

Comparison of the expression levels of molecular markers among the peripheral area and central area of primary tumor and metastatic lymph node tumor in patients with squamous cell carcinoma of the lung

メタデータ	言語: English 出版者: 公開日: 2016-03-20 キーワード (Ja): キーワード (En): 作成者: 宇田川, 響 メールアドレス: 所属:
URL	https://jair.repo.nii.ac.jp/records/2001915

Title: Comparison of the expression levels of molecular markers among the peripheral area and central area of primary tumor and metastatic lymph node tumor in patients with squamous cell carcinoma of the lung

Authors list:

Hibiki Udagawa, MD,^{1,2,5} Genichiro Ishii, MD, PhD,¹ Masahiro Morise, MD, PhD,²

Shigeki Umemura, MD, PhD,² Shingo Matsumoto, MD, PhD,² Kiyotaka Yoh, MD,²

Seiji Niho, MD, PhD,² Hironobu Ohmatsu, MD,² Masahiro Tsuboi, MD, PhD,⁴ Koichi

Goto, MD, PhD,² and Atsushi Ochiai, MD, PhD¹ Yuichiro Ohe, MD, PhD,^{3,5}

¹Department of Pathology, Research Center for Innovative oncology, National Cancer Center Hospital East, Kashiwa, Chiba, Japan

²Department of Thoracic Oncology, National Cancer Center Hospital East, Kashiwa, Chiba, Japan

³Department of Thoracic Oncology, National Cancer Center Hospital, Tsukiji, Tokyo, Japan

⁴Department of Thoracic Surgery, National Cancer Center Hospital East, Kashiwa, Chiba, Japan

⁵Advanced Clinical Research of Cancer, Juntendo University Graduate School of
Medicine, Tokyo, Japan

Address for Correspondence: Genichiro Ishii, MD, PhD, Pathology Division,
Research Center for Innovative Oncology, National Cancer Center Hospital East,
Kashiwa, Chiba, 277-8577, Japan, E-mail address: gishii@east.ncc.go.jp
Fax: +81-4-7131-4724, Tel: +81-4-7133-1111

Abstract:

Purpose: Immunohistochemical analysis for the identification of clinically relevant biomarkers is important. However, there have been no detailed reports about the heterogeneous expressions of the various markers in squamous cell carcinoma of the lung.

Methods: 113 patients with squamous cell carcinoma of the lung with lymph node metastasis were included. The expression levels of 9 molecules (E-cadherin, S100A4, CD44, ALDH1, SOX2, EGFR, HER2, FGFR1, VEGFR2) in the peripheral area, central area of primary tumor and metastatic lymph nodes were evaluated by immunohistochemistry. The differences in the staining scores of these molecules among the three areas were assessed. We also analyzed the relationships between the expression levels of these molecules and the recurrence-free survival.

Results: The E-cadherin expression was higher in the central area than in the peripheral area and metastatic lymph nodes (median staining score: 60 vs. 50, 30), the CD44 expression was higher in the central area than in the metastatic lymph nodes (117 vs. 90), and the EGFR expression was higher in the central area than in the peripheral area

and metastatic lymph nodes (163 vs. 130, 110). Low CD44 expression in the central area, low EGFR expression in the peripheral area and high SOX2 expression in the metastatic lymph nodes were associated with a shorter recurrence-free survival ($p < 0.01$, $p = 0.02$, $p = 0.03$, respectively).

Conclusions: Our findings confirmed that some molecular markers exhibited different expression levels in anatomically different areas, and suggested that immunohistochemical analysis for biomarkers area-by-area may provide useful information for more precise prediction of the recurrence.

Key Words: Squamous cell carcinoma, Lung cancer, Immunohistochemistry, Heterogeneity, Postoperative prognosis

Text:

Introduction

Lung cancer is among the most common of all cancers in the world today.(Siegel et al. 2013) Non-small cell lung cancer (NSCLC) accounts for approximately 85% of newly diagnosed lung cancers. Squamous cell carcinoma (SqCC) of the lung is the second most common histological type of NCSLC and accounts for approximately 30% of all cases of NSCLC.

Heterogeneity in cellular morphology, gene expression and protein expression among different areas of the tumor are encountered in a majority of tumors.(Marusyk and Polyak 2010) In the peripheral area of primary lung SqCC, the tumor cell infiltrate fills and destroys the alveolar spaces, associated with a weak desmoplastic reaction.(Funai et al. 2003; Watanabe et al. 2011) On the other hand, in the central area, the tumor cells form irregular-shaped nests associated with marked desmoplastic reaction. Although SqCC of the lung is well-known to exhibit morphological heterogeneity, as described above, few studies have focused on intratumoral heterogeneity of gene and protein expressions.

Some studies, including our own, indicated the existence of discrepancies in the expressions of various tumor markers between primary tumors and metastatic lymph

nodes, and the expression levels of the relevant molecules in the metastatic lymph nodes came to be recognized as more useful predictors of the tumor prognosis than those in the primary tumors.(Matsuwaki et al. 2014; Neri et al. 2012; Wen et al. 2013) Therefore, we considered that it would be important to consider the heterogeneous expressions of tumor markers in various areas of tumors, for example, in the peripheral and central areas of primary tumors and in primary tumors vs. metastatic lymph nodes.

The biological features based on the presence of cancer stem cells (CSC), epithelial-mesenchymal transition (EMT) and receptor tyrosine kinases (RTK) determine the malignancy grade of tumor cells.(Deeb et al. 2004; Sterlacci et al. 2014) The CSC theory was proposed to explain the heterogeneity in the biology of cancer cells.(Marusyk and Polyak 2010) CSCs are supposed to have metastatic potential and natural resistance to the chemotherapy and radiotherapy.(Donovan et al. 2013; Sterlacci et al. 2014) EMT is a process wherein epithelial cells gradually transform into mesenchymal-like cells and could also be one of the causes of morphological tumor heterogeneity.(Bartis et al. 2013) During the process of EMT, cancer cells acquire the capacity to migrate and invade the surrounding stroma. RTKs are transmembrane proteins that serve as receptors for many growth factors, and activation of the RTK signaling pathways lead to cell proliferation and anti-apoptotic activity. Heterogeneous

expression of RTKs in cancer cells has also been reported.(Andersson et al. 2004)

Taken together, determination of the expression heterogeneity of these biomarkers would be required for the identification of clinically relevant biomarkers. Thus, we performed this study to investigate whether discrepancies of biomarker expressions might exist within the same individual tumors and to determine the relationships between intratumoral expression heterogeneity of these markers and the postoperative prognosis in patients with SqCC of the lung.

Materials and Methods

Patients

The study cohort consisted of patients with SqCC of the lung with pathological T1a-2a (tumor size ≤ 5 cm) and lymph node metastasis (N1-2) who underwent complete resection between August 1992 and December 2010 in National Cancer Center Hospital East. Patients who received chemotherapy or/and radiotherapy for SqCC of the lung before the surgery, or had the history of other malignant diseases within 5 years were excluded. The data of the remaining 113 patients were analyzed in this study. The histologic classification was based on the World Health Organization

Classification (2004). The 7th edition of the TNM classification for lung cancer (International Association for the Study of Lung Cancer) was used for the pathological staging. Clinical characteristics were retrieved from the medical records.

This study was conducted with the approval of the Institutional Review Boards of the National Cancer Center, approval number 2013-331.

Pathological Studies

The surgical specimens were fixed with 10% formalin or absolute methyl alcohol and embedded in paraffin. We defined the peripheral areas of the primary tumor (PA) as the areas where were adjacent to normal lung parenchyma, and the central areas of the primary tumor (CA) as the middle of the tumor with marked desmoplastic reaction.

Three representative PA (Figure 1a) and three representative CA (Figure 1b) were randomly selected to construct the tissue microarrays (TMA), and a total of six tissue cores (2 mm in diameter each) were obtained from each tumor. To prevent the influence or bias of necrosis or fibrosis in the tumor, the representative areas containing

sufficient tumor cells were carefully selected. For obtaining tissue specimens from the metastatic lymph nodes (LN) surgically removed (Figure 1c), the largest metastatic lymph nodes were chosen.

Antibodies and Immunohistochemistry

Immunohistochemical stainings of E-cadherin, S100A4, clusters of differentiation 44 (CD44), aldehyde dehydrogenase 1 (ALDH1), sex-determining region Y-box 2 (SOX2), epidermal growth factor receptor (EGFR), human epidermal growth factor receptor 2 (HER2), fibroblast growth factor receptor 1 (FGFR1) and vascular endothelial growth factor receptor (VEGFR2) were performed according to the procedure described in our previous report.(Kirita et al. 2013) The primary antibodies used in this study are summarized in Supplemental-Table 1. For EGFR and HER2 staining, the automated staining system Benchmark ULTRA (Roche/Ventana Medical Systems, Tucson, AZ) was used.

Evaluation of Immunohistochemistry

Two pathologists (H. U. and G. I.) with no knowledge of the patients' clinical or pathological data evaluated the individual sections by light microscopy. The labeling scores were calculated by multiplying the percentage of positive cancer cells per lesion (0-100%) by the staining intensity (0=negative, 1=weak, 2=strong), the only immunohistochemical staining scores of viable tumor cells were evaluated. The staining score for the whole primary tumor was calculated as the average of the scores for the six tissue cores from the PA and CA. The staining score for the PA was calculated as the average of the scores for the three tissue cores from the PA. Similarly, the staining score for the CA was calculated as the average of the scores for the three tissue cores from the CA. The staining score for the LN was calculated as the score obtained for the whole LN. The median scores for each marker were chosen as the cutoff values to determine high and low expression levels.

Statistical Analysis

The differences in the median scores among the areas were assessed by the Mann-Whitney U-test. Recurrence-free survival (RFS) was defined as the time from the

date of surgery until the date of tumor recurrence, death, or the last follow-up. The survival curves were estimated by the Kaplan-Meier method, and the differences in RFS were compared by the log-rank test. Multivariate analysis was performed using the Cox proportional hazards model. All the P-values were two sided, and P-values of <0.05 were considered as denoting statistical significance. All the statistical analyses were performed using the JMP statistical software package for Windows, Ver.11 (SAS Institute, Cary, NC).

Results

Clinicopathological characteristics

The clinicopathological characteristics of all the patients are listed in Table 1. The median RFS of all the patients was 3.1 years. The median follow-up for the censored patients was 8.6 years (range: 0.3-16.3). N2 was significantly correlated with a shorter RFS ($p=0.0394$), however, none of the other factors evaluated was associated with the RFS.

Median staining scores for the PA, CA and LN

Figure 2 and Supplemental-Figure 1 shows the immunochemical staining scores for each molecule in the PA, CA and LN.

For E-cadherin, the median score in the PA was 50, that in the CA was 60, and that in the LN was 30. The median score in the CA was significantly higher than that in the PA ($p=0.0012$) or LN ($p<0.0001$). Also, the median score in the PA was significantly higher than that in the LN ($p=0.0069$) (Figure 2a).

The median score for CD44 in the PA, CA and LN were 123, 117 and 90, respectively. The median score in the LN was significantly lower than that in the PA ($p=0.0016$) or CA ($p=0.0019$) (Figure 2b).

The median score for EGFR in the PA, CA and LN were 130, 163 and 110, respectively. The median score in the CA was significantly higher than that in the PA ($p=0.0372$) or LN ($p=0.0002$) (Figure 2c).

No statistically significant differences in the median scores for S100A4, ALDH1, SOX2, HER2, FGFR1 or VEGFR2 were observed among the PA, CA and LN (Supplemental-Figure 1).

Figure 3 shows the typical immunohistochemical stainings for E-cadherin, CD44, SOX2 and EGFR in each of the areas.

Survival analysis on the basis of the staining scores in the PA, CA and LN

1) PA

Only EGFR expression was associated with recurrence (Supplemental-Table 2A). The patients with high EGFR expression showed significantly longer RFS than the patients with low EGFR expression ($p=0.0187$) (Figure 4-A).

2) CA

Only CD44 expression in the CA was associated with recurrence (Supplemental-Table 2B). The patients with high CD44 expression showed significantly longer RFS than the patients with low CD44 expression ($p=0.0073$) (Figure 4-B).

3) LN

Only SOX2 expression was associated with recurrence (Supplemental-Table 2C). The patients with high SOX2 expression showed significantly shorter RFS than the patients with low SOX2 expression ($p=0.0303$) (Figure 4-C).

Survival analysis on the basis of the staining score for the whole primary tumor

We also analyzed the association between the expression in the whole primary tumor and RFS. However, RFS was not associated with the expression of any of the molecules (Supplemental-Table 3).

Multivariate analysis to identify the predictors of RFS

All factors that were identified as being significant by univariate analysis (N factor, EGFR expression in PA, CD44 expression in CA, SOX2 expression in LN) were

included as the covariates in the multivariate analysis (Table 2). High SOX2 expression in the LN was identified as a significant independent predictor of a poor RFS [Hazard Ratio: 1.736 (95% confidence interval; 1.055-2.901), $p=0.0298$].

Discussion

Heterogeneity of molecular expression within individual tumors and between the primary and metastatic tumors has been reported in some types of cancers.(Jacobsen et al. 2013; Kroepil et al. 2013; Melchers et al. 2013) In gastric cancer, differential epithelial cell adhesion molecule (EpCAM) expression between the center and the invasive front of the tumor has been reported, and high EpCAM expression in the invasive front, but not that in the center of the tumor, has been reported as a poor prognostic factor.(Kroepil et al. 2013) In oral and oropharyngeal squamous cell carcinoma, heterogeneous expressions of E-cadherin, EpCAM and claudin-7 have been reported, and lack of claudin-7 expression in the center of the primary tumor was identified as a poor prognostic factor.(Melchers et al. 2013) In adenocarcinoma of the lung, a high expression level of C4.4A, a glycoposphatidylinositol (GPI)-anchored protein and a structural homolog of the urokinase-type plasminogen activator (uPAR)(Oshiro et al. 2012), in the central part of the tumor was identified as a poor prognostic factor, whereas that in the tumor periphery was not found to be associated

with the prognosis.(Jacobsen et al. 2013) On the other hand, C4.4A expression, neither in the tumor periphery nor tumor center, was associated with the prognosis in patients with SqCC of the lung.(Jacobsen et al. 2013) Taken together, it is important to realize that the heterogeneity profile of intratumoral expression may depend on the organ site and/or the histological type and evaluation of the expression profile area-by-area would be recommended for predicting the prognosis. This is the first study to employ this approach for patients with SqCC of the lung with lymph node metastasis.

In our study, the CD44 expression level in the LN was lower than that in the PA and CA of the primary tumor. Few studies have described the discrepancy in CD44 expression between the primary and metastatic tumors. In prostate cancer, CD44 expression in LN was significantly reduced, and loss of CD44 expression in the primary tumors was associated with a higher tumor malignancy grade.(De Marzo et al. 1998) Similarly, in breast cancer also, loss of CD44 expression was reported to be associated with a higher tumor malignancy grade(Friedrichs et al. 1995) and lymph node metastasis.(Schneider et al. 1999) Based on these studies, we consider that cancer cells showing decreased CD44 expression in the LN in patients of lung SqCC may represent an aggressive phenotype.

The current study clearly showed that the expression level of E-cadherin was significantly higher in the CA than in the PA of the primary tumor and LN. In digestive tract cancers(Lee et al. 2013) and oral SqCC(Wang et al. 2009), decrease of E-cadherin expression was reportedly found in the invasive front of the tumor. In SqCC of the lung, it may be postulated that PA may correspond to the “invasive front” of the primary tumor, as the morphological findings may fit. Similar to the finding in our study, another study reported that the E-cadherin expression in a metastatic site was significantly lower than that in primary colorectal carcinoma.(Elzagheid et al. 2006)

Taken together, it is considered that the EMT phenotype is maintained in the cancer cells at metastatic sites.

Previously, discordant EGFR expression within same tumor was reported in NSCLC(Grob et al. 2013; Italiano et al. 2006), colorectal cancer(Scartozzi et al. 2004; Scartozzi et al. 2007), gastric carcinoma(Nagatsuma et al. 2014) and vulvar carcinoma(de Melo Maia et al. 2014). In our study, EGFR expression was significantly higher in the CA than in the PA of the primary tumor and the LN. Evaluation of EGFR expression is important to decide on anti-EGFR antibody therapy. A retrospective analysis of the FLEX study suggested that chemotherapy with cetuximab, anti-EGFR-antibody, improved the overall survival of patients with tumors showing

high EGFR expression, but not of those with tumors showing low EGFR expression.(Pirker et al. 2012) However, analysis of EGFR expression is often performed on small biopsy specimens that contain a limited number of tumor cells. The results of evaluation of EGFR expression may be potentially inaccurate in 25% of small biopsies as compared to that in whole tumors based on the larger resected specimens.(Jakobsen et al. 2014) Our study findings may probably explain the discrepancy in EGFR expression between small biopsies and larger resected specimens. Thus, while evaluating EGFR expression in small biopsy specimens of primary tumor or metastatic lymph nodes obtained by procedures such as transbronchial biopsy or transbronchial needle aspiration, it may be necessary to bear in mind the existence of intratumoral heterogeneity in expression. Thus, we think that the second surgical biopsy is a useful method to evaluate the biomarker expression, when the small biopsy specimen is insufficient material. However, the invasiveness of the surgical biopsy is worried. Also, it may be useful to assess the several small biopsy specimens, if we can obtain them by transbronchial biopsy. Furthermore, if we can obtain the small biopsy specimens from primary tumor and metastatic lymph node by transbronchial biopsy or transbronchial needle aspiration, it may be informative to evaluate both.

Chen Y et al. reported a meta-analysis about the relationship between SOX2 expression in the primary tumor and the prognosis in patients with SqCC of the lung.(Chen et al. 2013) In the current multivariate analysis, high SOX2 expression in the LN was identified as a significant independent predictor of a shorter RFS. Our group previously reported that expression of some specific proteins in metastatic lymph nodes, but not in primary tumors, may serve as useful information for predicting the risk of recurrence in pathological N2 lung SqCC.(Matsuwaki et al. 2014) The current results also support this concept. It is well known that the tumor microenvironment of metastatic lymph nodes differs from that in the primary tumors. This difference may explain why the immunophenotypical characteristics of tumor cells in the LN were found to be more strongly predictive of recurrence than those of the primary tumors.

In conclusion, our study indicated heterogeneous expression of several proteins in various regions of lung cancer. In the future, if immunohistochemical evaluation of biomarker expression is performed for arriving at a decision on molecular-targeted therapy and the analysis of prognostic factors, careful consideration must be given to the potential discrepancy in expression among different areas of the tumor and between primary and metastatic tumors. The results of our study may be expected to serve as

useful information to predict the postoperative prognosis and for personalized cancer therapy in patients with SqCC of the lung.

Disclosure/Conflict of Interest

The authors have no conflict of interests to disclose.

Acknowledgments

We thank Yuka Nakamura for her help with the immunohistochemical analyses in this study. This work was supported by the National Cancer Center Research and Development Fund (23-A-12), Foundation for the Promotion of Cancer Research, 3rd-Term Comprehensive 10-Year Strategy for Cancer Control, Program for the Promotion of Fundamental Studies in Health Sciences of the National Institute of Biomedical Innovation, and JSPS KAKENHI (24659185).

References:

- Andersson J, Linderholm B, Bergh J, Elmberger G (2004) HER-2/neu (c-erbB-2) evaluation in primary breast carcinoma by fluorescent in situ hybridization and immunohistochemistry with special focus on intratumor heterogeneity and comparison of invasive and in situ components *Applied immunohistochemistry & molecular morphology : AIMM / official publication of the Society for Applied Immunohistochemistry* 12:14-20
- Bartis D, Mise N, Mahida RY, Eickelberg O, Thickett DR (2013) Epithelial-mesenchymal transition in lung development and disease: does it exist and is it important? *Thorax* doi:10.1136/thoraxjnl-2013-204608
- Chen Y, Huang Y, Huang Y, Chen J, Wang S, Zhou J (2013) The prognostic value of SOX2 expression in non-small cell lung cancer: a meta-analysis *PloS one* 8:e71140 doi:10.1371/journal.pone.0071140
- De Marzo AM, Bradshaw C, Sauvageot J, Epstein JI, Miller GJ (1998) CD44 and CD44v6 downregulation in clinical prostatic carcinoma: relation to Gleason grade and cytoarchitecture *The Prostate* 34:162-168
- de Melo Maia B et al. (2014) EGFR expression in vulvar cancer: clinical implications and tumor heterogeneity *Human pathology* 45:917-925 doi:10.1016/j.humpath.2014.01.015
- Deeb G, Wang J, Ramnath N, Slocum HK, Wiseman S, Beck A, Tan D (2004) Altered E-cadherin and epidermal growth factor receptor expressions are associated with patient survival in lung cancer: a study utilizing high-density tissue microarray and immunohistochemistry *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* 17:430-439 doi:10.1038/modpathol.3800041
- Donovan CA et al. (2013) Correlation of breast cancer axillary lymph node metastases with stem cell mutations *JAMA surgery* 148:873-878 doi:10.1001/jamasurg.2013.3028
- Elzagheid A et al. (2006) E-cadherin expression pattern in primary colorectal carcinomas and their metastases reflects disease outcome *World journal of gastroenterology : WJG* 12:4304-4309
- Friedrichs K et al. (1995) CD44 isoforms correlate with cellular differentiation but not with prognosis in human breast cancer *Cancer research* 55:5424-5433
- Funai K et al. (2003) Clinicopathologic characteristics of peripheral squamous cell carcinoma of the lung *The American journal of surgical pathology* 27:978-984

- Grob TJ et al. (2013) Frequent intratumoral heterogeneity of EGFR gene copy gain in non-small cell lung cancer *Lung cancer* 79:221-227
doi:10.1016/j.lungcan.2012.11.009
- Italiano A et al. (2006) Comparison of the epidermal growth factor receptor gene and protein in primary non-small-cell-lung cancer and metastatic sites: implications for treatment with EGFR-inhibitors *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO* 17:981-985
doi:10.1093/annonc/mdl038
- Jacobsen B et al. (2013) Ly6/uPAR-related protein C4.4A as a marker of solid growth pattern and poor prognosis in lung adenocarcinoma *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer* 8:152-160 doi:10.1097/JTO.0b013e318279d503
- Jakobsen JN, Santoni-Rugiu E, Sorensen JB (2014) Changes in epidermal growth factor receptor expression during chemotherapy in non-small cell lung cancer *Cancer chemotherapy and pharmacology* 73:131-137 doi:10.1007/s00280-013-2329-0
- Kirita K et al. (2013) Identification of biological properties of intralymphatic tumor related to the development of lymph node metastasis in lung adenocarcinoma *PloS one* 8:e83537 doi:10.1371/journal.pone.0083537
- Kroepil F et al. (2013) High EpCAM expression is linked to proliferation and lauren classification in gastric cancer *BMC research notes* 6:253
doi:10.1186/1756-0500-6-253
- Lee SJ et al. (2013) Combined aberrant expression of E-cadherin and S100A4, but not beta-catenin is associated with disease-free survival and overall survival in colorectal cancer patients *Diagnostic pathology* 8:99
doi:10.1186/1746-1596-8-99
- Marusyk A, Polyak K (2010) Tumor heterogeneity: causes and consequences *Biochimica et biophysica acta* 1805:105-117 doi:10.1016/j.bbcan.2009.11.002
- Matsuwaki R et al. (2014) Immunophenotypic features of metastatic lymph node tumors to predict recurrence in N2 lung squamous cell carcinoma *Cancer science*
doi:10.1111/cas.12434
- Melchers LJ et al. (2013) Lack of claudin-7 is a strong predictor of regional recurrence in oral and oropharyngeal squamous cell carcinoma *Oral oncology* 49:998-1005
doi:10.1016/j.oraloncology.2013.07.008
- Nagatsuma AK, Aizawa M, Kuwata T, Doi T, Ohtsu A, Fujii H, Ochiai A (2014) Expression profiles of HER2, EGFR, MET and FGFR2 in a large cohort of patients with gastric adenocarcinoma *Gastric cancer : official journal of the*

- International Gastric Cancer Association and the Japanese Gastric Cancer Association doi:10.1007/s10120-014-0360-4
- Neri S et al. (2012) Recruitment of podoplanin positive cancer-associated fibroblasts in metastatic lymph nodes predicts poor prognosis in pathological N2 stage III lung adenocarcinoma *Annals of surgical oncology* 19:3953-3962
doi:10.1245/s10434-012-2421-4
- Oshiro R et al. (2012) C4.4A is associated with tumor budding and epithelial-mesenchymal transition of colorectal cancer *Cancer science* 103:1155-1164 doi:10.1111/j.1349-7006.2012.02263.x
- Pirker R et al. (2012) EGFR expression as a predictor of survival for first-line chemotherapy plus cetuximab in patients with advanced non-small-cell lung cancer: analysis of data from the phase 3 FLEX study *The lancet oncology* 13:33-42 doi:10.1016/s1470-2045(11)70318-7
- Scartozzi M, Bearzi I, Berardi R, Mandolesi A, Fabris G, Cascinu S (2004) Epidermal growth factor receptor (EGFR) status in primary colorectal tumors does not correlate with EGFR expression in related metastatic sites: implications for treatment with EGFR-targeted monoclonal antibodies *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 22:4772-4778 doi:10.1200/JCO.2004.00.117
- Scartozzi M, Bearzi I, Berardi R, Mandolesi A, Pierantoni C, Cascinu S (2007) Epidermal growth factor receptor (EGFR) downstream signalling pathway in primary colorectal tumours and related metastatic sites: optimising EGFR-targeted treatment options *British journal of cancer* 97:92-97
doi:10.1038/sj.bjc.6603847
- Schneider J et al. (1999) Histologic grade and CD44 are independent predictors of axillary lymph node invasion in early (T1) breast cancer *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine* 20:319-330
- Siegel R, Naishadham D, Jemal A (2013) Cancer statistics, 2013 *CA: a cancer journal for clinicians* 63:11-30 doi:10.3322/caac.21166
- Sterlacci W, Savic S, Fiegl M, Obermann E, Tzankov A (2014) Putative stem cell markers in non-small-cell lung cancer: a clinicopathologic characterization *Journal of thoracic oncology : official publication of the International Association for the Study of Lung Cancer* 9:41-49
doi:10.1097/jto.0000000000000021

- Wang X, Zhang J, Fan M, Zhou Q, Deng H, Aisharif MJ, Chen X (2009) The expression of E-cadherin at the invasive tumor front of oral squamous cell carcinoma: immunohistochemical and RT-PCR analysis with clinicopathological correlation *Oral surgery, oral medicine, oral pathology, oral radiology, and endodontics* 107:547-554 doi:10.1016/j.tripleo.2008.11.021
- Watanabe Y et al. (2011) Alveolar space filling ratio as a favorable prognostic factor in small peripheral squamous cell carcinoma of the lung *Lung cancer* 73:217-221 doi:10.1016/j.lungcan.2010.12.001
- Wen J, Luo KJ, Hu Y, Yang H, Fu JH (2013) Metastatic lymph node CHIP expression is a potential prognostic marker for resected esophageal squamous cell carcinoma patients *Annals of surgical oncology* 20:1668-1675 doi:10.1245/s10434-012-2733-4

Titles and legends to figures:

Figure 1. Hematoxylin and eosin (H.E.) staining of a primary tumor and metastatic lymph node. a) Representative features of a peripheral area of the primary tumor (H.E.). b) Representative features of a central area of the primary tumor (H.E.). c) Representative features of a metastatic lymph node (H.E.).

Figure 2. Box plots of the immunochemical staining scores and the changes of those within the same case in the peripheral area (PA) and central area (CA) of the primary tumor and metastatic lymph node (LN) for a) E-cadherin, b) CD44, c) EGFR. Statistical analysis was performed by the Mann-Whitney U-test.

Figure 3. Immunohistochemical staining for E-cadherin, CD44, SOX2 and EGFR in a peripheral area and central area of the primary tumor and a metastatic lymph node. A-C, E-cadherin expression in each area of the same case. D-F, CD44 expression in each area of the same case. G-I, SOX2 expression in each area of the same case. J-K, EGFR expression in each area of the same case.

Figure 4. Correlations of the immunohistochemical staining scores with the recurrence-free survival (RFS) curves drawn by the Kaplan-Meier method. (A) RFS analysis for EGFR expression in the peripheral area of the primary tumor. (B) RFS analysis for CD44 expression in the central area of the primary tumor. (C) RFS analysis for SOX2 expression in a metastatic lymph node.

Supplemental-Figure 1. Box plots of the immunochemical staining scores in the peripheral area (PA) and central area (CA) of the primary tumor and metastatic lymph node (LN) for A) S100A4, B) ALDH1, C) SOX2, D) HER2, E) FGFR1 and F) VEGFR2.