Published in final edited form as:
Cancer Res. 2008 May 1; 68(9): 3099–3107. doi:10.1158/0008-5472.CAN-07-2113.

Potential Use of Quantitative Tissue Phenotype to Predict Malignant Risk for Oral Premalignant Lesions

Martial Guillaud, **Lewei Zhang**, **Catherine Poh**, **Miriam P. Rosin**, and **Calum MacAulay**
British Columbia Cancer Agency/Cancer Research Center, Vancouver, BC, Canada (MG, LZ, CP, MR, CMC). School of Kinesiology, Simon Fraser University, Burnaby, British Columbia, Canada (MR). Faculty of Dentistry, University of British Columbia, Vancouver, BC, Canada (CP, LZ).

Abstract

The importance of early diagnosis in improving mortality and morbidity rates of oral squamous cell carcinoma (SCC) has long been recognized. However, a major challenge for early diagnosis is our limited ability to differentiate oral premalignant lesions (OPLs) at high risk of progressing into invasive SCC from those at low risk. We investigated the potential of Quantitative Tissue Phenotype (QTP), measured by high-resolution image analysis, to recognize severe dysplasia/carcinoma *in situ* (*CIS*) (known to have an increased risk of progression) and to predict progression within hyperplasia or mild/moderate dysplasia (termed HMD). We generated a Nuclear Phenotypic Score (NPS), a combination of 5 nuclear morphometric features that best discriminate 4,027 “normal” nuclei (selected from 29 normal oral biopsies) from 4,298 “abnormal” nuclei (selected from 30 SCC biopsies). This NPS was then determined for a set of 69 OPLs. Severe dysplasia/*CIS*, showed a significant increase in NPS compared to HMD. However, within the latter group, elevated NPS was strongly associated with the presence of high-risk LOH patterns. There was a statistical difference between NPS of HMD that progressed to cancer and those that did not. Individuals with a high NPS had a 10-fold increase in relative risk of progression. In the multivariate Cox model, LOH and NPS together were the strongest predictors for cancer development. These data suggest that QTP could be used to identify lesions that require molecular evaluation and should be integrated with such approaches to facilitate the identification of HMD OPLs at high risk of progression.

Keywords

Risk prediction; oral malignancy; LOH; dysplasia; image analysis

Introduction

Although the importance of early diagnosis in improving the mortality and morbidity of oral squamous cell carcinoma (SCC) has long been recognized, the disease is still frequently diagnosed late and prognosis has not changed for the last 3 decades (1). A major challenge for early diagnosis of the at-risk tissue is our limited ability to differentiate oral premalignant lesions (OPLs) at high risk of progressing into invasive SCC from those at low risk. Oral cancer is believed to develop from OPLs: progressing from hyperplasia, through the increasing degrees of dysplasia to carcinoma *in situ* (*CIS*) and finally invasive squamous cell carcinoma (SCC). Histology, the current gold standard, is reasonably effective in predicting malignant risk for severe dysplasia/*CIS*, of which 30–40% are likely to recur or progress even with aggressive surgical treatment (1,2). However, overall, it is a poor predictor for the majority of

Request for reprints: Address Correspondence to: Dr. Martial Guillaud, 675 West 10th Ave, Vancouver, BC, V5Z1L3, Canada Tel.: (604) 675-8086 Fax: (604) 675-8099, E-mail: E-mail: mguillau@bccancer.bc.ca.

OPLs – those with hyperplasia or with mild/moderate dysplasia (hereinafter termed HMD) in that most do not progress into cancer. The literature reports a wide range (11% – 36%) for the overall risk of malignant transformation, depending on the type of lesions being followed and the length of follow-up (3,4). Leukoplakia with dysplasia are more likely to progress to oral SCC than those without dysplasia (4); however, many studies have reported that for the individual lesion, the grade of dysplasia provides little indication of whether or not it will progress to cancer (4,5,6).

Alternative approaches that facilitate the identification of high-risk OPLs, particularly those within HMD, need to be developed. The foremost development in such approaches over the last 5 – 10 years has been aimed at finding molecular markers (7,8,9). Among these, microsatellite analysis of loss of heterozygosity (LOH) has shown promise as a powerful adjunct tool for risk prediction (7,10–14). Retrospective studies from our research team have shown that LOH pattern can classify morphologically indistinguishable HMD OPLs into different risk categories. OPLs without LOH at 3p and 9p have the lowest risk of progression. In comparison those with LOH at 3p and/or 9p but not in other arms have a 3.8-fold increase in relative cancer risk. Those with LOH at 3p and/or 9p plus additional losses (at 4q, 8p, 11q, or 17p, high risk) have a 33-fold increase in relative cancer risk (14).

Given that the largest proportion of OPLs are HMD, there is a need for an inexpensive system to triage such samples for molecular analysis. There is some indication that imaging devices that characterize phenotype changes associated with cancer risk could play such a role by complementing and adding to our current knowledge of risk assessment (15,16,17). Moreover, since exploration of phenotype provides a window to the history of the cell, showing whether the genetic alterations within are associated with structural/functional changes, this approach could expand our understanding of the biology underlying malignant progression.

Histopathological evaluation of presence and degree of dysplasia is an appraisal of the tissue (and cell) phenotype using multiple parameters. Quantitative Tissue Pathology (QTP) breaks down the components of this phenotype into multiple quantifiable units that can be studied independently and in combination, allowing the investigator to examine associations of such change with progression risk (Fig. 1A–B). For example, increased proportion of heterochromatin condensation is a high-risk nuclear characteristic for cancer (18). Traditional pathology judges this by one criterion, hyperchromatism, whereas QTP measures multiple chromatin features, such as whether the increased DNA is distributed around the edge of the nucleus or clustered in the center (top row of Fig. 1B), whether the nucleus is dark with light areas or light with dark areas (middle row of Fig. 1B), and whether the increased chromatin is evenly distributed (euchromatin) or clumped locally (heterochromatin) (bottom row of Fig. 1B).

In this paper, we present the preliminary results of an in-house QTP Imaging System used to assess nuclear phenotype changes in OPLs. This is the first study to correlate QTP-detected **N**uclear **P**henotype change or **S**core (called **NPS**) with not only pathological grading, but also molecular changes and outcome of OPLs. The results showed that NPS correlates strongly with histology grading, genetic damage (LOH) and cancer progression.

Materials and Methods

Sample Source

A total of 128 oral premalignant biopsies were collected from 128 subjects. All these lesions were obtained from the British Columbia Oral Biopsy Service with selection based on the criteria below. These lesions were grouped into a training and a test set.

The training set contained two groups of specimens representing the two extreme ends of the intra-epithelium spectrum: 30 oral biopsies with areas of relatively normal oral mucosa (e.g., amalgam tattoo or melanotic macule) and a group of 29 SCC.

The test set contained 69 biopsies that can be separated into two sub-groups: a group of 25 biopsies diagnosed as severe or *CIS* which was denoted as high-histology risk group, since they are known to be histologically at risk for progression (1,2); and a group of 44 biopsies which were a subset of 83 well-characterized OPLs with sufficient material for QTP used in our previous study that explored the use of LOH to predict risk of progression for OPLs with HMD (14). These 44 lesions (one biopsy per patient) diagnosed as hyperplasia, mild or moderate dysplasia are denoted the low-histology risk group. They were from patients without a prior history of head and neck cancer. The only prerequisite for inclusion in the present study was the availability of two adjacent unstained slides that were serial to slides used for the previous LOH analysis. Of these lesions, 15 have progressed to cancer and the remaining have not. Among the 15 progressing lesions, there were 6 hyperplasias, 5 mild dysplasias and 4 moderate dysplasias. Among the non-progressing lesions, there were 15 hyperplasias, 8 mild dysplasias, and 6 moderate dysplasias. There was no difference between the progressing and the non-progressing lesions in terms of gender, age distribution, and smoking history (all $P < 0.05$). On average, the non-progressing cases have been monitored for over twice the duration of progressing cases (71 months vs. 30 months) to ensure that progression did not occur.

Sample preparation

All 128 samples were formalin-fixed and paraffin-embedded. The histological diagnoses of the samples were confirmed by two oral pathologists (CP and LZ). Serial sections, 4-μm in thickness, were cut from each sample and placed on two glass slides, one stained with Hematoxylin and Eosin (H&E) and the other with Feulgen-Thionin (19). Representative areas of the histological diagnosis for each sample were circled on the H&E slide by an oral pathologist (CP or LZ), and corresponding areas on the Feulgen-stained slide were examined with the QTP Imaging System.

QTP Imaging System

This System is a modified version of the Cyto-Savant automated quantitative system (Cancer Imaging, BC Cancer Agency) (19). The illumination wavelength was 600 ± 5 nm, corresponding to the absorption peak of the Thionin stain. The effective pixel sampling space within the plane of the sample was 0.34 μm^2 and the effective pixel sampling area was 0.116 μm^2. The software is specifically designed for interactive semi-automatic cellular and architectural analysis of tissue sections (19,20). A strict quality control procedure was implemented to ensure the stability of the imaging system for each analysis (21). The imaging system characteristics followed the recommendation of the European Society of Analytical Cellular Pathology (22).

QTP Image Analysis

An experienced cytotechnician under the guidance of an oral pathologist (CP or LZ) delineated the region of interest (ROI) from the Feulgen-stained slide under a 20X objective. The area selected represented the most abnormal tissue according to the pathologist. The analysis of collected nuclei in each biopsy was performed automatically by the system and consisted of four steps (a) delineation and focusing of the region of interest (Fig. 1A), (b) automatic thresholding and segmentation of the nuclei in the field (23), (c) interactive correction of segmentation errors and selection of only non-overlapping cells, and (d) automatic collection of individually focused images of each selected cell (Fig. 1C). For normal epithelium and OPLs, the full width of the epithelium had to be present. Only in-focus, intact nuclei were collected by the cytotechnician with nuclei obtained from all layers of the epithelium. The regions used

in each case of squamous cell carcinoma were selected using three criteria: 1) the region of tumor selected needed to show tumor differentiation that was representative of the case; 2) the region had to have minimal amounts of inflammation and tissue; and 3) the tumor region needed to be reasonably large, i.e., not a single line or layer of tumor cells.

On average, more than 100 cell nuclei were collected from the delineated area for each biopsy. For each nucleus, 110 nuclear phenotype features were calculated by the system, measuring the size, the shape, the amount and the distribution of stained DNA in the nucleus (24).

Calculation of Nuclear Phenotype Score (NPS)

To determine nuclear features that are predictors for OPLs with different cancer risk, we first trained the imaging system to separate normal specimens (30 samples) from cancer specimens (29 SCC samples) in the training set. This was done by establishing first a cell-by-cell NPS, which is a linear discriminant function of the nuclear features that are statistically significant, not collinear, that taken together, used as a predictor, can classify cells into separate categories. For this process, all epithelial cells (4,027 cells) from the 30 normal samples were grouped together and used to define the “normal-like” cell group, and all epithelial cells (4,298 cells) from the 29 SCC samples were grouped together to define the ‘cancer-like’ cell group. A forward stepwise discriminant function analysis of the 110 nuclear phenotypic features was performed to select which of these features could be used to generate a discriminant function that best separated the two groups of cells. All features were considered as continuous variables. In stepwise discriminant function analysis, a model of discrimination is built step-by-step. Specifically, at each step all variables are reviewed and evaluated to determine which one will contribute most to the discrimination between groups. That variable will then be included in the model, and the process repeats. The process stops when no new variable satisfies the “entry” criterion (F value greater than a specified threshold).

An interpretation of this discriminant function is that it represents a cell-by-cell phenotype classification procedure which assigns a score to each nucleus indicating similarity to normal or cancer cells.

Similar to the method described in a companion paper (15), the range of this cellular score was divided into 10 continuous regions, the first region *A* representing values of phenotypically ‘normal’ cells, and region *J* representing values of phenotypically ‘cancer-like’ cells.

From the percentage of cells in each of these 10 regions, the NPS, assigned to each biopsy, was calculated as followed:

NPS = 1*A + 2*B + 3*C + 4*D + 5*E + 6*F + 7*G + 8*H + 9*I + 10*J; with A the proportion of cells whose disciminant score falls in the region A, B the proportion of cells whose disciminant score falls in the region B, and so forth.

This NPS represents the weighted sum of the 10 regions. The possible values for the NPS range from 1 to 10. A NPS with a value of 1 corresponds to a specimen containing 100% of ‘normal-like’ cells, whereas a NPS with a value of 10 corresponds to a specimen containing 100% of ‘cancer-like’ cells.

In a previous study (15), we determined how sensitive the NPS process is to the interactive steps involved in the collection of the nuclei used to calculate NPS. In that study, 30 pre-neoplastic bronchial biopsies which spanned the ranges of pathology grades and NPS were selected and measured. All samples were measured twice by one experienced cytotechnologist and twice by another experienced cytotechnologist. The measurement of the NPS was highly

reproducible between individual technologists. The coefficient of correlation between two histopathology technologists was 0.98. The intra-technologist variability was comparable.

Statistical Analysis

For comparison of difference between groups, a non-parametric Mann-Whitney U test was performed. For comparison of difference between more than two groups, a non parametric Kruskal-Wallis ANOVA was performed. When the test showed a statistical difference across groups, a multiple comparison of mean ranks was performed to confirm significant differences across group medians (multiple comparisons of mean ranks for all groups). Differences between group frequencies were assessed with the Yates corrected Chi-square test. Time-to-progression curves were estimated by the Kaplan-Meier method and comparisons were performed using the log-rank test. Relative risks were determined using Cox regression analysis. All analyses were performed with Statistica6 (StatSoft, Inc., 2001, Tulsa, OK). All *P* values were two-sided. A two-sided *P* value of less than 0.05 was considered statistically significant.

Results

Nuclear Phenotype Scores (NPS)

We used a forward stepwise discriminant function analysis to identify 5 of the 110 nuclear features that could best discriminate cells from normal and cancer samples in the training set: (1) *Max_radius*, the maximum radius of the nucleus; (2) *Harm003_fft*, the third order harmonic of the nuclear boundary – measures how much of the nuclear boundary can be explained by 3 lobes; (3) *Fractal_area1*, the relative spatial distribution of high and low optical density variations in the nucleus – a measurement of heterochromatin vs. euchromatin organization; (4) *OD_skewness*, measures whether the nucleus is dark with light areas or light with dark areas; and (5) *Long90_runs*, a measurement of the fraction of the nuclear diameter one can travel before an intensity change is encountered (24). A total of 94% of the 4027 ‘normal-like’ cells and 77% of the 4298 ‘cancer-like’ cells were correctly classified using these 5 discriminant features. These features were used to generate the cell-by-cell discriminant scores which were amalgamated across all of the cells selected in the epithelium to generate the NPS for each specimen. Fig. 2 shows representative images of the region of interest of four oral mucosa specimens with the corresponding histogram distribution of texture feature *Fractal_area1*, used in the calculation of the NPS.

To determine the ability of NPS to classify oral lesions into different risk groups, we correlated NPS with (1) pathology classification; (2) molecular patterns; and (3) outcome, the ultimate yardstick for judging the validity of new diagnostic tools.

Correlation of NPS with Pathology

Fig. 3 shows the association between NPS and the pathology diagnoses. There was a monotonic increase in NPS with severity of pathology diagnosis (median NPS: 3.7 for normal, 4.14 for the low-histology risk group, i.e. the HMD group [hyperplasia, mild or moderate dysplasia], 5.24 for high-histology risk group [severe dysplasia or *C*IS] and 7.25 for SCC). There is a significant increase in NPS between the low-histology risk group (HMD) and the high-histology risk group (P = 0.036). There are significant differences between all logical histological groupings except normal versus HMD and severe dysplasia/CIS. The increase in NPS between the severe dysplasia/*CIS* group and the SCC was close to significance (*P* = 0.08).

The remaining analyses were performed using only the lesions classified as hyperplasia, mild dysplasia and moderate dysplasia (the HMD group).

Correlation of NPS with Progression to Cancer

We sought to determine whether NPS would identify hyperplasia, mild and moderate dysplasia (HMD) at high risk of developing into invasive SCC, by comparing NPS values in progressing and non-progressing lesions. Progressing cases showed significantly higher NPS value compared to non-progressing lesions (median NPS: 5.7 for progressing lesions vs. 3.8 for non-progressing lesions; $P < 0.0001$). As shown in Fig 4A, although the sample set was small, this trend remained statistically significant when the hyperplasia and dysplasia were examined separately: median NPS, 4.9 for progressing hyperplasia vs. 3.8 for non-progressing hyperplasia ($P = 0.016$); and median NPS, 6.3 for progressing low-grade dysplasia vs. 3.8 for non-progressing low-grade dysplasia ($P = 0.001$).

Cutoff NPS value and Time to Progression to Cancer

For the NPS to be used by clinicians in cancer prediction, a cutoff value is needed above which clinicians could expect increased cancer risk for an OPL. We have chosen a cutoff value of 4.5 for the NPS since it provided the best separation (area under the ROC curve, Figure 4B) between the non-progressing lesions and the progressing lesions. In the training set (normal and SCC, Fig. 2), only 5 of the 30 (17%) normal samples had a NPS value ≥ 4.5 (high-NPS); whereas 28 of the 29 (97%) of SCC had a high-NPS.

When the HMD lesions were categorized by this cutoff value, 17 of the 44 HMD test study cases had a high-NPS. Thirteen of the 17 (76%) cases with high-NPS progressed into invasive SCC in contrast to only 2 of the 27 (7%) cases with low-NPS ($P < 0.0001$). In other words, with this cutoff, we could correctly classify 86% of non-progressing cases (25/29) and 86% of progressing cases (13/15).

Time-to-progression curves were plotted as a function of the NPS classification into one of the two groups (high-NPS vs. low-NPS) (Fig. 4C). A significant difference was observed between the low-NPS group and the high-NPS group ($P < 0.00015$).

The proportion of lesions progressing to cancer at 5 years was 71% (12/17) for lesions with high-NPS and only 22% for lesions with low-NPS. There was a 10-fold (RR = 10.3 [CI: 2.9–60.0]) increase in the relative risk of progression to cancer for oral lesions with a high-NPS in comparison to those with a low-NPS.

Correlation of NPS with LOH

Table 1 shows association between NPS and different LOH patterns, indicating the proportion of cases within each LOH pattern that are above and below the NPS cutoff. NPS values were consistently higher for samples with LOH at all 7 chromosome regions examined. This increase was significant for 3p ($P = 0.03$), 4q ($P = 0.003$), 9p ($P = 0.003$) and 11q ($P = 0.05$). We also looked at associations with multiple losses, and for LOH at 3p &/or 9p plus LOH at any of the arms 4q, 8p, 11q, 13q and 17p, patterns previously identified as having markedly increased risk of progression (9,11). NPS was strongly associated with the presence of these high-risk LOH patterns: multiple losses ($P = 0.0005$), and LOH at 3p &/or 9p plus LOH at any of the arms 4q, 8p, 11q, 13q and 17p ($P < 0.0005$).

Seventeen of the 44 HMD lesions had a high-NPS. Molecularly, 16 of these 17 cases had LOH information. Of these 16, 11 (69%) showed LOH at 3p &/or 9p, a loss believed to be essential although not sufficient for cancer progression (8); in contrast, only 7 of the 27 (26%) cases with NPS < 4.5 (low-NPS) showing such a loss ($P = 0.008$).Similarly, 9 of these 16 (56%) lesions showed LOH at 3p &/or 9p and at any of the other 5 chromosomes, which corresponds to the highest risk pattern for cancer progression (14) whereas only 4 of the 27 (15%) cases with NPS < 4.5 (low-NPS) showed such a pattern ($P = 0.007$).

Regression Analysis with Multiple Covariates

We performed a regression analysis with multiple covariates for the HMD group looking at histology, LOH and NPS. For each of these covariates we assigned a value of either 0 or 1, denoting either low or high risk marker value, respectively. For histology, hyperplasia was coded 0 and mild or moderate dysplasia was coded 1. For LOH, cases with LOH at 3p/or 9p plus LOH at any of the arms 4q, 8p, 11q, 13q and 17p were coded 1, and 0 otherwise. For NPS, cases with a high-NPS (NPS higher than 4.5) were coded 1 and cases with Low-NPS were coded 0. The results of the Cox Proportional Hazard regression are given in Table 2. In this table, the first 3 models examine NPS, LOH and histology separately and show that only LOH and NPS are significant in predicting cancer (with similar *P* value: $P < 0.001$). For HMD, histology is not a predictor of cancer progression ($P = 0.88$). Categorization of hyperplasia with mild dysplasia *vs* moderate dysplasia was also not significant (data not shown). When LOH and NPS were entered together into the COX regression analysis (model 4), NPS is no longer as significant as it was when used as a single predictor (model 1: $P = 0.0007$), even though the borderline *P*-value of 0.07 does suggest that a larger sample size may be required for a definitive conclusion. The different models clearly show that conventional histology has no predictive effect and that NPS is a better “phenotypic” predictive marker for cancer risk than histology (models 5 and 6).

Discussion

There have been tremendous scientific advancements in many fields over the last decade, owing, in a significant part, to rapid computer technology development; such development is revolutionizing many areas including the health care system. In this study, we explored the value of a computer-driven imaging system as an adjunct tool to assist the pathologist in judging the progression risk of oral premalignant lesions. The importance of dysplasia phenotypes as cancer risk predictors is well recognized. The strong association of marked dysplasia with increased risk of cancer progression has been observed and confirmed in multiple sites (e.g., lung, esophagus, breast, cervix and skin) and is the reason why histopathological interpretation of dysplasia is the current gold standard. However, this standard is far from accurate because of a number of issues: 1) by nature it is subjective, and studies have reported low intra- and inter-observer agreement in grading epithelial dysplasia (25–26); 2) there is a lack of knowledge of the weight to be associated with each of the visual characteristics used to assess histopathological grades as well as the way in which they interact for risk assessment; and 3) reactive changes [with little malignant potential] can cause alterations that resemble low-grade dysplasia, making differential diagnosis in early disease particularly challenging. Finding a solution to the above issues involves at least two parts: first, the development of a tool that can be made to be high throughput, that is objective, quantitative and sensitive enough to unravel subtle differences in OPLs, such as between reactive and true low-grade premalignant changes, and between non-progressing and progressing low-grade dysplasia; and second, the collection of specimens with known risk factors including high-grade histology and high-risk molecular pattern, and ideally with outcome to test the system.

Our results showed that the in-house imaging system developed by our research team can produce data in high throughput fashion. We used two well-defined groups of cells (normal and cancer) to generate a Nuclear Phenotype Score for each specimen. This system was able to analyze 110 nuclear features and identify not only differences in nuclear size and chromatin density as noted by pathologists, but also differences in nuclear shape and chromatin texture, features not readily quantified by pathologists. The identification of different features of chromatin texture changes by the system is significant since such changes are now widely accepted as an indication of genetic or epigenetic changes (18). Furthermore, the system also quantified the changes measured, allowing an objective computation of a NPS for each sample.

The NPS was noted to continuously increase with severity of histological grading ($P < 0.0001$). A significant difference was observed between the low-histology risk (HMD) group and the high-risk histology group (severe dysplasia/*CIS*). In British Columbia, a histological diagnosis of severe dysplasia/CIS results in the majority of cases receiving treatment by surgery. Thus we were unable to examine a potential association of NPS in progression within this category. However, 72% of the severe dysplasia/*CIS* in this study had NPS > 4.5 indicating a strong association of NPS with this high-risk histological category.

This is the first study to investigate the relationship between NPS and outcome for HMD OPLs, which is ultimately the most important aspect of validation of any new test. High NPS was found to be strongly associated with cancer progression ($P < 0.0001$). This association was independent of histology since high NPS in oral lesions with low-grade or no dysplasia was still strongly associated with cancer progression, although the number of cases were small. These results were also assessed by multivariate Cox regression analysis and showed an absence of predictive value for histology classification among HMD lesions. In a mixed model, LOH is still the most significant predictor, but a larger sample size is required to determine whether QTP adds significantly to the correlation in a mixed model. It does support the use of QTP for triaging samples that require LOH analysis.

This is the first study to investigate the relationship between NPS and molecular changes of premalignant lesions. We employed the microsatellite analysis for LOH, a technique currently used by many research groups in the study of cancer risk of OPLs. Increased NPS was significantly associated with the presence of previously reported high-risk LOH patterns, including the presence of multiple losses and LOH at 3p &/or 9p plus additional losses (at 4q, 8p, 11q, or 17p) (11,12,13,14).

We have determined a NPS cutoff value of 4.5 based on the training set (normal and SCC, Fig. 2). Seventeen of the 34 HMD cases showed a high NPS. Sixteen of these 17 cases had LOH data available and all showed LOH at 3p &/or 9p, a loss believed to be essential although not sufficient for cancer progression (14). Eleven of these 16 cases progressed into invasive SCC. Individuals with a high NPS have a 10-fold increase in relative risk of developing a cancer within 5 years compared to individuals with a low NPS (CI: 2.05–46.6).. Similar results were obtained when the NPS threshold used to define the high-NPS and the low-NPS groups was varied (NPS = 4.0 through 5.0, data not shown).

NPS could therefore offer clinicians additional information on the possible malignant potential of lesions. For example, 6 of the 21 hyperplasia had a high NPS, 5 of these 6 had a molecular pattern available and 4 of them showed the high-risk LOH pattern. All 3 of the hyperplasia that progressed into cancer had these characteristics. Of the 23 low-grade dysplasia, 11 (48%) had a high NPS, and 9 of these had LOH at 3p &/or 9p (7 had LOH at 3p &/or 9p plus another loss). Eight of the 11 progressed into cancer. These results suggest that the QTP Image System could be a powerful adjunct tool and reproducible method to identify HMD lesions for molecular assessment.

These findings are similar to results obtained by our team in bronchial preneoplastic lesions (15). Using the same approach as the one described in this paper, we showed that QTP correlates better with genetic damage and cancer progression than conventional pathology. Bronchial NPS is currently being used as a surrogate endpoint biomarker in a chemoprevention trial in bronchial dysplasia (27). Similarly, using different morphometric features, quantitative analysis of endometrial hyperplasia correlates better with genetic clonality than conventional pathology (28). Similar results by other groups have been found in bladder, prostate, cervix and breast cancer for the association of QTP with histological grading (29–31). Regarding the nuclear features used in the calculation of our NPS, they are similar to the features used in

other studies for the calculation of morphometric scores such as Quantitative Nuclear Grade by Veltri (29) in transitional cell carcinoma; Discriminant score by Hanselaar et al (32) in cervical cytological specimens and Morphometric Nuclear Grade by Bacus et al (16) in breast and cervical specimens. The main consistency between all these studies, different in terms of tissue types and specimens (cytology /histology), is that they all used a combination of different types of features, i.e., features describing the nuclear shape and size as well as features measuring chromatin texture. Furthermore, nuclear features calculations can easily be reproduced as described in Rodenacker (33).

Recently, Montinori *et al* (17) has shown that these findings could be reproduced using haematoxylin and eosin stained material and fully automated analysis to predict recurrence in urothelial lesions. If this work is confirmed, it could be a definitive breakthrough for the implementation of quantitative histology in routine histopathology as both conventional and quantitative analysis would be performed on the same material, with the analysis fast and automated.

A large longitudinal study is needed to confirm the prognostic value of QTP, and such a longitudinal study is under way. The overall goal of that study is to evaluate 3 new innovative, potentially complementary devices for early detection and follow-up of oral premalignant lesions (OPLs) and oral cancer, and to determine their potential role in the triage of patients in a provincial network that is being established in British Columbia (34). One device is an *in vivo* clinical tissue evaluation tool; the others involve the development of quantitative pathology approaches.

In conclusion, the results presented in this paper show that NPS correlates strongly with genetic damage, cancer progression and histology grading and could have a potential value for triaging OPL for molecular analysis.

ACKNOWLEDGEMENTS

Special thanks go to Anita Carraro for her expert technical support, and to Dr Michelle Follen (MD Anderson Medical Center, Houston, Texas) for her valuable criticism.

Supported by: National Institute of Dental and Craniofacial Research Grants R01 DE13124 and ROI DE017013.

Abbreviations

CIS, carcinoma *in situ*; LOH, loss of heterozygosity; OPL, oral premalignant lesion; SCC, squamous cell carcinoma; QTP, quantitative tissue phenotype; NPS, nuclear phenotype score; HMD, hyperplasia or mild/moderate dysplasia.

References

1. Vedtofte P, Holmstrup P, Hjorting-Hansen E, Pindborg JJ. Surgical treatment of premalignant lesions of the oral mucosa. Int J Oral Maxillofac Surg 1987;16(6):656–664. [PubMed: 3125262]
2. Thomson PJ, Wylie J. Interventional laser surgery: an effective surgical and diagnostic tool in oral precancer management. Int J Oral Maxillofac Surg 2002;31(2):145–153. [PubMed: 12102411]
3. Silverman, S., Jr, editor. Oral Cancer. Vol. Ed. 3. Atlanta, GA: American Cancer Society; 1990.
4. Silverman S Jr, Gorsky M, Lozada F. Oral leukoplakia and malignant transformation: a follow-up study of 257 patients. Cancer 1984;53:563–568. [PubMed: 6537892]
5. Mincer HH, Coleman SA, Hopkins KP. Obswervations on the clinical characteristics of oral lesions showing histologic epithelial dysplasia. Oral Surg Oral Med Oral Pathol 1972;33(3):389–399. [PubMed: 4501168]

6. Cruz IB, Snigjders PJ, Meijer CJ, Braakhuis BJ, Snow GB, Walboomers JM, van der Waal I. p53 expression above the basal cell layer in oral mucosa is an early event of malignant transformation and has predictive value for developing oral squamous cell carcinoma. J Pathol 1998;184(4):360–368. [PubMed: 9664901]
7. Lippman SM, Hong WK. Molecular markers of the risk of oral cancer. N Engl J Med 2001;344:1323–1326. [PubMed: 11320393]
8. Sidransky D. Emerging molecular markers of cancer. Nat Rev Cancer 2002;2:210–219. [PubMed: 11990857]
9. Tong BC, Dhir K, Ha PK, et al. Use of single nucleotide polymorphism arrays to identify a novel region of loss on chromosome 6q in squamous cell carcinomas of the oral cavity. Head Neck 2004;26:345–352. [PubMed: 15054738]
10. Rosin, MP.; Zhang, L.; Poh, C. Molecular markers of oral premalignant lesion risk. California: Academic Press; 2003. Molecular markers of oral premalignant lesion risk; p. 245-259.
11. Califano J, van der Riet P, Westra W, et al. Genetic progression model for head and neck cancer: implications for field cancerization. Cancer Res 1996;56:2488–2492. [PubMed: 8653682]
12. Mao L, Lee JS, Fan YH, et al. Frequent microsatellite alterations at chromosomes 9p21 and 3p14 in oral premalignant lesions and their value in cancer risk assessment. Nat Med 1996;2:682–685. [PubMed: 8640560]
13. Partridge M, Emilion G, Pateromichelakis S, A'Hern R, Phillips E, Langdon J. Allelic imbalance at chromosomal loci implicated in the pathogenesis of oral precancer, cumulative loss and its relationship with progression to cancer. Oral Oncol 1998;34:77–83. [PubMed: 9682768]
14. Rosin MP, Cheng X, Poh C, et al. Use of allelic loss to predict malignant risk for low-grade oral epithelial dysplasia. Clin Cancer Res 2000;6:357–362. [PubMed: 10690511]
15. Guillaud M, le Riche JC, Dawe C, et al. Nuclear morphometry as a biomarker for bronchial intraepithelial neoplasia: correlation with genetic damage and cancer development. CytometryA 2005;63:34–40.
16. Bacus JW, Boone CW, Bacus JV, et al. Image Morphometric Nuclear Grading of Intraepithelial Neoplastic Lesions with Applications to Cancer Chemoprevention Trials1. Cancer Epidemiology, Biomarkers & Prevention 1999;9:1087–1094.
17. Montironi R, Scarpelli M, Lopez-Beltran, et al. Chromatin phenotype karyometry can predict recurrence in papillary urothelial neoplasms of low malignant potential. Cell Oncol 2007;29:47–58. [PubMed: 17429141]
18. Zink D, Fischer AH, Nickerson JA. Nuclear structure in cancer cells. Nat Rev Cancer 2004;4:677–687. [PubMed: 15343274]
19. Garner, D.; Fergusson, G.; Palcic, B. The Cyto-Savant system in Automated cervical Cancer Screening. Grohs, HK.; Husian, OAN., editors. Hong-Kong: Igaku-shoin Medical Publishers, Inc; 1994. p. 305-317.
20. Kamalov R, Guillaud M, Haskins D, et al. A Java application for tissue section image analysis. Comput Methods Programs Biomed 2005;77:99–113. [PubMed: 15652632]
21. Chiu D, Guillaud M, Cox D, Follen M, MacAulay C. Quality assurance system using statistical process control: an implementation for image cytometry. Cell Oncol 2004;26:101–117. [PubMed: 15371646]
22. Haroske G, Baak JP, Danielsen H, et al. Fourth updated ESACP consensus report on diagnostic DNA image cytometry. Anal Cell Pathol 2001;23:89–95. [PubMed: 11904464]
23. MacAulay C, Palcic B. An edge relocation segmentation algorithm. Anal Quant Cytol Histol 1990;12:165–171. [PubMed: 2369469]
24. Doudkine A, MacAulay C, Poulin N, Palcic B. Nuclear texture measurements in image cytometry. Pathologica 1995;87:286–299. [PubMed: 8570289]19
25. Karabulut A, Reibel J, Therkildsen MH, Praetorius F, Nielsen HW, Dabelsteen E. Observer variability in the histologic assessment of oral premalignant lesions. J Oral Pathol Md 1995;24:198–200.
26. Kujan O, Khattab A, Oliver RJ, Roberts SA, Thakker N, Sloan P. Why oral histopathology suffers inter-observer variability on grading oral epithelial dysplasia :An attempt to understand the sources of variation. Oral Oncology 2007;43:224–231. [PubMed: 16931119]

27. Lam S, Xu X, Parker-Klein H, et al. Surrogate end-point biomarker analysis in a retinol chemoprevention trial in current and former smokers with bronchial dysplasia. Int J Oncol 2003;23:1607–1613. [PubMed: 14612933]

28. Mutter GL, Baak JP, Crum CP, Richart RM, Ferenczy A, Faquin WC. Endometrial precancer diagnosis by histopathology, clonal analysis, and computerized morphometry. J Pathol 2000;190:462–469. [PubMed: 10699996]

29. Veltri RW, Partin AW, Miller MC. Quantitative nuclear grade (QNG): a new image analysis-based biomarker of clinically relevant nuclear structure alterations. J Cell Biochem 2000:151–157. [PubMed: 10797574]

30. Guillaud M, Cox D, Adler-Storthz K, et al. Exploratory analysis of quantitative histopathology of cervical intraepithelial neoplasia: objectivity, reproducibility, malignancy-associated changes, and human papillomavirus. Cytometry *A* 2004;60:81–89. [PubMed: 15229860]

31. Poulin N, Frost A, Carraro A, et al. Risk biomarker assessment for breast cancer progression: replication precision of nuclear morphometry. Anal Cell Pathol 2003;25:129–138. [PubMed: 12775917]

32. Hanselaar A, Poulin N, Pahlplatz M, et al. DNA-cytometry of progressive and regressive cervical intraepithelial DNA-cytometry of progressive and regressive cervical intraepithelial neoplasia. Anal Cell Pathol 1998;16:11–27. [PubMed: 9584897]

33. Rodenacker K, Bengtsson E. A feature set for cytometry on digitized microscopic images. Anal Cell Pathol 2003;25:1–36. [PubMed: 12590175]

34. Rosin MP, Poh CF, Guillaud M, William PM, Zhang L, MacAulay C. Visualization and other emerging technologies as change makers for oral cancer prevention. Ann NY Acad Sci 2007;1098:167–183. [PubMed: 17332080]

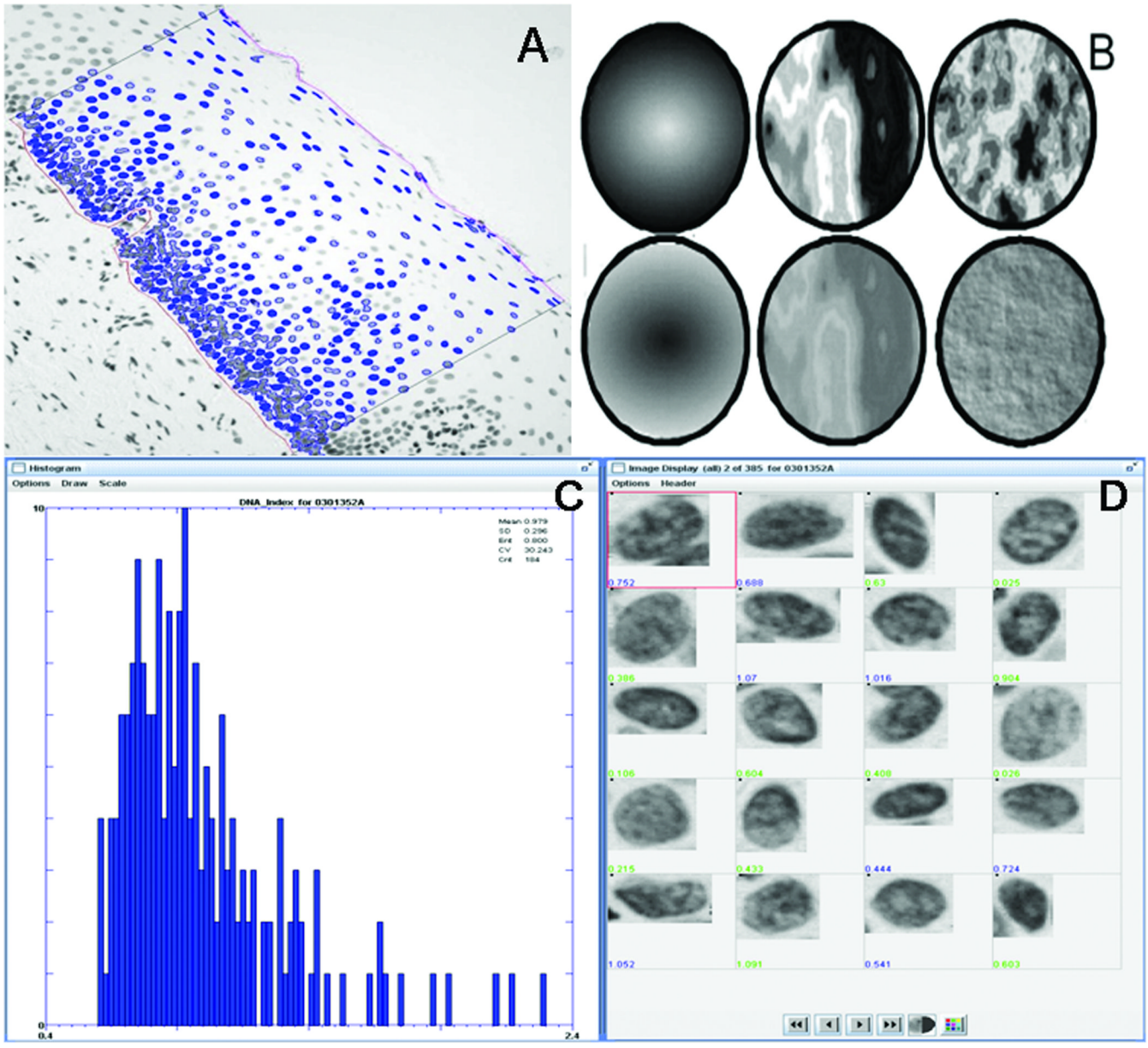

Fig. 1.

Quantitative Tissue Phenotype (QTP) of an oral dysplastic epithelium. A: A Region of Interest (ROI) is manually delineated according to pathologist diagnosis. Nuclei selected by the technician within the ROI are automatically segmented (in blue); B: Graphic representations of 3 of the 110 nuclear features assessed in this study, all of which measure DNA distribution in the nucleus. Top row: *OD-Skewness*_measures whether the nucleus is dark with light areas or light with dark areas; middle row: *Fractal_area1* measures heterochromatin vs. euchromatin organization, i.e. large intensity contrast between highly condensed chromatin and non-condensed chromatin, bottom row: *Long90_Run* measures the fraction of nuclear diameter one can travel before an intensity change is encountered. C: Nuclei images and corresponding features are displayed for graphical and statistical representation and analysis.

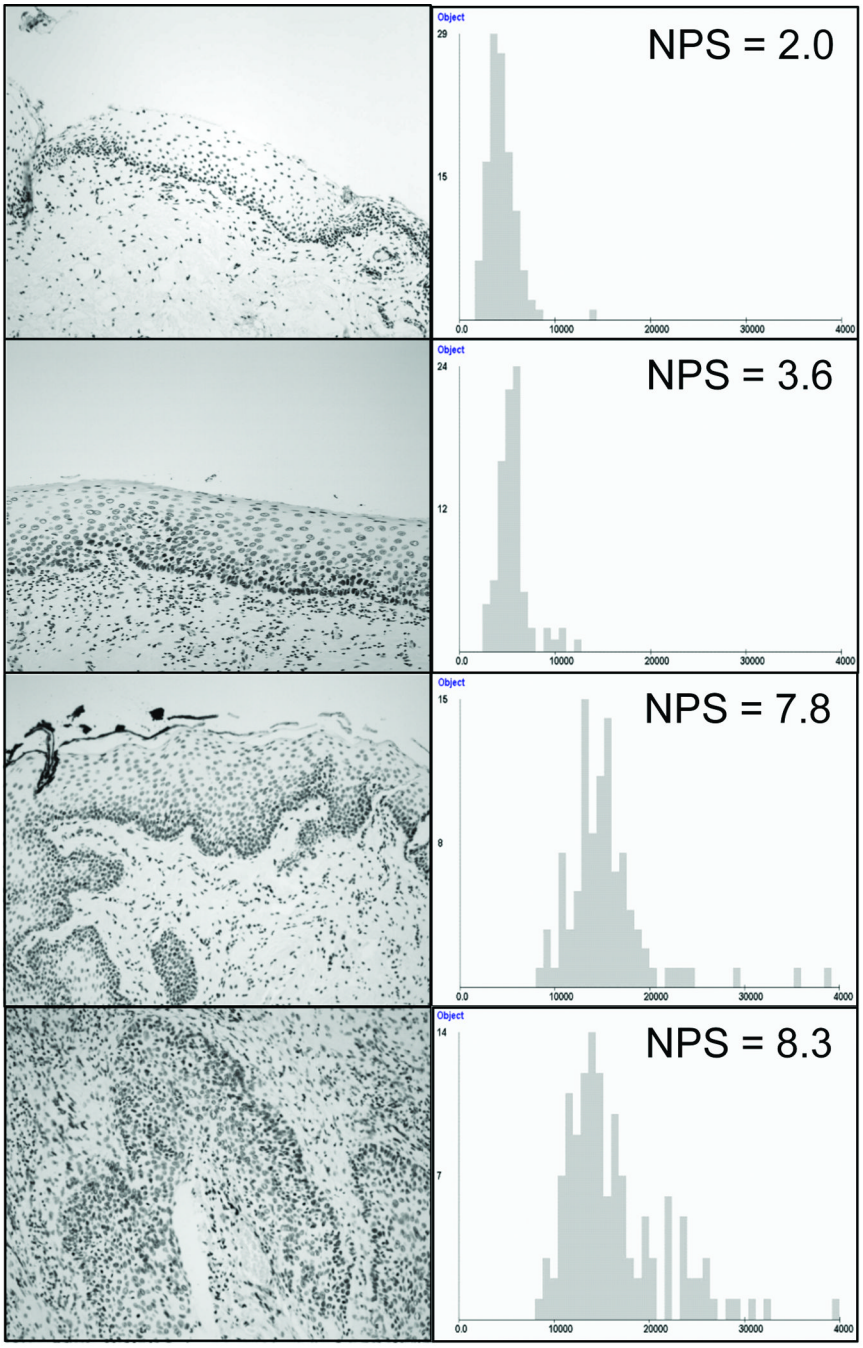

Fig. 2.

Images of the Region of Interest (ROI) of four oral mucosa specimens with the corresponding histogram distribution of texture feature *Fractal_area1*, used in the calculation of the NPS. Top row: normal epithelium; second row: non-progressing mild dysplasia; third row: progressing mild dysplasia; fourth row: SCC.

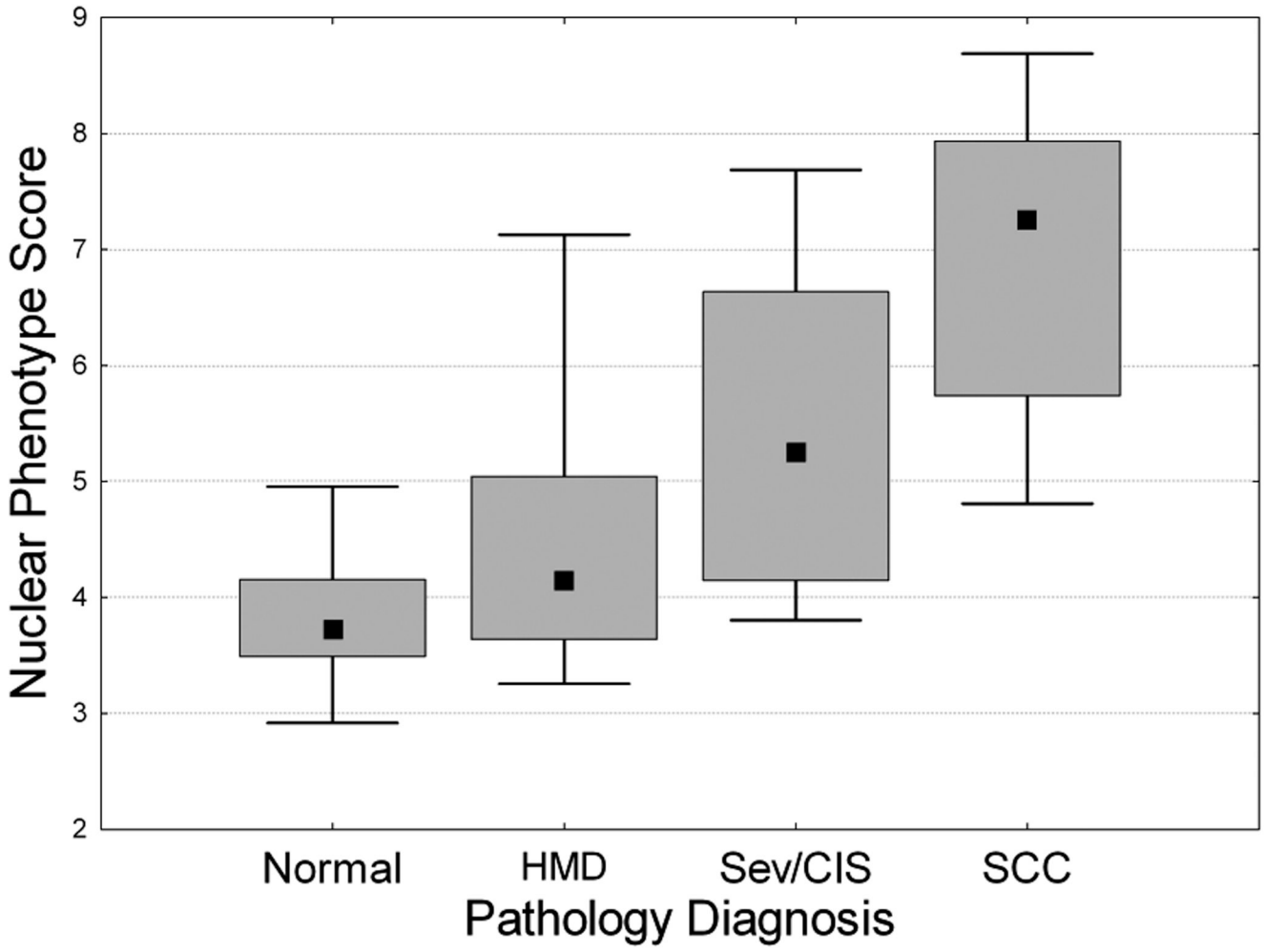

Fig. 3.

Correlation of Nuclear Phenotype Score (NPS) with histopathology grade. Error bar represents 5th and 95th percentiles, box represents central 50th percentile and black square represents the NPS Median. N: Normal; HMD: Hyperplasia, Mild, Moderate Dysplasia; Sev-*CIS*: Severe dysplasia/*CIS*; SCC: Squamous Cell Carcinoma.

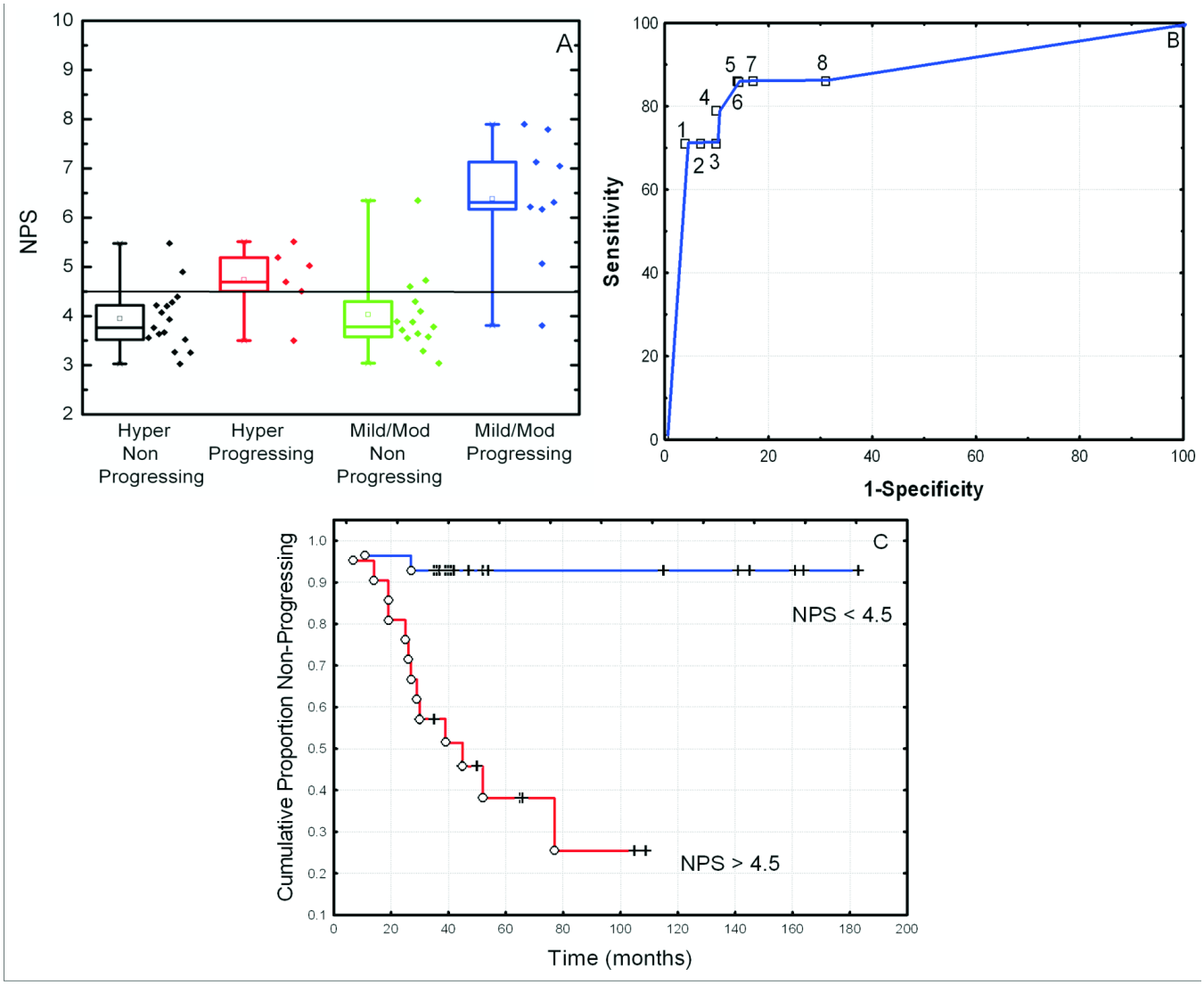

Fig. 4.

Correlation of NPS with progression to cancer for hyperplasia, mild and moderate dysplasia (HMD). (A) Box plots of NPS for progressing and non–progressing hyperplasia, and progressing and non-progressing mild/moderate dysplasia. Error bar represents 5th and 95th percentiles, box represents central 50th percentile and black square represents median NPS values. (B) Receiver Operating Characteristic Curve of NPS to identify progressing lesions with a cut-off value of 4.2 (1), 4.5 (2) and 4.7 (3). (C) Probability of having no progression to cancer, for specimens with a low NPS (NPS <4.5) and specimens with a High NPS (NPS > 4.5).

Table 1

Proportions of lesions with low and high NPS medians in groups with different LOH patterns.

Pattern	# Specimens Informative	Low NPS		High		*P*
		No loss	Loss	No Loss	Loss	
3p LOH	39	20	4	8	7	*0.03*
9p LOH	43	21	6	5	11	*0.003*
4q LOH	41	27	0	9	5	*0.003*
8p LOH	41	23	3	10	5	0.1
11q LOH	43	25	2	11	5	0.05
13q LOH	40	25	1	12	2	0.27
17p LOH	43	23	4	12	4	0.33
At least 1 loss	43	17	10	3	13	*0.006*
Multiple loss	43	23	4	5	11	*0.0005*
LOH at 3p or 9p	43	20	7	3	13	*0.0006*
High-risk LOH(*)	43	23	4	5	11	*0.0005*

* LOH at 3p &/or 9p plus at least 1 other arm (4q, 8p, 11q, 13q or 17p). Loci analyzed are given in ref 14.

Table 2

Analysis of cancer risk by combining NPS, LOH and Histology using the Cox proportional regression hazards model.

Model	N	Variable	B Value (Standard Error)	*P* alues
1	43	NPS	2.57 (0.76)	0.0007
2	43	LOH	3.5 (1.04)	0.0008
3	43	Histology	0.05 (0.35)	0.88
4	43	NPS	1.46 (0.81)	0.07
		LOH	2.81 (1.09	0.01
5	43	NPS	2.6 (0.77)	0.0006
		Histology	0.2 (0.34)	0.54
6	43	NPS	1.46 (0.81)	0.07
		LOH	2.82 (1.09)	0.01
		Histology	0.01 (0.36)	0.96

Models 1–3 correspond to univariate models. Model 4 and 5 represent models with combination of markers.