A Comprehensive Analysis of Human Gene Expression Profiles Identifies Stromal Immunoglobulin k C as a Compatible Prognostic Marker in Human Solid Tumors

Marcus Schmidt1, Birte Hellwig6, Seddik Hammad7, Amnah Othman7, Miriam Lohr6, Zonglin Chen1, Daniel Boehm1, Susanne Gebhard1, Ilka Petry1, Antje Lebrecht1, Cristina Cadenas7, Rosemarie Marchan7, Joanna D. Stewart7, Christine Solbach1, Lars Holmberg8,9,12, Karolina Edlund10, Hanna Göransson Kutima11, Achim Rody13, Anders Berglund8,14, Mats Lambe7,8, Anders Isaksson11, Johan Bottling10, Thomas Kann15, Hans Bojar18, Hans-Anton Lehr19, Ugur Sahin5, Heinz Koelbl1, Mathias Gehrmann20, Patrick Micke10, Jan G. Hengstler7

Abstract

Purpose: Although the central role of the immune system for tumor prognosis is generally accepted, a single robust marker is not yet available.

Experimental Design: On the basis of receiver operating characteristic analyses, robust markers were identified from a 60-gene B cell–derived metagene and analyzed in gene expression profiles of 1,810 breast cancer; 1,056 non–small cell lung carcinoma (NSCLC); 513 colorectal; and 426 ovarian cancer patients. Protein and RNA levels were examined in paraffin-embedded tissue of 330 breast cancer patients. The cell types were identified with immunohistochemical costaining and confocal fluorescence microscopy.

Results: We identified immunoglobulin k C (IGKC) which as a single marker is similarly predictive and prognostic as the entire B-cell metagene. IGKC was consistently associated with metastasis-free survival across different molecular subtypes in node-negative breast cancer (n = 965) and predicted response to anthracycline-based neoadjuvant chemotherapy (n = 845; P < 0.001). In addition, IGKC gene expression was prognostic in NSCLC and colorectal cancer. No association was observed in ovarian cancer. IGKC protein expression was significantly associated with survival in paraffin-embedded tissues of 330 breast cancer patients. Tumor-infiltrating plasma cells were identified as the source of IGKC expression.

Conclusion: Our findings provide IGKC as a novel diagnostic marker for risk stratification in human cancer and support concepts to exploit the humoral immune response for anticancer therapy. It could be validated in several independent cohorts and carried out similarly well in RNA from fresh frozen as well as from paraffin tissue and on protein level by immunostaining. Clin Cancer Res; 18(9); 2695–703. ©2012 AACR.

Authors' Affiliations: Departments of 1 Obstetrics and Gynecology, 2 Medical Biometry, 3 Pathology, and 4 Hematology and Oncology, University Hospital; TRON-Translational Oncology at the Medical University Mainz, Mainz; Department of Statistics, Technical University, Dortmund; Leibniz Research Centre for Working Environment and Human Factors (IfADo) at Dortmund TU, Germany; Regional Oncology Center; Department of Surgical Sciences; Department of Genetics and Pathology; Department of Obstetrics and Gynecology, University Hospital; King's College London, United Kingdom; Department of Obstetrics and Gynecology, University Hospital Hamburg, Germany; Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden; Department of Obstetrics and Gynecology, University Hospital; Düsseldorf University, Düsseldorf, Germany; Department of Pathology, University of Lausanne, Lausanne, Switzerland; and Bayer GmbH, Leverkusen, Germany

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).


Parts of this study were presented at the San Antonio Breast Cancer Symposium 2009 and at the Annual Meeting of the American Society of Clinical Oncology 2010.

The study contains parts of the doctoral theses of A. Gerhold-Ay and Z. Chen.

Corresponding Authors: Marcus Schmidt, Department of Obstetrics and Gynecology, University Hospital Mainz, Langenbeckstr. 1, Mainz 55131, Germany. Phone: 49-6131-1172-683; Fax: 49-6131-1175-673; E-mail: marcus.schmidt@unimedizin-mainz.de; Jorg Rahnenfuhrer, Department of Statistics, Technical University, Dortmund 44221, Germany. Phone: 49-231-7553-121; Fax: 49-231-7555-303; E-mail: rahnenfuhrer@statistik.tu-dortmund.de; and Jan G. Hengstler, Leibniz Research Centre for Working Environment and Human Factors (IfADo) at Dortmund TU, Ardeystr. 67, Dortmund 44139, Germany. Phone: 49-231-1084-348/349; Fax: 49-231-1084-403; E-mail: hengstler@ifado.de

doi: 10.1158/1078-0432.CCR-11-2210

©2012 American Association for Cancer Research.
Translational Relevance

This study reports that immunoglobulin κ C (IGKC) RNA levels robustly define prognosis in a comprehensive analysis of available breast cancer data sets and predict response to neoadjuvant anthracycline-based therapy. In addition, IGKC maintains its prognostic relevance also in non–small cell lung and colorectal cancer, suggesting a global mechanism in the biology of adenocarcinomas. Using real-time PCR and immunohistochemistry in formalin-fixed, paraffin-embedded (FFPE) tissue, we could validate the prognostic impact of IGKC. Furthermore, using confocal microscopy, we identified tumor-infiltrating plasmablasts and plasma cells as the source of IGKC expression. This study has major clinical implications: (i) IGKC as a prognostic and predictive marker that lends itself to systematic testing in FFPE tissue samples allows an improved prediction of prognosis and response to chemotherapy, and (ii) the protective effects of this naturally occurring humoral immune response support the concept of immunotherapy.

Patients and Methods

Patients

Our analysis includes gene array data from 1,810 breast cancer patients (965 node negative, without chemotherapy and 845 with anthracycline-based chemotherapy), 1,056 non–small cell lung carcinoma (NSCLC), 513 colorectal, and 426 ovarian cancer patients. In addition, paraffin-embedded tissue blocks of 330 node-negative breast cancer patients were analyzed (details, Supplementary Methods).

Gene expression analysis and immunostaining

HG-U133A arrays were used to analyze Uppsala lung cancer (n = 196) cohorts (Supplementary Table S1). All other gene array data are publicly available (Supplementary Methods). IGKC mRNA levels in formalin-fixed, paraffin-embedded (FFPE) tissue were quantified by quantitative RT-PCR (qRT-PCR). For both immunohistochemistry and confocal-fluorescence microscopy, antibodies against MUM1/IRF4, CD20, pan-cytokeratin, or immunoglobulin G (IgG) and IGKC were used as previously described (details, Supplementary Methods).

Statistical analysis

Survival was analyzed by univariate and multivariate Cox models and visualized by Kaplan–Meier plots. The Brier score was used to evaluate the ability to predict survival. Meta-analyses were conducted by fixed and random effect models and visualized with forest plots (details, Supplementary Methods).

Results

IGKC is a representative marker of the B-cell gene signature

To condense the previously described breast cancer B-cell signature (15) that consists of 60 genes, we analyzed microarray data from our own breast cancer cohort (Mainz) and 2 independent cohorts [Rotterdam (19); Transbig (16, 17)]. The bioinformatic strategy was based on the optimal combination of 2 criteria (Fig. 1): (i) the best average correlation of each of the 60 genes with all other members of the B-cell metagene as a measure of representativeness, and (ii) the largest area under the receiver operating characteristic curve, as a measure of the ability of each individual gene to discriminate between patients with and without metastasis during a 5-year follow-up period. Using these 2 criteria, IGKC was identified as one of the genes with the best average correlation and largest area under the curve (AUC; Fig. 1) and also showed a wide dynamic range with a unimodal distribution (Supplementary Fig. S1). The results obtained by microarrays were confirmed with qRT-PCR in archived FFPE tissue from the Mainz cohort. IGKC mRNA levels
determined by qRT-PCR, correlated very well with the levels measured by gene array in fresh-frozen samples of the same tumors (Fig. 2A) and similarly, IGKC mRNA levels in paraffin-embedded tissue showed a significant association with MFI both in univariate and multivariate analyses (Table 3; Kaplan–Meier plot: Fig. 2B).

**IGKC is associated with better prognosis in breast cancer**

To further validate the prognostic impact of IGKC, we analyzed mRNA expression as a single marker in 5 publicly accessible gene array data sets of node-negative breast cancer patients who did not receive chemotherapy: the Mainz (15), Rotterdam (19), Transbig (16, 17), Yu (18), and NKI (20, 21) cohorts. The meta-analysis revealed a highly significant association of IGKC RNA levels with better prognosis ($P < 0.0001$, Fig. 3). The expression of IGKC was further analyzed in the 3 molecularly and biologically different subtypes of breast cancer (14): (i) estrogen receptor (ER) status positive and HER2 status negative, (ii) ER status negative, and (iii) HER2 status positive and ER status positive or negative carcinomas. High IGKC expression correlated with good prognosis in all subgroups with a particularly strong association in the HER2-positive subgroup (Fig. 3). The univariate (Table 1) and multivariate Cox regression models (Table 2) adjusted to established clinical factors (Supplementary Fig. S2) confirmed the association of IGKC with MFI (Table 1,
IGKC predicts response to anthracycline-based chemotherapy

In addition to the prediction of survival, IGKC expression levels were evaluated with regard to response to cytostatic drugs. We selected all published gene array data of breast cancer patients who had received anthracycline-based neoadjuvant therapy (Fig. 4). High IGKC expression was associated with complete response (CR) in a meta-analysis that included 7 cohorts ($n = 845$; $P < 0.0001$, Fig. 4). Analysis of the subgroups according to Desmedt (14) showed that IGKC is predictive for response in the ER$^−$/HER2$^−$ and in the HER2$^+$ subgroups but not in the ER$^+$/HER2$^+$ subgroups. In particular, the association with CR
was pronounced for the ER-negative patients ($P < 0.0001$, Supplementary Fig. S7A). In conclusion, IGKC showed strong correlation with survival, but also predicts chemosensitivity in ER-negative patients in the neoadjuvant setting.

IGKC is also prognostic in NSCLC and colorectal cancer

Because the immune response represents a general mechanism in tumor biology, we analyzed the prognostic impact of IGKC expression in lung, colorectal, and ovarian carcinomas (Supplementary Fig. S9). For lung cancer, we evaluated a novel cohort of 196 NSCLC patients from Uppsala. Both the B-cell metagene as well as single IGKC mRNA expression were significantly associated with longer survival in the univariate ($P < 0.001$) and multivariate Cox regression model ($P = 0.032$) adjusted to established clinical factors (Supplementary Fig. S9). Interestingly, Kaplan-Meier analysis in the subgroups revealed that this prognostic relevance was restricted to lung adenocarcinoma and was not seen in squamous lung carcinomas (Supplementary Fig. S9A and S9B), possibly because of the smaller sample size ($n = 66$). To further validate these results, we conducted a meta-analysis of publicly available Affymetrix data sets, including a total of 1,056 lung carcinomas (Fig. 3E and F). Both the univariate ($P = 0.011$; Fig. 3E) and the bivariate meta-analysis, adjusted to the proliferation marker ubiquitin-conjugating enzyme 2C UBE2C ($P = 0.015$; Fig. 3F), showed a significant association of IGKC with long-term overall survival.

Furthermore, we confirmed a significant association between IGKC and relapse-free survival in a meta-analysis of gene expression data of 513 patients with adenocarcinoma of the colorectum (Supplementary Fig. S9D). For overall survival, the association did not show significance (Supplementary Fig. S9E). No association was seen in a meta-analysis of 426 patients with ovarian cancer (Supplementary Fig. S9F).

IGKC protein expression in archived breast cancer tissue

Valuable biomarkers should be applicable for routine diagnostics. A major obstacle for gene expression studies is the limited availability of fresh tumor tissue in clinical practice. Indeed, most prognostic markers in breast cancer,
for example, ER, PR, HER2, and Ki-67, are routinely determined by immunohistochemistry. Therefore, we tested a monoclonal antibody against IGKC in FFPE tumor samples from the Mainz breast cancer cohort and found that IGKC was expressed in lymphoid cells in the tumor stroma of breast cancer (Fig. 5A). Immunostaining intensities correlated with IGKC RNA levels isolated from the tissue slides ($P = 0.014$; Jonckheere-terpstra test comparing staining intensity groups $0$ vs. $1^+$ vs. $2^+/3^+$) as well as with MFI (Fig. 5B).

**IGKC is expressed in tumor-infiltrating plasma cells**

Finally, to identify the cell type that was responsible for IGKC expression, we carried out costaining with antibodies against IGKC and either CD20 (a B-lymphocyte marker expressed in mature B cells but not on plasma cells), pan-cytokeratin (a marker for epithelial cells), or MUM1/IRF4 (a marker for activated B cells, plasmablasts, and plasma cells). No colocalization between IGKC and CD20, or IGKC and cytokeratin was observed (Fig. 5C). However, more than 90% of all cells that stained positive for nuclear MUM1/IRF4 were also positive for cytoplasmic IGKC (Fig. 5C). In addition, costaining with anti-human IgG showed that IGKC is only expressed in IgG-positive cells. Collectively, our results indicate that IGKC is expressed in mature plasma cells.

**Discussion**

Here, we describe a B cell–related gene signature, best represented by IGKC, as a strong prognostic marker in human breast, lung, and colorectal adenocarcinomas. Tumor-infiltrating plasma cells were identified to be the source of IGKC expression, which supports the concept that the adaptive humoral immune response is responsible for this host-dependent protective effect.

Numerous studies have shown the association of infiltrating immune cells and prognosis and response to therapy in different cancer types. However, most often the clinical relevance was ascribed to the T-cell lineage, with predominance of CD8$^+$, and CD45RO$^+$ T lymphocytes in colorectal, lung, and ovarian cancer (22–25).
metagenes, in the natural course of medically untreated 
gates the role of B and T cells, as typified by their respective 
acterized recently. Our own group systematically investi-
table evidence of the association of IGKC with survival has been verified in univariate and multivariate analyses. In our study, IGKC was associated with a better survival rate in both node-negative and node-positive breast cancer patients (15). The robust reproduction of IGKC’s clinical relevance in other cancer types represents in general one of the sparse exception that gene signatures are compatible between different cancer types. Mainly proliferation-related signatures have been shown to be transferable (29). Likewise the immunohistochemical analysis of the proliferation marker Ki-67 is of clinical importance in a variety of cancer entities. (30). In near analogy, the B-cell metagene reflects a general beneficial biological mechanism, which can easily be measured by IGKC protein staining. The validation of the gene expression findings in 330 node-negative FFPE tumors by immunohistochemistry was therefore of particular importance because fresh-frozen tissue is logistically demanding to obtain on a routine basis and often only small biopsies are available. Thus, an antibody-based detection of IGKC is applicable in routine cancer diagnostics.

Our finding that IGKC in tumors arises from plasma cells contradicts the provocative assumption that tumor cells are capable of producing immunoglobulins to promote growth and survival (31). Rather, it supports a previous report that breast cancer specimens typically have tumor infiltration of IgG-positive plasma cells (32). Similarly, another study of Wang and colleagues described that the majority of tumor-infiltrating plasma cells in invasive-ductal breast carcinomas was of IgG isotype suggesting that a tumor-derived antigen response may lead to the maturation of systemic B cells (33). In accordance, in our study, coexisting for IGKC and IgG confirmed increased heavy class isotype switch to IgG. This antigen-dependent switch from immunoglobulin M and immunoglobulin D to IgG1 production is a well-known feature of B-cell maturation (34) and plasma cell

### Table 1. IGKC is associated with MFI in 3 independent cohorts of systemically untreated node-negative breast cancer (combined Mainz, Rotterdam, and Transbig cohorts, n = 766): univariate Cox analysis

<table>
<thead>
<tr>
<th></th>
<th>Mainz cohort (n = 200)</th>
<th>Rotterdam cohort (n = 286)</th>
<th>Transbig cohort (n = 280)</th>
<th>Combined cohorts (n = 766)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IGKC a</td>
<td>0.052</td>
<td>&lt;0.001</td>
<td>0.060</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>P</td>
<td>0.81 (0.65—1.00)</td>
<td>0.80 (0.71—0.90)</td>
<td>0.85 (0.72—1.01)</td>
<td>0.79 (0.72—0.86)</td>
</tr>
</tbody>
</table>

*IGKC was analyzed as a continuous variable.

### Table 2. Multivariate Cox analysis adjusted to established clinical factors (combined Mainz and Transbig cohorts, n = 480)

<table>
<thead>
<tr>
<th></th>
<th>P</th>
<th>HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (&lt;50 vs. &gt;50 y)</td>
<td>0.791</td>
<td>1.14 (0.74—1.73)</td>
</tr>
<tr>
<td>pT stage (&lt;2 vs. &gt;2 cm)</td>
<td>0.012</td>
<td>1.78 (1.13—2.78)</td>
</tr>
<tr>
<td>Histologic grade</td>
<td>0.001</td>
<td>2.27 (1.41—3.65)</td>
</tr>
<tr>
<td>(grade 1 and 2 vs. grade 3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ER and PR (negative vs. positive)</td>
<td>0.964</td>
<td>1.01 (0.61—1.67)</td>
</tr>
<tr>
<td>HER2 status (negative vs. positive)</td>
<td>0.231</td>
<td>1.42 (0.79—2.53)</td>
</tr>
<tr>
<td>IGKC (continuous variable)</td>
<td>0.005</td>
<td>0.81 (0.70—0.93)</td>
</tr>
</tbody>
</table>

### Table 3. Prognostic relevance of IGKC determined by qRT-PCR in paraffin-embedded tumor tissue of patients (n = 330) with node-negative breast cancer

<table>
<thead>
<tr>
<th></th>
<th>P</th>
<th>HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IGKC (continuous variable)</td>
<td>0.004</td>
<td>0.882 (0.809—0.960)</td>
</tr>
<tr>
<td>Multivariate Cox analysis of MFI adjusted to established clinical factors</td>
<td>0.307</td>
<td>0.944 (0.593—1.501)</td>
</tr>
<tr>
<td>Age (&lt;50 vs. &gt;50 y)</td>
<td>0.880</td>
<td>0.964 (0.601—1.547)</td>
</tr>
<tr>
<td>pT stage (&lt;2 vs. &gt;2 cm)</td>
<td>&lt;0.001</td>
<td>3.853 (2.386—6.238)</td>
</tr>
<tr>
<td>Histologic grade (grade 3 vs. Grades 1 and 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ER and PR (negative vs. positive)</td>
<td>0.136</td>
<td>1.533 (0.874—2.690)</td>
</tr>
<tr>
<td>ERBB2 status (positive vs. negative)</td>
<td>0.405</td>
<td>1.277 (0.718—2.270)</td>
</tr>
<tr>
<td>IGKC (continuous variable)</td>
<td>0.001</td>
<td>0.871 (0.805—0.944)</td>
</tr>
</tbody>
</table>
differentiation (35) after antigen encounter. Notably, several reports have characterized oligoclonal expansion of B cells in breast cancer (36–40). But none of these groups have yet shown a robust clinical impact of these intriguing findings.

Interestingly, the impact of the peritumoral immune system could be shown in other tumor entities, that is, in NSCLC and colorectal cancer, but not in ovarian cancer. We speculate that this may be explained by distinct growth pattern in different organs and subsequent different immunogenic properties. The biological roles of the IGKC signature have to be addressed in further studies. Nevertheless, the strong prognostic impact shared by breast, lung, and colorectal adenocarcinomas represents, to the best of our knowledge, the first robust comprehensive biomarker predicting the response of the immune system in a variety of cancer types.

We have to acknowledge the retrospective nature of our study, but currently prospective analyses of breast cancer without adjuvant treatment are not feasible considering current treatment recommendations (41). Also, a detailed evaluation of additional malignant tumor types is difficult because of limited clinical and pathologic data in the published expression array data sets. It should be considered that not only k but also l light chain–associated probe sets are among the top genes indicating an antitumor response (Supplementary Fig. S11). However, IGKC combines the advantages of not only belonging to the top genes indicating a favorable prognosis but also offers the possibility that RNA from paraffin tissue can be used, and the results could be validated by immunostaining with commercially available antibodies.

The novelty of our study is (i) the translation of our B-cell metagene approach (15) to other tumor types, (ii) the validation by independent methods, and (iii) the establishment of IGKC as a biomarker for clinical diagnostics on FFPE tissues. In conclusion, our findings strongly support the emerging role of the immune system as a clinically relevant hallmark of cancer biology (42).

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Acknowledgments
The authors thank Ms. Seebase for her most competent support with data processing and statistical analyses and Mrs. Holzer as well as Mrs. Pfeffer for excellent technical assistance. Dr. Simon Elkan, Michael Bergkvist, and Kristina Lamberg helped to establish the Uppsala lung cancer cohort. The authors also thank Dr. Friedrich Kommoss for helpful discussion and Ms. Susanne Lindemann for valuable bibliographic support.

Grant Support
The study was supported by the Federal Ministry of Education and Research (BMBF, NGFN project Oncoprofile), the Swedish Cancer foundation, the Uppsala Lions Cancer foundation, and by the German Research Council (DFG, contract numbers RA 870/4-1 and RA 870/5-1).

No potential conflicts of interest were disclosed.

References
1. LaMarco K. Dynamic duo: cancer drugs and the immune system. Sci Transl Med 2008;1:12ec44.


