

Keywords: lung carcinoid; gene expression profile; diagnostic algorithm; GC; CEACAM1

Gene expression profiling reveals GC and CEACAM1 as new tools in the diagnosis of lung carcinoids

F Toffalorio^{*1,10}, E Belloni^{*2,10}, M Barberis³, G Bucci⁴, L Tizzoni⁵, G Pruneri^{6,7}, C Fumagalli³, G Spitaleri¹, C Catania¹, F Melotti⁸, P G Pelicci^{2,7}, L Spaggiari^{7,9} and T De Pas¹

¹Division of Medical Oncology of the Respiratory Tract, European Institute of Oncology, Via Ripamonti 435, 20141 Milan, Italy;

²Department of Experimental Oncology, Molecular Medicine for Care Program, European Institute of Oncology, Via Ripamonti 435, 20141 Milan, Italy; ³Histopathology and Molecular Diagnostics Unit, Pathology Division, European Institute of Oncology, Milan, Italy; ⁴Center of Genomic Science of IIT@SEMM, Milan, Italy; ⁵Real Time PCR Service, FIRC Institute of Molecular Oncology Foundation, Milan, Italy; ⁶Pathology Division, European Institute of Oncology, Milan, Italy; ⁷University of Milan, School of Medicine, Milan, Italy; ⁸Department of Pathology and Laboratory Medicine, Fondazione IRCCS National Cancer Institute, Milan, Italy and ⁹Thoracic Surgery Division, European Institute of Oncology, Milan, Italy

Background: Classification of lung carcinoids into typical and atypical is a diagnostic challenge since no immunohistochemical tools are available to support pathologists in distinguishing between the two subtypes. A differential diagnosis is essential for clinicians to correctly discuss therapy, prognosis and follow-up with patients. Indeed, the distinction between the two typical and atypical subtypes on biopsies/cytological specimens is still unfeasible and sometimes limited also after radical surgeries. By comparing the gene expression profile of typical (TC) and atypical carcinoids (AC), we intended to find genes specifically expressed in one of the two subtypes that could be used as diagnostic markers.

Methods: Expression profiling, with Affymetrix arrays, was performed on six typical and seven atypical samples. Data were validated on an independent cohort of 29 tumours, by means of quantitative PCR and immunohistochemistry (IHC).

Results: High-throughput gene expression profiling was successfully used to identify a gene signature specific for atypical lung carcinoids. Among the 273 upregulated genes in the atypical vs typical subtype, GC (vitamin D-binding protein) and CEACAM1 (carcinoembryonic antigen family member) emerged as potent diagnostic markers. Quantitative PCR and IHC on a validation set of 17 ACs and 12 TCs confirmed their reproducibility and feasibility.

Conclusions: GC and CEACAM1 can distinguish between TC and AC, defining an IHC assay potentially useful for routine cytological and histochemical diagnostic procedures. The high sensitivity and reproducibility of this new diagnostic algorithm strongly support a further validation on a wider sample size.

Gene expression microarrays can be used to identify signatures specific for tumour subtypes, allowing the classification into categories with a potential impact on clinical practice. In lung cancer, expression profiling unravelled interesting aspects regarding early diagnosis

(Spira *et al*, 2007; Beane *et al*, 2008), prognosis (Bianchi *et al*, 2007), and treatment (Gyorffy *et al*, 2006; Gemma *et al*, 2006).

Lung carcinoids are rare malignant neuroendocrine tumours, accounting for the 1.2% of all lung cancer cases (SEER Research

*Correspondence: Dr F Toffalorio; E-mail: francesca.toffalorio@ieo.it or Dr E Belloni; E-mail: elena.belloni@ieo.eu

¹⁰These authors contributed equally to this work.

Received 2 August 2013; revised 19 December 2013; accepted 7 January 2014; published online 11 February 2014

© 2014 Cancer Research UK. All rights reserved 0007–0920/14

Data, 1973–2004). According to the World Health Organization (WHO) classification, lung neuroendocrine tumours comprise four different tumour types: the low-grade typical carcinoid (TC), the intermediate-grade atypical carcinoid (AC), the high-grade large-cell neuroendocrine carcinoma (LCNEC) and small-cell lung cancer (SCLC) (Arrighoni *et al*, 1972; Travis *et al*, 1998; Scott, 2003). The subclassification of lung carcinoid tumours into typical and atypical is of clinical importance. Typical carcinoid and AC show a different biological behaviour, accounting for the different prognosis: the 5-year survival rate is 40–69% for ACs and 87–100% for TCs (Scott, 2003). As a consequence, distinguishing AC from TC is necessary to correctly discuss patient's prognosis and follow-up, as well as to select patients suitable for clinical trials to test new drugs in adjuvant settings. Moreover, discovering a tool that allows a differential diagnosis between TC and AC in pre-operative cytological or histological specimens could be extremely useful for an individual clinical decision-making, whenever surgery is a high-risk procedure.

Histologically, it is possible to distinguish between TC and AC according to the number of mitosis and the presence of necrosis: TC is characterised by <2 mitoses per 2 mm² (10 HPF) with no necrosis, while AC is defined by a number of mitoses between 2 and 10 per 2 mm² and/or coagulative necrosis. Pathologists referent for these malignancies (Scott, 2003; Skov *et al*, 2008) acknowledge such diagnostic parameters as very difficult to assess, because depending on the size and quality of the specimens. Furthermore, prolonged or inadequate fixation, thick sections, suboptimal staining, or the presence of extensive necrosis, fibrous or inflammatory stroma, and crush artefacts can considerably hamper the evaluation, clearly suggesting that the definition of easily reproducible diagnostic tools is a major clinical need.

The identification of genes differently expressed in TC and AC may be able to identify an AC-specific signature, which, aside from revealing new insights into the nature of these tumours, could display a potential impact on diagnosis.

MATERIALS AND METHODS

Affymetrix microarray. The test set was composed by seven TC and nine AC (diagnosis from the Division of Pathology at the European Institute of Oncology). All the patients not chemo-radio naïve and with a second tumour were excluded from the study. A written informed consent for research use of biological samples was obtained from all the patients.

RNA was extracted from frozen samples with high tumour cellularity (>75%). RNA extraction was performed, using the Illustra RNAspin mini kit (GE Healthcare, Little Chalfont, UK) and following the manufacturer's indications. Tumour sections were prepared and stored in RNAspin Lysis solution at –80 °C until extraction. RNA integrity was assessed by using an Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA). Labelled fragmented cRNA (12 µg) was hybridised to oligonucleotide probes on Affymetrix HG-U133 Plus 2.0 arrays containing 54 675 sets of oligonucleotide probesets, which correspond to ≈30 000 unique human genes. Scanned GeneChip images were processed using the Affymetrix GCOS 1.4 software (Affymetrix, Santa Clara, CA, USA) to obtain an intensity value. Probeset intensities were imported in Partek GS v6.5 (Partek Incorporated, St Louis, MO, USA), and processed using the robust multi-array analysis (RMA) algorithm for background adjustment, normalisation, and log₂ transformation of perfect match (PM) values (Irizarry *et al*, 2006). Data were normalised to the median of all samples. We build an ANOVA one-way model with the two main categories comparing TC vs AC. We selected high confidence genes with a *P*-value of <0.01 (ANOVA one-way, Bonferroni Step-Up multiple-test correction) and fold change ≥ +2 or ≤ –2. We clustered the significant genes (hierarchical clustering, euclidean distances, average linkage

method) and plotted as a heatmap. Microarrays were quality checked for potential technical artefacts and three outliers were removed. More in detail, in order to improve classifiers prediction accuracy we used principal component analysis (PCA) as implemented in Partek GS. We detected 3 outliers (1 TC and 2 ACs) and removed them from further analysis (leaving a total of 13 samples, 6 TCs and 7 ACs).

The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (Edgar *et al*, 2002) and are accessible through GEO Series accession number GSE35679 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE35679>).

Diagnostic markers selection. Among 273 upregulated genes in ACs vs TCs, the top 10 scoring the highest fold changes (ranging from 38.766 to 534.063, ACs vs TCs) were selected and the absolute expression values for each sample were considered. A comparison between atypical vs typical tumour values allowed the selection of two markers (*GC* and *CEACAM1*) clearly discriminating the two subtypes (Supplementary Table 1).

Quantitative PCR and immunohistochemistry. qRT-PCR was carried out on the ABI/Prism 7900 HT Sequence Detector System (Applied Biosystems, Foster City, CA, USA), using a pre-PCR step of 10 min at 94.5 °C, followed by 40 cycles of 30 s at 97 °C and 60 s at 59.7 °C. TaqMan Arrays' 384 wells are pre-loaded with TaqMan Gene Expression Assays. Each TaqMan Array evaluates two cDNA samples in triplicate generated in a reverse transcription step using random primers. Selected Taqman assays for *GC* and *CEACAM1* were GC-Hs00167096_m1 and CEACAM1-Hs00266109_m1 by Applied Biosystems. Values relative to the *18s* housekeeping gene were used to normalise the data. A mix composed of RNA samples, obtained from the five typical tumours studied in the expression profiling experiment, was used as a calibrator. For each reaction, 100 ng of cDNA was added to 100 µl of reaction mix containing 50 µl of TaqMan PCR Mastermix 2x No UNG (Applied Biosystems) and loaded into each sample-loading port of 384-well TaqMan Low Density Arrays.

Immunohistochemistry (IHC) was performed according to the standard operating procedures applied at the Division of Pathology. Anti-*GC* and anti-*CEACAM1* antibodies were used according to the following specifications: anti-*GC* dilution 1:200 (cod. HPA001526; Sigma-Aldrich Corp., St Louis, MO, USA); anti-*CEACAM1* dilution 1:100 (cod. ab49510-100, clone 29H2; Abcam, Cambridge, UK).

Tumour samples. Frozen carcinoid samples were specifically collected at the European Institute of Oncology for this project starting from 2007. FFPE carcinoid samples were retrieved from the institute archive. As for the 16 tumour samples constituting the test set, all the 29 samples selected for the validation set were chemo-radio naïve.

RESULTS

The Affymetrix microarray technology was chosen to generate expression profiles from a total of 13 frozen samples (training set), 6 from typical and 7 from atypical surgically resected primary lung carcinoids (Figure 1; Supplementary Table 2). As represented in Supplementary Table 1, the derived expression profile revealed 584 differentially expressed genes, separating AC from TC. We initially focused on the group of 273 upregulated transcripts, specifically concentrating on those absolute expression values clearly differentiating ACs from TCs. This led to the identification of *GC* (vitamin D-binding protein) and *CEACAM1* (carcinoembryonic antigen family member), which scored among the highest fold changes (Material and methods; Supplementary Table 1). More precisely, we selected *GC* because showing the highest differences

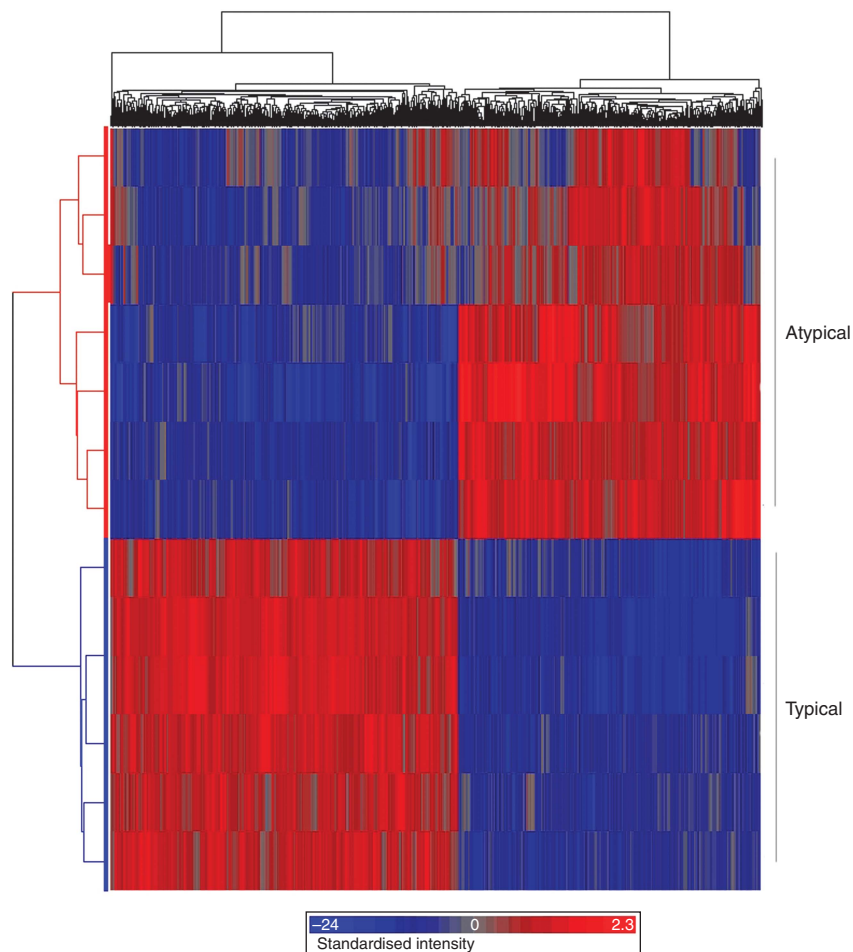


Figure 1. Hierarchical clustering of 584 differentially expressed genes in 13 lung carcinoid samples. Atypical samples are represented in the upper part of the dendrogram and typical samples in the lower. Expression levels are expressed according to the standardised intensity indicated in the scale at the bottom of the figure (higher levels in red and lower in blue).

in the expression levels, and *CEACAM1* that has already been described overexpressed in several tumour types (Beauchemin and Arabzadeh, 2013). The significant differences among the two tumour types prompted us to consider these two genes as interesting candidates for the definition of a new diagnostic strategy.

To confirm the different GC and *CEACAM1* levels in ACs vs TCs, we checked for their expression by (i) quantitative PCR (qRT-PCR) on an independent cohort of 12 patients (frozen samples from 6 TCs and 6 ACs) and (ii) IHC on the same 12 plus additional 17 tumours (FFPE samples from 17 TCs and 12 ACs), defining a validation set of 29 samples in total. The features of the patients and the tumours analysed are summarised in Supplementary Table 3.

In parallel, we completed our analysis with an independent review of the 29 cases constituting the validation set, performed by two pathologists dedicated to thoracic malignancies. As a result, and consistently with acknowledged diagnostic difficulties, the classification of the analysed samples was revised (Table 1): four cases of TCs (patients 2, 3, 4, and 21) were considered as ACs whereas one previous AC (patient 7) was classified as TC. The reasons for the change in the diagnosis consisted mainly in the strict application of the parameters proposed by the WHO classification. New 2- to 4- μ m-thick routine histological sections and consecutive Ki-67 stained sections consented to detect mitoses, discriminating them from nuclear picnosis or from regressive nuclear alterations related to apoptosis.

The results of qRT-PCR analysis recapitulated the new pathological assessment (Figure 2A vs B), identifying the 4 misdiagnosed

cases among the 12 analysed and detecting higher levels of both GC and *CEACAM1* in atypical tumours (Figure 2B). Similarly, IHC evidenced that all tumours showing *CEACAM1* expression were ACs (Figure 3). For tumours not showing *CEACAM1* IHC staining, instead, GC expression became relevant: <20% GC positivity was detected only in TCs. Samples with GC positivity up to 40%, but with low expression levels (1+) were also detected only in TCs. Conversely, samples showing GC positivity in >20% of cells with medium to high expression levels (2+; 3+) were only ACs (Table 2). Strikingly, the analysis of immunohistochemical levels confirmed all the five misdiagnoses evidenced in the entire validation set, although the GC/*CEACAM1* staining vs the pathologic evaluation remained discordant in four cases (18, 20, 22, and 24; Table 1). Of note, also the 13 initial cases (Affymetrix microarray analysis) underwent the pathologist's re-evaluation, confirming the diagnoses of 6 TCs and 7 ACs.

On the basis of our data, we defined a new immunohistochemical diagnostic algorithm, hypothesising that the availability of these ancillary markers could have avoided the diagnostic misinterpretation.

DISCUSSION

A differential diagnosis between TC and AC is relevant for clinicians to correctly discuss prognosis and follow-up with patients after surgery, as well as for surgical decision-making in

Table 1. Diagnoses of patients included into the validation set, according to routine evaluation, blind revision, and IHC markers

Patient	Routine diagnosis	Blind revision	Markers	CEACAM % cells (intensity)	GC % cells (intensity)
1	Typical	Typical	Typical	0	10 (weak 1 +)
2	Typical	Atypical	Atypical	80 (intense 3 +)	75 (intense 3 +)
3	Typical	Atypical	Atypical	75 (moderate 2 +)	20 (intense 3 +)
4	Typical	Atypical	Atypical	75 (moderate 2 +)	75 (intense 3 +)
5	Typical	Typical	Typical	0	0
6	Typical	Typical	Typical	0	20 (weak 1 +)
7	Atypical	Typical	Typical	0	10 (weak 1 +)
8	Atypical	Atypical	Atypical	0	80 (moderate 2 +)
9	Atypical	Atypical	Atypical	30 (intense 3 +)	75 (intense 3 +)
10	Atypical	Atypical	Atypical	20 (weak 1 +)	75 (moderate 2 +)
11	Atypical	Atypical	Atypical	75 (intense 3 +)	30 (moderate 2 +)
12	Atypical	Atypical	Atypical	75 (moderate 2 +)	60 (intense 3 +)
13	Atypical	Atypical	Atypical	0	40 (moderate 2 +)
14	Typical	Typical	Typical	0	10 (weak 1 +)
15	Typical	Typical	Typical	0	20 (moderate 2 +)
16	Typical	Typical	Typical	0	40 (weak 1 +)
17	Typical	Typical	Typical	0	0
18	Typical	Typical	ATYPICAL	10 (weak 1 +)	40 (intense 3 +)
19	Typical	Typical	Typical	0	0
20	Typical	Typical	ATYPICAL	10 (weak 1 +)	40 (intense 3 +)
21	Typical	Atypical	Atypical	20 (weak 1 +)	40 (intense 3 +)
22	Typical	Typical	ATYPICAL	50 (moderate 2 +)	40 (intense 3 +)
23	Typical	Typical	Typical	0	10 (weak 1 +)
24	Atypical	Atypical	TYPICAL	0	20 (moderate 2 +)
25	Atypical	Atypical	Atypical	0	25 (moderate 2 +)
26	Atypical	Atypical	Atypical	10 (weak 1 +)	10 (intense 3 +)
27	Atypical	Atypical	Atypical	0	80 (weak 1 +)
28	Atypical	Atypical	Atypical	0	25 (moderate 2 +)
29	Typical	Typical	Typical	0	5 (weak 1 +)

Abbreviation: IHC = immunohistochemistry. Misdiagnosed samples are evidenced in grey. Discordances between pathological and immunohistochemical evaluations are indicated in capital letters. Patients' number corresponding to the 12 cases analysed also by qRT-PCR are in bold.

borderline cases and as a main inclusion criteria in experimental studies of systemic treatment.

Differently from the classification of gastrointestinal carcinoids, which has incorporated the Ki-67 index, pulmonary carcinoids are subdivided into typical and atypical only on the base of the mitotic count and/or the presence of necrosis. Recently, the use of the Ki-67 index has been explored also for lung carcinoids but, unfortunately, even if the two subtypes show different mean Ki-67 values, the distribution of Ki-67 indices in the two groups frequently overlaps, making this marker only partially useful for differential diagnosis (Walts *et al*, 2012). Therefore, up to now, no immunohistochemical tools are available to support pathologists in distinguishing between TC and AC except those parameters defined by the WHO classification, which are sometimes difficult to interpret, especially in cases of poor quality or small quantities of material.

We performed an Affymetrix microarray gene expression profiling to identify a gene signature specific for atypical lung carcinoids. Among the most overexpressed genes in AC *vs* TC, we selected GC (vitamin D-binding protein or Graves Diseases Susceptibility To-gene) and CEACAM1 (carcinoembryonic antigen family member) that we validated by means of qRT-PCR and IHC

as new potential diagnostic markers. GC, the vitamin D-binding protein (VDBT) has a relevant immunomodulatory function in the lung and protein variations have been associated with airways diseases, such as chronic obstructive pulmonary disease (COPD), asthma, bronchiectasias, and lung cancer (Chishimba *et al*, 2010). CEACAM1 encodes a member of the carcinoembryonic antigen (CEA) gene family, which belongs to the immunoglobulin superfamily. Being involved in intercellular and intracellular signalling, CEACAM1 has a pleiotropic role in normal homeostasis and tumour progression (Ergun *et al*, 2000; Beauchemin and Arabzadeh, 2013). In particular, the expression of CEACAM1 in the tumour microenvironment significantly contributes to angiogenesis and tumour metastasis. A further contribution to tumour progression derives from CEACAM1 involvement in immune response: its expression on melanoma cells prevents attacks from T and NK cells (Ortenberg *et al*, 2012). For these reasons, it is considered as a promising therapeutic target, because inhibitory antibodies targeting CEACAM1 could render tumour cells more susceptible to the immune system attack. Moreover, because overexpressed in several tumours, CEACAM1 could represent a novel therapeutic target as well as a biomarker to monitor disease progression. This concept could also apply to ACs, since our data

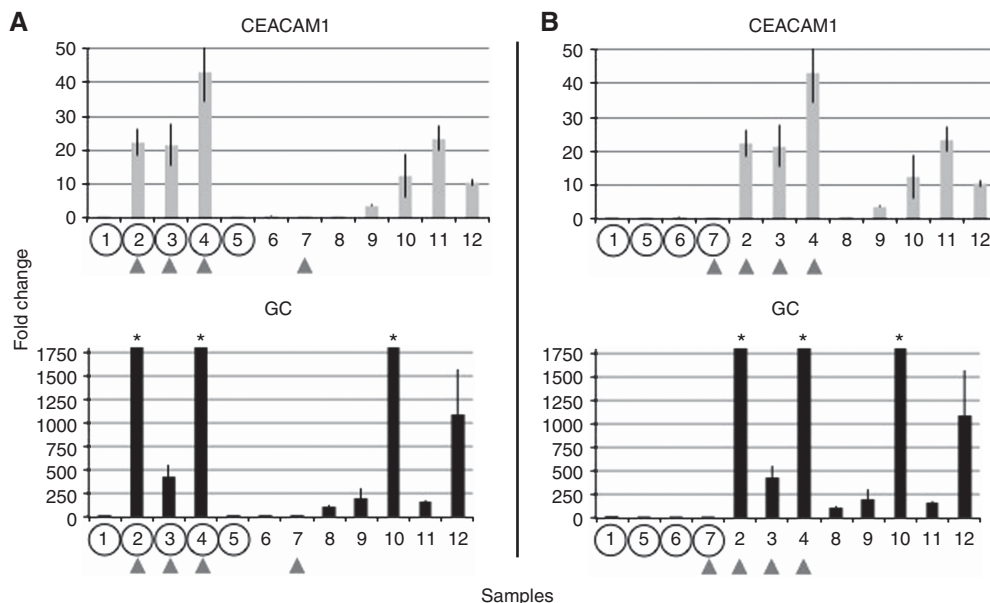


Figure 2. (A and B) Results of RQ-PCR for CEACAM1 and GC on 12 lung carcinoid samples. Results are shown on samples with the original diagnosis (A) and following the final evaluation (B). Grey bars represent results for CEACAM1 and black bars for GC. Samples have been numbered according to Table 1. Circled numbers represent typical carcinoids and plain numbers stand for atypical. Samples have been numbered according to Table 1. Misdiagnosed cases are marked with grey arrowheads. Asterisks indicate samples with out-of-scale fold changes: values were 18 266.71 for sample #2, 4042.27 for sample #4, and 4510.04 for sample #10. Please note: the lack of uniformity in the scales of the two charts reporting CEACAM1 and GC expression is due to the strong difference in the levels detected for the two genes.

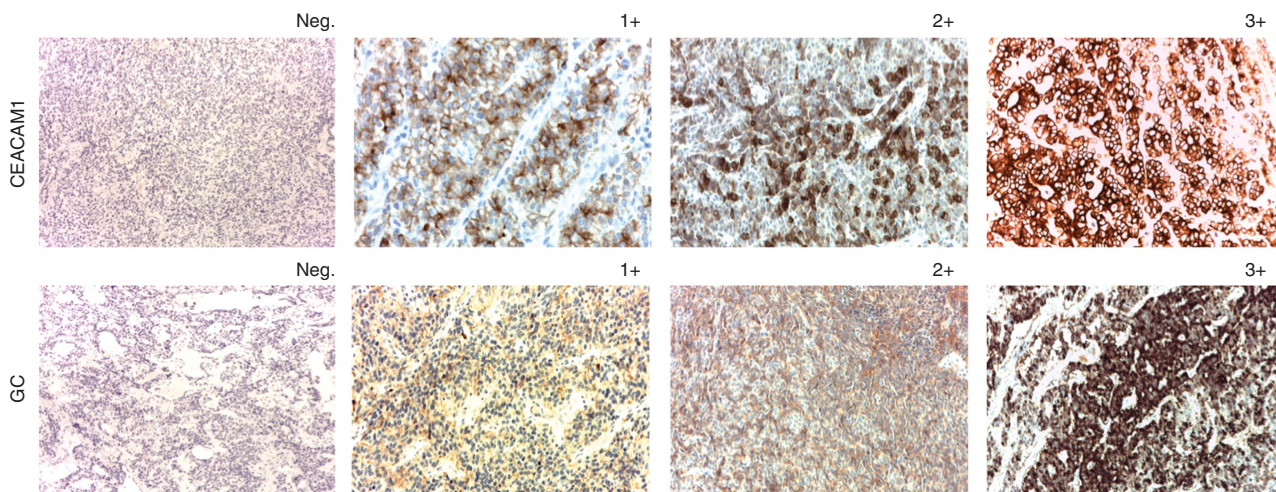


Figure 3. Representative examples of CEACAM and GC immunostaining. Immunoreactivity was rendered semiquantitatively on a scale from negative to 3+. Tumours were considered as negative if staining was completely absent; 1+ cases showed weak reactivity, 2+ moderate, and 3+ cases showed strong reactivity. The percentage of positive cells was taken into account and registered in each case.

indicate high CEACAM1 expression in this subtype. In terms of correlation with the clinical outcome, a role for GC has not definitely been addressed, while an association between CEACAM1 expression and negative prognosis has already been proposed. Indeed, lung adenocarcinomas expressing CEACAM1 showed a worst overall and disease-free survival compared with negative tumours (Laack *et al*, 2002).

When GC and CEACAM1 are evaluated together, they proved to be a potent tool able to discriminate ACs from TCs, resulting in the definition of an IHC assay possibly suitable for routine diagnostic procedures. On the basis of our data, we derived an algorithm, represented in Table 2, with a sensitivity ranging from

80% for TC to 90% for AC, as determined in our sample of 29 carcinoid tumours and assuming the double blind revision as a reference (Supplementary Table 4).

The development of this new tool is of great importance, considering that the distinction between the two tumour subtypes on biopsies/cytological specimens is still unfeasible and sometimes limited also after radical surgeries. Our algorithm, instead, could potentially be used also in cytological samples, as illustrated in Supplementary Figure 1.

Given the rarity of this kind of lung tumours, we defined the algorithm on a relatively low number of samples. However, the high sensitivity and reproducibility of this new diagnostic

Table 2. Definition of a new diagnostic algorithm

Marker	Typical carcinoid	Atypical carcinoid
CEACAM > 0	—	X
CEACAM = 0	X	X
GC ≤ 20	X	—
20 < GC < 40 (1 +)	X	—
GC > 20 (2 +, 3 +)	—	X

Abbreviations: AC = atypical carcinoid; TC = atypical carcinoid. Tumours showing CEACAM1 expression are classified as AC. Samples without detectable immunostaining are further analysed with GC. When GC levels are scored below 20% or included between 20% and 40% with a low intensity (1 +), the tumour is classified as TC. Finally, tumours with GC levels above 20% and showing medium or high intensity (2 +, 3 +) are classified as AC.

algorithm strongly support a further validation effort on a wider sample size, to investigate whether this potential tool could become a part of the routine diagnostic procedure.

ACKNOWLEDGEMENTS

This study was supported by 'Fondazione Umberto Veronesi' and 'Istituto Superiore di Sanità, Progetto Malattie Rare'. We would like to thank Valentina Dall'Olio, Loris Bernard, Elvira Gerbino, Simone Minardi, and the Microarray Unit for technical assistance; Piergiorgio Solli, Giulia Veronesi, Francesco Petrella, Roberto Gasparri, Domenico Galetta, and Alessandro Borri (IEO Thoracic Surgery) for clinical assistance and help in samples collection, Giuseppe Pelosi for helping in the pathologic evaluation; Fabrizio Bianchi for helpful discussion.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Arrigoni MG, Woolner LB, Bernatz PE (1972) Atypical carcinoid tumors of the lung. *J Thoracic Cardiovasc Surg* **64**: 413–421.
- Beane J, Sebastiani P, Whitfield TH, Steiling K, Dumas YM, Lenburg ME, Spira A (2008) A prediction model for lung cancer diagnosis that integrates genomic and clinical features. *Cancer Prev Res* **1**(1): 56–64.
- Beauchemin N, Arabzadeh A (2013) Carcinoembryonic antigen-related cell adhesion molecules (CEACAMs) in cancer progression and metastasis. *Cancer Metastasis Rev* **32**(3–4): 643–671.
- Bianchi F, Nuciforo P, Vecchi M, Bernard L, Tizzoni L, Marchetti A, Buttita F, Felicioni L, Nicassio F, Di Fiore PP (2007) Survival prediction of stage I

NSCLC lung adenocarcinomas by expression of 10 genes. *J Clin Invest* **117**(11): 3436–3444.

Chishimba L, Thickett DR, Stockley RA, Wood AM (2010) The vitamin D axis in the lung: a key role for vitamin D-binding protein. *Thorax* **65**: 456–462.

Edgar R, Domrachev M, Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* **30**(1): 207–210.

Ergun S, Kilic N, Ziegeler G, Hansen A, Nollau P, Gotze J (2000) CEA-related cell adhesion molecule 1: a potent angiogenic factor and a major effector of vascular endothelial growth factor. *Mol Cell* **5**: 311–320.

Gemma A, Li C, Sugiyama Y, Matsuada K, Seike Y, Kosaihiira S, Minegishi Y, Noro R, Nara M, Seike M, Yoshimura A, Shionoya A, Kawakami A, Ogawa N, Uesaka H, Kudoh S (2006) Anticancer drug clustering in lung cancer based on gene expression profiles and sensitivity database. *BMC Cancer* **6**: 174.

Gyorffy B, Surowiak P, Kiesslich O, Denkert C, Schafer R, Dietel M, Lage H (2006) Gene expression profiling of 30 cancer cell lines predicts resistance towards 11 anticancer drugs at clinically achieved concentrations. *Int J Cancer* **118**(7): 1699–1712.

Irizarry RA, Wu Z, Jaffee HA (2006) Comparison of Affymetrix GeneChip expression measures. *Bioinformatic* **22**(7): 789–794.

Laack E, Nikbakht H, Peters A, Kugler C, Jasiewicz Y (2002) Expression of CEACAM1 in adenocarcinoma of the lung: a factor of independent prognostic significance. *J Clin Oncol* **20**: 4279–4284.

Ortenberg R, Sapir Y, Raz L, Hershkovitz L, Ben Arav A, Sapoznik S, Barshack I, Avivi C, Berkun Y, Besser MJ, Ben-Moshe T, Schachter J, Markel G (2012) Novel immunotherapy for malignant melanoma with a monoclonal antibody that blocks CEACAM1 homophilic interactions. *Mol Cancer Ther* **11**(6): 1300–1310.

Scott WJ (2003) Surgical treatment of other bronchial tumors. *Chest Surg Clin N Am* **13**: 111–128.

Skov BG, Krasnik M, Lantuejoul S, Skov T, Brambilla E (2008) Reclassification of neuroendocrine tumors improves the separation of carcinoids and the prediction of survival. *J Thorac Oncol* **12**: 1410–1415.

Spira A, Beane J, Shah V, Steiling K, Liu G, Schembri F, Gilman S, Dumas YM, Calner P, Sebastiani P, Sridhar S, Beamis J, Lamb C, Anderson T, Gerry N, Keane J, Lenburg ME, Brody JS (2007) Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer. *Nat Med* **13**(3): 361–366.

SEER Research Data (1973–2004) The US National Cancer Institute. Surveillance Epidemiology and End Results (SEER) database. <http://seer.cancer.gov>.

Travis WD, Rush W, Flieder DB, Falk R, Fleming MV, Gal AA, Koss MN (1998) Survival analysis of 200 pulmonary neuroendocrine tumors with clarification of criteria for atypical carcinoid and its separation from typical carcinoid. *Am J Surg Pathol* **22**: 934–944.

Walts A, Ines D, Marchevsky AM (2012) Limited role of Ki-67 proliferative index in predicting overall short-term survival in patients with typical and atypical pulmonary carcinoid tumor. *Mol Pathol* **25**(9): 1258–1264.

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 3.0 Unported License.

Supplementary Information accompanies this paper on British Journal of Cancer website (<http://www.nature.com/bjc>)