Specific changes of chromatin structure in nuclei of normal epithelium adjacent to laryngeal squamous cell carcinoma

A preliminary study of 82 cases

Thomas Dreyer [a,*], Iris Knoblauch [a],
David Garner [b], Alexei Doudkine [b],
Calum MacAulay [b], Branko Palcic [b] and
Christian Popella [c]

[a] *Institute of Pathology, Justus Liebig University of Giessen, Germany*
[b] *Cancer Imaging Department, British Columbia Cancer Agency, Vancouver BC, Canada*
[c] *Department of Otorhinolaryngology, Justus Liebig University of Giessen, Germany*

The aim of this study was to confirm the existence of specific nuclear texture feature alterations of histologically normal epithelial borders nearby invasive laryngeal cancer (NC).

Paraffin sections of NC and of chronic inflammations unrelated to cancer (CI) were analysed for nuclear texture and for integrated optical density (IOD-index) and were compared to normal epithelium of patients without evidence of cancer (NE). Several discriminant functions based on nuclear texture features were trained to separate different subgroups.

As the most important result, specific nuclear texture feature shifts were only found in NC with high-density lymphocytic stroma infiltrate (NC+). Classification of nuclei of NE versus NC+ was correct in 70%. The same classifier was correct in only 58% when nuclei of NE were classified versus CI. We also found lower values of IOD-Index within the NC+ group when compared to NE ($p < 0.001$).

*Corresponding author: Dr. Thomas Dreyer, Institut für Pathologie, Justus Liebig Universität, Langhansstrasse 10, D-35392 Giessen, Germany. Tel.: +49 641 99 41114; Fax: +49 641 99 41109; E-mail: thomas.dreyer@patho.med.uni-giessen.de.

1. Introduction

Laryngeal cancer is the most frequent malignant tumour of the upper aerodigestive tract in Europe [28]. This is probably attributable to an increased use of tobacco in many countries [9] including the female population during the past decades [43]. As laryngeal cancer of advanced stage can only be cured by laryngectomy, it is mandatory to evade such therapy by earlier cancer detection and by microsurgical treatment [24]. Studies performed in the past revealed that at the vicinity of cancer normal epithelium exhibits specific shifts in nuclear texture feature expression [27]. We recently found this phenomenon at the vicinity of laryngeal cancer [12]. Such alterations of chromatin distribution may be related to paracrine influences of cancer tissue on surrounding normal epithelium [27].

The aim of this study was to analyse the characteristics of such nuclear changes by careful microscopic selection of tumour borders without dysplastic or cancerous changes. As the amount of inflammatory stroma infiltrate was very different, we also correlated these features to the density of stroma lymphocytes. In order to estimate proliferation activity, we also measured the integrated optical density (IOD)-Index.

2. Material and methods

We used paraffin sections of surgical specimens of 20 cases of normal epithelium (NE), 20 cases of chronic inflammation unrelated to cancer (CI), and of 42 cases of histologically normal borders (NC) of invasive laryngeal cancer treated at the Department of Otorhinolaryngology of the Justus Liebig University in Giessen. The specimens were collected from the files of the Institute of Pathology of the Justus Liebig University of Giessen between 1994 and 1997.

The density of lymphocytic stroma infiltrate was quantified by counting the number of lymphocytes within 10 measuring fields. By setting a threshold at the median, which was equivalent to 50 lymphocytes per 10 measuring fields, the NC group was then split into subgroups of high-density (NC+) and of low-density lymphocytic stroma infiltrate (NC−).

Sections were stained according to a Feulgen–Thionin protocol [18,20]. Quantitative measurements were performed using a high-resolution image cytometer Cyto-Savant[TM] [33] and a Nikon $20 \times$ PlanApo objective with a numerical aperture of 0.75. Bright field images were acquired using a 10 nm bandpass filter centred at 600 nm corresponding to the spectral region of the peak absorption of the nuclear stain. The imaging unit was equipped with a high-resolution digital camera (MicroImager 1400, Xillix Technologies Corp.; pixel size: $6.8 \times 6.8 \ \mu m^2$; spatial resolution: 0.34 μm). At least 20 lymphocytes per section were taken as internal standard in order to normalise features depending on optical density. To obtain a maximum of nuclear texture information, on average 150 nuclei per section only of the lower half of the epithelium were selected. All nuclei per group were then pooled to superfiles.

In order to train discriminant functions to separate nuclei of NE versus NC, a linear stepwise procedure was performed based on a set of 65 source features of nuclear texture. These functions were trained to optimise the classification rate of NE nuclei versus nuclei of NC− as well as of NE nuclei versus NC+. The function trained at NE versus NC+ and based on 10 out of the 65 source features is designated as classifier 1. To detect non-specific nuclear alterations related to lymphocytic stroma response (epithelial healing effects), we also classified nuclei of NE versus CI. All classification processes were followed by jack-knife testing.

IOD-Index, which is equivalent to integrated optical density normalised against integrated optical density of reference cells, was calculated as mean value per section. The slide mean values of different subgroups were compared (Mann–Whitney U-test). So far, discriminant functions were designed and tested at the same material and only jack-knife testing was applied to estimate their performance. In order to obtain data also from completely independent test material, an additional test was performed, where all groups were split into sets of training (tr) and test cases (te) (Table 1). Only the training sets were then used for development of new functions, which were then applied to the test sets.

Table 1
Number of cases and of nuclei per subgroup investigated

Subgroup	Number of cases	Number of nuclei
NEtr	10	1448
NEte	10	1353
Citr	10	1443
CIte	10	1446
NCtr+	10	1417
NCte+	11	1569
NCtr−	11	1512
NCte−	10	1385

NE = normal epithelium; CI = chronic inflammation; NC = normal epithelium at the vicinity of cancer; tr = training set; te = test set; +/− = density of lymphocytic infiltrate above (+) or below (−) the threshold as defined above.
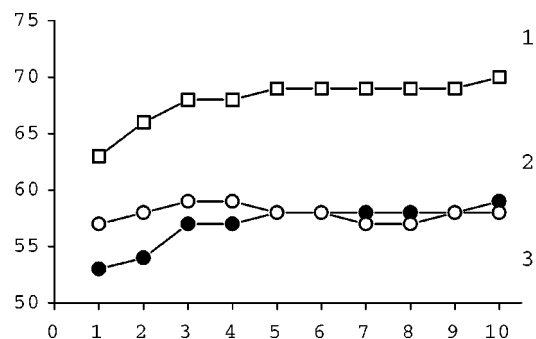

Fig. 1. Cell-by-cell classification data of discriminant functions based on 65 source features. Ten runs of stepwise discriminant function analysis were performed, in which 1 to 10 of the source features were allowed to build functions for classification of normal epithelium versus normal-appearing borders of invasive laryngeal cancer with dense lymphocytic stroma infiltrate (1). These functions were then applied to classification of normal epithelium versus normal borders of invasive cancer with low-density lymphocytic infiltrate (2) as well as to classification of normal epithelium versus chronic inflammation unrelated to cancer (3).

The function trained at ten features to classify NEtr versus NCtr+ (classifier 2) was then analysed in greater detail. Thereby, each of the selected single features was used to classify NEte versus NCte+. The respective classification rates are given and their level of significance was calculated by chi-square values.

3. Results

Using stepwise discriminant functions based on 65 source features, classification was only successful for NE versus NC+. When only 5 of these features were selected, the classification rate was 69%. When the number of selected features was extended to 10, classi-

Table 2
List of features selected by classifiers 1 and 2

List of the features selected by classifier 1, created to optimise classification of normal epithelium (NE) versus histologically normal epithelial cells nearby invasive cancer with high-density lymphocytic stroma infiltrate (NC+)			List of features selected by classifier 2, created to optimise classification of the same groups, but trained at the training set only		
Feature	abbreviation	F value	Feature	abbreviation	F value
High DNA area	[hiDNAar]	775.2	High DNA Area	[hiDNAar]	363.1
Short 45 runs	[sh45rns]	154.2	DNA component	[lwDNAcp]	313.2
Medium density object	[medenob]	118.8	Low DNA amount	[lwDNAat]	171.3
Medium DNA amount	[meDNAat]	116.0	Medium density object	[medenob]	158.1
Cluster shade	[clshd]	103.0	Short 0 runs	[sh0rns]	37.7
Low DNA amount	[lwDNAat]	49.8	Cluster prominence	[clprm]	36.9
Run 90 percent	[rn90pt]	39.4	Energy	–	3.42
Low density object	[lwdenob]	37.6	Low density object	[lwdenob]	1.9
Cluster prominence	[clprm]	27.0	Run 135 percent	[rn135pt]	1.0
Low DNA component	[lwDNAcp]	25.7	Cluster shade	[clshd]	0.1

fication rate could be increased to 70%. Using the same function to classify nuclei of NE versus NC−, the classification rate based on 5 features was only 58% and could not be increased further when 10 features were involved. In order to check the specificity of the functions trained to classify NE versus NC+, we also applied these functions to the classification of NE versus CI. This again resulted in only 58% correctly classified nuclei when 5 features were selected and could not be further improved after inclusion of the other 5 features (Fig. 1). All classifications were confirmed by jack-knife testing, which led to the same results. Features selected by discriminant function, which was based on 10 features to classify NE versus NC+ (Table 2), are explained in the appendix, greater details are given elsewhere [11]. Interestingly, two features with high classification rates, as indicated by their F-values, describe the relative amount of highly condensed chromatin. While "high DNA area" calculates the area fraction of high optical density and thereby represents a direct measure of relative amount of condensed chromatin, "short 45 runs" describes the relative amount of short pixel arrays of the same optical density. If the relative amount of short pixel arrays is high, this indicates that within a given nucleus frequent changes of optical density predominate, which is the case when disconnected areas of high-density chromatin are present (Fig. 2).

Measurements of IOD-Index revealed lower values within the NC+ group when compared to the NE group ($p < 0.001$, Whitney U-test), whereas all other groups did not reveal significant differences.

Finally, a second training procedure restricted to the training sets was performed. These functions were then
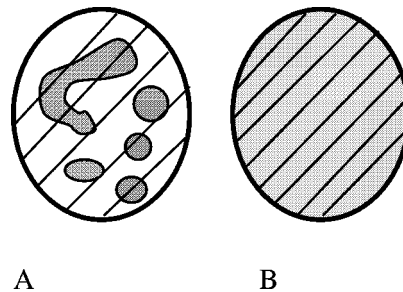

A   B

Fig. 2. Nucleus containing disconnected areas of high-density chromatin. This nucleus would reveal a high value of "high DNA area" and "short 45 runs" (A). Nuclei with disperse low-density chromatin reveal lower values for both features (B). As features describing pixel arrays depend on directions, they cannot be regarded as rotationally invariant.

applied to the respective test sets, resulting in nearly the same data (Fig. 3). When 10 features were used to classify NEtr versus NCtr+ (classifier 2), this function was still correct in 67% of the corresponding test nuclei (NEte versus NCte+). Again, poor classification data were obtained when this function was applied to the classification of NEte versus NCte− as well as to the classification of NEte versus CIte.

Mean values and variation per group of the features selected by classifier 1 or 2 are given in Table 3, single feature classification rates are given in Table 4. These data demonstrate significant differences between nuclei of NEte and NCte+ when features are used which describe the spatial distribution of chromatin at different condensation states, whereas features describing the length of pixel arrays of similar density running into specified directions are not rotationally invariant,

Table 3

Group mean values and coefficients of variation (%) of features selected by classifier 1 or 2

| Feature	NEtr	NEte	CItr	CIte	NCtr+	NCtr−	NCte+	NCte−
Energy	0.09 (14.05)	0.08 (13.86)	0.09 (13.97)	0.09 (14.89)	0.09 (13.61)	0.08 (14.82)	0.08 (15.72)	0.08 (15.48)
HiDNAar	0.52 (52.23)	0.61 (43.48)	0.4 (72.01)	0.54 (59.24)	0.33 (81.86)	0.46 (60.4)	0.41 (69.43)	0.56 (48.23)
LwDNAat	0.04 (138.2)	0.03 (177.0)	0.05 (148.9)	0.06 (210.9)	0.08 (132.6)	0.05 (145.5)	0.07 (159.1)	0.04 (175.1)
MeDNAat	0.33 (79.0)	0.25 (98.73)	0.46 (62.88)	0.31 (92.27)	0.51 (53.75)	0.38 (70.51)	0.42 (66.47)	0.29 (90.86)
Lwdenob	4.68 (76.11)	3.41 (96.09)	4.81 (72.8)	3.24 (97.08)	4.86 (67.99)	4.88 (72.6)	4.57 (71.63)	3.65 (87.65)
Medenob	2.46 (78.48)	4.15 (62.89)	1.85 (85.56)	2.9 (77.85)	1.68 (78.92)	2.42 (78.68)	2.2 (83.73)	2.74 (81.21)
LwDNAcp	2.15 (75.16)	1.52 (91.08)	2.36 (69.17)	1.79 (90.35)	3.32 (57.85)	2.41 (69.03)	2.51 (65.09)	1.96 (86)
Clshd	0.08 (432.2)	0.03 (1180)	0.05 (702.6)	0.1 (345.1)	0.07 (459.1)	0.04 (722.0)	0.02 (1927)	0.12 (256.7)
Clprm	2.48 (16.88)	2.46 (16.6)	2.54 (18.97)	2.49 (18.83)	2.52 (19.21)	2.45 (17.83)	2.44 (19.52)	2.39 (18.38)
Sh0rns	0.69 (9.5)	0.71 (9.4)	0.68 (9.25)	0.68 (9.25)	0.67 (8.92)	0.69 (9.06)	0.69 (9.15)	0.70 (9.23)
Sh45rns	0.69 (9.38)	0.7 (8.33)	0.69 (8.0)	0.7 (9.31)	0.69 (8.99)	0.69 (10.03)	0.71 (9.59)	0.71 (9.82)
Rn90pt	0.62 (10.57)	0.66 (10.04)	0.63 (10.06)	0.62 (10.9)	0.63 (9.91)	0.64 (10.23)	0.66 (11.49)	0.64 (11.3)
Rn135pt	0.64 (10.85)	0.66 (9.88)	0.63 (10.15)	0.64 (10.8)	0.64 (10.61)	0.65 (11.15)	0.66 (10.78)	0.65 (11.38)
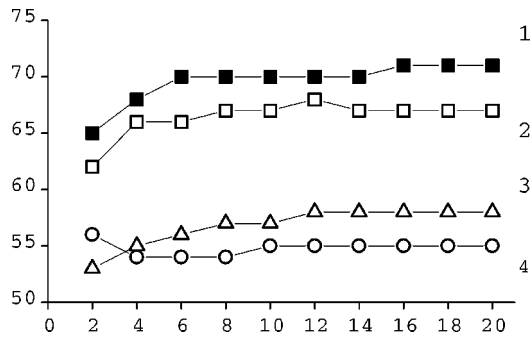
Fig. 3. Cell-by-cell classification data obtained from a discriminant function trained at the training set of normal epithelium (NEtr) versus the training set of normal-appearing epithelium with dense lymphocytic infiltrate nearby cancer (NCtr+) (1). These functions were then applied to classification of the test sets of normal epithelium (NEte) versus normal-appearing borders of invasive laryngeal cancer with dense lymphocytic stroma infiltrate (NCte+) (2), to NEte versus normal-appearing epithelium nearby cancer of low density lymphocytic infiltrate (NCte−) (3) as well as to classification of NEte versus chronic inflammation (CIte) (4).

Table 4

Classification rate of single features selected by classifier 1 or 2 for nuclei of NEte versus NCte+

Feature	Classification rate [%]	Level of significance
Medenob	68.7	$p < 0.0001$
LwDNAcp	64.2	$p < 0.0001$
HiDNAar	63.5	$p < 0.0001$
MeDNAat	62.1	$p < 0.0001$
LwDNAat	59.0	$p < 0.0001$
Lwdnob	58.0	$p < 0.0001$
Sh0rns	56.2	$p < 0.0001$
Clshd	51.9	$p < 0.05$
Rn135pt	51.3	n.s.
Clprm	50.4	n.s.
Sh45rns	49.9	n.s.
Rn90pt	49.8	n.s.
Energy	49.0	n.s.

Significance was tested by the chi-square value. n.s. = not significant.

and therefore, were not confirmed at independent test set material.

While the slide means of IOD were significantly lower within the NCte+ group in comparison to the NEte group ($p < 0.02$) such difference was not observed at the corresponding training sets. Additionally, IOD values of all nuclei of these groups were directly compared separately for training and test sets. Thereby, both sets of NC+ revealed significantly smaller values in comparison to their corresponding NE sets ($p < 0.001$) (Fig. 4).

4. Discussion

Most importantly, these results indicate that specific shifts in nuclear features of histologically normal epithelial cells occur under the influence of adjacent laryngeal cancer [12]. These features predominantly describe changes in spatial chromatin distribution measured as texture feature alterations. The current observations are consistent with findings of others that also indicate that shifts in nuclear texture features of normal cells can be found at the vicinity of cancers [6,7, 17,41].

In this preliminary study, classifier 1 was trained and tested at the same material and only jack-knife testing
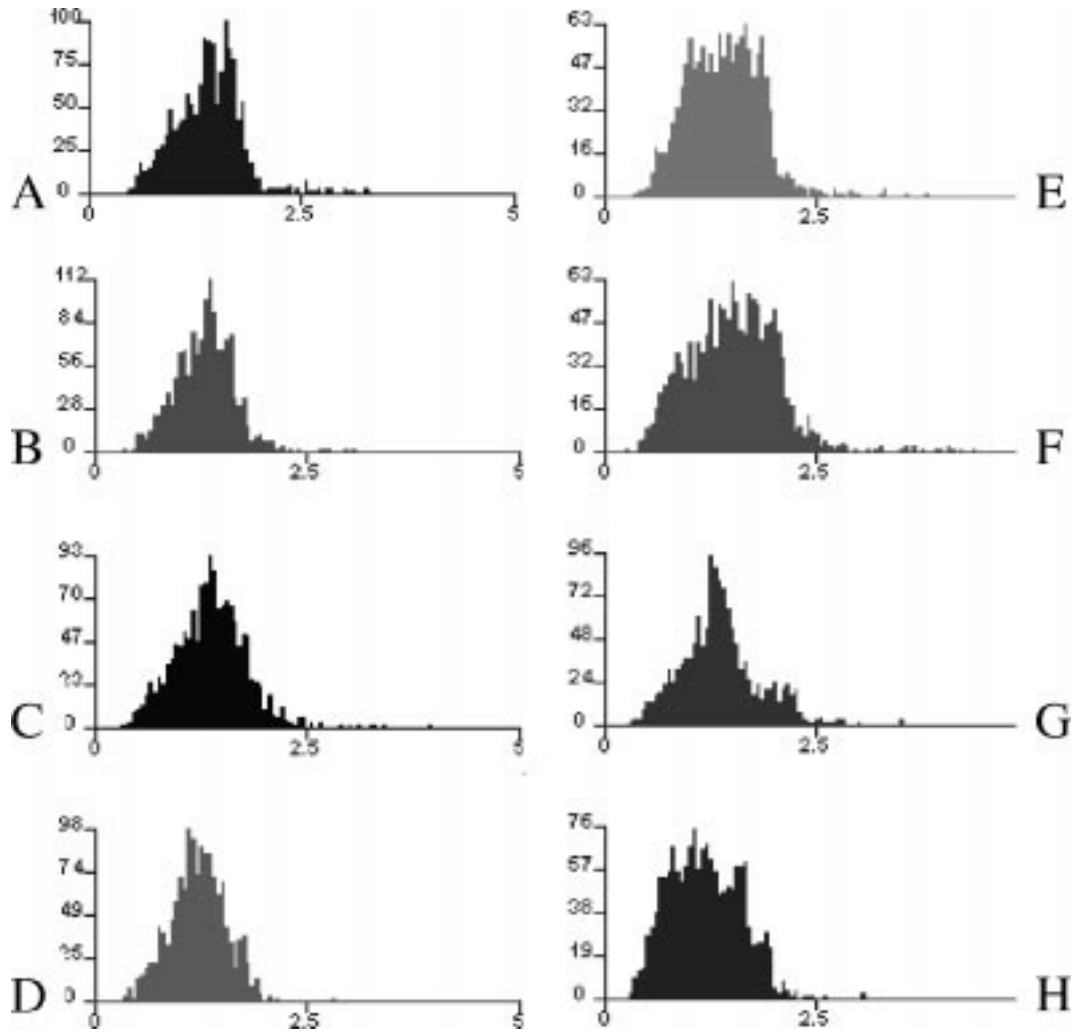
Fig. 4. IOD histograms of all measured nuclei per subgroup. (A) and (B): training set of normal epithelium (A), and of benign epithelium with inflammatory stroma infiltrate (B), (C) and (D): training set of normal-appearing epithelium with low-density (C) and high-density (D) inflammatory infiltrate of the stroma nearby invasive cancer, (E) and (F): test set of normal epithelium (E) and of benign epithelium with inflammatory infiltrate (F), (G) and (H): benign- appearing epithelium nearby invasive cancer with low-density (G) and high-density (H) inflammatory stroma infiltrate.

was used for estimation of classification rate. While this test performs well in feature estimation, it does not compensate for training errors based on feature selection. Hence, data of jack-knife classification still may be too optimistic in comparison to testing the same classifier at completely independent material [26,34]. Furthermore, training procedures based on nuclei increase the number of objects in comparison to slide-based classification. However, nuclei of the same section were processed in the same way. Therefore, features sensitive to staining and fixation variability may reveal highly correlated data at such nuclei. Hence, the number of independent objects cannot be regarded

as equivalent to the total number of nuclei measured. Total independence can only be expected from different specimens. Thereby, the number of sections (specimens) is also limiting the independence of nuclei [37, 38].

Nevertheless, our study clearly shows that classification rates obtained from independent test material were similar to those obtained from all cases. Therefore, we conclude that our finding is similar, if not identical, to the phenomenon described as Malignancy Associated Changes (MAC) [3,8,14,21,23,30,32,47].

To date the biological background of this particular phenomenon remains poorly understood.

In the head and neck area very early stages of field cancerisation responsible for the development of synchronous [10,36] or metachronous primaries [39] are the place where the origin of common chromosomal aberrations is expected [4]. Following this concept, MAC might be interpreted as a very early stage of preneoplastic change, where chromosomal alterations induce changes in nuclear texture expression identified as MAC but not visible by conventional histology. However, this concept is not in agreement with the observation that MAC-feature expression is reversible after tumour removal [27].

Another possible explanation is that specific immune response to tumour antigens mediated by different cell types specialised in antigen presentation and mediatd by T-lymphocytes [25] might also cause specific alterations of normal epithelial cells at the tumour vicinity identified as MAC. This could explain the absence of MAC feature expression in the absence of lymphocytic immune response.

Extracellular signals mediated by cytokines and growth factors may also be involved in MAC-feature expression [40]. Their influence on activation and repression of transcription of various genes may also explain the observed shifts in nuclear texture features of histologically normal epithelium at the vicinity of cancer [1,5,35].

Another possible concept is that the MAC-phenomenon reflects very early and thereby subvisible stages of apoptotic cell death. This concept would be consistent with two observations of our study. Firstly, the features detecting texture differences between benign changes versus epithelium at the vicinity of invasive cancer describe chromatin clumping as it is typically the case in the late-stage apoptosis, which is only then clearly visible by means of conventional light microscopy. This especially refers to "high DNA area" calculating the relative amount of condensed chromatin. This feature revealed the highest F-values for classification of NE versus NC+ indicating that texture alterations of this feature group are typical for normal-appearing cells at the vicinity of laryngeal cancer. Moreover, some of the features selected in this study are identical to those that are already suggested to be specific for apoptotic changes of nuclear chromatin arrangement. This is typically the case for "run length" features as well as for "cluster shade" [29]. This is even more interesting as both studies rely on different material. While Matthews and co-workers used cell cultures of a Chinese hamster ovary, our study was based on paraffin sections. Secondly, in this study MAC-feature

expression was observed in the NC+ group whereas the NC− group was clearly MAC-negative. In contrast to the other groups, NC+ also revealed a significantly reduced group mean value for IOD-Index. Because DNA measurement at tissue sections cannot be used for single cell or single slide classification for several reasons [2,15,16,19,46], this method has to be avoided for diagnostic purposes. However, in this study we only compared the group mean values of IOD-Index. Even these data are affected by sectioning artefacts, nuclear overlapping and other sources of errors typically related to this procedure. Nevertheless, despite of the technical restrictions, the group mean difference between NE and NC+ was highly significant and therefore cannot be readily explained by any of the known sources of errors. These findings lead to different possible interpretations. As apoptosis is known to be related to the cell cycle, one possible explanation would be that in this study apoptosis occurred predominantly as a post-mitotic event resulting in diploid nuclei. Alternatively, if apoptosis occurs during the S-phase, the DNA content of these cells may be increased but the proliferation activity may be down-regulated. Both interpretations would explain our observation of reduced DNA-Index in comparison to normal epithelium.

If the MAC-phenomenon is related to apoptotic cell death the contradiction of reversibility of MAC-feature expression after successful tumour removal on the one hand versus irreversibility of cell death on the other hand has to be explained. One possible approach to this problem is that behaviour of cell populations cannot be projected to behaviour of single cells. While MAC-feature analysis usually requires analysis of cell populations, apoptosis typically is an observation at single cell level [22]. Thereby, the reversibility of the MAC-feature expression after complete tumour removal only means that this feature cannot be identified any longer. It is therefore conceivable that the MAC-phenomenon precedes cell death and thus can be irreversible. After successful tumour treatment these cells would die, then be removed and only normal cells would be left. In this case the MAC-phenomenon would be reversible at population level [27,40] but irreversible at the cellular level. Thereby, specific shifts in texture feature expression might reflect early stages of apoptotic DNA digestion [45].

Another phenomenon must also be explained. Apoptosis usually appears in the absence of inflammatory response, while in this study the MAC-feature expression was found exclusively in the presence of lympho-

cytic stroma infiltrate. For further explanation the nature of the lymphocytic infiltrate would have to be analysed in greater detail, which needs different sample preparation and therefore was beyond the scope of this study. However, one possible explanation might be that the lymphocytic infiltrate is primarily directed against the tumour cells but not against the epithelial cells at the tumour vicinity. Malignancies are capable of inducing apoptosis of anti-tumour effector cells (activated T-lymhocytes) thus escaping the host's immune response to tumour cells [31]. The presence of a dense lymphocytic infiltrate thereby might only reflect the conflict between the immune system attacking the tumour and the tumour counterattacking the immune response. In this case the lymphocytic stroma infiltrate also present at the vicinity of cancer does only indicate that the tumor is able to induce apoptose not only at the epithelium but also at the lymphocytes of the vicinity.

As the MAC-phenomenon per definition describes subvisible changes, it has to be explained why apoptosis was not detected visually. This can be explained by the fact that the visible features typically observed in apoptotic cells only last for less than 1 hour [44]. Thus, even if the frequency of apoptotic cells is increased, visible changes still remain a rare event.

These data do not provide a final concept regarding relationship between MAC phenomenon and apoptosis. Nevertheless, a considerable amount of evidences does indicate that the MAC-phenomenon is related to apototic cell death.

As fresh material was not available for this retrospective investigation additional prospective studies are necessary to analyse the role of apoptosis at the vicinity of cancer.

Acknowledgement

Appendix

The following description refers to features selected by discriminant functions. The images are represented by a matrix of pixels, $P_{i,j}$, spanning $i = 1, L$ columns and $j = 1, M$ rows and with the upper left hand pixel

as the co-ordinate system origin, $i = j = 1$. The binary object mask function defines the group of image pixels belonging to the object:

$$\Omega = (\Omega_{1,1}, \Omega_{1,2}, \ldots, \Omega_{i,j}, \ldots, \Omega_{L,M}),$$

where

$$\Omega_{i,j} = \begin{cases} 1 & \text{if } (i,j) \in \text{object}, \\ 0 & \text{if } (i,j) \notin \text{object} \end{cases}$$

and where, "$(i,j) \in$ object" means pixels at co-ordinates i and j are part of the object, and "$(i,j) \notin$ object" means pixels at co-ordinates i and j are not part of the object.

A.1. IOD Index

This feature is the normalised measure of the integrated optical density of the object:

$$\text{IOD Index} = \frac{IOD\ Amount}{iod_{norm}},$$

where iod_{norm} is the mean value of the *IOD Amount* for a particular object population from the slide (e.g., lymphocytes). *IOD Amount* is the raw (unnormalised) measure of IOD.

A.2. Low/medium/and high DNA area

These discrete texture features classify optical density into classes of low, medium (med), and high (hi). Here, only the ratio of the area of high optical density regions of the object to the total object area is measured [hiDNAar]:

$$\text{hiDNAar} = \frac{\sum_{i=1}^{L} \sum_{j=1}^{M} \Omega_{i,j}^{\text{hi}}}{\sum_{i=1}^{L} \sum_{j=1}^{M} \Omega_{i,j}} = \frac{A^{\text{hi}}}{A},$$

where A is the object area.

A.3. Low/medium/and high DNA amount

These discrete texture features represent the total extinction ratio for low-, medium-, and high- optical density regions of the object, calculated as the value of the integrated optical density of the low-, medium-, and high-density regions, respectively, divided by the total integrated optical density. Out of this group only low

DNA amount [lwDNAat] and medium DNA amount [meDNAat] were selected:

$$\mathrm{lwDNAat} = \frac{\sum_{i=1}^{L}\sum_{j=1}^{M} OD_{i,j}\Omega_{i,j}^{\mathrm{low}}}{\sum_{i=1}^{L}\sum_{j=1}^{M} OD_{i,j}\Omega_{i,j}},$$

$$\mathrm{meDNAat} = \frac{\sum_{i=1}^{L}\sum_{j=1}^{M} OD_{i,j}\Omega_{i,j}^{\mathrm{med}}}{\sum_{i=1}^{L}\sum_{j=1}^{M} \mathrm{OD}_{i,j}\Omega_{i,j}},$$

where *OD* is the optical pixel density.

A.4. Low/medium/high/and medium-high DNA component

These discrete texture features are characteristic of the compactness of low-, medium-, high-, and combined medium- and high-density regions, respectively, treated as single disconnected objects. In this study we only used "low DNA component" [lwDNAcp], which is defined as follows:

$$\mathrm{lwDNAcp} = \frac{(P^{\mathrm{low}})^2}{4\pi\, A^{\mathrm{low}}},$$

A = area; P = perimeter.

A.5. Low-/medium-/and high-density object

These discrete texture features are the numbers of connected subcomponents of the objects consisting of more than one pixel of low, medium, and high density.

A.6. Entropy/energy

Only energy was selected in this study which represents a measure of "order" in object grey level organisation: large values correspond to large regions of constant grey level:

$$\mathrm{energy} = \sum_{l}(H_l^s)^2 + \sum_{m}(H_m^d)^2,$$

where H_l^s is the probability of neighbouring pixels having grey levels which sum to l, and H_m^d is the probability of neighbouring pixels having grey level differences of m, where an 8-connected neighbourhood is assumed [42]. Values of grey levels, l, m, used in the sum and difference histogram are obtained by quantisation of the dynamic range of each individual object into 40 levels.

A.7. Cluster shade [clshd]

This feature gives large absolute values for objects with a few distinct clumps of uniform intensity having large contrast with the rest of the object. Negative values correspond to dark clumps against a light background while positive values indicate light clumps against a dark background:

$$\mathrm{clshd} = \frac{\sum_l(l - 2\overline{I^q})^3 H_l^s}{(\sum_l(l - 2\overline{I^q})^2 H_l^s)^{3/2}},$$

H_l^s, H_m^d, l, m are defined above, $\overline{I^q}$ represents the mean intensity of the object calculated for the grey scale quantized to 40 levels.

A.8. Cluster prominence [clprm]

This feature measures the darkness of clusters:

$$\mathrm{clprm} = \frac{\sum_l(l - 2\overline{I^q})^4 H_l^s}{(\sum_l(l - 2\overline{I^q})^2 H_l^s)^2},$$

H_l^s, $\overline{I^q}$, and l are defined above.

A.9. Short 0/45/90/and 135 runs

These features belong to a group called run length. This group describes texture in terms of grey level runs, representing sets of consecutive, collinear pixels having the same grey level value. The length of the run is the number of pixels in the run. These features are calculated over the image with intensity function values transformed into 8 levels. The run length texture features are defined using grey level length matrices, $\Re_{p,q}^{\Theta}$ for each of the four principal directions: $\theta = 0°$, $45°$, $90°$, $135°$, where the directions are defined clockwise with respect to the positive x-axis [13]. Each element of matrix $\Re_{p,q}^{\Theta}$ specifies the number of times that the object contains a run of length q, in a given direction, $é$, consisting of pixels lying in the grey level range, p (out of 8 grey levels). Let $N^g = 8$ be the number of grey levels, and N^r be the number of different run lengths that occur in the object; then this feature is described as follows: large values for objects are obtained if short runs, oriented at $0°$, $45°$, $90°$, or $135°$ dominate. In this study only the features "short 0 runs" [sh0rns] and "short 45 runs" [sh45rns] were selected:

$$\mathrm{sh}\theta\_\mathrm{rns} = \frac{\sum_{p=1}^{N^g}\sum_{q=1}^{N^r}(\Re_{p,q}^{\Theta}/q^2)}{\sum_{p=1}^{N^g}\sum_{q=1}^{N^r}\Re_{p,q}^{\Theta}}.$$

A.10. run 0/45/90/and 135 percent

These features are calculated as the ratio of the total number of possible runs to the object's area, having its lowest value for pictures with the most linear structure. In this study, only "run 90 percent" [rn90pt] and "run 135 percent" [rn135pt] were selected.

$$\text{rn}\theta\text{pt} = \frac{\sum_{p=1}^{N^g} \sum_{q=1}^{N^r} \Re_{p,q}^{\Theta}}{A},$$

where A is the object area.

References

[1] S.A. Aaronson, Growth factors and cancer, *Science* **254** (1991), 1146.

[2] J.P. Baak, E. Noteboom and J.J. Koevoets, The influence of fixatives and other variations in tissue processing on nuclear morphometric features, *Anal. Quant. Cytol. Histol.* **11** (1989), 219–224.

[3] P.H. Bartels, M. Bibbo, H.E. Dytsch, F.T. Pishotta, K. Yamauchi and G.L. Wied, Diagnostic marker display for intermediate cells from the uterine cervix, *Acta Cytol.* **26** (1982), 29–34.

[4] G.C. Bedi, W.H. Westra, E. Gabrielson, W. Koch and D. Sidransky, Multiple head and neck tumours: Evidence of a common clonal origin, *Cancer Res.* **56** (1996), 2484–2487.

[5] J. Blenis, Signal transduction via the MAP kinasis: proceed at your own RSK, *Proc. Natl. Acad. Sci.* **90** (1993), 5889.

[6] M. Bibbo, A.G. Montag, E. Lerma-Puertas, H.E. Dytsch, S. Leelakusolvong and P.H. Bartels, Karyometric marker features in normal appearing glands adjacent to human colonic adenocarcinoma, *Anal. Quant. Cytol. Histol.* **11** (1989), 281–285.

[7] M. Bibbo, A.G. Montag, E. Lerma-Puertas, H.E. Dytsch, S. Leelakusolvong and P.H. Bartels, Karyometric marker features in normal appearing glands adjacent to human colonic adenocarcinoma, *Cancer Res.* **50** (1990), 147–151.

[8] G. Burger, U. Jütting and K. Rodinacker, Changes in the benign cell population in cases of cervical cancer and its precursors, *Anal. Quant. Cytol.* **3** (1981), 261–271.

[9] M.S. Cattaruzza, P. Maisonneuve and P. Boyle, Epidemiology of laryngeal cancer, *Eur. J. Cancer B. Oral. Oncol.* **32B** (1996), 293–305.

[10] I.J. Dhooge, M. De-Vos and P.B. Van-Cauwenberge, Multiple primary malignant tumors in patients with head and neck cancer: results of a prospective study and future perspectives, *Laryngoscope* **108** (1998), 250–256.

[11] A. Doudkine, C.E. MacAulay, N. Poulin and B. Palcic, Nuclear texture measurements in image cytometry, *Pathologica* **87** (1995), 286–299.

[12] T. Dreyer, A. Doudkine, C.E. MacAulay, B. Palcic and H. Glanz, Quantitative analysis of laryngeal epithelial lesions using image cytometry, *Anal. Cell. Pathol.* **13** (1997), 64.

[13] M.M. Galloway, Texture analysis using gray level run lengths, *Computer Graphics, Image Processing* **4** (1975), 172–179.

[14] O.C. Gruner, Study of the changes met with the leucocytes in certain cases of malignant disease, *Br. J. Surg.* **3** (1916), 506–522.

[15] A. Gschwendtner and T. Mairinger, How thick is your section? The influence of section thickness on DNA-cytometry on histological sections, *Anal. Cell. Pathol.* **9** (1995), 29–37.

[16] A. Gschwendtner, A. Kreczy and K. Prein, DNA cytometry: diploid standard and section thickness [letter; comment], *Am. J. Clin. Pathol.* **100** (1994), 451–455.

[17] M. Guillaud, A. Doudkine, D. Garner and C.E. MacAulay, Malignancy associated changes in cervical smears: systematic changes in cytometric features with the grade of dysplasia, *Anal. Cell. Pathol.* **9** (1995), 191–204.

[18] A.G. Hanselaar, N. Poulin, M.M. Pahlplatz, D.M. Garner, C. MacAulay, J. Matisic, J. LeRiche and B. Palcic, DNA cytometry of progressive and regressive cervical intraepithelial neoplasia, *Anal. Cell. Pathol.* **16** (1998), 11–27.

[19] G. Haroske, W. Meyer, V. Dimmer, K.D. Kunze and F. Theissig, Analysis procedure in tissue sections, *Zentralbl. Pathol.* **139** (1994), 407–417.

[20] N. Ikeda, C.E. MacAulay, S. Lam, J. LeRiche, P. Payne, D.M. Garner, C. Konaka, H. Kato and B. Palcic, Malignancy associated changes in the bronchial epithelial cells and clinical application as a biomarker, *Lung Cancer* **19** (1998), 161–166.

[21] R.A. Kemp, C.E. MacAulay, D. Garner and B. Palcic, Detection of malignancy associated changes in cervical cell nuclei using feed-forward neural networks, *Anal. Cell. Pathol.* **14** (1997), 31–40.

[22] J.F.R. Kerr, A.H. Wyllie and A.R. Currie, Apoptosis: a basic biological phenomenon with wide ranging implications in tissue kinetics, *Br. J. Cancer* **26** (1972), 239–257.

[23] H. Klawe and J. Rowinski, Malignancy associated changes (MAC) in cells of buccal smears detected by objective image analysis, *Acta Cytologica* **18** (1974), 33–33.

[24] Kleinsasser, *Mikrolaryngoskopie und endolaryngeale Mikrochirurgie, Technik und typische Befunde*, FK Schattauer, Stuttgart, 1968.

[25] M. Kronenberg, G. Siu, L.E. Hood and N. Shastri, The molecular genetics of the T-cell antigen receptor and T-cell antigen recognition, *Annu. Rev. Immunol.* **4** (1986), 529.

[26] L.E. Larsen, D.O. Walter, J.J. MacNew and W.R. Adey, On the problem of bias in error rate estimation for discriminant analysis, *Pattern Recog.* **3** (1971), 217–223.

[27] C.E. MacAulay, S. Lam, P.W. Payne, J.C. LeRiche and B. Palcic, Malignancy associated changes in bronchial epithelial cells in biopsy specimens, *Anal. Quant. Cytol. Histol.* **17** (1995), 55–61.

[28] H. Maier and M. Tisch, Epidemiology of laryngeal cancer: results of the Heidelberg case-control study, *Acta Otolaryngol. Suppl. (Stockh.)* **527** (1997), 160–164.

[29] J.B. Matthews, A. Harrison, B. Palcic and K. Skov, Automated fluorescence microscopic measurement of apoptosis frequency following ionizing radiation exposure in cultured mammalian cells, *Int. J. Radiat. Biol.* **73** (1998), 629–639.

[30] H.E. Nieburgs, B.E. Herman and H. Reisman, Buccal cell changes in patients with malignant tumors, *Lab. Invest.* **11** (1962), 80–188.

[31] J. O'Connell, G.C. O'Sullivan, J.K. Collins and F. Shanahan, The Fas counterattack: Fas-mediated T cell killing by colon cancer expressing Fas ligand, *J. Exp. Med.* **184** (1996), 1075–1082.

[32] B. Palcic, D.M. Garner and C.E. MacAulay, Image cytometry and chemoprevention in cervical cancer, *J. Cell. Biochem.* Suppl. **23** (1995), 43–54.

[33] B. Palcic, D.M. Garner, C.E. MacAulay, J. Matisic and G.H. Anderson, Oncometrics Imaging Corporation and Xillix Technologies Corporation. Use of the Cyto Savant in quantitative cytology, *Acta Cytol.* **40** (1996), 67–72.

[34] S. Raudys and A. Jain, Small sample size effect in statistical pattern recognition: recommendations for practitioners, *IEEE Trans. Pattern. Anal. Mach. Intell.* **13** (1991), 252–264.

[35] J. Schlessinger and D. Bar-Sagi, Activation of RAS and other signaling pathways by receptor tyrosine kinases, *Cold Spring Harbor Symp. Quant. Biol.* **59** (1994), 173.

[36] A.G. Scholes, J.A. Woolgar, M.A. Boyle, J.S. Brown, E.D. Vaughan, C.A. Hart, A.S. Jones and J.K. Field, Synchronous oral carcinomas: independent or common clonal origin?, *Cancer Res.* **58** (1998), 2003–2006.

[37] H. Schulerud, G.B. Kristensen, L. Vlatkovic, F. Albregtsen, K. Liestol and H.E. Danielsen, Prognosis of cervical cancer using image analysis of cell nuclei, in: *Proc. of 10th Scandinavian Conference on Image Analysis*, Vol. 2, 1997, pp. 651–658.

[38] H. Schulerrud, G.B. Kristensen, K. Liestoel, L. Vlatkovic, A. Reith, F. Albregtsen and H.E. Danielsen, A review of caveats in statistical nuclear image analysis, *Anal. Cell. Pathol.* **16** (1998), 63–82.

[39] D.L. Slaughter, H.W. Southwick and W. Smejkal, "Field cancerization" in oral stratified squamous epithelium: clinical implications of multicentric origin, *Cancer* **6** (1953), 963–968.

[40] X.R. Sun, C.E. MacAulay and B. Palcic, An in vitro model for studying malignancy associated changes, *Anal. Cell. Pathol.* **13** (1997), 66.

[41] B. Susnik, A. Worth, J. LeRiche and B. Palcic, Malignancy associated changes (MAC) in breast: changes in chromatin distribution of epithelial cells in normal appearing tissue adjacent to carcinoma, *Anal. Quant. Cytol. Histol.* **17** (1995), 62–68.

[42] M. Unser, Sum and difference histograms for texture classification, *IEEE Trans. Pattern. Anal. Mach. Intell. (PAMI)* **8** (1996), 118–125.

[43] J.A.H. Waterhouse, Epidemiology, in: *Neoplasms of the Larynx*, A. Ferlito, ed., Churchill Livingstone, London, 1993, pp. 49–64.

[44] C. Vidair, C. Chen, C. Ling and W. Dewey, Apoptosis induced by X-irradiation of rec-myc cells is postmitotic and not predicted by the time after irradiation or behaviour of sister cells, *Cancer Res.* **56** (1996), 4116–4118.

[45] A. Wyllie, Glucocorticoid induced thymocyte apoptosis is associated with endogenous endonuclease activation, *Nature* **284** (1980), 555.

[46] X. Xin, J. Sudbi, M. Boysen and A. Reith, Caveats in the use of paraffin sections for DNA assessment in nuclei as demonstrated by confocal laser scanning microscopy, *Anal. Quant. Cytol. Histol.* **18** (1996), 104.

[47] D.J. Zahniser, K.L. Wong, J.F. Brenner, H.G. Ball and M.L. Hutchinson, Contextual analysis and intermediate cell markers enhance high resolution cell image analysis for automated cervical smear diagnosis, *Cytometry* **12** (1991), 10–14.