DNA copy number aberrations in endobronchial lesions:

A validated predictor for cancer

Robert AA van Boerdonk
Johannes MA Daniels
Peter JF Snijders
Katrien Grünberg
Erik Thunnissen
Mark A van de Wiel
Bauke Ylstra
Pieter E Postmus
Chris JLM Meijer
Gerrit A Meijer
Egbert F Smit
Thomas G Sutedja
Daniëlle AM Heideman

Thorax 2014;69(5);464-470
CHAPTER 4 VALIDATION OF CNA-BASED CLASSIFIER TO PREDICT LUNG CANCER

ABSTRACT

Introduction
We recently identified a DNA copy number aberration (CNA)-based classifier, including changes at 3p26.3-p11.1, 3q26.2-29, and 6p25.3-24.3, as a risk predictor for cancer in individuals presenting with endobronchial squamous metaplasia. The current study was set out to validate the prediction accuracy of this classifier in an independent series of endobronchial squamous metaplastic and dysplastic lesions.

Materials and Methods
The study included 36 high-risk subjects who had endobronchial lesions of various histological grades that were identified and biopsied by autofluorescence bronchoscopy and were subjected to arrayCGH in a nested case-control design. Of the 36 patients, 12 had a carcinoma in situ or invasive carcinoma at the same site at follow-up (median 11 months, range 4-24), while 24 controls remained cancer-free (78 months, range 21-142).

Results
The previously defined CNA-based classifier demonstrated 92% (95%CI: 77-98%) accuracy for cancer (in situ) prediction. All nine subjects with CNA-based classifier-positive endobronchial lesions at baseline experienced cancer outcome, whereas all 24 controls and 3 cases were classified as being low-risk.

Conclusions
In conclusion, CNAs prove to be a highly accurate biomarker for assessing the progression risk of endobronchial squamous metaplastic and dysplastic lesions. This classifier could assist in selecting subjects with endobronchial lesions who might benefit from more aggressive therapeutic intervention or surveillance.

INTRODUCTION
Lung cancer is a major cause of cancer-related mortality worldwide. This is largely due to the fact that the majority of patients is diagnosed with advanced stage disease when treatment options with curative intent are generally unfeasible. The ability to detect lung cancer at early or pre-invasive stages together with adequate curative treatment could have a favorable impact on disease outcome. Early detection is currently available given sensitive clinical tools such as low-dose computed tomography (LDCT) and autofluorescence bronchoscopy (AFB). With the advent of these imaging techniques, the challenge has become to differentiate true premalignant lesions with high-risk of progression to cancer from the many suspicious, though regressive lesions to better target interventions and avoid overtreatment.

In case of AFB, endobronchial sites showing abnormal autofluorescence can be biopsied for histopathological examination. Histological grading according to the WHO classification, however, does not provide an accurate estimate of cancer risk. Whereas endobronchial lesions with squamous metaplasia and dysplasia are at risk of developing into lung cancer, with the highest risk being associated with the most severe grades, these lesions behave erratically and only a minority truly progresses towards lung cancer. As such, the prediction of the progression risk of an individual endobronchial lesion by means of histological classification is imprecise and subject to considerable inter-observer variation. Therefore, biomarkers are needed that can aid in assessing an individual’s risk of subsequent cancer to tailor intervention strategies in case these lesions are detected.

In the search for such biomarkers, we recently analyzed a series of endobronchial squamous metaplastic lesions with known clinicopathological outcome (i.e., location-matched endobronchial cancer in cases, and cancer-free outcome in controls) by means of arrayCGH. This study discovered two distinct clusters at the molecular level with significant association to lesion outcome, that is, one cluster with lesions showing multiple DNA copy number aberrations (CNAs), including exclusively progressive lesions of cases, and one with lesions showing no or relatively few CNAs, including all non-progressive lesions of controls and one lesion of a case. In this discovery series of endobronchial squamous metaplastic lesions, a molecular classifier was built based on CNAs at 3p26.3-p11.1 (loss), 3q26.2-29 (gain), and 6p25.3-24.3 (loss), that predicted cancer diagnosed within 44 months with 97% accuracy.

In the current study, we validate the prediction accuracy of the CNA-based classifier using an independent set of endobronchial lesions of various histological grades. A case-control arrayCGH profiling study was performed nested within a cohort of subjects at risk of lung cancer, but without evidence of prevalent carcinoma in situ (CIS) or invasive carcinoma at baseline bronchoscopy, who underwent AFB at regular time intervals. The series comprised endobronchial lesions of 12 patients who were identified with CIS or invasive carcinoma (2CIS) at the initial site on follow-up bronchoscopy (further referred to as cases), and 24 controls who remained cancer-free. Accuracy of the previously defined CNA-based classifier in predicting endobronchial cancer in this series at baseline was determined.
METHODS

Patient population
This study was nested within a cohort of 479 subjects (365 male and 114 female subjects) at risk of lung cancer based on smoking habits (i.e., more than 20 pack years), chronic obstructive pulmonary disease (COPD), signs and symptoms, and/or a history of ear-nose-throat (ENT) or lung cancer, who visited the Department of Pulmonary Diseases, VU University Medical Center (Amsterdam, The Netherlands) for baseline AFB examination between October 1995 and November 2011. Clinical and histological characteristics of individual AFB-visualized lesions were carefully documented as subjects underwent AFB surveillance at regular time intervals (on average 3-6 month time intervals) with biopsy collection for histopathological examination from all sites showing abnormal autofluorescence at the respective or previous visits. For each site, paired formalin-fixed paraffin-embedded (FFPE) and frozen biopsies guided by AFB were obtained after informed consent from subjects. Study approval was obtained by the Institutional Review Board of VU University Medical Center. Follow-up data of subjects were available through the Dutch nationwide network and registry of histo- and cytopathology (PALGA, Bunnik, the Netherlands).

A total of 42 subjects without evidence of prevalent ≥CIS were documented with location-matched CIS or invasive carcinoma at the initial site within a follow-up period exceeding 90 days. Of these, 12 case subjects (10 men, 2 women) were eligible for this study based on the following: 1) confirmed absence of ≥CIS in the location-matched baseline endobronchial biopsy by comprehensive histopathological review of H&E-stained FFPE and corresponding cryo-sections (see below), 2) availability of sufficient DNA quality and quantity following laser-capture microdissection of frozen biopsy for molecular analysis, and 3) not included in the previous study12.

Twenty-four control subjects were chosen in such a way that relative to cases a similar distribution was ensured for those parameters that in theory might confound the analyses, that is, baseline routine histopathological diagnosis, age, gender, smoking habits, COPD and history of cancer, but who remained cancer-free for a follow-up period that exceeded that of respective cases. At the individual level, each case had two matched controls. For matching, continuous variables were dichotomized and controls were chosen to fall within the same category as their matching case for at least 4 of the 6 abovementioned parameters. The number of controls was chosen to be twice that of the cases because a further increase in sample size was considered to render only modest improvement of power given the fixed number of the case group and the expected homogeneous nature of arrayCGH profiles in the control group12. Follow-up data, supplemented through PALGA, supported the cancer-free status of control subjects for a median of 78 months (range 21-142 months).

Histopathological evaluation
Two designated pathologists (ET, KG) who were blinded for the clinicopathological information and molecular findings performed an independent review of all baseline and follow-up biopsies using Haematoxylin & Eosin (H&E)-stained FFPE and cryo-preserved tissue sections. Histopathological assessment of FFPE tissue sections was performed in accordance with the World Health Organization/International Association for the Study of Lung Cancer (WHO/IASLC) histological classification system of pre-invasive and invasive squamous lesions of the bronchus13,14. Cryo-sections were categorized using the following composite grouping scheme: (I) normal histology (N); (II) low-grade disease (LGD), including squamous metaplasia (SqM), mild dysplasia (miD), and moderate dysplasia (moD); (III) high-grade disease (HGD), including severe dysplasia (SD); and (IV) CIS or invasive carcinoma. Despite the limitations faced in accurate histological examination of H&E-stained frozen sections, this approach is reasonably accurate for assessing sample representativeness, and the research purpose of correlating histology with molecular analysis15. In case the grade of dysplasia could not be further defined, the term ‘dysplasia-indefinite for grade’ (D-ind) was used. Lesions were ultimately classified into three categories: LGD, HGD and indeterminate on the basis of the highest histomorphological abnormality scored on H&E-stained FFPE and frozen sections.

ArrayCGH analysis
A designated molecular biologist (RvB) performed sample preparation and genomic DNA isolation of frozen biopsy specimens as described before12. To specifically enrich for epithelial cells (≥70%), all samples were laser-capture microdissected using a Leica ASLMD microscope (Leica, Wetzlar, Germany) prior to genomic DNA extraction, as described before12. Subsequent linear whole genome amplification, Cy3-/Cy5-labeling (both from Enzo Life Sciences, Farmingdale, NY, USA), and hybridization of laser-capture microdissected biopsy-derived DNAs onto the Agilent 4x44K arrayCGH platform (Agilent Technologies, Santa Clara, CA, USA) were performed as previously described12,16,17. Samples were co-hybridized against a Megapool Reference DNA (Kreatech Diagnostics, Amsterdam, The Netherlands), a homogeneous pool of DNA isolates from hundred healthy male or female individuals (XY-mix and XX-mix, respectively). Across arrayCGH comparisons were made to measure test signals12. Array scan and imaging data acquisition were performed as described before (Agilent CGH 107_Sep09 protocol)12. Oligonucleotides were mapped according to the human genome build NCBI 36 (March 2006). The entire dataset was available through the Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/), under GSE accession number: GSE45287. Accompanying data analyses were done as previously described12, including smoothing, segmentation, calling and regioning of the current dataset.

Measuring PIK3CA copy number by quantitative real-time PCR
PIK3CA gene copy number was determined by quantitative real-time PCR (qPCR) on a 7500 Fast Real-Time PCR System (Applied Biosystems, Nieuwerkerk a/d IJssel, The Netherlands) using non-amplified genomic DNA as essentially described before12. All qPCR reactions were carried out in duplicate, and the threshold cycle numbers were averaged.
Data and statistical analysis

The molecular analyses were performed in a blinded fashion and results were correlated to clinicopathological data afterwards. Unsupervised hierarchical clustering analysis of arrayCGH data was performed to determine the association between genomic profiles and lesion outcome\(^\text{18}\). The association between clustering results and case/control status was determined by chi-square testing. Cancer risk was predicted by using the previously defined statistical ‘endobronchial cancer risk’ model based on CNAs at 3p26.3-p11.1, 3q26.2-29, and 6p25.3-24.3\(^\text{12}\). When the predicted probability is larger than 0.5, classification of the sample is ‘high-risk’ for endobronchial cancer. Prediction accuracy, sensitivity and specificity of the CNA-based classifier for endobronchial cancer (in situ) were calculated in this series. Given the skewed study population, the denominators for positive and negative predictive value calculations are essentially unknown and therefore these parameters were not calculated. Strength of the association between copy number classification based on arrayCGH and real-time PCR analyses was examined using McNemar testing. Two-sided p-values below 0.05 were considered statistically significant. Statistical analyses were performed using IBM SPSS Statistics v20.0 software package (New York, NY, USA).

RESULTS

Baseline characteristics of study subjects and their endobronchial lesions

Table 1 shows demographics, histology and clinical variables for the individuals of the current study stratified for lesion outcome (i.e., cases and controls). Detailed histological diagnoses at study entry and molecular findings are listed per lesion analyzed in Supplementary Table 1. The arrayCGH DNA copy number profiles of the endobronchial lesions of all controls (24/24; 100%) and 3 (25%) cases were quiescent with only few genomic aberrations (average percentage of altered features 0.2%, range 0.0-2.4%; cluster 1, Figure 1) while the DNA copy number profiles of the remaining cases (9/12, 75%) displayed multiple genomic aberrations (average percentage of altered features 51.8%, range 27.6-78.2%; cluster 2, Figure 1). A significant association between cluster assignment by unsupervised hierarchical clustering analysis and lesion outcome was found (p<0.001). To determine most common CNAs in our series, the frequency of gains and losses were plotted per genomic probe (Figure 2). Highly frequent genomic aberrations (≥50% frequency) in cases included gains at 3q23-q29 and 19q13.12-q13.2 and losses at 3p26.3-p11.1, 5q11.1-q35.3, 8p23.3-p12 and 9p24.3-p21.1.

Validation of the risk classifier

We next applied the output from the model built in our previous study\(^\text{12}\), that is, a risk classifier based on specific CNAs at 3p26.3-p11.1 (loss), 3q26.2-q29 (gain) and 6p25.3-p24.3 (loss), using a predicted probability cutoff of 0.5 to assign individuals to high- or low-risk groups, and tested whether the assignments are related to lesion outcome (Table 2). All subjects with CNA classifier-
CHAPTER 4
VALIDATION OF CNA-BASED CLASSIFIER TO PREDICT LUNG CANCER

FIGURE 1 Unsupervised hierarchical clustering analysis.

Unsupervised hierarchical clustering analysis of 36 baseline endobronchial lesions revealed two clusters: cluster 1, in which the complete set of controls congregated, and cluster 2, in which the majority of cases were assigned. Genomic loci are marked by corresponding chromosome numbers (Y-axis). Loci that were lost, gained, or amplified are shown in red, green, and white, respectively. Cases (black boxes) and controls (white boxes) are represented on the X-axis.

Positive endobronchial lesions at baseline experienced cancer outcome, whereas all 24 controls and the remaining 3 cases were classified as low-risk. No intermediate test results with a probability around the threshold level of 0.5 were observed in this series. The CNA-classifier had a 92% (95%CI: 77-98%) prediction accuracy, 75% (95%CI: 46-92%) sensitivity and 100% (95%CI: 84-100%) specificity for endobronchial carcinoma (in situ) in the current series (Table 2).

FIGURE 2 Frequency plots.

A) Frequency plots of CNAs are shown for A) baseline endobronchial lesions of cases (n=12), and B) baseline endobronchial lesions of controls (n=24). Percentages of gains (in green) and losses (in red) as determined by 4x44K arrayCGH analysis are shown per spotted oligonucleotide on the positive and negative Y-axis, respectively. Dashed lines within the figures indicate 50% frequency of gain or loss.

Predominant value of chromosomal region 3q

Aiming to fine-tune the CNA-based classifier, logistic regression analysis with forward feature selection using both arrayCGH datasets combined (GSE23644 and GSE45287; total of 18 cases and 47 controls) was performed. This confirmed CNAs at 3p26.3-p11.1, 3q26.2-q29 and 6p25.3-p24.3 as parameters of the classifier, and could not reveal additional predictive DNA copy number feature(s) for refinement of the risk model. When we examined the relative contribution of the individual chromosomal regions to our risk model, the gain at 3q26.2-q29 contributed most. Gain
at 3q26.2-29 was present in virtually all lesions classified as high-risk by the model (13/14, 93%). We further assessed the value of a gain at 3q26.2-q29 for risk prediction by independent qPCR analysis of the PIK3CA gene located within this chromosomal region (Figure 3). A 95% concordance between arrayCGH and qPCR analyses was obtained, with one case sample revealing a 3q26.32 gain by arrayCGH while not by qPCR, and two control samples showing solely a gain by qPCR. The qPCR had an 85% (95%CI: 74-92%) prediction accuracy, 61% (95%CI: 39-80%) sensitivity and 94% (95%CI: 82-98%) specificity for endobronchial carcinoma (in situ).

**TABLE 2 Risk classifier**

<table>
<thead>
<tr>
<th>Outcome risk classifier</th>
<th>Lesion outcome</th>
<th>cancer (in situ)</th>
<th>cancer-free</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-risk</td>
<td>3</td>
<td>24</td>
<td>27</td>
<td>Prediction accuracy: 92% (95%CI: 77-98%)</td>
</tr>
<tr>
<td>High-risk</td>
<td>9</td>
<td>-</td>
<td>9</td>
<td>Sensitivity: 75% (95%CI: 46-92%)</td>
</tr>
<tr>
<td>Total</td>
<td>12</td>
<td>24</td>
<td>36</td>
<td>Specificity: 100% (95%CI: 84-100%)</td>
</tr>
</tbody>
</table>

*) Risk was predicted by using the previously defined model based on CNAs at 3p26.3-p11.1, 3q26.2-29, and 6p25.3-24.3. When the predicted probability is larger than 0.5, classification of the sample is ‘high-risk’ for endobronchial cancer (in situ). Prediction accuracy, sensitivity and specificity of the CNA classifier for endobronchial cancer (in situ) were calculated in this series.

**DISCUSSION**

In the present work, we validated our previously defined CNA-based classifier12 to assess an individual’s risk for subsequent endobronchial cancer among subjects that present with endobronchial squamous metaplastic and dysplastic lesions. Our pre-defined classifier of CNAs at 3p26.3-p11.1 (loss), 3q26.2-29 (gain), and 6p25.3-24.3 (loss) demonstrated 92% (95%CI: 77-98%) accuracy for predicting cancer outcome in this independent validation set. All subjects with CNA classifier-positive endobronchial lesions experienced cancer outcome. Of note, the classifier did not only perform well in lesions graded as squamous metaplasia, as shown here and in the previous study12, but had predictive value along the full spectrum of pre-invasive disease at the histological level while retaining 100% specificity. Thus, CNAs prove to be a highly accurate biomarker for assessing the progression risk of endobronchial squamous metaplastic and dysplastic lesions. The validated classifier proposed here holds great promise for improving the clinical management of subjects diagnosed with endobronchial squamous metaplastic or dysplastic lesions.

During recent years, numerous molecular studies have attempted to unravel whether specific (epi)genetic aberrations underlying lung carcinogenesis are detectable in the premalignant respiratory epithelium and would allow for assessing a person’s risk of developing lung cancer19-27. Yet, molecular biomarkers that confer accurate risk stratification for subjects harboring one or more endobronchial pre-invasive lesions have not been validated. The lack of validated markers may be explained by the fact that the majority of discovery studies involved cross-sectional sample series, thereby not taking into account individual lesion’s outcome20-24;26. Since only few endobronchial squamous metaplastic and dysplastic lesions eventually progress to carcinoma (in situ).
CHAPTER 4  VALIDATION OF CNA-BASED CLASSIFIER TO PREDICT LUNG CANCER

situated, longitudinal cohort studies are desired to identify molecular events that drive progression of these lesions into a malignant phenotype. However, collecting these longitudinal series requires an extensive period of close surveillance of large numbers of individuals at risk of lung cancer. This underscores the uniqueness of the series examined in this study.

The results presented here, together with our previous findings\textsuperscript{21-22}, provide a consistent pattern of CNAs in endobronchial squamous metaplastic and dysplastic lesions that are predictive of endobronchial cancer. This is in agreement with the few longitudinal studies that previously reported amplifications of 3q\textsuperscript{19-20} and regional losses at chromosomes 3p\textsuperscript{19-21}, 5q, and 9p\textsuperscript{21}. Though these studies mainly defined markers within a very limited sample set or by low-resolution analyses, they further confirm the robustness of CNAs as classifiers. The almost uniform finding of chromosome 3q26.2-2q9 gain/amplification in our case series pinpoints the key importance of this chromosomal aberration in squamous lung carcinogenesis\textsuperscript{10,25-28}, as was also recognized in other tumor types\textsuperscript{26,29}. Putative target oncogenes within this region are PIK3CA, SOX2 and TP63\textsuperscript{11}. Interestingly, the two regions that added up to our classifier (i.e., losses at 3p26.3-p11.1 and 6p25.3-24.3) were also (partly) identified as areas of significant CNAs in a large cohort of squamous cell lung cancers analyzed in The Cancer Genome Atlas project, and include putative target tumor suppressor genes such as FOXP1 and FHIT\textsuperscript{12}.

Regarding the clinical performance of the classifier, the specificity of 100% (95%CI: 84-100%) is proficient; allowing early interventional decisions in subjects with endobronchial lesions who were classified as high-risk by the presence of CNAs. Intermediate test results with a probability around the threshold level of 0.5 were infrequently observed (i.e., not in this series, and once in the previous study\textsuperscript{21}), supporting the robustness and utility of the CNA classifier. However, the sensitivity of the CNA classifier could be subject of improvement. In its current format, the classifier lacks sufficient sensitivity for dismissal of subjects with endobronchial lesions that were classified as low-risk by absence of CNAs from further surveillance. The three cases of the current series that were inaccurately classified included case 3 with at maximum the clinical endpoint diagnosis of CIS, case 9 with inadequate quality measures for accurate histological diagnosis of the cryo-biopsy obtained 16 months prior to cancer diagnosis, and case 8 with an indefinite (i.e., no consensus) diagnosis in the cryo-biopsy 4 months prior to cancer diagnosis. For the latter case, raw data processing suggested a gained chromosomal segment at 3q, however, values did not reach the threshold level for being finalized as a gain by the calling algorithm used. Overall, these case data may point to sampling bias, limitation of biopsy quality and/or histopathological examination\textsuperscript{40}, or inaccuracy of the classifier at early onset of (pre)malignant disease as critical factors in the clinical sensitivity of the classifier. Particularly, the contribution of 3q gain to the classifier offers possibilities for a simple diagnostic method, for example qPCR assay, to be applied in biomaterials such as sputum\textsuperscript{41} or bronchial brushings\textsuperscript{42}, however at the cost of some diagnostic accuracy as shown in this study.

The strengths of this study include the validation of the CNA-based classifier as an objective, molecular determinant for lung cancer risk assessment in an independent sample set and the uniqueness of the longitudinal sample series. A limitation of this study may be the number of lesions included. Although we adjusted for potential confounders, some influence of demographic or clinical parameters on lesion outcome cannot completely be ruled out. Furthermore, the prediction accuracy, sensitivity, and specificity figures of the CNA-based classifier may be imprecise due to the small numbers. The latter illustrates the relative infrequency of central airway lesions and underscores the uniqueness of the cohort described in this study.

In summary, our data validate that the presence of specific CNAs in endobronchial squamous metaplastic and dysplastic lesions predicts endobronchial cancer. The validated classifier holds great promise, as alternative or additive to histological grading, for stratification of subjects with endobronchial lesions for risk of subsequent cancer. The classifier may have important practical implications providing guidance to clinicians on early interventional decisions for a generally equivocal set of endobronchial lesions, thereby allowing one to tailor treatment choices and avoid unnecessary follow-up procedures. The classifier may further be of great value as a surrogate endpoint reflecting cancer development or risk in therapeutic and/or chemopreventive trials. A challenge for future research is to assess the effect of earlier intervention of endobronchial squamous metaplastic and dysplastic lesions as guided by the classifier onto lung cancer-specific mortality and the quality of life.
REFERENCES


### SUPPLEMENTARY TABLE 1  
Details of baseline endobronchial lesions: histology and corresponding molecular features

<table>
<thead>
<tr>
<th>Case subjects</th>
<th>Study ID</th>
<th>Routine histology diagnosis FFPE biopsy</th>
<th>Review histology diagnosis FFPE biopsy</th>
<th>Histology diagnosis frozen biopsy</th>
<th>Aberrant genomic features (%)</th>
<th>CNA-based classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>N</td>
<td>ntd</td>
<td>LGD</td>
<td>27.6</td>
<td>pos</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>SqM</td>
<td>miD</td>
<td>LGD</td>
<td>60.6</td>
<td>pos</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>SqM</td>
<td>SqM</td>
<td>LGD</td>
<td>0.0</td>
<td>neg</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>SqM</td>
<td>D-ind</td>
<td>no consensus (HGD/LGD)</td>
<td>78.2</td>
<td>pos</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>moD</td>
<td>miD</td>
<td>no consensus (HGD/LGD)</td>
<td>71.0</td>
<td>pos</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>SD</td>
<td>no consensus (ntd/SD)</td>
<td>HGD</td>
<td>58.2</td>
<td>pos</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>SD</td>
<td>D-ind</td>
<td>D-ind</td>
<td>56.9</td>
<td>pos</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>SD</td>
<td>moD</td>
<td>no consensus (D-ind/N)</td>
<td>0.2</td>
<td>neg</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>SD</td>
<td>D-ind</td>
<td>ntd*</td>
<td>0.1</td>
<td>neg</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>SD</td>
<td>SD</td>
<td>HGD</td>
<td>31.7</td>
<td>pos</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>SD</td>
<td>ntd</td>
<td>D-ind</td>
<td>44.3</td>
<td>pos</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>D-ind</td>
<td>SD</td>
<td>HGD</td>
<td>38.0</td>
<td>pos</td>
<td></td>
</tr>
</tbody>
</table>

Histology review diagnoses are reported as the consensus diagnosis, or “no consensus” (i.e., no agreement between the two expert pathologists) with individual assignment indicated between brackets. Abbreviations used: D-ind, dysplasia-indefinite for grade; HGD, high-grade disease; LGD, low-grade disease; miD, mild dysplasia; moD, moderate dysplasia; N, normal histology; neg, negative (CNAs not present); ntd, not able to determine (i.e., specimen of inadequate quality measures (as described by *)); pos, positive (CNAs present); SD, severe dysplasia; SqM, squamous metaplasia. * absence of CIS confirmed.