded cervical specimens. Previous studies have tested Sakura Finetek's Tissue-Tek® Xpress® Molecular Fixative on a variety of human tissues (353, 354), but not cervical tissue. Additionally, instead of using entire sections, only microdissected samples were used for microarray analysis in the present work. The use of MFPE specimens could potentially revolutionize how clinical samples are handled due to the immense opportunity afforded by the ability to perform molecular studies on the same sample as that used for clinical diagnosis. The current practice of using adjacent blocks is very approximate at best. At worst, we have already observed cases where the lesions present at one end of the tissue block are not present at the other end, making it entirely possible that an adjacent block contains none of the lesions present in an examined diagnostic block. More work, however, will still be needed to confirm the validity of using MFPE sections for clinical work. Even then, pathologists trained to interpret formalin-fixed sections, including all their inherent artifacts, might still be reluctant to adjust to a new fixative.

In addition to the technical advances presented here, the investigations in this thesis underscore a couple key approaches to cancer biomarker discovery that are often overlooked. First, by studying MACs in sputum, we made use of non-malignant cells that are present in far greater abundance than malignant cells. A similar approach could be taken with gene expression

microarrays. In fact, stroma data was collected but has yet to be analyzed. Just as non-malignant cells in the sputum displayed MAC features that betrayed the presence of malignant cells elsewhere in the lung, stroma underlying malignant lesions in the cervix might exhibit characteristic alterations in gene expression.

A second commonly overlooked approach to biomarker discovery is the heterogeneity present within stratified squamous epithelium, like that found in the cervix. Many studies continue to treat the entire thickness of the epithelium as a homogeneous unit, especially in gene expression studies like the one conducted here. However, even morphologically from an H&E stained slide, it is clear that cells near the basement membrane do not behave like cells near the surface. With microdissection and increasingly sensitive assays, it is becoming feasible to study stratified epithelium in all its complexity and this thesis demonstrates a bit of what can be uncovered through such an analysis.

Although the various assays were used in specific combinations in this thesis, it may be possible that many of them could be recombined. For example, a combined thionin and immunocytochemistry double staining approach might improve ploidy analysis in lung cancer, assuming an appropriate immunocytochemical marker was found. MAC and ICC could even be simultaneously layered upon ploidy analysis. This would provide the most biological information, but it is unresolved as to whether antigen retrieval would have a detrimental impact on the ability to measure MAC features. Assuming antigen retrieval does not adversely affect the measurement of MACs, double staining might also be useful in trying to understand the biochemical mechanisms underlying MAC features. In turn, this might point to other novel candidate biomarkers for cancer.

Ploidy-based image cytometry continues to serve us well as a tool for early detection and screening of cancer. Advances in molecular biology and our understanding of the biochemical

mechanisms of carcinogenesis have motivated the pursuit of novel biomarkers. By attempting to layer additional biological information on top of a ploidy analysis, this thesis has demonstrated that the Combined Score, based on ploidy and MAC features, can be used for risk assessment of lung cancer. Meanwhile, this thesis has introduced a number of technical advancements, including combined Feulgen-thionin and immunocytochemistry double staining with antigen retrieval, gene expression analysis of molecular fixative preserved paraffin-embedded cervical specimens, as well as microdissection of cervical layers so that they may be treated and understood separately in expression analyses. These advances should unlock new avenues of research into the underlying molecular mechanisms of disease and new biomarkers that will alert us when these processes are occurring. It is hoped that future research will continue to bear fruit, with improved combination ploidy tests translating into improved outcomes and care for generations of cancer patients to come.

References

1. Statistics Canada (2012) Leading causes of death in Canada, 2009. Statistics Canada Catalogue no. 84-215-x2012001. <http://www.statcan.gc.ca/pub/84-215-x/84-215-x2012001-eng.htm>. Accessed August 7, 2012 (Statistics Canada, Ottawa).
2. Jemal A, Bray F, Center MM, Ferlay J, Ward E, & Forman D (2011) Global cancer statistics. *CA Cancer J Clin* 61(2):69-90.
3. Siegel R, Naishadham D, & Jemal A (2012) Cancer statistics, 2012. *CA Cancer J Clin* 62(1):10-29.
4. Kelloff GJ, Lippman SM, Dannenberg AJ, Sigman CC, Pearce HL, Reid BJ, Szabo E, Jordan VC, Spitz MR, Mills GB, Papadimitrakopoulou VA, Lotan R, Aggarwal BB, Bresalier RS, Kim J, Arun B, Lu KH, Thomas ME, Rhodes HE, Brewer MA, Follen M, Shin DM, Parnes HL, Siegfried JM, Evans AA, Blot WJ, Chow WH, Blount PL, Maley CC, Wang KK, Lam S, Lee JJ, Dubinett SM, Engstrom PF, Meyskens FL, Jr., O'Shaughnessy J, Hawk ET, Levin B, Nelson WG, Hong WK, & the AACR Task Force on Cancer Prevention (2006) Progress in chemoprevention drug development: The promise of molecular biomarkers for prevention of intraepithelial neoplasia and cancer – A plan to move forward. *Clin Cancer Res* 12(12):3661-3697.
5. Mathew A & George PS (2009) Trends in Incidence and Mortality Rates of Squamous Cell Carcinoma and Adenocarcinoma of Cervix - Worldwide. *Asian Pac J Cancer Prev* 10(4):645-650.
6. Vizcaino AP, Moreno V, Bosch FX, Munoz N, Barros-Dios XM, Borras J, & Parkin DM (2000) International trends in incidence of cervical cancer: II. Squamous-cell carcinoma. *Int J Cancer* 86(3):429-435.
7. BC Cancer Agency (2005) Screening for Cancer. <http://www.bccancer.bc.ca/HPI/CancerManagementGuidelines/ScreeningforCancer.htm>. Accessed 13 August 2012.
8. Strong K, Wald N, Miller A, & Alwan A (2005) Current concepts in screening for noncommunicable disease: World Health Organization Consultation Group Report on methodology of noncommunicable disease screening. *J Med Screen* 12(1):12-19.
9. Health Canada (1998) Cervical cancer screening in Canada: 1998 surveillance report. Catalogue no. H39-616/1998E. <http://www.phac-aspc.gc.ca/publicat/ccsic-dccuac/pdf/cervical-e3.pdf>. Accessed August 16, 2012 (Minister of Public Works and Government Services Canada, Ottawa).
10. Gotzsche PC & Nielsen M (2011) Screening for breast cancer with mammography. *Cochrane Database Syst Rev* (1):CD001877.
11. Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK, & Humphrey L (2009) Screening for breast cancer: an update for the U.S. Preventive Services Task Force. *Ann Intern Med* 151(10):727-737.

12. Qaseem A, Denberg TD, Hopkins RH, Jr., Humphrey LL, Levine J, Sweet DE, & Shekelle P (2012) Screening for colorectal cancer: a guidance statement from the American College of Physicians. *Ann Intern Med* 156(5):378-386.
13. Zauber AG, Lansdorp-Vogelaar I, Knudsen AB, Wilschut J, van Ballegooijen M, & Kuntz KM (2008) Evaluating test strategies for colorectal cancer screening: a decision analysis for the U.S. Preventive Services Task Force. *Ann Intern Med* 149(9):659-669.
14. Stamey TA, Yang N, Hay AR, McNeal JE, Freiha FS, & Redwine E (1987) Prostate-specific antigen as a serum marker for adenocarcinoma of the prostate. *N Engl J Med* 317(15):909-916.
15. Potosky AL, Feuer EJ, & Levin DL (2001) Impact of screening on incidence and mortality of prostate cancer in the United States. *Epidemiol Rev* 23(1):181-186.
16. Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, Kwiatkowski M, Lujan M, Lilja H, Zappa M, Denis LJ, Recker F, Paez A, Maattanen L, Bangma CH, Aus G, Carlsson S, Villers A, Rebillard X, van der Kwast T, Kujala PM, Blijenberg BG, Stenman UH, Huber A, Taari K, Hakama M, Moss SM, de Koning HJ, & Auvinen A (2012) Prostate-cancer mortality at 11 years of follow-up. *N Engl J Med* 366(11):981-990.
17. Andriole GL, Crawford ED, Grubb RL, 3rd, Buys SS, Chia D, Church TR, Fouad MN, Isaacs C, Kvale PA, Reding DJ, Weissfeld JL, Yokochi LA, O'Brien B, Ragard LR, Clapp JD, Rathmell JM, Riley TL, Hsing AW, Izmirlian G, Pinsky PF, Kramer BS, Miller AB, Gohagan JK, & Prorok PC (2012) Prostate cancer screening in the randomized Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial: mortality results after 13 years of follow-up. *J Natl Cancer Inst* 104(2):125-132.
18. Moyer VA (2012) Screening for Prostate Cancer: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med* 157(2):120-134.
19. Jemal A, Siegel R, Xu J, & Ward E (2010) Cancer statistics, 2010. *CA Cancer J Clin* 60(5):277-300.
20. Fontana RS, Sanderson DR, Woolner LB, Taylor WF, Miller WE, & Muhm JR (1986) Lung cancer screening: The Mayo program. *J Occup Med* 28(8):746-750.
21. Henschke CI, Yankelevitz DF, Libby DM, Pasmantier MW, Smith JP, & Miettinen OS (2006) Survival of patients with Stage I lung cancer detected on CT screening. *N Engl J Med* 355(17):1763-1771.
22. Swensen SJ, Jett JR, Hartman TE, Midthun DE, Mandrekar SJ, Hillman SL, Sykes AM, Aughenbaugh GL, Bungum AO, & Allen KL (2005) CT screening for lung cancer: Five-year prospective experience. *Radiology* 235(1):259-265.
23. Edell E, Lam S, Pass H, Miller YE, Sutedja T, Kennedy T, Loewen G, Keith RL, & Gazdar A (2009) Detection and localization of intraepithelial neoplasia and invasive carcinoma using fluorescence-reflectance bronchoscopy: An international, multicenter clinical trial. *J Thorac Oncol* 4(1):49-54.

24. Moghissi K, Dixon K, & Stringer MR (2008) Current indications and future perspective of fluorescence bronchoscopy: A review study. *Photodiagnosis Photodyn Ther* 5(4):238-246.
25. McWilliams A, MacAulay C, Gazdar AF, & Lam S (2002) Innovative molecular and imaging approaches for the detection of lung cancer and its precursor lesions. *Oncogene* 21(45):6949-6959.
26. Yee J, Sadar MD, Sin DD, Kuzyk M, Xing L, Kondra J, McWilliams A, Man SF, & Lam S (2009) Connective tissue-activating peptide III: A novel blood biomarker for early lung cancer detection. *J Clin Oncol* 27(17):2787-2792.
27. National Lung Screening Trial Research Team (2011) Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 365(5):395-409.
28. Ganti AK & Mulshine JL (2006) Lung cancer screening. *Oncologist* 11(5):481-487.
29. Lam S, MacAulay C, Hung J, LeRiche J, Profio AE, & Palcic B (1993) Detection of dysplasia and carcinoma in situ with a lung imaging fluorescence endoscope device. *J Thorac Cardiovasc Surg* 105(6):1035-1040.
30. Lam S, Macaulay C, Leriche JC, Ikeda N, & Palcic B (1994) Early localization of bronchogenic carcinoma. *Diagn Ther Endosc* 1(2):75-78.
31. McWilliams A, Lam B, & Sutedja T (2009) Early proximal lung cancer diagnosis and treatment. *Eur Respir J* 33(3):656-665.
32. Croswell JM, Baker SG, Marcus PM, Clapp JD, & Kramer BS (2010) Cumulative incidence of false-positive test results in lung cancer screening: A randomized trial. *Ann Intern Med* 152(8):505-512.
33. McWilliams A & Lam S (2005) Lung cancer screening. *Curr Opin Pulm Med* 11(4):272-277.
34. Bach PB, Mirkin JN, Oliver TK, Azzoli CG, Berry DA, Brawley OW, Byers T, Colditz GA, Gould MK, Jett JR, Sabichi AL, Smith-Bindman R, Wood DE, Qaseem A, & Detterbeck FC (2012) Benefits and harms of CT screening for lung cancer: a systematic review. *JAMA* 307(22):2418-2429.
35. Canadian Cancer Society's Steering Committee on Cancer Statistics (2012) Canadian cancer statistics 2012. <http://www.cancer.ca/statistics>. Accessed 16 August 2012 (Canadian Cancer Society, Toronto, ON).
36. Sankaranarayanan R, Esmy PO, Rajkumar R, Muwonge R, Swaminathan R, Shanthakumari S, Fayette JM, & Cherian J (2007) Effect of visual screening on cervical cancer incidence and mortality in Tamil Nadu, India: a cluster-randomised trial. *Lancet* 370(9585):398-406.
37. Qiao Y-L, Sellors JW, Eder PS, Bao Y-P, Lim JM, Zhao F-H, Weigl B, Zhang W-H, Peck RB, Li L, Chen F, Pan Q-J, & Lorincz AT (2008) A new HPV-DNA test for

- cervical-cancer screening in developing regions: a cross-sectional study of clinical accuracy in rural China. *Lancet Oncol* 9(10):929-936.
38. Denny L, Kuhn L, Hu CC, Tsai WY, & Wright TC, Jr. (2010) Human papillomavirus-based cervical cancer prevention: long-term results of a randomized screening trial. *J Natl Cancer Inst* 102(20):1557-1567.
 39. Cuzick J, Arbyn M, Sankaranarayanan R, Tsu V, Ronco G, Mayrand MH, Dillner J, & Meijer CJ (2008) Overview of human papillomavirus-based and other novel options for cervical cancer screening in developed and developing countries. *Vaccine* 26 Suppl 10:K29-41.
 40. Naucler P, Ryd W, Tornberg S, Strand A, Wadell G, Elfgrén K, Radberg T, Strander B, Forslund O, Hansson BG, Hagmar B, Johansson B, Rylander E, & Dillner J (2009) Efficacy of HPV DNA testing with cytology triage and/or repeat HPV DNA testing in primary cervical cancer screening. *J Natl Cancer Inst* 101(2):88-99.
 41. Schiffman M, Wentzensen N, Wacholder S, Kinney W, Gage JC, & Castle PE (2011) Human papillomavirus testing in the prevention of cervical cancer. *J Natl Cancer Inst* 103(5):368-383.
 42. Sankaranarayanan R, Rajkumar R, Theresa R, Esmy PO, Mahe C, Bagyalakshmi KR, Thara S, Frappart L, Lucas E, Muwonge R, Shanthakumari S, Jeevan D, Subbarao TM, Parkin DM, & Cherian J (2004) Initial results from a randomized trial of cervical visual screening in rural south India. *Int J Cancer* 109(3):461-467.
 43. Villa LL (2008) Assessment of new technologies for cervical cancer screening. *Lancet Oncol* 9(10):910-911.
 44. Schiffman M, Castle PE, Jeronimo J, Rodriguez AC, & Wacholder S (2007) Human papillomavirus and cervical cancer. *Lancet* 370(9590):890-907.
 45. Saslow D, Solomon D, Lawson HW, Killackey M, Kulasingam SL, Cain J, Garcia FA, Moriarty AT, Waxman AG, Wilbur DC, Wentzensen N, Downs LS, Jr., Spitzer M, Moscicki AB, Franco EL, Stoler MH, Schiffman M, Castle PE, & Myers ER (2012) American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer. *CA Cancer J Clin* 62(3):147-172.
 46. Moyer VA & U.S. Preventive Services Task Force (2012) Screening for cervical cancer: U.S. Preventive services task force recommendation statement. *Ann Intern Med* 156(12):880-891.
 47. Dillner J, Rebolj M, Birembaut P, Petry KU, Szarewski A, Munk C, de Sanjose S, Naucler P, Lloveras B, Kjaer S, Cuzick J, van Ballegooijen M, Clavel C, & Iftner T (2008) Long term predictive values of cytology and human papillomavirus testing in cervical cancer screening: joint European cohort study. *BMJ* 337:a1754.
 48. Hanahan D & Weinberg Robert A (2011) Hallmarks of cancer: the next generation. *Cell* 144(5):646-674.

49. Sporn MB (1993) Chemoprevention of cancer. *Lancet* 342(8881):1211-1213.
50. Kinzler KW & Vogelstein B (1996) Lessons from hereditary colorectal cancer. *Cell* 87(2):159-170.
51. Nordling CO (1953) A new theory on cancer-inducing mechanism. *Br J Cancer* 7(1):68-72.
52. Soria JC, Kim ES, Fayette J, Lantuejoul S, Deutsch E, & Hong WK (2003) Chemoprevention of lung cancer. *Lancet Oncol* 4(11):659-669.
53. Hanahan D & Weinberg RA (2000) The hallmarks of cancer. *Cell* 100(1):57-70.
54. Rajagopalan H & Lengauer C (2004) Aneuploidy and cancer. *Nature* 432(7015):338-341.
55. Duesberg P, Li R, & Rasnick D (2004) Aneuploidy approaching a perfect score in predicting and preventing cancer: highlights from a conference held in Oakland, CA in January, 2004. *Cell Cycle* 3(6):823-828.
56. Cohen C (1996) Image cytometric analysis in pathology. *Hum Pathol* 27(5):482-493.
57. Slaughter DP, Southwick HW, & Smejkal W (1953) "Field cancerization" in oral stratified squamous epithelium; clinical implications of multicentric origin. *Cancer* 6(5):963-968.
58. Knudson AG, Jr. (1971) Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A* 68(4):820-823.
59. O'Shaughnessy JA, Kelloff GJ, Gordon GB, Dannenberg AJ, Hong WK, Fabian CJ, Sigman CC, Bertagnolli MM, Stratton SP, Lam S, Nelson WG, Meyskens FL, Alberts DS, Follen M, Rustgi AK, Papadimitrakopoulou V, Scardino PT, Gazdar AF, Wattenberg LW, Sporn MB, Sakr WA, Lippman SM, & Von Hoff DD (2002) Treatment and prevention of intraepithelial neoplasia: An important target for accelerated new agent development. *Clin Cancer Res* 8(2):314-346.
60. Sporn MB & Suh N (2000) Chemoprevention of cancer. *Carcinogenesis* 21(3):525-530.
61. Sporn MB & Suh N (2002) Chemoprevention: an essential approach to controlling cancer. *Nat Rev Cancer* 2(7):537-543.
62. Mukhtar H (2012 Preprint) Chemoprevention: Making it a success story for controlling human cancer. *Cancer Lett*.
63. Chemoprevention Working Group (1999) Prevention of cancer in the next millennium: Report of the Chemoprevention Working Group to the American Association for Cancer Research. *Cancer Res* 59(19):4743-4758.
64. Winterhalder RC, Hirsch FR, Kotantoulas GK, Franklin WA, & Bunn PA, Jr. (2004) Chemoprevention of lung cancer – from biology to clinical reality. *Ann Oncol* 15(2):185-196.

65. Lam S, leRiche JC, McWilliams A, MacAulay C, Dyachkova Y, Szabo E, Mayo J, Schellenberg R, Coldman A, Hawk E, & Gazdar A (2004) A randomized Phase IIb trial of pulmicort turbuhaler (budesonide) in people with dysplasia of the bronchial epithelium. *Clin Cancer Res* 10(19):6502-6511.
66. Lam S, MacAulay C, le Riche JC, Dyachkova Y, Coldman A, Guillaud M, Hawk E, Christen M-O, & Gazdar AF (2002) A randomized Phase IIb trial of anethole dithiolethione in smokers with bronchial dysplasia. *J Natl Cancer Inst* 94(13):1001-1009.
67. Lam S, McWilliams A, leRiche J, MacAulay C, Wattenberg L, & Szabo E (2006) A Phase I study of myo-inositol for lung cancer chemoprevention. *Cancer Epidemiol Biomarkers Prev* 15(8):1526-1531.
68. Saito M, Kato H, Tsuchida T, & Konaka C (1994) Chemoprevention effects on bronchial squamous metaplasia by folate and vitamin B12 in heavy smokers. *Chest* 106(2):496-499.
69. Govindan R, Page N, Morgensztern D, Read W, Tierney R, Vlahiotis A, Spitznagel EL, & Piccirillo J (2006) Changing epidemiology of small-cell lung cancer in the United States over the last 30 years: analysis of the surveillance, epidemiologic, and end results database. *J Clin Oncol* 24(28):4539-4544.
70. Oken MM, Marcus PM, Hu P, Beck TM, Hocking W, Kvale PA, Cordes J, Riley TL, Winslow SD, Peace S, Levin DL, Prorok PC, & Gohagan JK (2005) Baseline chest radiograph for lung cancer detection in the randomized Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial. *J Natl Cancer Inst* 97(24):1832-1839.
71. Anonymous (2010) Lung. *AJCC Cancer Staging Manual*, eds Edge SB, Byrd DR, Compton CC, Fritz AG, Greene FL, & Trotti A (Springer, New York, NY), 7th Ed, pp 253-270.
72. Mountain CF (1997) Revisions in the International System for Staging Lung Cancer. *Chest* 111(6):1710-1717.
73. Franklin WA, Wistuba II, Geisinger K, Lam S, Hirsch FR, Muller KM, Sozzi G, Brambilla E, & Gazdar A (2004) Squamous dysplasia and carcinoma in situ. in *World Health Organization classification of tumours. Pathology and genetics of tumours of the lung, pleura, thymus and heart*, eds Travis WD, Brambilla E, Müller-Hermelink HK, & Harris CC (IARC Press, Lyon), pp 68-72.
74. Brambilla E, Travis WD, Colby TV, Corrin B, & Shimosato Y (2001) The new World Health Organization classification of lung tumours. *Eur Respir J* 18(6):1059-1068.
75. Saccomanno G, Archer VE, Auerbach O, Saunders RP, & Brennan LM (1974) Development of carcinoma of the lung as reflected in exfoliated cells. *Cancer* 33(1):256-270.
76. Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger KR, Yatabe Y, Beer DG, Powell CA, Riely GJ, Van Schil PE, Garg K, Austin JH, Asamura H, Rusch VW, Hirsch FR, Scagliotti G, Mitsudomi T, Huber RM, Ishikawa Y, Jett J, Sanchez-Cespedes M, Sculier JP, Takahashi T, Tsuboi M, Vansteenkiste J, Wistuba I, Yang PC, Aberle D,

- Brambilla C, Flieder D, Franklin W, Gazdar A, Gould M, Hasleton P, Henderson D, Johnson B, Johnson D, Kerr K, Kuriyama K, Lee JS, Miller VA, Petersen I, Roggli V, Rosell R, Saijo N, Thunnissen E, Tsao M, & Yankelewitz D (2011) International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. *J Thorac Oncol* 6(2):244-285.
77. Sato M, Shames DS, Gazdar AF, & Minna JD (2007) A translational view of the molecular pathogenesis of lung cancer. *J Thorac Oncol* 2(4):327-343.
 78. Blagosklonny MV (2000) p53 from complexity to simplicity: mutant p53 stabilization, gain-of-function, and dominant-negative effect. *FASEB J* 14(13):1901-1907.
 79. Benedet JL, Bender H, Jones H, 3rd, Ngan HY, & Pecorelli S (2000) FIGO staging classifications and clinical practice guidelines in the management of gynecologic cancers. FIGO Committee on Gynecologic Oncology. *Int J Gynaecol Obstet* 70(2):209-262.
 80. Richart RM (1990) A modified terminology for cervical intraepithelial neoplasia. *Obstet Gynecol* 75(1):131-133.
 81. Solomon D, Davey D, Kurman R, Moriarty A, O'Connor D, Prey M, Raab S, Sherman M, Wilbur D, Wright T, Jr., & Young N (2002) The 2001 Bethesda System: terminology for reporting results of cervical cytology. *JAMA* 287(16):2114-2119.
 82. Doorbar J (2006) Molecular biology of human papillomavirus infection and cervical cancer. *Clin Sci (Lond)* 110(5):525-541.
 83. Smith JS, Lindsay L, Hoots B, Keys J, Franceschi S, Winer R, & Clifford GM (2007) Human papillomavirus type distribution in invasive cervical cancer and high-grade cervical lesions: a meta-analysis update. *Int J Cancer* 121(3):621-632.
 84. Mantovani F & Banks L (2001) The human papillomavirus E6 protein and its contribution to malignant progression. *Oncogene* 20(54):7874-7887.
 85. Munger K, Basile JR, Duensing S, Eichten A, Gonzalez SL, Grace M, & Zacny VL (2001) Biological activities and molecular targets of the human papillomavirus E7 oncoprotein. *Oncogene* 20(54):7888-7898.
 86. Yildiz IZ, Usubutun A, Firat P, Ayhan A, & Kucukali T (2007) Efficiency of immunohistochemical p16 expression and HPV typing in cervical squamous intraepithelial lesion grading and review of the p16 literature. *Pathol Res Pract* 203(6):445-449.
 87. Kalof AN, Evans MF, Simmons-Arnold L, Beatty BG, & Cooper K (2005) p16INK4A immunoexpression and HPV in situ hybridization signal patterns: potential markers of high-grade cervical intraepithelial neoplasia. *Am J Surg Pathol* 29(5):674-679.
 88. Pierzchalski A, Mittag A, & Tarnok A (2011) Introduction A: recent advances in cytometry instrumentation, probes, and methods--review. *Methods Cell Biol* 102:1-21.

89. Zuba-Surma EK, Kucia M, Abdel-Latif A, Lillard JW, Jr., & Ratajczak MZ (2007) The ImageStream System: a key step to a new era in imaging. *Folia Histochem Cytobiol* 45(4):279-290.
90. McGrath KE, Bushnell TP, & Palis J (2008) Multispectral imaging of hematopoietic cells: Where flow meets morphology. *J Immunol Methods* 336(2):91-97.
91. Li X & Darzynkiewicz Z (1999) The Schrödinger's Cat Quandary in Cell Biology: Integration of Live Cell Functional Assays with Measurements of Fixed Cells in Analysis of Apoptosis. *Exp Cell Res* 249(2):404-412.
92. Remmerbach TW, Meyer-Ebrecht D, Aach T, Würflinger T, Bell AA, Schneider TE, Nietzke N, Frerich B, & Böcking A (2009) Toward a multimodal cell analysis of brush biopsies for the early detection of oral cancer. *Cancer Cytopathol* 117(3):228-235.
93. Würflinger T, Stockhausen J, Meyer-Ebrecht D, & Böcking A (2004) Robust automatic coregistration, segmentation, and classification of cell nuclei in multimodal cytopathological microscopic images. *Comput Med Imaging Graph* 28(1-2):87-98.
94. Laffers W, Mittag A, Lenz D, Tárnok A, & Gerstner AOH (2006) Iterative restaining as a pivotal tool for n-color immunophenotyping by slide-based cytometry. *Cytometry A* 69A(3):127-130.
95. Gaiser T, Berroa-Garcia L, Kemmerling R, Dutta A, Ried T, & Heselmeyer-Haddad K (2011) Automated analysis of protein expression and gene amplification within the same cells of paraffin-embedded tumour tissue. *Cell Oncol* 34(4):337-342.
96. Preston K, Jr. (1986) High-resolution image analysis. *J Histochem Cytochem* 34(1):67-74.
97. Ingram M & Preston K, Jr. (1970) Automatic analysis of blood cells. *Sci Am* 223(5):72-82.
98. Bartels PH, Olson GB, Layton JM, Anderson RE, & Wied GL (1975) Computer discrimination of T and B lymphocytes. *Acta Cytol* 19(1):53-57.
99. Prewitt JM & Mendelsohn ML (1966) The analysis of cell images. *Ann N Y Acad Sci* 128(3):1035-1053.
100. Green DK (1974) Machine to find metaphase cells. *Exp Cell Res* 86(1):170-174.
101. Green DK & Neurath PW (1974) The design, operation and evaluation of a high speed automatic metaphase finder. *J Histochem Cytochem* 22(7):531-535.
102. Verhest AP (1994) Scanning devices for cytogenetic analysis. *Compendium on the Computerized Cytology and Histology Laboratory*, eds Wied GL, Bartels PH, Rosenthal PH, & Schenck U (Tutorials of Cytology, Chicago), pp 278-283.
103. Vooijs GP, van't Hof Grootenboer BE, van Aspert van Erp AJM, de Schipper FA, Pahlplatz MMM, & Hanselaar TGJM (1994) Accuracy of cytologic diagnosis: Light microscopic and cytometric criteria. *Compendium on the Computerized Cytology and*

- Histology Laboratory*, eds Wied GL, Bartels PH, Rosenthal PH, & Schenck U (Tutorials of Cytology, Chicago), pp 32-42.
104. Bahr GF, Bartels PH, Bibbo M, De Nicolas M, & Wied GL (1973) Evaluation of the Papanicolaou stain for computer assisted cellular pattern recognition. *Acta Cytol* 17(2):106-112.
 105. Bartels PH, Bibbo M, Bahr GF, Taylor J, & Wied GL (1973) Cervical cytology: descriptive statistics for nuclei of normal and atypical cell types. *Acta Cytol* 17(5):449-453.
 106. Bartels PH & Wied GL (1974) Performance testing for automated prescreening devices in cervical cytology. *J Histochem Cytochem* 22(7):660-662.
 107. Bibbo M, Bartels PH, Bahr GF, Taylor J, & Wied GL (1973) Computer recognition of cell nuclei from the uterine cervix. *Acta Cytol* 17(4):340-350.
 108. Birdsong GG (1996) Automated screening of cervical cytology specimens. *Hum Pathol* 27(5):468-481.
 109. Blomster J & Stenkvist B (1976) Classification of cells in vaginal smears with television microscopy. *Virchows Arch B Cell Pathol* 20(3):205-216.
 110. Burger G, Jutting U, & Rodenacker K (1981) Changes in benign cell populations in cases of cervical cancer and its precursors. *Anal Quant Cytol* 3(4):261-271.
 111. Dawson IM, Heanley CP, Heber-Percy AC, & Tylko JK (1967) Cesar: cervical smear analyser and reader. A new approach to evaluating cells in cytological preparations. *J Clin Pathol* 20(5):724-730.
 112. Diacumakos EG, Day E, & Kopac MJ (1962) Exfoliated cell studies and the cytoanalyzer. *Ann N Y Acad Sci* 97:498-513.
 113. Holmquist J, Bengtsson E, Eriksson O, Nordin B, & Stenkvist B (1978) Computer analysis of cervical cells. Automatic feature extraction and classification. *J Histochem Cytochem* 26(11):1000-1017.
 114. Holmquist J, Bengtsson E, Olsen B, Stenkvist B, & Noguchi Y (1976) Analysis of grey-level histograms as a method of classifying cells from the uterine cervix. *Comput Biol Med* 6(3):213-223.
 115. McMaster GW (1968) A measurement of the pattern of normal and malignant cervical cells. *Acta Cytol* 12(1):9-14.
 116. Mellors RC & Silver R (1951) A microfluorometric scanner for the differential detection of cells; application to exfoliative cytology. *Science* 114(2962):356-360.
 117. Oliver LH, Poulsen RS, Toussaint GT, & Louis C (1979) Classification of Atypical Cells in the Automatic Cyto-Screening for Cervical-Cancer. *Pattern Recognit* 11(3):205-212.

118. Rosenberg SA, Ledeen KS, & Kline T (1969) Automatic identification and measurement of cells by computer. *Science* 163(871):1065-1067.
119. Stenkvis B, Bergstrom R, Brinne U, Hesselius I, Kiviranta A, Nordgren H, Schnurer L, Stendahl U, Stenson S, Soderstrom J, & Sorensen CS (1987) Automatic analysis of Papanicolaou smears by digital image processing. *Gynecol Oncol* 27(1):1-14.
120. Tanaka N, Ikeda H, Ueno T, Mukawa A, Watanabe S, Okamoto K, Hosoi S, & Tsunekawa S (1987) Automated cytologic screening system (CYBEST model 4): an integrated image cytometry system. *Appl Opt* 26(16):3301-3307.
121. Tanaka N, Ikeda H, Ueno T, Watanabe S, & Imasato Y (1977) Fundamental study of automatic cyto-screening for uterine cancer. III. New system of automated apparatus (CYBEST) utilizing the pattern recognition method. *Acta Cytol* 21(1):85-89.
122. Tolles WE, Horvath WJ, & Bostrom RC (1961) A study of the quantitative characteristics of exfoliated cells from the female genital tract. II. Suitability of quantitative cytological measurements for automatic prescreening. *Cancer* 14:455-468.
123. Tolles WE, Horvath WJ, & Bostrom RC (1961) A study of the quantitative characteristics of exfoliated cells from the female genital tract. I. Measurement methods and results. *Cancer* 14:437-454.
124. Tucker JH (1976) CERVISCAN: an image analysis system for experiments in automatic cervical smear prescreening. *Comput Biomed Res* 9(2):93-107.
125. Taylor J, Bahr GF, Bartels PH, Bibbo M, Richards DL, & Wied GL (1975) Development and evaluation of automatic nucleus finding routines: thresholding of cervical cytology images. *Acta Cytol* 19(3):289-298.
126. Bengtsson E, Eriksson O, Holmquist J, & Stenkvis B (1977) A software system to record and analyze digitized cell images. *Comput Programs Biomed* 7(4):233-246.
127. Bengtsson E, Holmquist J, Olsen B, & Stenkvis B (1976) SCANCANS--an interactive scanning cell analysis system. *Comput Programs Biomed* 6(1):39-49.
128. Bacus JV (1994) The CAS 200™ MultiScan™ automated pathology workstation. *Compendium on the Computerized Cytology and Histology Laboratory*, eds Wied GL, Bartels PH, Rosenthal PH, & Schenck U (Tutorials of Cytology, Chicago), pp 360-367.
129. Koss LG, Bartels PH, Sychra JJ, & Wied GL (1978) Diagnostic cytologic sample profiles in patients with bladder cancer using TICAS system. *Acta Cytol* 22(5):392-397.
130. Koss LG, Bartels PH, Bibbo M, Freed SZ, Taylor J, & Wied GL (1975) Computer discrimination between benign and malignant urothelial cells. *Acta Cytol* 19(4):378-391.
131. Koss LG, Bartels PH, & Wied GL (1980) Computer-based diagnostic analysis of cells in the urinary sediment. *J Urol* 123(6):846-849.
132. Mango LJ (1994) Computer-assisted cervical cancer screening using neural networks. *Cancer Lett* 77(2-3):155-162.

133. Ashfaq R, Saliger F, Solares B, Thomas S, Liu G, Liang Y, & Saboorian MH (1997) Evaluation of the PAPNET system for prescreening triage of cervicovaginal smears. *Acta Cytol* 41(4):1058-1064.
134. Bergeron C, Masseroli M, Ghezi A, Lemarie A, Mango L, & Koss LG (2000) Quality control of cervical cytology in high-risk women. PAPNET system compared with manual rescreening. *Acta Cytol* 44(2):151-157.
135. Wilbur DC & Norton MK (2002) The Primary Screening Clinical Trials of the TriPath AutoPap® System. *Epidemiology* 13(3):S30-S33.
136. Schiffman M & Solomon D (2009) Screening and Prevention Methods for Cervical Cancer. *JAMA* 302(16):1809-1810.
137. Zahniser DJ (1994) The ThinPrep® processor and the CDS-1000™ cytology workstation. *Compendium on the Computerized Cytology and Histology Laboratory*, eds Wied GL, Bartels PH, Rosenthal PH, & Schenck U (Tutorials of Cytology, Chicago), pp 389-398.
138. Siebers AG, Klinkhamer PJJM, Grefte JMM, Massuger LFAG, Vedder JEM, Beijers-Broos A, Bulten J, & Arbyn M (2009) Comparison of Liquid-Based Cytology With Conventional Cytology for Detection of Cervical Cancer Precursors. *JAMA* 302(16):1757-1764.
139. Arbyn M, Bergeron C, Klinkhamer P, Martin-Hirsch P, Siebers AG, & Bulten J (2008) Liquid Compared With Conventional Cervical Cytology: A Systematic Review and Meta-analysis. *Obstet Gynecol* 111(1):167-177.
140. Miller FS, Nagel LE, & Kenny-Moynihan MB (2007) Implementation of the ThinPrep imaging system in a high-volume metropolitan laboratory. *Diagn Cytopathol* 35(4):213-217.
141. Lozano R (2007) Comparison of computer-assisted and manual screening of cervical cytology. *Gynecol Oncol* 104(1):134-138.
142. Biscotti CV, Dawson AE, Dziura B, Galup L, Darragh T, Rahemtulla A, & Wills-Frank L (2005) Assisted primary screening using the automated ThinPrep Imaging System. *Am J Clin Pathol* 123(2):281-287.
143. Kjellstrand P (1980) Mechanisms of the Feulgen acid hydrolysis. *J Microsc* 119(3):391-396.
144. Schulte EK & Wittekind DH (1990) Standardization of the Feulgen reaction: the influence of chromatin condensation on the kinetics of acid hydrolysis. *Anal Cell Pathol* 2(3):149-157.
145. Reith A & Danielsen H (1994) Assessment of DNA ploidy in tumor material: Preparation and measurement by image cytometry. *Compendium on the Computerized Cytology and Histology Laboratory*, eds Wied GL, Bartels PH, Rosenthal PH, & Schenck U (Tutorials of Cytology, Chicago), pp 185-193.

146. Biesterfeld S, Beckers S, Del Carmen Villa Cadenas M, & Schramm M (2011) Feulgen Staining Remains the Gold Standard for Precise DNA Image Cytometry. *Anticancer Res* 31(1):53-58.
147. Haroske G, Baak JP, Danielsen H, Giroud F, Gschwendtner A, Oberholzer M, Reith A, Spieler P, & Bocking A (2001) Fourth updated ESACP consensus report on diagnostic DNA image cytometry. *Anal Cell Pathol* 23(2):89-95.
148. Böcking A, Adler CP, Common HH, Hilgarth M, Granzen B, & Auffermann W (1984) Algorithm for a DNA-cytophotometric diagnosis and grading of malignancy. *Anal Quant Cytol* 6(1):1-8.
149. Böcking A, Striepecke E, Auer H, & Füzesi L (1994) Static DNA cytometry: Biological background, technique and diagnostic interpretation. *Compendium on the Computerized Cytology and Histology Laboratory*, eds Wied GL, Bartels PH, Rosenthal PH, & Schenck U (Tutorials of Cytology, Chicago), pp 107-128.
150. Auer GU, Caspersson TO, & Wallgren AS (1980) DNA content and survival in mammary carcinoma. *Anal Quant Cytol* 2(3):161-165.
151. Tong H, Shen R, Wang Z, Kan Y, Wang Y, Li F, Wang F, Yang J, Guo X, & the Mass Cervical Cancer Screening Regimen Group (2009) DNA ploidy cytometry testing for cervical cancer screening in China (DNACIC Trial): a prospective randomized, controlled trial. *Clin Cancer Res* 15(20):6438-6445.
152. Guillaud M, Benedet JL, Cantor SB, Staerke G, Follen M, & MacAulay C (2006) DNA ploidy compared with human papilloma virus testing (Hybrid Capture II) and conventional cervical cytology as a primary screening test for cervical high-grade lesions and cancer in 1555 patients with biopsy confirmation. *Cancer* 107(2):309-318.
153. Sun XR, Wang J, Garner D, & Palcic B (2005) Detection of cervical cancer and high grade neoplastic lesions by a combination of liquid-based sampling preparation and DNA measurements using automated image cytometry. *Cell Oncol* 27(1):33-41.
154. Lorenzato M, Caudroy S, Nou JM, Dalstein V, Joseph K, Bellefqih S, Durlach A, Thil C, Dez F, Bouttens D, Clavel C, & Birembaut P (2008) Contribution of DNA ploidy image cytometry to the management of ASC cervical lesions. *Cancer Cytopathol* 114(4):263-269.
155. Hanselaar AG, Vooijs GP, Oud PS, Pahlplatz MM, & Beck JL (1988) DNA ploidy patterns in cervical intraepithelial neoplasia grade III, with and without synchronous invasive squamous cell carcinoma. Measurements in nuclei isolated from paraffin-embedded tissue. *Cancer* 62(12):2537-2545.
156. Hanselaar AG, Vooijs GP, Van't Hof-Grootenboer AE, Gemmink JH, De Leeuw H, & Pahlplatz MM (1991) Cytophotometric analysis of corresponding cytological and histological cervical intraepithelial neoplasia grade III specimens. *Cytometry* 12(1):1-9.
157. Hanselaar AG, Vooijs GP, Van't Hof-Grootenboer AE, & Pahlplatz MM (1990) Cytophotometric analysis of cervical intraepithelial neoplasia grade III, with and without synchronous invasive squamous cell carcinoma. *Cytometry* 11(8):901-906.

158. Mariuzzi G, Santinelli A, Valli M, Sisti S, Montironi R, Mariuzzi L, Alberti R, & Pisani E (1992) Cytometric evidence that cervical intraepithelial neoplasia I and II are dysplasias rather than true neoplasias. An image analysis study of factors involved in the progression of cervical lesions. *Anal Quant Cytol Histol* 14(2):137-147.
159. Böcking A, Hilgarth M, Auffermann W, Hack-Werdier C, Fischer-Becker D, & von Kalkreuth G (1986) DNA-cytometric diagnosis of prospective malignancy in borderline lesions of the uterine cervix. *Acta Cytol* 30(6):608-615.
160. Tammemagi MC, Lam SC, McWilliams AM, & Sin DD (2011) Incremental value of pulmonary function and sputum DNA image cytometry in lung cancer risk prediction. *Cancer Prev Res (Phila)* 4(4):552-561.
161. Xing S, Khanavkar B, Nakhosteen JA, Atay Z, Jöckel K-H, Marek W, & the Research Institute for Diagnosis Treatment of Early Lung Cancer (RIDTELC) Lung Study Group (2005) Predictive value of image cytometry for diagnosis of lung cancer in heavy smokers. *Eur Respir J* 25(6):956-963.
162. Schramm M, Wrobel C, Born I, Kazimirek M, Pomjanski N, William M, Kappes R, Gerharz CD, Biesterfeld S, & Böcking A (2011) Equivocal cytology in lung cancer diagnosis. *Cancer Cytopathol* 119(3):177-192.
163. Auffermann W & Böcking A (1985) Early detection of precancerous lesions in dysplasias of the lung by rapid DNA image cytometry. *Anal Quant Cytol Histol* 7(3):218-226.
164. Santos-Silva AR, Ribeiro ACP, Soubhia AMP, Miyahara GI, Carlos R, Speight PM, Hunter KD, Torres-Rendon A, Vargas PA, & Lopes MA (2011) High incidences of DNA ploidy abnormalities in tongue squamous cell carcinoma of young patients: an international collaborative study. *Histopathology* 58(7):1127-1135.
165. Bremmer JF, Brakenhoff RH, Broeckaert MA, Beliën JA, Leemans CR, Bloemena E, van der Waal I, & Braakhuis BJ (2011) Prognostic value of DNA ploidy status in patients with oral leukoplakia. *Oral Oncol* 47(10):956-960.
166. Remmerbach TW, Weidenbach H, Pomjanski N, Knops K, Mathes S, Hemprich A, & Böcking A (2001) Cytologic and DNA-cytometric early diagnosis of oral cancer. *Anal Cell Pathol* 22(4):211-221.
167. Stenkivist B & Olding-Stenkivist E (1990) Cytological and DNA characteristics of hyperplasia/inflammation and cancer of the prostate. *Eur J Cancer* 26(3):261-267.
168. Zetterberg A & Forsslund G (1991) Ploidy level and tumor progression in prostatic carcinoma. *Acta Oncol* 30(2):193-199.
169. Hedley DW, Friedlander ML, Taylor IW, Rugg CA, & Musgrove EA (1983) Method for analysis of cellular DNA content of paraffin-embedded pathological material using flow cytometry. *J Histochem Cytochem* 31(11):1333-1335.
170. Boardman LA, Johnson RA, Petersen GM, Oberg AL, Kabat BF, Slusser JP, Wang L, Morlan BW, French AJ, Smyrk TC, Lindor NM, & Thibodeau SN (2007) Higher

- Frequency of Diploidy in Young-Onset Microsatellite-Stable Colorectal Cancer. *Clin Cancer Res* 13(8):2323-2328.
171. Iwaizumi M, Shinmura K, Mori H, Yamada H, Suzuki M, Kitayama Y, Igarashi H, Nakamura T, Suzuki H, Watanabe Y, Hishida A, Ikuma M, & Sugimura H (2009) Human Sgo1 downregulation leads to chromosomal instability in colorectal cancer. *Gut* 58(2):249-260.
 172. Moura MM, Cavaco BM, Pinto AE, & Leite V (2011) High Prevalence of RAS Mutations in RET-Negative Sporadic Medullary Thyroid Carcinomas. *J Clin Endocrinol Metab* 96(5):E863-E868.
 173. Hedley DW (1989) Flow cytometry using paraffin-embedded tissue: five years on. *Cytometry* 10(3):229-241.
 174. Hedley DW, Friedlander ML, & Taylor IW (1985) Application of DNA flow cytometry to paraffin-embedded archival material for the study of aneuploidy and its clinical significance. *Cytometry* 6(4):327-333.
 175. Cornelisse CJ, van de Velde CJ, Caspers RJ, Moolenaar AJ, & Hermans J (1987) DNA ploidy and survival in breast cancer patients. *Cytometry* 8(2):225-234.
 176. Zimmerman PV, Hawson GA, Bint MH, & Parsons PG (1987) Ploidy as a prognostic determinant in surgically treated lung cancer. *Lancet* 2(8558):530-533.
 177. MacAulay C, Lam S, Payne PW, LeRiche JC, & Palcic B (1995) Malignancy-associated changes in bronchial epithelial cells in biopsy specimens. *Anal Quant Cytol Histol* 17(1):55-61.
 178. Gruner OC (1916) A study of the changes met with in the leucocytes in certain cases of malignant disease. *Brit J Surg* 3(11):506-525.
 179. Nieburgs HE (1968) Recent progress in the interpretation of malignancy associated changes (MAC). *Acta Cytol* 12(6):445-453.
 180. Palcic B, Garner DM, Beveridge J, Sun XR, Doudkine A, MacAulay C, Lam S, & Payne PW (2002) Increase of sensitivity of sputum cytology using high-resolution image cytometry: Field study results. *Cytometry* 50(3):168-176.
 181. Klawe H & Rowinski J (1974) Malignancy associated changes (MAC) in cells of buccal smears detected by means of objective image analysis. *Acta Cytol* 18(1):30-33.
 182. Us-Krasovec M, Erzen J, Zganec M, Stojan-Fleazar M, Lavrencak J, Garner D, Doudkine A, & Palcic B (2005) Malignancy associated changes in epithelial cells of buccal mucosa: a potential cancer detection test. *Anal Quant Cytol Histol* 27(5):254-262.
 183. Mommers ECM, Poulin N, Sangulin J, Meijer CJLM, Baak JPA, & van Diest PJ (2001) Nuclear cytometric changes in breast carcinogenesis. *J Pathol* 193(1):33-39.
 184. Anderson G, MacAulay C, Matisic J, Garner D, & Palcic B (1997) The use of an automated image cytometer for screening and quantitative assessment of cervical lesions

- in the British Columbia Cervical Smear Screening Programme. *Cytopathology* 8(5):298-312.
185. Guillaud M, Cox D, Adler-Storthz K, Malpica A, Staerke G, Matisic J, Van Niekerk D, Poulin N, Follen M, & MacAulay C (2004) Exploratory analysis of quantitative histopathology of cervical intraepithelial neoplasia: Objectivity, reproducibility, malignancy-associated changes, and human papillomavirus. *Cytometry A* 60A(1):81-89.
 186. Garner DM, Harrison A, MacAulay C, & Palcic B (1994) Cyto-Savant™ and its use in automated screening of cervical smears. *Compendium on the Computerized Cytology and Histology Laboratory*, eds Wied GL, Bartels PH, Rosenthal PH, & Schenck U (Tutorials of Cytology, Chicago), pp 346-352.
 187. Palcic B & MacAulay C (1994) Malignancy associated changes: Can they be employed clinically? *Compendium on the Computerized Cytology and Histology Laboratory*, eds Wied GL, Bartels PH, Rosenthal PH, & Schenck U (Tutorials of Cytology, Chicago), pp 157-165.
 188. Shirata NK, Longatto Filho A, Roteli-Martins C, Espoladore LM, Pittoli JE, & Syrjanen K (2003) Applicability of liquid-based cytology to the assessment of DNA content in cervical lesions using static cytometry. *Anal Quant Cytol Histol* 25(4):210-214.
 189. Bibbo M, Michelassi F, Bartels PH, Dytch H, Bania C, Lerma E, & Montag AG (1990) Karyometric marker features in normal-appearing glands adjacent to human colonic adenocarcinoma. *Cancer Res* 50(1):147-151.
 190. Montag AG, Bartels PH, Dytch HE, Lerma-Puertas E, Michelassi F, & Bibbo M (1991) Karyometric features in nuclei near colonic adenocarcinoma. Statistical analysis. *Anal Quant Cytol Histol* 13(3):159-167.
 191. Kemp RA, Reinders DM, & Turic B (2007) Detection of lung cancer by automated sputum cytometry. *J Thorac Oncol* 2(11):993-1000.
 192. Payne PW, Sebo TJ, Doudkine A, Garner D, MacAulay C, Lam S, LeRiche JC, & Palcic B (1997) Sputum screening by quantitative microscopy: A reexamination of a portion of the National Cancer Institute Cooperative Early Lung Cancer Study. *Mayo Clin Proc* 72(8):697-704.
 193. Poh CF, MacAulay CE, Laronde DM, Williams PM, Zhang L, & Rosin MP (2011) Squamous cell carcinoma and precursor lesions: diagnosis and screening in a technical era. *Periodontol* 2000 57(1):73-88.
 194. Guillaud M, Adler-Storthz K, Malpica A, Staerke G, Matisic J, Van Niekerk D, Cox D, Poulin N, Follen M, & MacAulay C (2005) Subvisual chromatin changes in cervical epithelium measured by texture image analysis and correlated with HPV. *Gynecol Oncol* 99(3 Suppl 1):S16-23.
 195. Scheurer ME, Guillaud M, Tortolero-Luna G, MacAulay C, Follen M, & Adler-Storthz K (2007) Human papillomavirus-related cellular changes measured by cytometric analysis of DNA ploidy and chromatin texture. *Cytometry B Clin Cytom* 72B(5):324-331.

196. Lam S, Kennedy T, Unger M, Miller YE, Gelmont D, Rusch V, Gipe B, Howard D, LeRiche JC, Coldman A, & Gazdar AF (1998) Localization of bronchial intraepithelial neoplastic lesions by fluorescence bronchoscopy. *Chest* 113(3):696-702.
197. Guillaud M, le Riche JC, Dawe C, Korbelik J, Coldman A, Wistuba II, Park I-W, Gazdar A, Lam S, & MacAulay CE (2005) Nuclear morphometry as a biomarker for bronchial intraepithelial neoplasia: Correlation with genetic damage and cancer development. *Cytometry A* 63(1):34-40.
198. Smitha T, Sharada P, & Girish HC (2011) Morphometry of the basal cell layer of oral leukoplakia and oral squamous cell carcinoma using computer-aided image analysis. *J Oral Maxillofac Pathol* 15(1):26-33.
199. Song J & Shea CR (2010) Benign versus malignant parakeratosis: a nuclear morphometry study. *Mod Pathol* 23(6):799-803.
200. Priya SS & Sundaram S (2011) Morphology to morphometry in cytological evaluation of thyroid lesions. *J Cytol* 28(3):98-102.
201. Gratzner HG, Young IT, & Sher SE (1979) An immunocytochemical approach to cell kinetics automation. *J Histochem Cytochem* 27(1):496-499.
202. Schätzle P, Wuttke R, Ziegler U, & Sonderegger P (2011) Automated quantification of synapses by fluorescence microscopy. *J Neurosci Methods* 204(1):144-149.
203. Gross DS & Rothfeld JM (1985) Quantitative immunocytochemistry of hypothalamic and pituitary hormones: validation of an automated, computerized image analysis system. *J Histochem Cytochem* 33(1):11-20.
204. Natarajan S, Mahajan S, Boaz K, & George T (2010) Prediction of lymph node metastases by preoperative nuclear morphometry in oral squamous cell carcinoma: a comparative image analysis study. *Indian J Cancer* 47(4):406-411.
205. Trachtenberg AJ, Robert JH, Abdalla AE, Fraser A, He SY, Lacy JN, Rivas-Morello C, Truong A, Hardiman G, Ohno-Machado L, Liu F, Hovig E, & Kuo WP (2012) A primer on the current state of microarray technologies. *Methods Mol Biol* 802:3-17.
206. Schena M, Shalon D, Davis RW, & Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270(5235):467-470.
207. DeRisi J, Penland L, Brown PO, Bittner ML, Meltzer PS, Ray M, Chen Y, Su YA, & Trent JM (1996) Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat Genet* 14(4):457-460.
208. Leung YF & Cavalieri D (2003) Fundamentals of cDNA microarray data analysis. *Trends Genet* 19(11):649-659.
209. Kroll TC & Wolf S (2002) Ranking: a closer look on globalisation methods for normalisation of gene expression arrays. *Nucleic Acids Res* 30(11):e50.

210. Quackenbush J (2002) Microarray data normalization and transformation. *Nat Genet* 32 Suppl:496-501.
211. Smyth GK & Speed T (2003) Normalization of cDNA microarray data. *Methods* 31(4):265-273.
212. Dudoit S, Yang YH, Callow MJ, & Speed TP (2002) Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. *Stat Sin* 12(1):111-139.
213. Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J, & Speed TP (2002) Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* 30(4):e15.
214. Smoot ME, Ono K, Ruscheinski J, Wang PL, & Ideker T (2011) Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 27(3):431-432.
215. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, & Mesirov JP (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102(43):15545-15550.
216. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstrale M, Laurila E, Houstis N, Daly MJ, Patterson N, Mesirov JP, Golub TR, Tamayo P, Spiegelman B, Lander ES, Hirschhorn JN, Altshuler D, & Groop LC (2003) PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* 34(3):267-273.
217. Huang DW, Sherman BT, & Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4(1):44-57.
218. Huang DW, Sherman BT, & Lempicki RA (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 37(1):1-13.
219. Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, Guo N, Ladunga I, Ulitsky-Lazareva B, Muruganujan A, Rabkin S, Vandergriff JA, & Doremieux O (2003) PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucleic Acids Res* 31(1):334-341.
220. Werner M, Chott A, Fabiano A, & Battifora H (2000) Effect of formalin tissue fixation and processing on immunohistochemistry. *Am J Surg Pathol* 24(7):1016-1019.
221. Battifora H & Kopinski M (1986) The influence of protease digestion and duration of fixation on the immunostaining of keratins. A comparison of formalin and ethanol fixation. *J Histochem Cytochem* 34(8):1095-1100.
222. Curran RC & Gregory J (1977) The unmasking of antigens in paraffin sections of tissue by trypsin. *Experientia* 33(10):1400-1401.

223. Shi SR, Key ME, & Kalra KL (1991) Antigen retrieval in formalin-fixed, paraffin-embedded tissues: an enhancement method for immunohistochemical staining based on microwave oven heating of tissue sections. *J Histochem Cytochem* 39(6):741-748.
224. Shi SR, Cote RJ, & Taylor CR (2001) Antigen retrieval techniques: current perspectives. *J Histochem Cytochem* 49(8):931-937.
225. Werner M, Von Wasielewski R, & Komminoth P (1996) Antigen retrieval, signal amplification and intensification in immunohistochemistry. *Histochem Cell Biol* 105(4):253-260.
226. Morgan JM, Navabi H, & Jasani B (1997) Role of calcium chelation in high-temperature antigen retrieval at different pH values. *J Pathol* 182(2):233-237.
227. Shi SR, Cote RJ, Hawes D, Thu S, Shi Y, Young LL, & Taylor CR (1999) Calcium-induced modification of protein conformation demonstrated by immunohistochemistry: What is the signal? *J Histochem Cytochem* 47(4):463-470.
228. Straus W (1971) Inhibition of peroxidase by methanol and by methanol-nitroferricyanide for use in immunoperoxidase procedures. *J Histochem Cytochem* 19(11):682-688.
229. Shaipanich T, McWilliams A, & Lam S (2006) Early detection and chemoprevention of lung cancer. *Respirology* 11(4):366-372.
230. Hirsch FR, Merrick DT, & Franklin WA (2002) Role of biomarkers for early detection of lung cancer and chemoprevention. *Eur Respir J* 19(6):1151-1158.
231. Spitz MR, Etzel CJ, Dong Q, Amos CI, Wei Q, Wu X, & Hong WK (2008) An expanded risk prediction model for lung cancer. *Cancer Prev Res (Phila)* 1(4):250-254.
232. Bremnes RM, Sirera R, & Camps C (2005) Circulating tumour-derived DNA and RNA markers in blood: a tool for early detection, diagnostics, and follow-up? *Lung Cancer* 49(1):1-12.
233. Wang YC, Hsu HS, Chen TP, & Chen JT (2006) Molecular diagnostic markers for lung cancer in sputum and plasma. *Ann N Y Acad Sci* 1075:179-184.
234. Roy HK, Subramanian H, Damania D, Hensing TA, Rom WN, Pass HI, Ray D, Rogers JD, Bogojevic A, Shah M, Kuzniar T, Pradhan P, & Backman V (2010) Optical detection of buccal epithelial nanoarchitectural alterations in patients harboring lung cancer: Implications for screening. *Cancer Res* 70(20):7748-7754.
235. Katz RL, Zaidi TM, Fernandez RL, Zhang J, He W, Acosta C, Daniely M, Madi L, Vargas MA, Dong Q, Gao X, Jiang F, Caraway NP, Vaporciyan AA, Roth JA, & Spitz MR (2008) Automated detection of genetic abnormalities combined with cytology in sputum is a sensitive predictor of lung cancer. *Mod Pathol* 21(8):950-960.
236. Varella-Garcia M, Schulte AP, Wolf HJ, Feser WJ, Zeng C, Braudrick S, Yin X, Hirsch FR, Kennedy TC, Keith RL, Baron AE, Belinsky SA, Miller YE, Byers T, & Franklin WA (2010) The detection of chromosomal aneusomy by fluorescence *in situ*

- hybridization in sputum predicts lung cancer incidence. *Cancer Prev Res (Phila)* 3(4):447-453.
237. Schreiber G & McCrory DC (2003) Performance characteristics of different modalities for diagnosis of suspected lung cancer: Summary of published evidence. *Chest* 123(1 Suppl):115S-128S.
 238. Tercelj M, Ales A, Rott T, Sever N, Prodnik L, Erzen J, Primic-Zakelj M, & Rylander R (2008) DNA-based sputum cell image analysis for lung cancer in a clinical setting. *Acta Cytol* 52(5):584-590.
 239. Doudkine A, MacAulay C, Poulin N, & Palcic B (1995) Nuclear texture measurements in image cytometry. *Pathologica* 87(3):286-299.
 240. Chiu D, Guillaud M, Cox D, Follen M, & MacAulay C (2004) Quality assurance system using statistical process control: an implementation for image cytometry. *Cell Oncol* 26(3):101-117.
 241. Guillaud M, Cox D, Malpica A, Staerckel G, Maticic J, Van Niekirk D, Adler-Storthz K, Poulin N, Follen M, & MacAulay C (2004) Quantitative histopathological analysis of cervical intra-epithelial neoplasia sections: methodological issues. *Cell Oncol* 26(1-2):31-43.
 242. MacAulay CE (1989) Development, implementation and evaluation of segmentation algorithms for the automatic classification of cervical cells. Ph.D. Dissertation (The University of British Columbia, Vancouver).
 243. Palcic B, MacAulay C, Shlien S, Treurniet W, Tezcan H, & Anderson G (1992) Comparison of three different methods for automated classification of cervical cells. *Anal Cell Pathol* 4(6):429-441.
 244. Travis WD, Colby TV, Corrin B, Shimosato Y, & Brambilla E (1999) Histologic and graphical text slides for the histological typing of lung and pleural tumors. in *International histological classification of tumors*, ed World Health Organization Pathology Panel (Springer Verlag, Berlin), p 5.
 245. Cassidy A, Duffy SW, Myles JP, Liloglou T, & Field JK (2007) Lung cancer risk prediction: A tool for early detection. *Int J Cancer* 120(1):1-6.
 246. Kiyohara C & Ohno Y (2010) Sex differences in lung cancer susceptibility: A review. *Gend Med* 7(5):381-401.
 247. Gray J, Mao JT, Szabo E, Kelley M, Kurie J, & Bepler G (2007) Lung cancer chemoprevention: ACCP evidence-based clinical practice guidelines (2nd Edition). *Chest* 132(3 Suppl):56S-68S.
 248. Brambilla E, Gazzeri S, Lantuejoul S, Coll JL, Moro D, Negoescu A, & Brambilla C (1998) p53 mutant immunophenotype and deregulation of p53 transcription pathway (Bcl2, Bax, and Waf1) in precursor bronchial lesions of lung cancer. *Clin Cancer Res* 4(7):1609-1618.

249. Scholzen T & Gerdes J (2000) The Ki-67 protein: From the known and the unknown. *J Cell Physiol* 182(3):311-322.
250. Brown DC & Gatter KC (2002) Ki67 protein: The immaculate deception? *Histopathology* 40(1):2-11.
251. Zhu C-Q, Shih W, Ling C-H, & Tsao M-S (2006) Immunohistochemical markers of prognosis in non-small cell lung cancer: A review and proposal for a multiphase approach to marker evaluation. *J Clin Pathol* 59(8):790-800.
252. Meert AP, Feoli F, Martin B, Verdebout JM, Mascaux C, Verhest A, Ninane V, & Sculier JP (2004) Ki67 expression in bronchial preneoplastic lesions and carcinoma *in situ* defined according to the new 1999 WHO/IASLC criteria: A preliminary study. *Histopathology* 44(1):47-53.
253. Lee JJ, Liu D, Lee JS, Kurie JM, Khuri FR, Ibarguen H, Morice RC, Walsh G, Ro JY, Broxson A, Hong WK, & Hittelman WN (2001) Long-term impact of smoking on lung epithelial proliferation in current and former smokers. *J Natl Cancer Inst* 93(14):1081-1088.
254. Khuri FR, Lee JS, Lippman SM, Lee JJ, Kalapurakal S, Yu R, Ro JY, Morice RC, Hong WK, & Hittelman WN (2001) Modulation of proliferating cell nuclear antigen in the bronchial epithelium of smokers. *Cancer Epidemiol Biomarkers Prev* 10(4):311-318.
255. Szulakowski P, Crowther AJ, Jimenez LA, Donaldson K, Mayer R, Leonard TB, MacNee W, & Drost EM (2006) The effect of smoking on the transcriptional regulation of lung inflammation in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 174(1):41-50.
256. Brenner DJ (2004) Radiation risks potentially associated with low-dose CT screening of adult smokers for lung cancer. *Radiology* 231(2):440-445.
257. Bibbo M, Bartels PH, Dytch HE, & Wied GL (1985) Ploidy patterns in cervical dysplasia. *Anal Quant Cytol Histol* 7(3):213-217.
258. Aleskandarany M, Rakha E, Macmillan R, Powe D, Ellis I, & Green A (2011) MIB1/Ki-67 labelling index can classify grade 2 breast cancer into two clinically distinct subgroups. *Breast Cancer Res Treat* 127(3):591-599.
259. Yerushalmi R, Woods R, Ravdin PM, Hayes MM, & Gelmon KA (2010) Ki67 in breast cancer: prognostic and predictive potential. *Lancet Oncol* 11(2):174-183.
260. Zellweger T, Günther S, Zlobec I, Savic S, Sauter G, Moch H, Mattarelli G, Eichenberger T, Curschellas E, Rüfenacht H, Bachmann A, Gasser TC, Mihatsch MJ, & Bubendorf L (2009) Tumour growth fraction measured by immunohistochemical staining of Ki67 is an independent prognostic factor in preoperative prostate biopsies with small-volume or low-grade prostate cancer. *International Journal of Cancer* 124(9):2116-2123.
261. Sahebali S, Depuydt CE, Segers K, Vereecken AJ, Van Marck E, & Bogers JJ (2003) Ki-67 immunocytochemistry in liquid based cervical cytology: useful as an adjunctive tool? *J Clin Pathol* 56(9):681-686.

262. Van Niekerk D, Guillaud M, Matisic J, Benedet JL, Freeberg JA, Follen M, & MacAulay C (2007) p16 and MIB1 improve the sensitivity and specificity of the diagnosis of high grade squamous intraepithelial lesions: methodological issues in a report of 447 biopsies with consensus diagnosis and HPV HCII testing. *Gynecol Oncol* 107(1 Suppl 1):S233-240.
263. McCluggage WG (2007) Immunohistochemistry as a diagnostic aid in cervical pathology. *Pathology* 39(1):97-111.
264. Gupta N, Srinivasan R, & Rajwanshi A (2010) Functional biomarkers in cervical precancer: an overview. *Diagn Cytopathol* 38(8):618-623.
265. Landberg G, Tan EM, & Roos G (1990) Flow cytometric multiparameter analysis of proliferating cell nuclear antigen/cyclin and Ki-67 antigen: a new view of the cell cycle. *Exp Cell Res* 187(1):111-118.
266. Xiaoyan S, Xianglin Y, Deding T, Jianping G, & Guoqing H (2005) Analysis of DNA ploidy, cell cycle and Ki67 antigen in nasopharyngeal carcinoma by flow cytometry. *J Huazhong Univ Sci Technolog Med Sci* 25(2):198-201.
267. Singh M, Mehrotra S, Kalra N, Singh U, & Shukla Y (2008) Correlation of DNA ploidy with progression of cervical cancer. *J Cancer Epidemiol* 2008:298495.
268. Rovera G, Santoli D, & Damsky C (1979) Human promyelocytic leukemia cells in culture differentiate into macrophage-like cells when treated with a phorbol diester. *Proc Natl Acad Sci U S A* 76(6):2779-2783.
269. Cecic IK, Li G, & MacAulay C (2012) Technologies supporting analytical cytology: clinical, research and drug discovery applications. *J Biophotonics* 5(4):313-326.
270. Li G, Guillaud M, leRiche J, McWilliams A, Gazdar A, Lam S, & MacAulay C (2012) Automated sputum cytometry for detection of intraepithelial neoplasias in the lung. *Anal Cell Pathol (Amst)* 35(3):187-201.
271. Kolles H, Forderer W, Bock R, & Feiden W (1993) Combined Ki-67 and Feulgen stain for morphometric determination of the Ki-67 labelling index. *Histochemistry* 100(4):293-296.
272. Oud PS, Bauwens A, & Nauwelaers FA (1997) Multiparameter absorption measurements in automated microscopy. Simultaneous quantitative determination of DNA and nuclear antigen. *Acta Cytol* 41(1):188-196.
273. Fleskens SJHM, Takes RP, Otte-Höller I, Van Doesburg L, Smeets A, Speel E-JM, Slootweg PJ, & Van Der Laak JAWM (2010) Simultaneous assessment of DNA ploidy and biomarker expression in paraffin-embedded tissue sections. *Histopathology* 57(1):14-26.
274. Kruse A-J, Baak JPA, de Bruin PC, Jiwa M, Snijders WP, Jan Boodt P, Fons G, Houben PWH, & Sien The H (2001) Ki-67 immunoquantitation in cervical intraepithelial neoplasia (CIN): a sensitive marker for grading. *J Pathol* 193(1):48-54.

275. Baak JPA, Kruse A-J, Robboy SJ, Janssen EAM, van Diermen B, & Skaland I (2006) Dynamic behavioural interpretation of cervical intraepithelial neoplasia with molecular biomarkers. *J Clin Pathol* 59(10):1017-1028.
276. Hendzel MJ, Wei Y, Mancini MA, Van Hooser A, Ranalli T, Brinkley BR, Bazett-Jones DP, & Allis CD (1997) Mitosis-specific phosphorylation of histone H3 initiates primarily within pericentromeric heterochromatin during G2 and spreads in an ordered fashion coincident with mitotic chromosome condensation. *Chromosoma* 106(6):348-360.
277. Lorenzato M, Bory JP, Cucherousset J, Nou JM, Bouttens D, Thil C, Dez F, Evrard G, Quereux C, Birembaut P, & Clavel C (2002) Usefulness of DNA ploidy measurement on liquid-based smears showing conflicting results between cytology and high-risk human papillomavirus typing. *Am J Clin Pathol* 118(5):708-713.
278. Chen M (2004) Evaluation of applying Feulgen stain for DNA analysis on destained hematoxylin-eosin-stained cytologic smears. *Anal Quant Cytol Histol* 26(5):255-258.
279. Zeng HS, MacAulay C, Mclean DI, & Palcic B (1993) Novel Microspectrophotometer and Its Biomedical Applications. *Optical Engineering* 32(8):1809-1814.
280. Rocha R, Vassallo J, Soares F, Miller K, & Gobbi H (2009) Digital slides: Present status of a tool for consultation, teaching, and quality control in pathology. *Pathol Res Pract* 205(11):735-741.
281. Brachtel E & Yagi Y (2011 Preprint) Digital imaging in pathology - current applications and challenges. *J Biophotonics*.
282. Pantanowitz L, Valenstein P, Evans A, Kaplan K, Pfeifer J, Wilbur D, Collins L, & Colgan T (2011) Review of the current state of whole slide imaging in pathology. *J Pathol Inform* 2(1):36.
283. Wiltling SM, de Wilde J, Meijer CJ, Berkhof J, Yi Y, van Wieringen WN, Braakhuis BJ, Meijer GA, Ylstra B, Snijders PJ, & Steenbergen RD (2008) Integrated genomic and transcriptional profiling identifies chromosomal loci with altered gene expression in cervical cancer. *Genes Chromosomes Cancer* 47(10):890-905.
284. Manavi M, Hudelist G, Fink-Retter A, Gschwandtler-Kaulich D, Pischinger K, & Czerwenka K (2007) Gene profiling in Pap-cell smears of high-risk human papillomavirus-positive squamous cervical carcinoma. *Gynecol Oncol* 105(2):418-426.
285. Sun DJ, Liu Y, Lu DC, Kim W, Lee JH, Maynard J, & Deisseroth A (2007) Endothelin-3 growth factor levels decreased in cervical cancer compared with normal cervical epithelial cells. *Hum Pathol* 38(7):1047-1056.
286. Wong YF, Cheung TH, Tsao GS, Lo KW, Yim SF, Wang VW, Heung MM, Chan SC, Chan LK, Ho TW, Wong KW, Li C, Guo Y, Chung TK, & Smith DI (2006) Genome-wide gene expression profiling of cervical cancer in Hong Kong women by oligonucleotide microarray. *Int J Cancer* 118(10):2461-2469.

287. Ahn WS, Bae SM, Lee JM, Namkoong SE, Han SJ, Cho YL, Nam GH, Seo JS, Kim CK, & Kim YW (2004) Searching for pathogenic gene functions to cervical cancer. *Gynecol Oncol* 93(1):41-48.
288. Shim C, Zhang W, Rhee CH, & Lee JH (1998) Profiling of differentially expressed genes in human primary cervical cancer by complementary DNA expression array. *Clin Cancer Res* 4(12):3045-3050.
289. Gius D, Funk MC, Chuang EY, Feng S, Huettner PC, Nguyen L, Bradbury CM, Mishra M, Gao S, Buttin BM, Cohn DE, Powell MA, Horowitz NS, Whitcomb BP, & Rader JS (2007) Profiling microdissected epithelium and stroma to model genomic signatures for cervical carcinogenesis accommodating for covariates. *Cancer Res* 67(15):7113-7123.
290. Kendrick JE, Conner MG, & Huh WK (2007) Gene expression profiling of women with varying degrees of cervical intraepithelial neoplasia. *J Low Genit Tract Dis* 11(1):25-28.
291. Hudelist G, Czerwenka K, Singer C, Pischinger K, Kubista E, & Manavi M (2005) cDNA array analysis of cytobrush-collected normal and malignant cervical epithelial cells: a feasibility study. *Cancer Genet Cytogenet* 158(1):35-42.
292. Sopov I, Sorensen T, Magbagbeolu M, Jansen L, Beer K, Kuhne-Heid R, Kirchmayr R, Schneider A, & Durst M (2004) Detection of cancer-related gene expression profiles in severe cervical neoplasia. *Int J Cancer* 112(1):33-43.
293. Chen Y, Miller C, Mosher R, Zhao X, Deeds J, Morrissey M, Bryant B, Yang D, Meyer R, Cronin F, Gostout BS, Smith-McCune K, & Schlegel R (2003) Identification of cervical cancer markers by cDNA and tissue microarrays. *Cancer Res* 63(8):1927-1935.
294. Lewis F, Maughan NJ, Smith V, Hillan K, & Quirke P (2001) Unlocking the archive--gene expression in paraffin-embedded tissue. *J Pathol* 195(1):66-71.
295. Lehmann U & Kreipe H (2001) Real-time PCR analysis of DNA and RNA extracted from formalin-fixed and paraffin-embedded biopsies. *Methods* 25(4):409-418.
296. von Ahlfen S, Missel A, Bendrat K, & Schlumpberger M (2007) Determinants of RNA quality from FFPE samples. *PLoS One* 2(12):e1261.
297. Masuda N, Ohnishi T, Kawamoto S, Monden M, & Okubo K (1999) Analysis of chemical modification of RNA from formalin-fixed samples and optimization of molecular biology applications for such samples. *Nucleic Acids Res* 27(22):4436-4443.
298. Vincek V, Nassiri M, Nadji M, & Morales AR (2003) A tissue fixative that protects macromolecules (DNA, RNA, and protein) and histomorphology in clinical samples. *Lab Invest* 83(10):1427-1435.
299. Florell SR, Coffin CM, Holden JA, Zimmermann JW, Gerwels JW, Summers BK, Jones DA, & Leachman SA (2001) Preservation of RNA for functional genomic studies: a multidisciplinary tumor bank protocol. *Mod Pathol* 14(2):116-128.

300. Mutter GL, Zahrieh D, Liu C, Neuberg D, Finkelstein D, Baker HE, & Warrington JA (2004) Comparison of frozen and RNALater solid tissue storage methods for use in RNA expression microarrays. *BMC Genomics* 5:88.
301. Paska C, Bogi K, Szilak L, Tokes A, Szabo E, Sziller I, Rigo J, Jr., Sobel G, Szabo I, Kaposi-Novak P, Kiss A, & Schaff Z (2004) Effect of formalin, acetone, and RNALater fixatives on tissue preservation and different size amplicons by real-time PCR from paraffin-embedded tissue. *Diagn Mol Pathol* 13(4):234-240.
302. Gugic D, Nassiri M, Nadji M, Morales A, & Vincek V (2007) Novel tissue preservative and tissue fixative for comparative pathology and animal research. *J Exp Anim Sci* 43(4):271-281.
303. Ergin B, Meding S, Langer R, Kap M, Viertler C, Schott C, Ferch U, Riegman P, Zatloukal K, Walch A, & Becker KF (2010) Proteomic analysis of PAXgene-fixed tissues. *J Proteome Res* 9(10):5188-5196.
304. Stanta G, Mucelli SP, Petrera F, Bonin S, & Bussolati G (2006) A novel fixative improves opportunities of nucleic acids and proteomic analysis in human archive's tissues. *Diagn Mol Pathol* 15(2):115-123.
305. Delfour C, Roger P, Bret C, Berthe ML, Rochaix P, Kalfa N, Raynaud P, Bibeau F, Maudelonde T, & Boulle N (2006) RCL2, a new fixative, preserves morphology and nucleic acid integrity in paraffin-embedded breast carcinoma and microdissected breast tumor cells. *J Mol Diagn* 8(2):157-169.
306. Turashvili G, Yang W, McKinney S, Kalloger S, Gale N, Ng Y, Chow K, Bell L, Lorette J, Carrier M, Luk M, Aparicio S, Huntsman D, & Yip S (2012) Nucleic acid quantity and quality from paraffin blocks: defining optimal fixation, processing and DNA/RNA extraction techniques. *Exp Mol Pathol* 92(1):33-43.
307. Nassiri M, Ramos S, Zohourian H, Vincek V, Morales AR, & Nadji M (2008) Preservation of biomolecules in breast cancer tissue by a formalin-free histology system. *BMC Clin Pathol* 8:1.
308. Shadeo A, Chari R, Lonergan KM, Pusic A, Miller D, Ehlen T, Van Niekerk D, Matisic J, Richards-Kortum R, Follen M, Guillaud M, Lam WL, & MacAulay C (2008) Up regulation in gene expression of chromatin remodelling factors in cervical intraepithelial neoplasia. *BMC Genomics* 9:64.
309. Shadeo A, Chari R, Vatcher G, Campbell J, Lonergan KM, Matisic J, van Niekerk D, Ehlen T, Miller D, Follen M, Lam WL, & MacAulay C (2007) Comprehensive serial analysis of gene expression of the cervical transcriptome. *BMC Genomics* 8:142.
310. Eisen MB, Spellman PT, Brown PO, & Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 95(25):14863-14868.
311. van Niekerk D. Personal communication.

312. Penland SK, Keku TO, Torrice C, He X, Krishnamurthy J, Hoadley KA, Woosley JT, Thomas NE, Perou CM, Sandler RS, & Sharpless NE (2007) RNA expression analysis of formalin-fixed paraffin-embedded tumors. *Lab Invest* 87(4):383-391.
313. Masir N, Ghoddoosi M, Mansor S, Abdul-Rahman F, Florence CS, Mohamed-Ismael NA, Tamby MR, & Md-Latar NH (2012) RCL2, a potential formalin substitute for tissue fixation in routine pathological specimens. *Histopathology* 60(5):804-815.
314. Kalantari M, Garcia-Carranca A, Morales-Vazquez CD, Zuna R, Montiel DP, Calleja-Macias IE, Johansson B, Andersson S, & Bernard HU (2009) Laser capture microdissection of cervical human papillomavirus infections: copy number of the virus in cancerous and normal tissue and heterogeneous DNA methylation. *Virology* 390(2):261-267.
315. Chew GK, Cruickshank ME, Rooney PH, Miller ID, Parkin DE, & Murray GI (2005) Human papillomavirus 16 infection in adenocarcinoma of the cervix. *Br J Cancer* 93(11):1301-1304.
316. Domazet B, MacLennan GT, Lopez-Beltran A, Montironi R, & Cheng L (2008) Laser capture microdissection in the genomic and proteomic era: targeting the genetic basis of cancer. *Int J Clin Exp Pathol* 1(6):475-488.
317. Uhlen M, Bjorling E, Agaton C, Szigartyo CA, Amini B, Andersen E, Andersson AC, Angelidou P, Asplund A, Asplund C, Berglund L, Bergstrom K, Brumer H, Cerjan D, Ekstrom M, Elobeid A, Eriksson C, Fagerberg L, Falk R, Fall J, Forsberg M, Bjorklund MG, Gumbel K, Halimi A, Hallin I, Hamsten C, Hansson M, Hedhammar M, Hercules G, Kampf C, Larsson K, Lindskog M, Lodewyckx W, Lund J, Lundberg J, Magnusson K, Malm E, Nilsson P, Odling J, Oksvold P, Olsson I, Oster E, Ottosson J, Paavilainen L, Persson A, Rimini R, Rockberg J, Runeson M, Sivertsson A, Skollermo A, Steen J, Stenvall M, Sterky F, Stromberg S, Sundberg M, Tegel H, Tourle S, Wahlund E, Walden A, Wan J, Wernerus H, Westberg J, Wester K, Wrethagen U, Xu LL, Hober S, & Ponten F (2005) A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol Cell Proteomics* 4(12):1920-1932.
318. Uhlen M, Oksvold P, Fagerberg L, Lundberg E, Jonasson K, Forsberg M, Zwahlen M, Kampf C, Wester K, Hober S, Wernerus H, Bjorling L, & Ponten F (2010) Towards a knowledge-based Human Protein Atlas. *Nat Biotechnol* 28(12):1248-1250.
319. Agarwal SM, Raghav D, Singh H, & Raghava GP (2011) CCDB: a curated database of genes involved in cervix cancer. *Nucleic Acids Res* 39(Database issue):D975-979.
320. Hammes LS, Tekmal RR, Naud P, Edelweiss MI, Kirma N, Valente PT, Syrjanen KJ, & Cunha-Filho JS (2008) Up-regulation of VEGF, c-fms and COX-2 expression correlates with severity of cervical cancer precursor (CIN) lesions and invasive disease. *Gynecol Oncol* 110(3):445-451.
321. Kneller JM, Ehlen T, Matisic JP, Miller D, Van Niekerk D, Lam WL, Marra M, Richards-Kortum R, Follen M, Macaulay C, & Jones SJ (2007) Using LongSAGE to Detect Biomarkers of Cervical Cancer Potentially Amenable to Optical Contrast Agent Labelling. *Biomark Insights* 2:447-461.

322. Narayan G, Bourdon V, Chaganti S, Arias-Pulido H, Nandula SV, Rao PH, Gissmann L, Durst M, Schneider A, Pothuri B, Mansukhani M, Basso K, Chaganti RS, & Murty VV (2007) Gene dosage alterations revealed by cDNA microarray analysis in cervical cancer: identification of candidate amplified and overexpressed genes. *Genes Chromosomes Cancer* 46(4):373-384.
323. Sova P, Feng Q, Geiss G, Wood T, Strauss R, Rudolf V, Lieber A, & Kiviat N (2006) Discovery of novel methylation biomarkers in cervical carcinoma by global demethylation and microarray analysis. *Cancer Epidemiol Biomarkers Prev* 15(1):114-123.
324. Soufla G, Sifakis S, Baritaki S, Zafiropoulos A, Koumantakis E, & Spandidos DA (2005) VEGF, FGF2, TGFB1 and TGFBR1 mRNA expression levels correlate with the malignant transformation of the uterine cervix. *Cancer Lett* 221(1):105-118.
325. Xu XC, Mitchell MF, Silva E, Jetten A, & Lotan R (1999) Decreased expression of retinoic acid receptors, transforming growth factor beta, involucrin, and cornifin in cervical intraepithelial neoplasia. *Clin Cancer Res* 5(6):1503-1508.
326. Aasen T, Hodgins MB, Edward M, & Graham SV (2003) The relationship between connexins, gap junctions, tissue architecture and tumour invasion, as studied in a novel in vitro model of HPV-16-associated cervical cancer progression. *Oncogene* 22(39):7969-7980.
327. Belletti B, Nicoloso MS, Schiappacassi M, Berton S, Lovat F, Wolf K, Canzonieri V, D'Andrea S, Zucchetto A, Friedl P, Colombatti A, & Baldassarre G (2008) Stathmin activity influences sarcoma cell shape, motility, and metastatic potential. *Mol Biol Cell* 19(5):2003-2013.
328. Termini L, Maciag PC, Soares FA, Nonogaki S, Pereira SM, Alves VA, Longatto-Filho A, & Villa LL (2010) Analysis of human kallikrein 7 expression as a potential biomarker in cervical neoplasia. *Int J Cancer* 127(2):485-490.
329. Santin AD, Cane S, Bellone S, Bignotti E, Palmieri M, De Las Casas LE, Roman JJ, Anfossi S, O'Brien T, & Pecorelli S (2004) The serine protease stratum corneum chymotryptic enzyme (kallikrein 7) is highly overexpressed in squamous cervical cancer cells. *Gynecol Oncol* 94(2):283-288.
330. Lee JW, Lee SJ, Seo J, Song SY, Ahn G, Park CS, Lee JH, Kim BG, & Bae DS (2005) Increased expressions of claudin-1 and claudin-7 during the progression of cervical neoplasia. *Gynecol Oncol* 97(1):53-59.
331. Sobel G, Paska C, Szabo I, Kiss A, Kadar A, & Schaff Z (2005) Increased expression of claudins in cervical squamous intraepithelial neoplasia and invasive carcinoma. *Hum Pathol* 36(2):162-169.
332. Hewitt KJ, Agarwal R, & Morin PJ (2006) The claudin gene family: expression in normal and neoplastic tissues. *BMC Cancer* 6:186.
333. Morin PJ (2005) Claudin proteins in human cancer: promising new targets for diagnosis and therapy. *Cancer Res* 65(21):9603-9606.

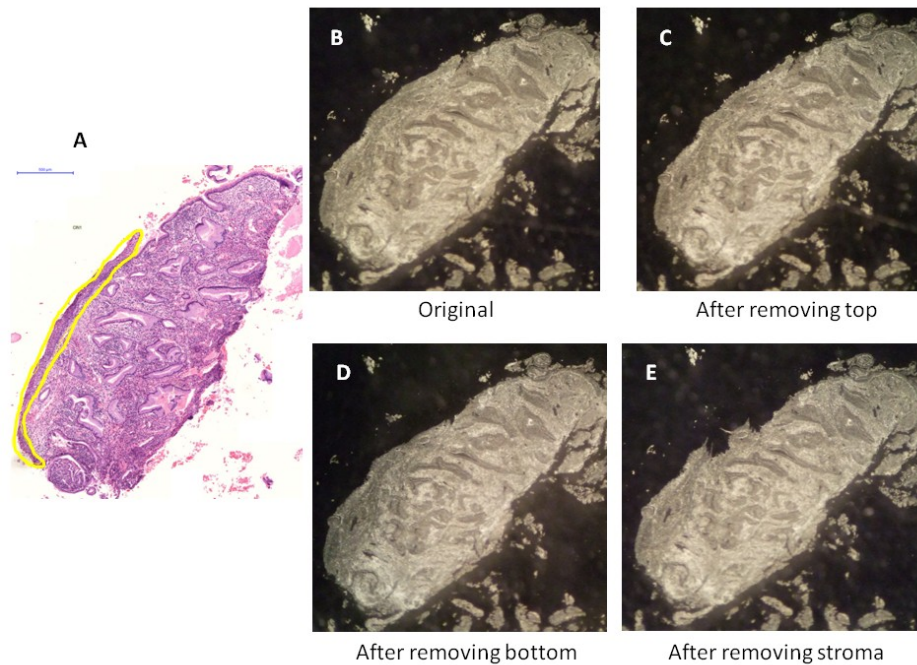
334. Kramer F, White K, Kubbies M, Swisshelm K, & Weber BH (2000) Genomic organization of claudin-1 and its assessment in hereditary and sporadic breast cancer. *Hum Genet* 107(3):249-256.
335. Huo Q, Kinugasa T, Wang L, Huang J, Zhao J, Shibaguchi H, Kuroki M, Tanaka T, Yamashita Y, Nabeshima K, & Iwasaki H (2009) Claudin-1 protein is a major factor involved in the tumorigenesis of colorectal cancer. *Anticancer Res* 29(3):851-857.
336. Resnick MB, Konkin T, Routhier J, Sabo E, & Pricolo VE (2005) Claudin-1 is a strong prognostic indicator in stage II colonic cancer: a tissue microarray study. *Mod Pathol* 18(4):511-518.
337. Miwa N, Furuse M, Tsukita S, Niikawa N, Nakamura Y, & Furukawa Y (2001) Involvement of claudin-1 in the beta-catenin/Tcf signaling pathway and its frequent upregulation in human colorectal cancers. *Oncol Res* 12(11-12):469-476.
338. Li D, Peng Z, Tang H, Wei P, Kong X, Yan D, Huang F, Li Q, Le X, & Xie K (2011) KLF4-mediated negative regulation of IFITM3 expression plays a critical role in colon cancer pathogenesis. *Clin Cancer Res* 17(11):3558-3568.
339. Andreu P, Colnot S, Godard C, Laurent-Puig P, Lamarque D, Kahn A, Perret C, & Romagnolo B (2006) Identification of the IFITM family as a new molecular marker in human colorectal tumors. *Cancer Res* 66(4):1949-1955.
340. Daniel-Carmi V, Makovitzki-Avraham E, Reuven EM, Goldstein I, Zilkha N, Rotter V, Tzehoval E, & Eisenbach L (2009) The human 1-8D gene (IFITM2) is a novel p53 independent pro-apoptotic gene. *Int J Cancer* 125(12):2810-2819.
341. Salas S, Jezequel P, Campion L, Deville JL, Chibon F, Bartoli C, Gentet JC, Charbonnel C, Gouraud W, Voutsinos-Porche B, Brouchet A, Duffaud F, Figarella-Branger D, & Bouvier C (2009) Molecular characterization of the response to chemotherapy in conventional osteosarcomas: predictive value of HSD17B10 and IFITM2. *Int J Cancer* 125(4):851-860.
342. Merkley MA, Weinberger PM, Jackson LL, Podolsky RH, Lee JR, & Dynan WS (2009) 2D-DIGE proteomic characterization of head and neck squamous cell carcinoma. *Otolaryngol Head Neck Surg* 141(5):626-632.
343. Arnouk H, Merkley MA, Podolsky RH, Stoppler H, Santos C, Alvarez M, Mariategui J, Ferris D, Lee JR, & Dynan WS (2009) Characterization of Molecular Markers Indicative of Cervical Cancer Progression. *Proteomics Clin Appl* 3(5):516-527.
344. Sano T, Oyama T, Kashiwabara K, Fukuda T, & Nakajima T (1998) Expression status of p16 protein is associated with human papillomavirus oncogenic potential in cervical and genital lesions. *Am J Pathol* 153(6):1741-1748.
345. Tozawa-Ono A, Yoshida A, Yokomachi N, Handa R, Koizumi H, Kiguchi K, Ishizuka B, & Suzuki N (2012) Heat shock protein 27 and p16 immunohistochemistry in cervical intraepithelial neoplasia and squamous cell carcinoma. *Hum Cell* 25(1):24-28.

346. Klaes R, Friedrich T, Spitkovsky D, Ridder R, Rudy W, Petry U, Dallenbach-Hellweg G, Schmidt D, & von Knebel Doeberitz M (2001) Overexpression of p16(INK4A) as a specific marker for dysplastic and neoplastic epithelial cells of the cervix uteri. *Int J Cancer* 92(2):276-284.
347. Regauer S & Reich O (2007) CK17 and p16 expression patterns distinguish (atypical) immature squamous metaplasia from high-grade cervical intraepithelial neoplasia (CIN III). *Histopathology* 50(5):629-635.
348. Tusher VG, Tibshirani R, & Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 98(9):5116-5121.
349. Sund M & Kalluri R (2009) Tumor stroma derived biomarkers in cancer. *Cancer Metastasis Rev* 28(1-2):177-183.
350. Barth PJ, Ramaswamy A, & Moll R (2002) CD34(+) fibrocytes in normal cervical stroma, cervical intraepithelial neoplasia III, and invasive squamous cell carcinoma of the cervix uteri. *Virchows Arch* 441(6):564-568.
351. Davidson B, Goldberg I, Gotlieb WH, Ben-Baruch G, & Kopolovic J (1998) Expression of matrix proteins in uterine cervical neoplasia using immunohistochemistry. *Eur J Obstet Gynecol Reprod Biol* 76(1):109-114.
352. Gaiotto MA, Focchi J, Ribalta JL, Stavale JN, Baracat EC, Lima GR, & Guerreiro da Silva ID (2004) Comparative study of MMP-2 (matrix metalloproteinase 2) immune expression in normal uterine cervix, intraepithelial neoplasias, and squamous cells cervical carcinoma. *Am J Obstet Gynecol* 190(5):1278-1282.
353. Nadji M, Nassiri M, Vincek V, Kanhoush R, & Morales AR (2005) Immunohistochemistry of tissue prepared by a molecular-friendly fixation and processing system. *Appl Immunohistochem Mol Morphol* 13(3):277-282.
354. Morales AR, Nassiri M, Kanhoush R, Vincek V, & Nadji M (2004) Experience with an automated microwave-assisted rapid tissue processing method: validation of histologic quality and impact on the timeliness of diagnostic surgical pathology. *Am J Clin Pathol* 121(4):528-536.

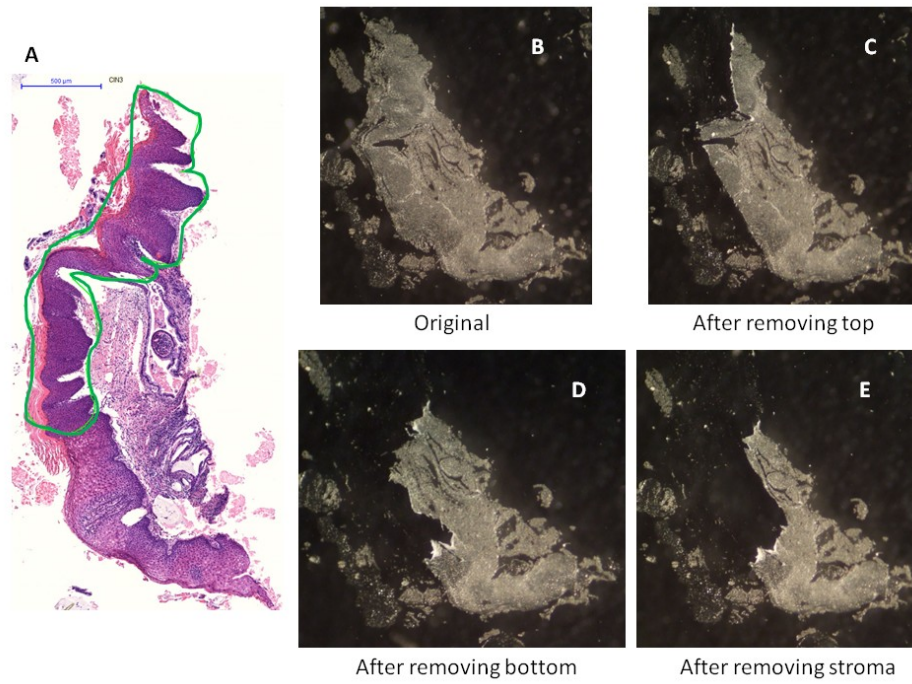
Appendix: Microdissection of Molecular Fixative

Samples

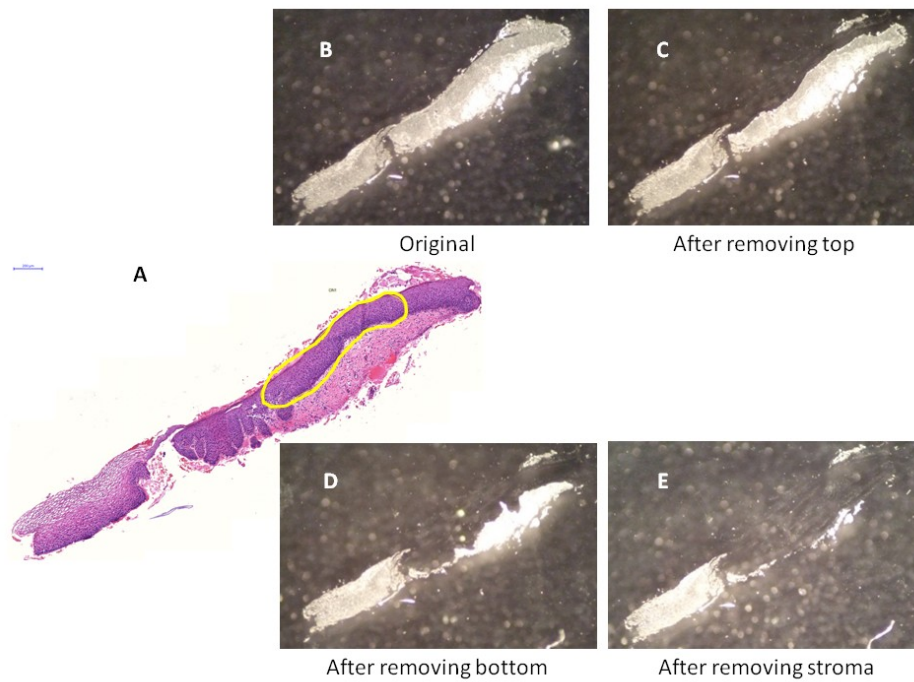
This appendix consists of a series of composite figures documenting the microdissection of each of the molecular fixative samples used in Chapters 4 and 5. Each figure is composed of five panels: A reference H&E section (A) is shown to the left beside photographs of an adjacent section before microdissection (B), after removing the top layer (C), after removing the bottom layer (D), and after removing the stroma (E). Microdissection was always performed in this order.



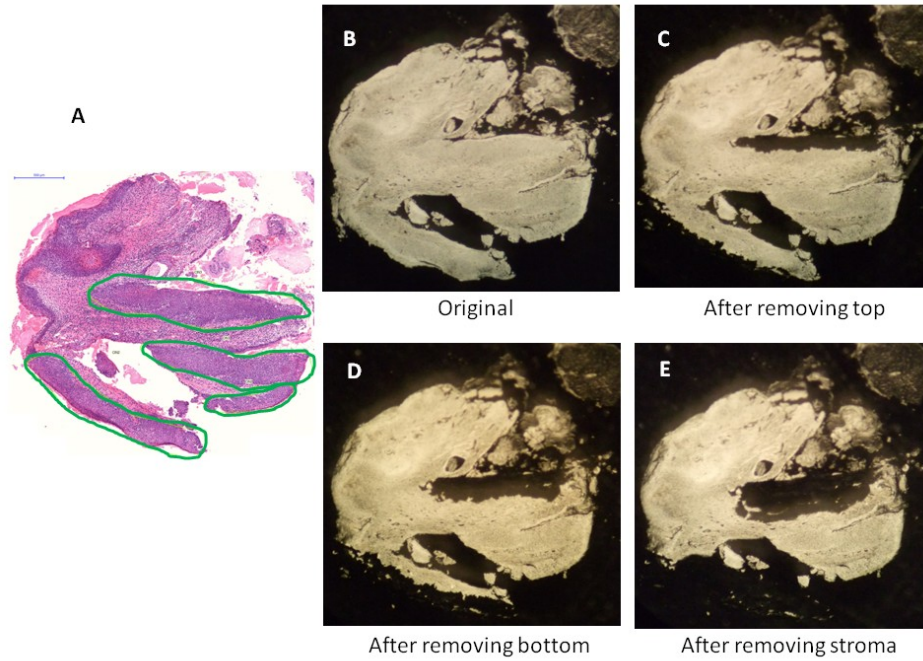
Appendix Figure 1: Case 0027. A region of CIN I is circled in yellow.



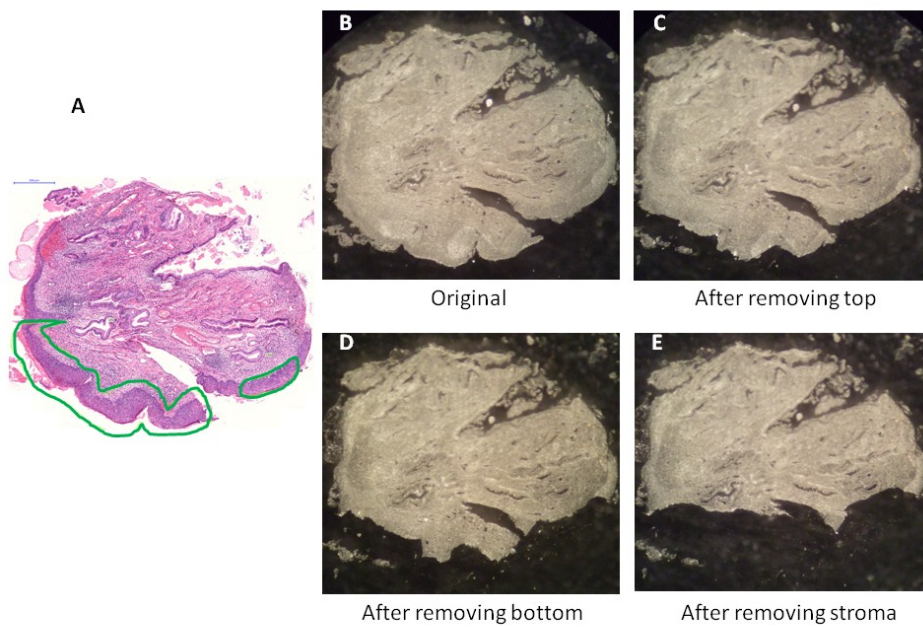
Appendix Figure 2: Case 0028. A region of CIN III is circled in green.



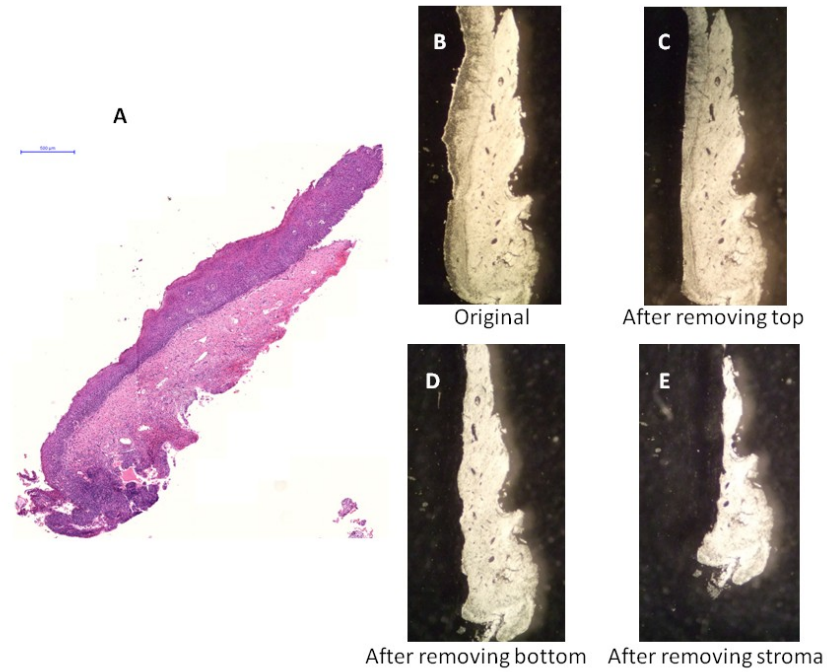
Appendix Figure 3: Case 0030. A region of CIN I is circled in yellow.



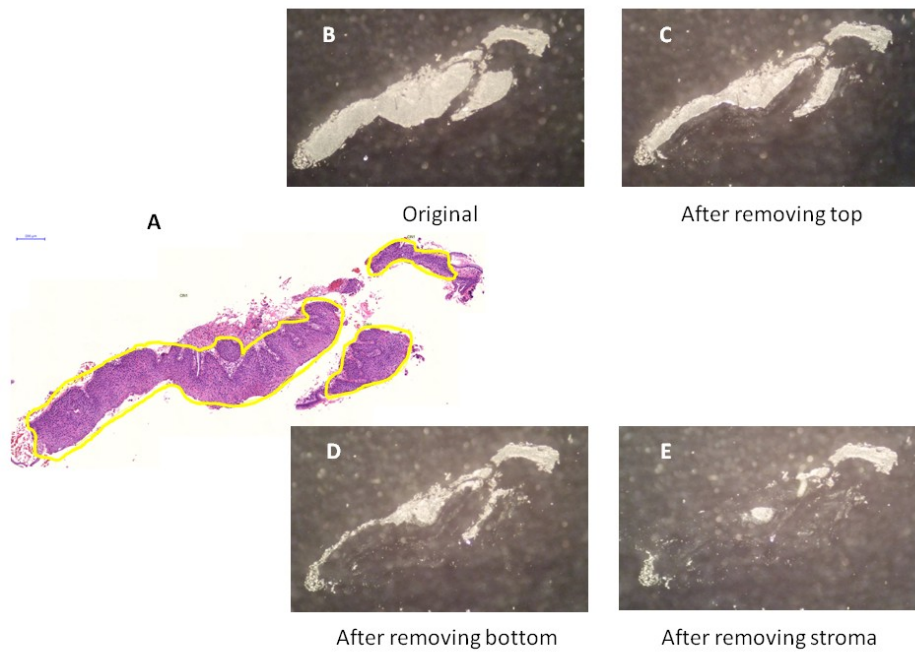
Appendix Figure 4: Case 0033A. High-grade regions are circled in green. The two horizontal regions across the middle are CIN III. The regions along the bottom are CIN II.



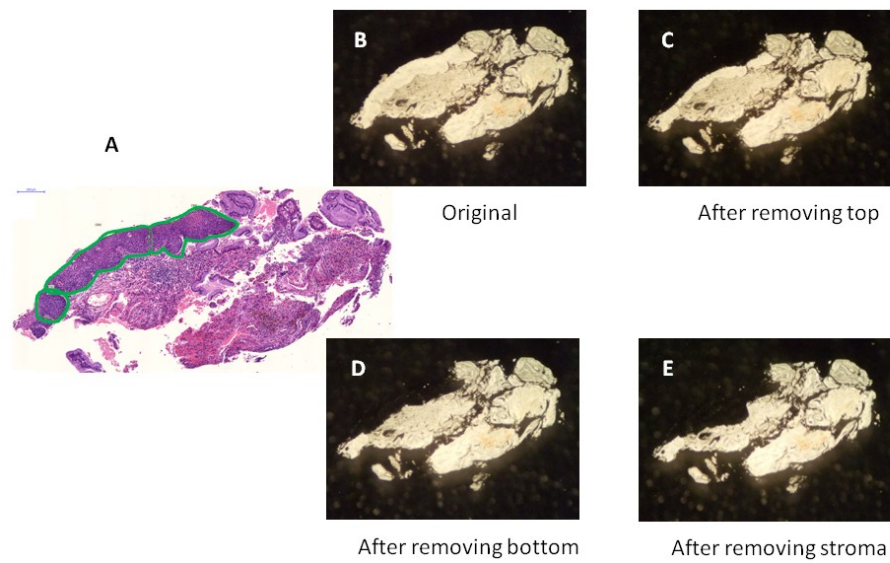
Appendix Figure 5: Case 0033B. Regions of CIN II are circled in green.



Appendix Figure 6: Case 0043. Epithelium of normal histopathology was microdissected from this case.



Appendix Figure 7: Case 0044. Regions of CIN I are circled in yellow.



Appendix Figure 8: Case 0053. Regions of high-grade dysplasia are circled in green. The first few sections microdissected contained more CIN II and were collected as 0053A. The latter sections contained more CIN III (small leftmost region) and were collected as 0053B.