Discovery of Lung Cancer Biomarkers by Profiling the Plasma Proteome with Monoclonal Antibody Libraries*⍟

Mariana Guergova-Kuras‡§, István Kurucz¶, William Hempel‡, Nadège Tardieu‡, János Kádas¶, Carole Malderez-Bloes‡, Anne Jullien‡, Yann Kieffer‡, Marina Hincapie‖, András Guttman‖, Eszter Csánky‡‡, Balázs Dezső§§, Barry L. Karger‖, and László Takács‡¶ ¶¶**

A challenge in the treatment of lung cancer is the lack of early diagnostics. Here, we describe the application of monoclonal antibody proteomics for discovery of a panel of biomarkers for early detection (stage I) of non-small cell lung cancer (NSCLC). We produced large monoclonal antibody libraries directed against the natural form of protein antigens present in the plasma of NSCLC patients. Plasma biomarkers associated with the presence of lung cancer were detected via high throughput ELISA. Differential profiling of plasma proteomes of four clinical cohorts, totaling 301 patients with lung cancer and 235 healthy controls, identified 13 lung cancer-associated ($p < 0.05$) monoclonal antibodies. The monoclonal antibodies recognize five different cognate proteins identified using immunoprecipitation followed by mass spectrometry. Four of the five antigens were present in non-small cell lung cancer cells *in situ*. The approach is capable of generating independent antibodies against different epitopes of the same proteins, allowing fast translation to multiplexed sandwich assays. Based on these results, we have verified in two independent clinical collections a panel of five biomarkers for classifying patient disease status with a diagnostics performance of 77% sensitivity and 87% specificity. Combining CYFRA, an established cancer marker, with the panel resulted in a performance of 83% sensitivity at 95% specificity for stage I NSCLC.   *Molecular & Cellular Proteomics 10: 10.1074/mcp.M111.010298, 1–14, 2011.*

From the ‡Biosystems International SAS, 91000 Evry, France, ¶Biosystems International Kft., H-4012 Debrecen, Hungary, the ‖Barnett Institute and Department of Chemistry and Chemical Biology, Northeastern University, Boston, Massachusetts 02115, the **Department of Pulmonology, Medical and Healthcare Center of University Debrecen, H-4012 Debrecen, Hungary, the ‡‡Department of Pulmonology, Semmelweis Hospital and Healthcare Center, H-3500 Miskolc, Hungary, and the §§Departments of Pathology & Oral Pathology, Medical & Health Science Center & Faculty of Dentistry of University Debrecen, H-4012 Debrecen, Hungary

Lung cancer (LC)[1] is the most common cause of cancer-related death, and the low overall 5-year survival rate, ranging from 10 to 14% (1), at least in part, reflects the problem of late diagnosis of the disease. On the other hand, early detection significantly improves LC survival (2). Although recent attempts for early detection of lung cancer via computer tomography (CT) screening of nonsymptomatic subjects at high risk were effective (3) and showed improved survival (4), because of the relatively low specificity of CT imaging, there is an imminent need for better identification of LC patients in the group of individuals presenting with small solitary pulmonary nodules. Currently used plasma protein biomarkers for LC, such as carcinoembryonic antigen (CEA), cytokeratin-19 fragment (CYFRA), squamous cell carcinoma antigen (SCC), and neuron-specific enolase exhibit insufficient sensitivity and specificity (5) for population screening and routine diagnosis of LC and are used only in specific cases to monitor the efficacy of radio- and chemotherapy (6). Unbiased gene expression experiments (7), proteomics (8), or auto-antibody profiling (9) efforts so far also failed to discover reliable and clinically useful markers for the early detection of LC (10).

Global proteome profiling holds promise for testing the hypothesis of whether protein species exist that are NSCLC-specific. However, global proteomic analysis of plasma has been hampered by methodological issues such as the lack of appropriate affinity reagents for downstream verification and development of clinically useful diagnostic tests. Because of this bottleneck, only very few reported biomarkers were subsequently validated, and practically none made it to the clinic (11).

Global antibody proteomics approaches (12, 13) have aimed to generate libraries of antibodies to cover most or all

---

[1] The abbreviations used are: LC, lung cancer; ACT, α-1 antichymotrypsin; AUC, area under the curve; C9, complement factor 9; CEA, carcinoembryonic antigen; CFH, complement factor H; CYFRA, cytokeratin-19 fragment; COPD, chronic obstructive pulmonary disease; Hpt, haptoglobin; LD CT, low dose spiral computer tomography; LRG1, leucine-rich α-2 glycoprotein 1; mAb, monoclonal antibody; NSCLC, non-small cell lung cancer; ROC, receiver operator characteristics; SCC, squamous cell carcinoma antigen.

individual proteins in the human proteome but were targeted against recombinant proteins as immunogens. Although an important development, the limitation using recombinant proteins as antigens is that their antigenic epitopes are not represented in the natural state. Another issue with these initiatives is the inherent limitation by our current understanding of the natural human genome and proteome. Our efforts focused on the development of a new technology coined as mAb proteomics that leads to fit-for-purpose affinity reagents against native epitopes allowing for rapid translation of research results into clinically useful immunoassays without any *a priori* knowledge and biases. Here we present its application for the discovery of plasma protein biomarkers in LC. We produced complex mAb libraries in the form of hybridoma supernatants harvested from cultures of somatic fusion and termed them nascent mAb libraries. The nascent mAb libraries were targeted to the immunogenic epitome of the NSCLC cancer plasma proteome. Differential screening (cancer *versus* control) of the libraries identified mAbs detecting NSCLC-associated plasma protein epitope markers, some of which were also present in the cancer tissue samples. Ultimately, we identified five biomarkers whose levels were statistically different in the plasma of NSCLC patients and healthy controls. Among them, four proteins α-1 antichymotrypsin (ACT), leucine-rich α-2 glycoprotein 1 (LRG1), haptoglobin (Hpt), and complement factor H (CFH) were previously associated with LC (14–17), whereas complement factor nine (C9) is a biomarker for which no quantitative studies demonstrating an association with cancer have been previously reported. Screening of cloned and complex plasma proteome-specific mAb libraries with the cognate antigens led to the detection of antibody partners, allowing the development of sandwich immunoassays. Combination of the biomarkers with CYFRA (18) resulted in a diagnostic performance that may provide sufficient specificity to complement CT imaging in population screening of asymptomatic subjects with a high risk of LC.

EXPERIMENTAL PROCEDURES

Clinical Samples—Plasma samples from patients with newly diagnosed lung cancer and no previous treatment were obtained from informed patients and apparently healthy individuals after obtaining their written consent by a clinical protocol approved by the regional/local ethics committee and the institutional review board of the clinic/company (see Table I) from Proteogenex (Culver City, CA) under clinical protocol PG-ONC 2003/1, Asterand (Royston, UK) under clinical protocol AST-FB-003 and from the Department of Pulmonology of the University of Debrecen in Hungary under clinical protocol RKEB/IKEB:2422-2005. Plasma specimens for cohorts I and III were obtained using K2-EDTA as anticoagulant, whereas specimens for cohorts II and IV were obtained using citrate as anticoagulant. Lung cancer staging was done according to the American Joint Committee on Cancer and was based on information in the final histopathology report having the LC-histotype according to the World Health Organization classification (19). Clinical data including stage at diagnosis, histology, additional pulmonary pathologies, smoking habits, and general patient demographics are presented in Table I for each cohort.

Complex Immunogen Preparation

Depletion of Abundant Proteins—Depletion of 12 abundant proteins was performed using a commercially available SEPPRO IgY12 LC10® (12.7 × 79.0 mm) column from Beckman Coulter (Fullerton, CA) on a BioCad chromatography HPLC work station (Applied Biosystems, Foster City, CA). Chromatography was performed according to the protocol supplied by the vendor, with minor buffer modifications. Briefly, a plasma sample (250 μl) was thawed and diluted by the addition of 750 μl of buffer A (25 mM Tris, 0.5 M NaCl, 1 mM MnCl$_2$, 1 mM CaCl$_2$, and 0.05% sodium azide, pH 7.4). The diluted plasma was loaded onto the SEPPRO IgY12 column at a flow rate of 0.5 ml/min for 30 min; the flow rate was then increased to 2 ml/min for the remainder of the run. The unbound proteins (depleted fraction) were washed off with binding buffer, and the depleted fraction was collected into a 15-ml centrifugal filter (Amicon) with a cut-off at 5kDa. The depleted plasma was concentrated by centrifugation at 3,500 × *g*. The bound proteins were eluted from the column with stripping buffer (100 mM glycine, pH 2.5) and collected into a separate tube. The column packing material was neutralized with 100 mM Tris-HCl, pH 8.0, for 10 min and re-equilibrated with binding buffer, before performing protein depletion from the next plasma sample. A total of 20 runs for the lung cancer plasma from cohort I (see Table I) were carried out. The concentrated and depleted fractions were pooled before further processing.

Glycoprotein Enrichment—Glycoprotein enrichment was performed using a multi-lectin affinity chromatography column containing a mixture of three lectins: concavalin A, wheat germ agglutinin, and jacalin lectin from Vector Laboratories (Burlingame, CA), as previously described (20).

Normalization—Protein normalization intended to reduce the dynamic range of protein concentration of the depleted plasma and the glycosylated protein fraction was performed as previously described (21). Briefly, an immunoaffinity column was prepared by covalently linking rabbit polyclonal antibodies raised against normal human serum from Sigma to HiTrap protein G HP columns (4.6 × 100) from GE Healthcare (Chalfont St. Giles, UK) using 15 mM dimethyl pimelimidate and 15 mM dimethyl suberimidate. For enrichment of the depleted plasma and glycosylated protein enriched fraction, 1 mg of protein was loaded onto the column and incubated for 5 min before washing with 1× PBS buffer at pH 7.0. The flow-through represents the enriched protein fraction, whereas the bound proteins were eluted with stripping buffer B from Agilent Technologies (Santa Clara, CA). The column was equilibrated with 1× PBS buffer, pH 7.0, between runs.

Hybridoma Library Generation

Immunization of BALB/c Mice—Two groups of four female BALB/c mice, at least 8 weeks of age from Charles River Laboratories (Lyon, France) were injected subcutaneously in the rear footpads and at the base of the tail, each group with one of the complex antigen protein mixtures in the presence of Freund's adjuvant. Each mouse received 10 μg of either glycoprotein-enriched depleted plasma (group A) or normalized glycoprotein-enriched depleted plasma (group B) on days 1, 15, and 29. At least 3 weeks after the third immunization, an intravenous injection with 10 μg of the complex antigen mixture in PBS was performed to boost the immune response. The mice were housed and treated in an animal facility accredited by the Regional Department of Veterinary Services (accreditation number 91-228-107), and all of the procedures were approved by the local ethical committee.

Hybridoma Technology—Four days following the final intravenous injection, the mice were sacrificed by cervical dislocation, the spleens were removed by dissection, and a cell suspension was prepared.

Somatic hybridization of spleen cells with SP2/00-Ag-14 myeloma cells was performed using polyethylene glycol, and the fused cells were seeded into 96-well microplates under conditions that achieve oligo- or monoclonality for hybrid cell lines. Cultures were maintained under hypoxanthine-aminopterin-thymidine selection according to standard procedures (22). After allowing the cells to grow for 2 weeks, the IgG antibody production in each well was measured by ELISA using the protocol described below for HTS direct ELISA on 25 μl of hybridoma supernatant and horseradish peroxidase-coupled goat anti-mouse IgG as the detection antibody. Immunoglobulin secreting hybridoma supernatants derived from this initial stage are termed nascent mAb libraries, for discrimination from libraries of cloned and characterized mAbs. Cloning of selected hybridomas was performed using hypoxanthine-aminopterin-thymidine containing semisolid medium, based on methylcellulose Clonacell-HY medium D from Stem-Cell technologies (Vancouver, Canada), in 4-cm culture dishes. After 4–10 days, hybridoma cell colonies were picked and transferred into 96-well plates filled with hypoxanthine-aminopterin-thymidine-containing complete medium.

Antibody Purification—Ascites fluid was produced by intraperitoneally injecting 1–5×10^6 cloned hybridoma cells into incomplete Freund's adjuvant-primed BALB/c mice, using a protocol that was approved by the animal care committee of the animal facility. Ascitic fluid was cleared of cells and cell debris, and IgG was purified using a two-step affinity chromatography procedure consisting of prepurification on a Pierce thiophilic adsorbent column from Thermo Fisher Scientific (Waltham, MA) followed by further purification using Protein G-Sepharose 4 FF from GE Healthcare as recommended by the suppliers. Following concentration by ultrafiltration, the purity of the preparations was assessed by SDS-PAGE, and the protein concentration was determined using the BCA protein assay kit from Pierce. The purified antibodies were stored as working aliquots at -78 °C.

High Throughput Direct-ELISA Screening

Plasma Processing—Depletion of the most abundant plasma proteins was performed using commercially available columns according to the protocols supplied by the vendors. The Human-7 Multiple Affinity Removal System from Agilent Technologies (Santa Clara, CA) was used to deplete the pooled plasma used in the first step of the screening process, as well as the individual plasma samples used in the second screening (see Fig. 1). This depletion system was chosen because it provided the most efficient and reproducible depletion of abundant plasma proteins with no apparent carry over. Single-use SpinTrap depletion columns from GE Healthcare were used to deplete individual plasma samples in the third screening as well as in the experiments used to determine and verify the performance of the antibody panel, removing any possibility of sample carryover that might occur from multi-use columns. The quality and reproducibility of the depletion was determined and controlled by ELISA using commercially available mAbs against human albumin (Zymed Laboratories Inc., South San Francisco, CA) and IgG (Southern Biotechnology Associates, Inc., Birmingham, AL). The depleted plasma (pooled or individual) was labeled with bifunctional NHS-biotin containing a long alkyl chain as a spacer, EZ-Link Sulfo-NHS-LC-Biotin from Pierce according to the manufacturer's instructions. The biotinylated depleted plasma samples were used as tracers in the high throughput ELISA experiments. Biotinylation of purified proteins (listed in supplemental Table 1) was also performed according to the manufacturer's instructions.

HTS Direct ELISA—384-well high protein-binding plates (Corning Inc., Lowell, MA) were coated with goat anti-mouse Igγ chain specific polyclonal antibody (goat anti-mouse IgG) (Southern Biotechnology Associates, Inc.) for 2 h at room temperature. Following washing, the wells were blocked with 0.5% BSA in PBS at 4 °C overnight. Undi-

luted culture supernatants from nascent hybridomas or from cloned cell lines were added in quadruplicates. Mouse anti-human mAb against albumin (Zymed Laboratory Inc.) was spotted in eight wells on each plate and used as positive control. Tissue culture medium was also added to eight wells as a negative control. The plates were incubated overnight at 4 °C, and following washing, all wells were incubated with tracers (biotinylated depleted plasma). Biotinylated human serum albumin (supplemental Table 1) was added to the wells containing the positive controls, and the incubation continued for 90 min at room temperature. The unbound proteins were then removed by washing, and horseradish peroxidase-coupled avidin (Vectastain Elite ABC peroxidase kit from Vector Laboratories) was added to each well for detection of bound biotinylated proteins, as specified by the vendor's protocol. Following washing, color development was carried out by adding freshly prepared substrate solution (H_2O_2; Sigma) and chromogen (*o*-phenylenediamine; Sigma) to each well. The reaction was followed by kinetic reading for 4 min at 450 nm at 37 °C, at 32-s intervals using ELISA readers from Molecular Devices (Toronto, Canada). All of the steps amenable to robotization were automated.

Screening Data Analysis—The V_{max} was calculated from the linear portion of the curves using SoftMax Pro software from Molecular Devices. The Z′ factor, a metric to quantify the quality of the screening experiment with respect to reproducibility and data scatter (23), was calculated for each plate using the positive and negative control wells. Plates with a Z′ factor below 0.5 (usually less than 10% in a screening experiment) were excluded and repeated. The positive and negative controls were used to normalize the data across the plates.

Outliers for each group of replicates were removed using an automated procedure based on the mean and standard deviation values of the multiple measurements. In total, 9.2% of all generated data were considered as aberrant. The data obtained after normalizing and averaging the replicates (supplemental Fig. 1) were further analyzed using statistical methods.

Statistical Methods

The normality of the distribution of the results was estimated using the Wilks-Shapiro test. The distribution of the results for each hybridoma was calculated separately for the control and lung cancer samples. Nonparametric statistical analyses were applied: differences between two independent groups were determined using the Mann-Whitney U test; differences between more than two groups were assessed using the Kruskal-Wallis one-way analysis of variance test. All statistical tests were two-sided and were performed using R statistical software (www.cran.r-project.org). A predictive model for discriminating lung cancer cases from healthy controls using the panels of mAbs was produced using the freely available machine learning toolkit Weka (http://www.cs.waikato.ac.nz/~ml/) with a linear support vector machine and sequential minimal optimization algorithm. The model was established on the entire data set using 10-fold cross-validation. Logistic regression based on the sequential minimal optimization algorithm predictions was calculated to produce probabilities of class values (here lung cancer *versus* control) and to generate the receiver operating characteristics (ROC) curves.

Immunoprecipitation

Immunoaffinity magnetic beads were prepared for each antibody by mixing the antibody with protein G-coupled Dynabeads (Invitrogen) followed by covalent cross-linking of the bound antibody using dimethoxypropane, according to the manufacturer's instructions. The antigen was immunoprecipitated by mixing the beads with either total or depleted plasma according to the specific objective of the experiment. Following overnight incubation at 4 °C to allow antigen binding, the beads with bound antigen were washed, and the antigen was

eluted using the elution buffer supplied with the kit. Proteins from the resulting eluate as well as the unbound fraction were separated by SDS-PAGE and the gels subjected to Coomassie Blue or silver staining. The band corresponding to the specifically immunoprecipitated protein was excised, subjected to in-gel trypsin digestion, and processed for mass-spectrometric analysis.

Analysis was performed at different MS laboratories. At the Laboratory of Proteomics Research at Biological Research Center of the Hungarian Academy of Sciences (Szeged, Hungary) analysis was done using a Thermo LCQ Fleet three-dimensional ion trap coupled with Eldex nanoHPLC system. Mascot Distiller (version 2.2.1.0) was used as a peak list-generating software and default peak picking parameters were used for the triple-play method except the precursor charge state was considered both 2 and 3 (default charges) in all cases and the peak profile parameters were changed: minimum peak with 0.1 Da, expected peak with 0.5 Da, and maximum peak with 2.5 Da. Mascot Daemon (version 2.2.04.) software was used as a search engine on the NCBInr (20080718) database with 6,833,826 actual protein entries and without species restriction. At the Proteomics Core Facility of the Department of Biochemistry and Molecular Biology of the Medical and Health Science Center of University of Debrecen (Debrecen, Hungary), analysis was done using ABI Q-Trap 4000 MS/MS system coupled with Agilent 1100 nanoLC. ProID (ABSciex) (version 1.4) was used as a peak list-generating software and search engine with default parameters. The SwissProt database (version 20090201) with 512,205 actual protein entries was searched without species restriction. At the Plate-Forme Spectrométrie de Masse et Protéomique at the Université Pierre et Marie Curie, Paris, France, analysis was done using the 4700 Proteomics Analyzer MALDI-TOF/TOF from Applied Biosystems. Peak lists were generated with Data Explorer 4.6 with default parameters and were submitted to the Mascot search engine with the NCBInr 20091010 database (9,868,855 sequences) restricted to human sequences.

Western Blot Analysis

Ten μg of each total and depleted protein sample was separated by SDS-PAGE on 4–20% Tris-glycine gels from Invitrogen and transferred to a nitrocellulose membrane (Whatman). The membranes were blocked for 1 h in PBS-Tween (pH 7.2) containing 5 mg/ml polyvinylpyrrolidone (Sigma) and incubated with the specific primary antibody (0.2 or 0.4 μg/ml) overnight. Immunoblots were developed by peroxidase-conjugated anti-mouse antibody from Southern Biotechnology Associates, Inc. and an ECL detection kit (Pierce). Protein bands were detected using different exposure times with a Gel Logic 1500 imaging system and Kodak MI software from Eastman Kodak (Rochester, NY). PageRuler Prestained Protein Ladder (Fermentas; Thermo Fischer Scientific) and MagicMarkXP Western protein ladder (Invitrogen) were used for molecular weight determination.

Mammalian Expression of Recombinant Human LRG1

HEK293T cells at ~60% confluency were transfected with the plasmid, ORIG-SC321789, from OriGene (Rockville, MD) containing the True-Clone (Human NM_052972) for human LRG1 using turbofectin 8.0 (OriGene) according to the manufacturer's description. Cell supernatants were collected 24, 48, and 72 h after transfection and stored at −20 °C.

ELISA with Recombinant Human LRG1

The experiment was based on the affinity of cytochrome c for LRG1 described in Ref. 24. The 96-well high-binding plates (Corning Inc.) were coated with cytochrome c from horse heart (Sigma) for 1 h at room temperature. Following washing, unbound sites were blocked with 0.5% BSA in PBS at room temperature for 30 min. Supernatants containing recombinant human LRG1, added to the wells at different dilutions, were incubated for 1 h at room temperature. After washing, specific antibodies were added at a concentration of 1 μg/ml and incubated for 1 h at room temperature. Following washing, all of the wells were incubated with horseradish peroxidase-coupled anti-mouse IgG (Southern Biotechnology Associates) for 30 min at room temperature. Finally, unbound proteins were removed by washing, and the plates were developed by adding freshly prepared substrate (H_2O_2) solution to each well containing tetramethylbenzidine (Sigma) as a chromogen. The reaction was terminated after 10 min of incubation, and the end point was read at 450 nm.

ELISA for CYFRA, CEA, and SCC were performed according to the manufacturer's instructions using respectively CYFRA 21–1 EIA, CanAg ACE (CEA) EIA, and CanAg SCC EIA kits from Fujirebio Diagnostics, Inc. (Malvern, PA).

Immunohistochemistry

All frozen and formalin-fixed paraffin-embedded surgical specimens used in this study were obtained from the research file of the Department of Pathology, University of Debrecen, utilizing protocols approved by the institutional review boards. Following hematoxylin-eosin staining of sections (4 μm), immunohistochemical labeling was carried out with the use of the Envision (biotin-free) peroxidase-based detection kit (Dako, Glostrup, Denmark) for mouse monoclonals using the purple VIP™ substrate-chromogen visualization from Vector Laboratories followed by either methyl-green or hematoxylin nuclear counterstaining, as previously described (25, 26). All of the antibodies were diluted in protein-free antibody diluent (DAKO, Denmark) and found to be reactive on tissues within the range of 1–4 μg/ml on native/cryostate and paraffin sections, respectively. To check the staining specificity for the test mAbs, each run included sections for negative control staining where isotype-specific irrelevant control mAbs (Dako, Denmark) were used in place of the primary antibodies. Tissue reactivities for the mAbs were then scored as previously described (26).

RESULTS

Application of mAb Proteomics for Lung Cancer Plasma Profiling—In our reverse approach for biomarker discovery (monoclonal antibody proteomics) (27), we first generated nascent and cloned libraries of affinity reagents (mAb) against native plasma protein epitopes, and then via unbiased high throughput direct ELISA profiling, we identified mAbs that discriminate plasma samples of LC patients from plasma samples of healthy controls (Fig. 1a). To generate the appropriate immunogen (Fig. 1b), we depleted the 12 most abundant plasma proteins representing ~90% (w/w) of the pooled samples from twenty NSCLC cases (cohort I in Table I) by means of a commercially available immunoaffinity depletion column. The small clinical cohort included patients with advanced cancer disease from the two main histological subtypes of NSCLC, adenocarcinoma, and squamous cell carcinoma. Two strategies were used to obtain protein fractions of high complexity from the depleted plasma proteins (Fig. 1). In the first strategy, plasma was depleted of abundant proteins. As a second strategy, glycosylated proteins, known as a source of potential cancer biomarkers (28), were enriched using multi-lectin affinity chromatography (20). Both the de-
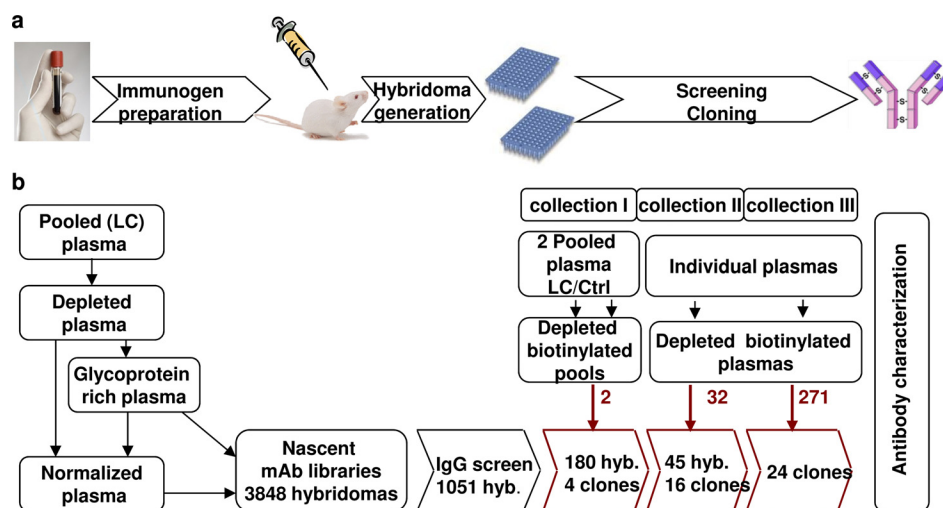
Fig. 1. **Monoclonal antibody proteomics strategy.** *a*, major steps. *b*, lung cancer biomarker discovery via mAb libraries. The plasma samples used in each screening step correspond to the cohort indicated on the *top* of the figure, and the number of samples are indicated next to the *red arrows*. The number of qualified hybridoma candidates (original or cloned cell line) showing statistical discrimination between LC and control after each screening step are indicated in the *red boxes*.

pleted plasma and glycosylated protein enriched fractions were then subjected to immunoaffinity chromatography based normalization (21) aimed to reduce representational differences among the remaining plasma proteins and to maximize the number of protein species reaching the immunogenic threshold during immunization. Hybridoma supernatants containing at least 50 ng/ml IgG were then tested with biotinylated plasma tracers using direct ELISA in a series of screening experiments designed to identify antibodies that detect plasma biomarkers for LC (Fig. 1*b*). The technical variability of the direct ELISA assay, in six identical experiments, was 10.6% across the full range of signal intensities, equivalent to the performance of protein microarrays and significantly better than that of shotgun mass spectrometry protein profiling (29). The reproducibility of the tracer preparation evaluated by the independent processing and screening of plasma aliquots with a panel of mAbs was above 90% (supplemental Fig. 1).

In the initial screening step, all 1051 IgG-producing nascent hybridomas from two somatic cell fusions were tested with depleted and biotinylated pooled plasma from 21 healthy control subjects and the 20 NSCLC patients used for immunogen preparation (cohort I) (Fig 2*a*). Plasma pooling was necessary to reduce the need for biological material (hybridoma supernatants and human plasma) and to reduce the biological variability of this step. A total of 184 nascent hybridoma supernatants were found to detect a normalized signal ratio higher than 1.5-fold between the disease and normal pools. The majority of antibodies (87%) detected a higher signal in the pooled LC plasma. The next screening steps involved testing the antibodies with individual plasma samples. First, the 184 candidate mAbs were examined with plasma from a small clinical cohort of 32 subjects (20 LC cases and 12 control subjects; Table I). The screening iden-

tified 61 hybridomas, discriminating ($p < 0.05$) the group of control subjects from subgroups of LC cases. The next screening step aimed to select those mAb candidates that separate mostly early stage LC from healthy subjects, thus securing the identification of biomarkers for early lung cancer detection. To this end, all of the 61 mAb candidates were screened with a larger and independent clinical cohort (cohort III; Table I and Fig. 2*b*) heavily weighted toward early stage (stage I and II) lung cancer. In this screening step, we identified 24 hybridomas that detected a difference of at least 15% between the apparent abundance of the cognate protein in plasma of the control and LC cases (based on the median group values, $p < 0.05$). The 13 best discriminating hybridomas were further characterized (supplemental Table 2). At this stage, all hybridomas were cloned at least two times. We sequenced the variable regions of the IgG heavy chain cDNAs, and the results confirmed the expectation, that all mAbs were from independent clones (supplemental Fig. 2).

Identification and Characterization of the Protein Biomarkers—Because the mAb libraries were generated against highly complex and quasi-representative mixtures of immunogenic proteins, mAb proteomics was not biased by *a priori* knowledge of cognate biomarker antigens. Cognate protein identification was performed at the distal phase of the process (30).

Ten mAbs that displayed clonal stability and high mAb yield were subjected to protein identification wherein immunoprecipitates were analyzed by shotgun mass spectrometry. The results revealed four different plasma proteins: Hpt beta chain for mAbs Bsi0071, CFH for mAb Bsi0271, Complement C9 (C9) for mAb Bsi0270 and Bsi0272, and leucine-rich α-2 glycoprotein 1 (LRG1) for mAb Bsi0351, Bsi0352, and Bsi0392. Peptide sequences from the MS experiments and the protein identifiers are summarized in supplemental Table 3. Visual pattern analysis of Western blots of SDS-PAGE gels of im-

TABLE I
Clinical cohorts with plasma samples used in the study

Cohort	Status	Number of subjects	Age	Gender[a]	Smoking status	packs/day (average)	Stage LC	Histology[b]	Other LD[c]
Cohort I (Proteogenex)	LC	20	46–75 (average 59)	M: 20	Current use: 15 Never used: 5	1.2	II: 3 II–III: 3 III: 12 IV: 1 Unknown: 1	SQC: 16 AC: 3 Other NSCLC: 1 SCLC: 0	COPD: 2
	Healthy	21	45–66 (average 55)	M: 21	Current use: 9 Never used: 12	1.0			NA
Cohort II (University of Debrecen)	LC	20	49–78 (average 61)	M: 17 F: 3	Current use: 20	NA	III–IV: 8 IV: 2 Unknown: 10	SQC: 5 AC: 10 Other NSCLC: 0 SCLC: 5	COPD: 5
	Healthy	12	50–64 (average 57)	M: 12	Current use: 12	NA			COPD: 6
Cohort III (Asterand)	LC	219	26–86 (average 61)	M: 158 F: 61	Current use: 116 Occasional use: 5 Previous use: 63 Never used: 32 Unknown: 3	1.3	I: 128 II: 39 III: 31 IV: 5 Unknown: 16	SQC: 104 AC: 85 Other NSCLC: 25 SCLC: 5	COPD: 6 Emphysema: 5 Chronic bronchitis: 21 Asthma: 3
	Healthy	169	18–70 (average 38)	M: 114 F: 55	Current use: 49 Occasional use: 4 Previous use: 25 Never used: 91	1.0			
Cohort IV (University of Debrecen)	LC	45	44–77 (average 59)	M: 32 F: 13	Current use: 45	1.7	I: 1 III: 1 III–IV: 12 IV: 18 Unknown: 13	SQC: 14 AC: 10 Other NSCLC: 4 SCLC: 17	COPD: 22
	Healthy	63	34–72 (average 54)	M: 23 F: 40	Current use: 63	1.1			COPD: 0
	Other LD	112	35–76 (average 57)	M: 58 F: 54	Current use: 112	1.3			COPD: 39 Pneumonia: 21 Fibrosis: 29 Sarcoidosis: 23

[a] M, male; F, female.
[b] AC, adenocarcinoma; SCLC, small cell lung cancer.
[c] Other LD indicates other noncancerous fibrotic, inflammatory, and autoimmune lung diseases; NA, information not available.
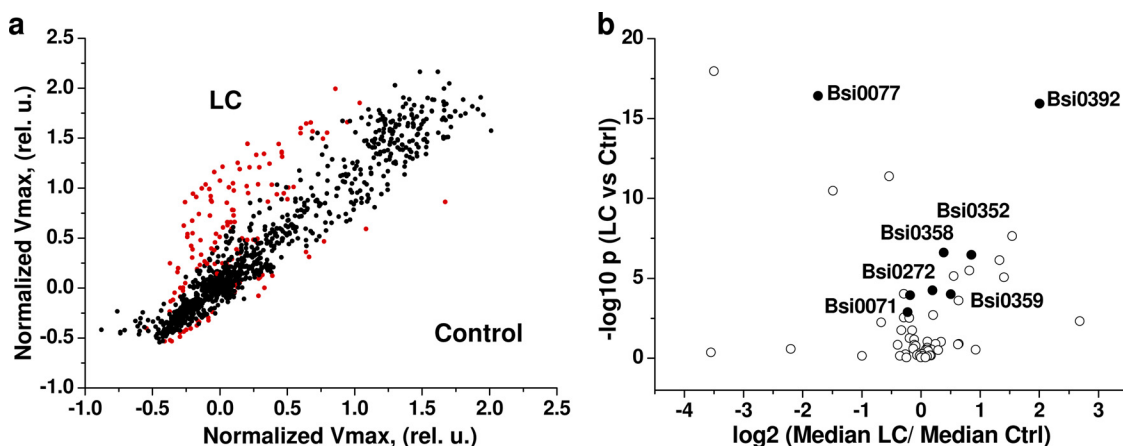
FIG. 2. **Screening for LC specific antibodies.** *a*, results from HTS ELISA of hybridoma supernatants with pooled plasma (*n* = 20) from LC patients (*y axis*) are plotted against the results obtained with pooled plasma (*n* = 21) from the matched control subjects (cohort I). Hybridomas considered as primary hits are shown in *red*. *b*, results from the screening of 61 candidate hybridomas with individual plasma samples from cohort III. The log2 of the fold difference between the median of the cancer and control groups as measured with each antibody is plotted on the *x axis*, whereas the *p* value from the Mann-Whitney *U* test of the comparison between the two groups is expressed on the *y axis*.

munoprecipitates of mAbs from a cloned mAb library and subsequent sandwich ELISA screening suggested that mAbs Bsi0358 and Bsi0359 react with ACT, whereas Bsi0077 reacts with CFH. Protein identification was validated by ELISA experiments with purified natural plasma proteins (Fig. 3*a*) and recombinant LRG1 expressed in HEK293 cells (Fig. 3*b*).

Our results show that the 13 leading candidate mAbs recognized at least five different biomarkers, whose concentrations in the plasma range from 40 μg/ml (C9) to a few mg/ml (Hpt). Mimotope redundancy testing with phage display suggested that antigenic epitopes were all independent (supplemental Table 4).

To address the question of whether the protein biomarkers are detectable in the cancer cells and to assess their specificity for LC tissue, five mAbs (one for each analyte) were tested on 286 archived tissue microarray cores (each 3 mm in diameter) from 85 cases with histologically and immunohistochemically well characterized stage I and II NSCLC. In addition, small cell LC and 42 normal tissue samples representing 16 different tissue types including intact lung were also examined (supplemental Table 5). The analysis of LC and other normal tissue specimens indicated that four of the mAbs, Bsi0033 (Hpt), Bsi0077 (CFH), Bsi0272 (C9), and Bsi0358 (ACT) distinctly and specifically reacted with LC cells (Fig. 3*c*), whereas normal tissues including lung had no or low reactivity. Bsi0392 (LRG1) did not react with the tissue sections (Table II). Similar to previously published reports (31), the anti-Hpt-specific mAb (Bsi0033) was shown to preferentially stain adenocarcinoma as compared with squamous cell carcinoma (Table II).

Clinical Utility of the Identified Biomarkers—Although direct ELISAs of the labeled plasma samples with the single mAb candidates do not provide absolute measurements of the biomarker concentrations, we have used the relative values (normalized rates from kinetic readings) obtained in the

screening experiments to estimate the diagnostic capacity of each biomarker. The values measured in the plasma of 219 LC cases and 169 healthy controls of clinical cohort III are shown in Fig. 4 (*a–e*). An increase of biomarker concentration in LC cases is clear for LRG1, ACT, C9, and Hpt. In the case of CFH, we observed a slight decrease. There was no association of the level of the biomarkers with gender neither in the group of cases nor in the group of controls. Similarly, smoking habits evaluated in two groups (current and occasional smokers *versus* previous smokers and nonsmoking individuals) did not correlate with any of the biomarker levels (not shown).

The discrimination capacity of the selected mAbs as potential diagnostics tools was determined from the area under the receiver operating characteristics curves (AUC and ROC, respectively). Biomarkers with the best performance were determined by the AUC as LRG1 and ACT, followed by Hpt, C9, and CFH (Fig. 4*f*). The sensitivity at 95% specificity was different among the different biomarkers and ranged from 45% for LRG1 to 13% for CFH when estimated for all LC stages. The performance of the biomarkers was evaluated for early stage LC patients (stage I) and compared with three current biomarkers for LC CYFRA, CEA, and SCC (32). From the biomarkers reported here, LRG1 had again the best performance with an AUC of 0.78, better than that of CEA (0.70) and SCC (0.67) but lower than CYFRA (0.84) (Fig. 4*g*).

Translation to an Immunodiagnostic Tool—The apparent protein redundancy observed at the biomarker candidate level prompted us to screen a complex (>600 mAbs) cloned plasma proteome-specific mAb library for the identification of partner mAbs to develop robust quantitative sandwich assays for each biomarker protein (Table III). Testing multiple candidates for each mAb resulted in assays for four of the biomarkers (Hpt, C9, ACT, and CFH). The library did not contain matched pair antibodies for LRG1 (24). We have to note here that the initially identified mAbs were also capable of reacting
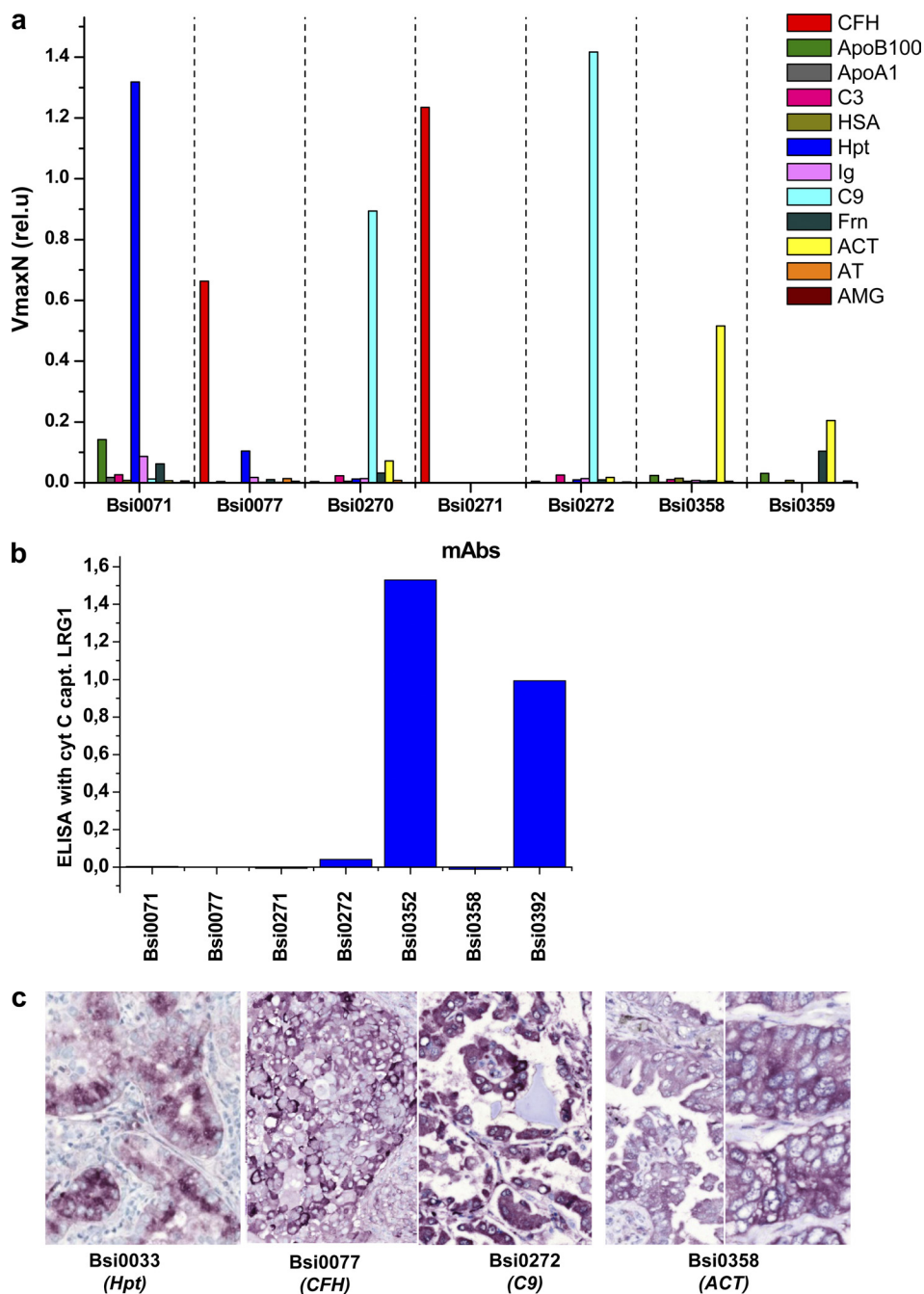
FIG. 3. **Antigen specificity of the selected mAbs.** *a*, reactivity of the selected antibodies with plasma proteins, measured using direct capture assay of the purified proteins (listed in supplemental Table 1) that were biotinylated, and the reaction with the antibodies was measured using ELISA. The signal is expressed relative V_{max} units of kinetic readings. *b*, reactivity of the selected antibodies with cytochrome *c*-captured recombinant human LRG1 protein in sandwich ELISA. Binding of the antibodies was detected using mouse-IgG specific horseradish peroxidase-enzyme antibody conjugate. Signal is expressed in relative optical density units. *c*, *in situ* immunohistochemical detection of various primary lung carcinoma tissue sections (paraffin sections after antigen retrieval) with the selected mAbs (BSI0033 is a LC- and Hpt-specific mAb that we described earlier (28); all other mAbs are described in this paper). All mAbs show specific immunoreaction with purple VIP™ chromogen detection. Note that the neoplastic cells are selectively labeled, in intracytoplasmic fashion, by the indicated mAbs (*purple-red staining*), whereas the stromal mesenchymal elements remain essentially negative (original magnification, 40×).

in sandwich format for three of the biomarkers (Hpt, CFH, and C9), suggesting recognition of structurally independent protein epitopes for the respective mAbs. In the case of ACT and LRG1, the mAbs were probably generated against structurally overlapping epitopes because they were not able to form matched pairs. The epitope mapping experiment using pep-

TABLE II

Primary NSCLC cases that were positively immunostained with the selected mAbs

Tissue specimens tested[a]	Bsi0033 (Hptβ)	Bsi0077 (CFH)	Bsi0272 (C9)	Bsi0358 (ACT)	Bsi0392 (LRG1)
NSCLC	64/80 (80%)	69/80 (86%)	59/80 (74%)	74/80 (92%)	0/81
AC	57/64 (89%)	54/64 (84%)	49/64 (77%)	58/64 (91%)	0/64
SQC	2/10 (20%)	10/10 (100%)	5/10 (50%)	10/10 (100%)	0/11
LCC	5/6 (83%)	5/6 (83%)	5/6 (83%)	6/6 (100%)	0/6
SCLC	3/18 (17%)	2/19 (11%)	1/16 (6%)	3/17 (18%)	0/15
Normal[b]	3/16 (19%)	3/16 (19%)	1/16 (6%)	6/16 (37%)	0/16

The number of positively stained tissue specimens (A) *versus* the total number of tested specimens (B) with each antibody is shown as ratio (A/B) and as a percentage.

[a] ADC, adenocarcinoma; LCC, large cell carcinoma; SCLC, small cell lung carcinoma. Normal included brain, spleen, heart, liver, skeletal muscle, smooth muscle, prostate, adrenal, kidney glomeruli, kidney tubule, thymus, lung alveoli and bronchi, and thyroid.

[b] The staining intensity in normal tissues noted as positive here, was only mild (2 in a intensity scale of 4 levels).

tide libraries showed that two of the mAbs against LRG1 (Bsi0351 and Bsi0352) had some sequence similarity (22%) between the peptide sets (supplemental Table 4), also confirming similar mimotop affinities. The two assays developed for this biomarker, either using cytochrome *c* or a previously published antibody 2F5.A2 (24) as a sandwich partner, provided similar results (Table III). Because of the better dynamic range of the sandwich assays, all biomarkers except CFH showed improvement of the discrimination between LC and healthy samples. In contrast to the single mAb-driven capture assays, the sandwich ELISA assays do not require a sample preparation step, and thus they are much more suitable for clinical applications as *in vitro* immunodiagnostics tools.

Using support vector machines and 10-fold cross-validation of the results generated on clinical cohort III (all LC stages included), we have determined a panel classifier composed of the five markers reported here. As expected, the panel increased the accuracy of the prediction (AUC of 0.88) compared with individual markers and provided specificity of 87% with a sensitivity of 77% (Fig. 4h, *solid red line*). Interestingly, from the known biomarkers, only CYFRA provided additional performance to the panel for discriminating stage I NSCLC patients and healthy controls. The results indicate that the combined performance with CYFRA is superior (Fig. 4h, *solid black line*) with an AUC of 0.93 and a sensitivity of 84% at 95% specificity across all stages of NSCLC. For stage I NSCLC, we observed slightly lower performance than that for more advanced stages (Fig. 4h, *inset*) but could still achieve a sensitivity of 83% at 95% specificity. To verify the performance of the panel on an independent collection, we have further measured plasma samples from an independent clinical cohort of 45 LC cases and 63 age-matched healthy controls (clinical cohort IV). The plasma were collected at a single clinical center and processed using a different analytical procedure as compared with clinical cohort III. The estimated performances of both panels were very similar both in terms of AUC and diagnostic performance (Fig. 4h, *dotted lines*), suggesting that the biomarker panel is not strongly

dependent on the type of analytical preprocessing as reported for other plasma biomarkers panels (33).

Hpt, LRG1, ACT, and C9 have been reported as acute phase proteins (34, 35). As a regulatory protein of the complement cascade, CFH also plays an important role in inflammation. Therefore, we tested acute pneumonia patients with elevated C-reactive protein levels and a group of patients with other inflammatory and auto-immune benign pulmonary diseases (clinical cohort IV, other LD). Although the COPD, fibrosis, and sarcoidosis patient groups were not distinguishable from controls, the calculated lung cancer index was slightly elevated in pneumonia patients (Fig. 4i).

DISCUSSION

We have developed a widely applicable approach for the generation of fit-for-purpose affinity reagents in the form of mAbs directed against disease-associated biomarkers requiring no *a priori* knowledge of disease biology and have demonstrated its utility using lung cancer as an example. Four of the five lung cancer-associated biomarkers reported here have been previously associated with cancer with varying levels of experimental evidence, whereas quantitative studies demonstrating complement C9 as a cancer biomarker in plasma were not reported. Haptoglobin in general, as well as the differentially glycosylated forms of this abundant plasma protein, has the strongest clinical evidence as a general cancer biomarker (36, 37). In LC, the levels of serum haptoglobin and its different glycoforms have been demonstrated as potential biomarkers for both small cell lung cancer (38) and NSCLC (14). We have also published results previously obtained with antibody Bsi0033, for which we showed a higher affinity for glycosylated forms of haptoglobin (30). There is also strong evidence for an association of CFH with bladder (39) and ovarian cancers (40). Autoantibodies to CFH were reported in sera of patients with early stage NSCLC (41). The association of the increased level of ACT in plasma/serum with pancreatic cancer was demonstrated in two proteomics studies (42, 43). In LC, the specific staining of NSCLC tumors
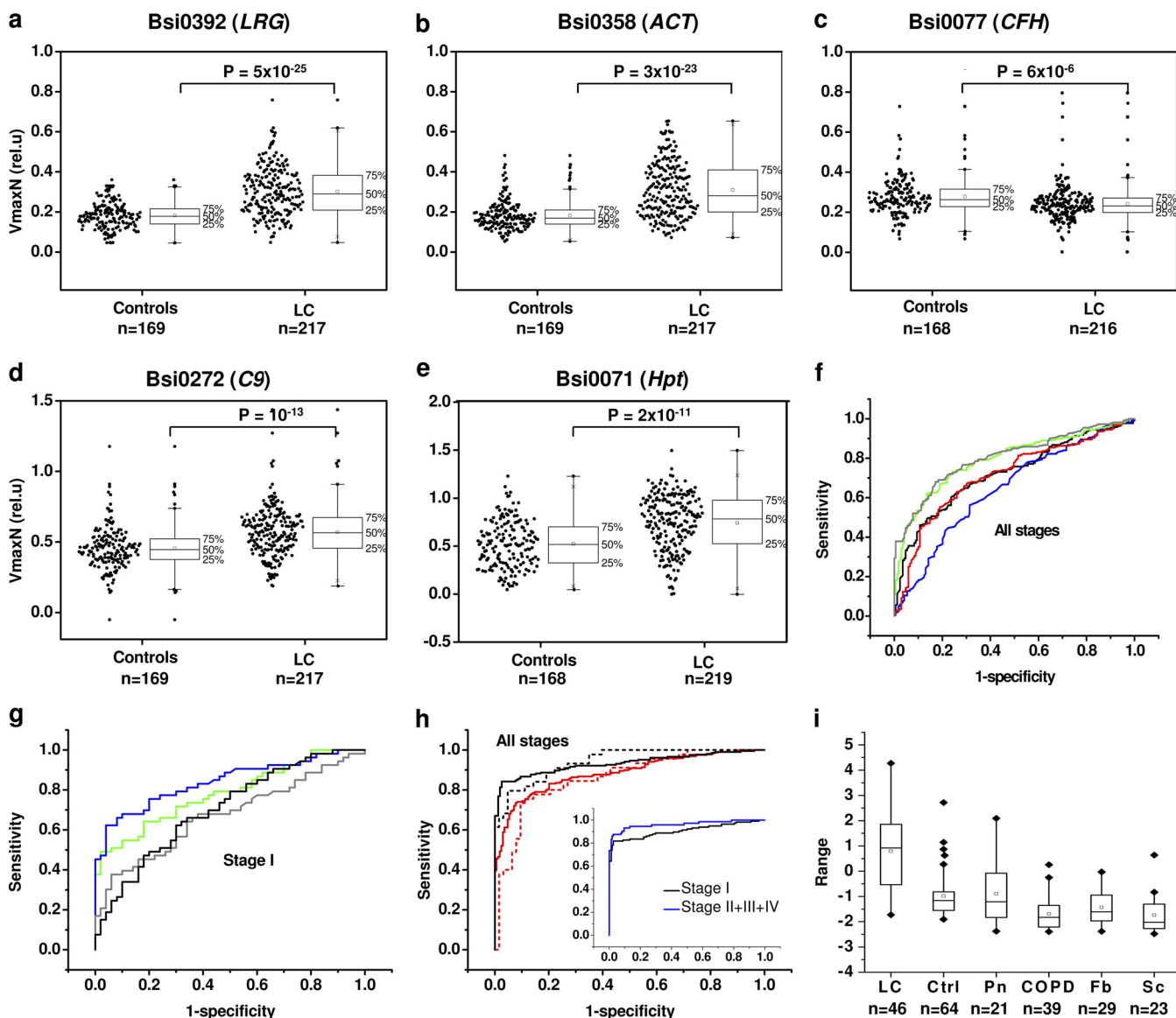
FIG. 4. **Clinical significance of the identified biomarkers.** *a−e*, statistical analysis of the ELISA results for the five biomarkers with capture ELISA. Nonparametric statistical analyses were applied: differences between two independent groups were determined by means of the Mann-Whitney *U* test. For each biomarker, the results obtained with the respective mAb and the individual plasma samples from clinical cohort III in the ELISA experiments are plotted on the *y axis*. The results are grouped by the disease status of the patients (healthy *versus* LC cases) on the *x axis*. The descriptive statistics for each group are shown as box plots on the *right side* of the individual measurements. *f*, ROC analysis for single biomarkers. *Gray*, Bsi0392 (LRG1), AUC = 0.8; *green*, Bsi0358 (ACT), AUC = 0.79; *black*, Bsi0071 (Hpt), AUC = 0.72; *red*, Bsi0272 (C9), AUC = 0.71; *blue*, Bsi0077 (CFH), AUC = 0.64. *g*, ROC analysis for LRG1 and three known biomarkers for the group of stage I LC as compared with healthy controls. *Blue*, CYFRA, AUC = 0.84; *green*, LRG1, AUC = 0.78; *gray*, SCC, AUC = 0.67; *black*, CEA, AUC = 0.70. *h*, ROC analysis for panel of five biomarkers (LRG1, ACT, C9, HPT, and CFH) using a linear combination of the individual measurements with capture ELISA. *Red solid line*, clinical cohort III (AUC = 0.88); *red dotted line*, clinical cohort IV (AUC = 0.86) and panel of five biomarkers (CYFRA, LRG1, ACT, C9, and Hpt) with sandwich ELISA. *Black solid line*, clinical cohort III (AUC = 0.93); *black dotted line*, clinical cohort IV (AUC = 0.94). The performance of the panel (CYFRA, LRG1, ACT, C9, and Hpt) in stage I patients (*black line*) as compared with stages II, III, and IV (*blue line*) in cohort III is shown in the *inset*. *i*, the index obtained from the combined measurements of four biomarkers (LRG1, ACT, C9, and Hpt) in patients with benign lung diseases as compared with healthy controls and lung cancer patients (clinical cohort IV). *Pn*, pneumonia; *Fb*, fibrosis; *Sc*, sarcoidosis. The results are presented as box plots.

with a monoclonal antibody that recognized ACT was reported and demonstrated that the tumor cells synthesized ACT *de novo* (16). Our results confirmed this observation and also provide evidence that early stage LC is associated with

an increased plasma concentration of ACT. The increase in serum/plasma levels of LRG1 has been associated with the presence of ovarian cancer (44, 45) as well as lung cancer in two semi-quantitative proteomics studies (15, 46) with nine

TABLE III
Identification of matched paired antibodies for sandwich assays

Protein	mAb	Direct ELISA			Matched pair identification (number of mAb partners)	Sandwich ELISA			
		Number of samples		Fold change[a]		mAbs[b]	Number of samples		Fold change[a]
		Control	LC				Control	LC	
C9	Bsi0272	169	217	1.27	6[c]	Bsi0272, Bsi0452	166	212	1.55
LRG	Bsi0352	166	209	1.88	0[d]	Bsi0352, 2F5.A2	166	190	2.13
LRG	Bsi0392	169	217	1.63	0[d]	Cytochrome c, Bsi0392	166	211	2.22
Hpt	Bsi0071	168	219	1.52	2[c]	Bsi1709, Bsi0033	154	106	1.75
ACT	Bsi0358	169	217	1.66	9[c]	Bsi0186, Bsi0358	166	212	1.87
CFH	Bsi0077	168	216	0.88	28[c]	Bsi0271, Bsi0885	166	208	0.87

[a] The fold change was calculated from the median values of the LC group *versus* the healthy controls.

[b] The capture mAb is shown first, and the detection labeled antibody second.

[c] The full library was first screened in direct ELISA with the biotinylated protein. The mAbs showing a positive reaction were then biotinylated and screened in SW format using total plasma and/or protein as antigen and the mAb identified from the direct screening as a partner.

[d] The full library was screened with the biotinylated mAbs identified in the direct screening.

and six patients, respectively. We are thus providing the first independent validation of this biomarker using a quantitative immunoassay on two independent clinical cohorts with 264 LC cases and 232 healthy controls. The C9 component of the complement system has not been associated as a plasma marker with cancers of any type.

Three of the biomarkers identified by our study (C9, Hpt, and ACT) are seemingly acute phase proteins (34), and there is some recent evidence that LRG1 is also induced by human IL-6 and synergistically up-regulated with either IL-1β or TNF-α in a pattern similar to those of type 1 acute phase proteins (35). The combination of increased plasma levels of Hpt, ACT, and LRG1 was recently shown to occur in patients with ovarian carcinoma (44). The increase of the levels of other acute phase proteins, such as C-reactive protein and serum amyloid protein, was associated with LC in several independent reports (47–49). It is now well accepted that cancer induces an inflammatory reaction, the nature and the extent of which is variable and depends on the cancer type (50). Interestingly, generalized inflammation reaction as deduced from nonspecific inflammation markers appears to play a particularly important role in lung cancer, and it is a strong negative prognostic factor for overall survival (51). Anti-inflammatory medication, such as cortisosteroid treatment of COPD patients, has been reported to decrease LC incidence (52), supporting the notion that the inflammatory reaction promotes cancer even at the initial stages. Another recent study showed that inflammatory responses to environmental exposures such as tobacco smoke also correlate with an increased risk of LC (53). However, the profiles of up-regulated acute phase proteins were reported to be specific both in the case of ovarian cancer (54) and in the case of LC, because other inflammatory lung diseases such as COPD did not reveal increases of the same proteins (47). In the prototype acute inflammatory clinical condition, bacterial pneumonia, associated with elevated levels of C-reactive protein, we detected some up-regulation of Hpt, C9, LRG1, and ACT, suggesting that the mechanisms, which induce transcriptional activation

and production rates of these biomarkers, share some but not all regulatory elements with the inflammatory response. The fact that we also detected four of the biomarkers in the cancer cells by immunohistology strongly suggests that proinflammatory stimuli, *e.g.* cytokines in concert with yet unknown cancer-specific mechanisms actually induce ACT, C9, HPT, and CFH expression by the tumor tissue.

Considering the circulating levels ($\sim$0.1–0.5 mg/ml) of the markers and the size of a stage I LC mass (1–2-cm diameter), it is unlikely that the tumor alone produces the markers detected in the plasma. Instead, it is more likely that as a result of the local inflammatory response, the surrounding lung tissue or more distant sites also contribute to the increase in the levels of the circulating biomarkers. Inclusion of CYFRA (55) together with the biomarkers reported here increased the diagnostic performance. In addition, keratin fragments detected by the CYFRA assay are released from the cancer cells (56), which is likely to eliminate the potential confounding effect of inflammation. Results from the National Lung Screening Trial of over 54,000 current and former smokers using low dose spiral computer tomography (LD CT) demonstrate that LC is detectable at early asymptomatic stages (3), leading to a 20% reduction in cancer deaths (4). However, LD CT suffers from low specificity because roughly 95% of the patients with solitary pulmonary nodules detected by LD CT actually did not have detectable cancer. To reduce unnecessary interventions because of overdiagnosis, there is an imminent need for a clinical test that provides the required specificity in conjunction with LD CT screening. Although the panel identified in this study needs further validation on a larger clinical cohort and may not be suitable as a stand-alone test, its combination with CYFRA may provide sufficient specificity to complement imaging in routine use and population screening for lung cancer.

In conclusion, our hypothesis-free approach of screening the plasma proteome with hybridoma libraries not only delivered results consistent with previously reported lung cancer biomarkers but also uncovered new ones as well. Increasing

the size of the initial libraries as well as the sensitivity of the screening assay is expected to contribute to the number of identified biomarkers because it is unlikely that the majority of specific lung cancer markers are among the high and medium abundant proteins. It is necessary to point out here that the screening steps performed in the study were designed to separate cancer from control cases. Other monoclonal antibodies that would detect not only the presence of the disease but the histological subtype as well can be identified by generating and screening the libraries with alternative hypotheses. The observed redundancy of antibodies generated against the same biomarkers also suggested that the global number of lung cancer-specific biomarker epitopes may be higher than reported here. Further characterization of the specificity of the LC-associated mAbs toward potential lung cancer-specific forms of the identified biomarker proteins should provide more insight into the biology of the disease. In principle, the development of antibodies for diagnostics is a difficult process with several limitations. One of the major strengths of our approach is that it inherently couples biomarker discovery with *in vitro* diagnostic development. Because the discovery phase is already driven by mAbs, translation, including the generation of reagents suitable for the development of specific *in vitro* diagnostic grade immunoassays, could be fast and effective.

A. G., B. L. K., M. G. K., I. K., W. H., N. T., C. M. B., A. J., Y. K., and L. T. hold shares and warrants of Biosystems International SAS. M. G. K., I. K., W. H., N. T., J. K., C. M. B., A. J., Y. K., and L. T. are employees of Biosystems International.

L. T. conceived the experiments; M. G. K., B. L. K., and L. T. designed the experiments; I. K. and A. G. contributed to data interpretation of immunologic and separation experiments, respectively; W. H., I. K., and J. K. contributed to immunization, antibody production and characterization; A. J. contributed to specific experimental design and execution of high throughput hybridoma generation and cloning; N. T. contributed to specific experimental design and execution of high throughput screening; C. M. B. and M. H. contributed to specific experimental design and execution of sample and immunogen preparation; C. M. B. contributed to specific experimental design and execution of library screening for antibody partners; Y. K. and M. G. K. contributed to specific design and execution of statistical analysis; E. C. contributed to specific clinical protocol design and execution of sample collection; B. D. executed and analyzed the IHC experiments; and M. G. K. and L. T. wrote the manuscript.

REFERENCES

1. Stewart, B. W., and Kleihues, P. (Eds): World Cancer Report. IARCPress. Lyon 2003, 181–188
2. Henschke, C. I., Yankelevitz, D. F., Libby, D. M., Pasmantier, M. W., Smith, J. P., and Miettinen, O. S. (2006) Survival of patients with stage I lung cancer detected on CT screening. *N. Engl. J. Med.* **355,** 1763–1771
3. van Klaveren, R. J., Oudkerk, M., Prokop, M., Scholten, E. T., Nackaerts, K., Vernhout, R., van Iersel, C. A., van den Bergh, K. A., van 't Westeinde, S., van der Aalst, C., Thunnissen, E., Xu, D. M., Wang, Y., Zhao, Y., Gietema, H. A., de Hoop, B. J., Groen, H. J., de Bock, G. H., van Ooijen, P., Weenink, C., Verschakelen, J., Lammers, J. W., Timens, W., Willebrand, D., Vink, A., Mali, W., and de Koning, H. J. (2009) Management of lung nodules detected by volume CT scanning. *N. Engl. J. Med.* **361,** 2221–2229
4. Aberle, D. R., Adams, A. M., Berg, C. D., Black, W. C., Clapp, J. D., Fagerstrom, R. M., Gareen, I. F., Gatsonis, C., Marcus, P. M., and Sicks, J. D. (2011) Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening. *N. Engl. J. Med.* **365,** 395–409
5. Kulpa, J., Wójcik, E., Reinfuss, M., and Kołodziejski, L. (2002) Carcinoembryonic antigen, squamous cell carcinoma antigen, CYFRA 21-1, and neuron-specific enolase in squamous cell lung cancer patients. *Clin. Chem.* **48,** 1931–1937
6. Ebert, W., Hoppe, M., Muley, T., and Drings, P. (1997) Monitoring of therapy in inoperable lung cancer patients by measurement of CYFRA 21–1, TPA-TP CEA, and NSE. *Anticancer Res.* **17,** 2875–2878
7. Beane, J., Spira, A., and Lenburg, M. E. (2009) Clinical impact of high-throughput gene expression studies in lung cancer. *J. Thorac. Oncol.* **4,** 109–118
8. Ocak, S., Chaurand, P., and Massion, P. P. (2009) Mass spectrometry-based proteomic profiling of lung cancer. *Proc. Am. Thorac. Soc.* **6,** 159–170
9. Zhong, L., Coe, S. P., Stromberg, A. J., Khattar, N. H., Jett, J. R., and Hirschowitz, E. A. (2006) Profiling tumor-associated antibodies for early detection of non-small cell lung cancer. *J. Thorac. Oncol.* **1,** 513–519
10. Brower, V. (2009) Biomarker studies abound for early detection of lung cancer. *J. Natl. Cancer Inst.* **101,** 11–13
11. Rifai, N., Gillette, M. A., and Carr, S. A. (2006) Protein biomarker discovery and validation: The long and uncertain path to clinical utility. *Nat. Biotechnol.* **24,** 971–983
12. Uhlén, M., Björling, E., Agaton, C., Szigyarto, C. A., Amini, B., Andersen, E., Andersson, A. C., Angelidou, P., Asplund, A., Asplund, C., Berglund, L., Bergström, K., Brumer, H., Cerjan, D., Ekström, M., Elobeid, A., Eriksson, C., Fagerberg, L., Falk, R., Fall, J., Forsberg, M., Björklund, M. G., Gumbel, K., Halimi, A., Hallin, I., Hamsten, C., Hansson, M., Hedhammar, M., Hercules, G., Kampf, C., Larsson, K., Lindskog, M., Lodewyckx, W., Lund, J., Lundeberg, J., Magnusson, K., Malm, E., Nilsson, P., Odling, J., Oksvold, P., Olsson, I., Oster, E., Ottosson, J., Paavilainen, L., Persson, A., Rimini, R., Rockberg, J., Runeson, M., Sivertsson, A., Sköllermo, A., Steen, J., Stenvall, M., Sterky, F., Strömberg, S., Sundberg, M., Tegel, H., Tourle, S., Wahlund, E., Waldén, A., Wan, J., Wernérus, H., Westberg, J., Wester, K., Wrethagen, U., Xu, L. L., Hober, S., and Pontén, F. (2005) A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol. Cell Proteomics* **4,** 1920–1932
13. Blow, N. (2007) Antibodies: The generation game. *Nature* **447,** 741–744
14. Hoagland, L. F., 4th, Campa, M. J., Gottlin, E. B., Herndon, J. E., 2nd, and Patz, E. F., Jr. (2007) Haptoglobin and posttranslational glycan-modified derivatives as serum biomarkers for the diagnosis of nonsmall cell lung

cancer. *Cancer* **110,** 2260–2268

15. Okano, T., Kondo, T., Kakisaka, T., Fujii, K., Yamada, M., Kato, H., Nishimura, T., Gemma, A., Kudoh, S., and Hirohashi, S. (2006) Plasma proteomics of lung cancer by a linkage of multi-dimensional liquid chromatography and two-dimensional difference gel electrophoresis. *Proteomics* **6,** 3938–3948

16. Higashiyama, M., Doi, O., Yokouchi, H., Kodama, K., Nakamori, S., and Tateishi, R. (1995) Alpha-1-antichymotrypsin expression in lung adenocarcinoma and its possible association with tumor progression. *Cancer* **76,** 1368–1376

17. Ajona, D., Castaño, Z., Garayoa, M., Zudaire, E., Pajares, M. J., Martinez, A., Cuttitta, F., Montuenga, L. M., and Pio, R. (2004) Expression of complement factor H by lung cancer cells: Effects on the activation of the alternative pathway of complement. *Cancer Res.* **64,** 6310–6318

18. Pujol, J. L., Grenier, J., Daurès, J. P., Daver, A., Pujol, H., and Michel, F. B. (1993) Serum fragment of cytokeratin subunit 19 measured by CYFRA 21–1 immunoradiometric assay as a marker of lung cancer. *Cancer Res.* **53,** 61–66

19. The World Health Organization (1982) Histological typing of lung tumours. *Neoplasma* **29,** 111–123

20. Kullolli, M., Hancock, W. S., and Hincapie, M. (2008) Preparation of a high-performance multi-lectin affinity chromatography (HP-M-LAC) adsorbent for the analysis of human plasma glycoproteins. *J. Sep. Sci.* **31,** 2733–2739

21. Takacs, L., Guttman, A., and Kuras, M. (2007) *Normalization of complex analyte mixtures*, Published Patent Application, PCT/IB2006/003161

22. Pontecorvo, G. (1975) Production of mammalian somatic cell hybrids by means of polyethylene glycol treatment. *Somatic Cell Genet.* **1,** 397–400

23. Zhang, J. H., Chung, T. D., and Oldenburg, K. R. (1999) A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J. Biomol. Screen* **4,** 67–73

24. Weivoda, S., Andersen, J. D., Skogen, A., Schlievert, P. M., Fontana, D., Schacker, T., Tuite, P., Dubinsky, J. M., and Jemmerson, R. (2008) ELISA for human serum leucine-rich alpha-2-glycoprotein-1 employing cytochrome c as the capturing ligand. *J. Immunol. Methods* **336,** 22–29

25. Gogolak, P., Rethi, B., Szatmari, I., Lanyi, A., Dezso, B., Nagy, L., and Rajnavolgyi, E. (2007) Differentiation of CD1a- and CD1a+ monocyte-derived dendritic cells is biased by lipid environment and PPARgamma. *Blood* **109,** 643–652

26. Tsakiris, I., Soos, G., Nemes, Z., Kiss, S. S., Andras, C., Szantó, J., and Dezso, B. (2008) The presence of carboxypeptidase-M in tumour cells signifies epidermal growth factor receptor expression in lung adenocarcinomas: the coexistence predicts a poor prognosis regardless of EGFR levels. *J. Cancer Res. Clin. Oncol.* **134,** 439–451

27. Csanky, E., Olivova, P., Rajnavolgyi, E., Hempel, W., Tardieu, N., Katalin, E. T., Jullien, A., Malderez-Bloes, C., Kuras, M., Duval, M. X., Nagy, L., Scholtz, B., Hancock, W., Karger, B., Guttman, A., and Takacs, L. (2007) Monoclonal antibody proteomics: Discovery and prevalidation of chronic obstructive pulmonary disease biomarkers in a single step. *Electrophoresis* **28,** 4401–4406

28. Peracaula, R., Barrabés, S., Sarrats, A., Rudd, P. M., and de Llorens, R. (2008) Altered glycosylation in tumours focused to cancer diagnosis. *Dis. Markers* **25,** 207–218

29. Kingsmore, S. F. (2006) Multiplexed protein measurement: Technologies and applications of protein and antibody arrays. *Nat. Rev. Drug Discov.* **5,** 310–320

30. Wang, D., Hincapie, M., Guergova-Kuras, M., Kadas, J., Takacs, L., and Karger, B. L. (2010) Antigen identification and characterization of lung cancer specific monoclonal antibodies produced by mAb proteomics. *J. Proteome Res.* **9,** 1834–1842

31. Abdullah, M., Schultz, H., Kähler, D., Branscheid, D., Dalhoff, K., Zabel, P., Vollmer, E., and Goldmann, T. (2009) Expression of the acute phase protein haptoglobin in human lung cancer and tumor-free lung tissues. *Pathol. Res. Pract.* **205,** 639–647

32. Molina, R., Filella, X., Augé, J. M., Fuentes, R., Bover, I., Rifa, J., Moreno, V., Canals, E., Viñolas, N., Marquez, A., Barreiro, E., Borras, J., and Viladiu, P. (2003) Tumor markers (CEA, CA 125, CYFRA 21–1, SCC and NSE) in patients with non-small cell lung cancer as an aid in histological diagnosis and prognosis: Comparison with the main clinical and pathological prognostic factors. *Tumour Biol* **24,** 209–218

33. Ostroff, R. M., Bigbee, W. L., Franklin, W., Gold, L., Mehan, M., Miller, Y. E., Pass, H. I., Rom, W. N., Siegfried, J. M., Stewart, A., Walker, J. J., Weissfeld, J. L., Williams, S., Zichi, D., and Brody, E. N. (2010) Unlocking biomarker discovery: Large scale application of aptamer proteomic technology for early detection of lung cancer. *PLoS One* **5,** e15003

34. Gabay, C., and Kushner, I. (1999) Acute-phase proteins and other systemic responses to inflammation. *N. Engl. J. Med.* **340,** 448–454

35. Shirai, R., Hirano, F., Ohkura, N., Ikeda, K., and Inoue, S. (2009) Up-regulation of the expression of leucine-rich alpha(2)-glycoprotein in hepatocytes by the mediators of acute-phase response. *Biochem. Biophys. Res. Commun.* **382,** 776–779

36. Okuyama, N., Ide, Y., Nakano, M., Nakagawa, T., Yamanaka, K., Moriwaki, K., Murata, K., Ohigashi, H., Yokoyama, S., Eguchi, H., Ishikawa, O., Ito, T., Kato, M., Kasahara, A., Kawano, S., Gu, J., Taniguchi, N., and Miyoshi, E. (2006) Fucosylated haptoglobin is a novel marker for pancreatic cancer: A detailed analysis of the oligosaccharide structure and a possible mechanism for fucosylation. *Int. J. Cancer* **118,** 2803–2808

37. Yoon, S. J., Park, S. Y., Pang, P. C., Gallagher, J., Gottesman, J. E., Dell, A., Kim, J. H., and Hakomori, S. I. (2010) *N*-Glycosylation status of beta-haptoglobin in sera of patients with prostate cancer vs. benign prostate diseases. *Int. J. Oncol.* **36,** 193–203

38. Bharti, A., Ma, P. C., Maulik, G., Singh, R., Khan, E., Skarin, A. T., and Salgia, R. (2004) Haptoglobin alpha-subunit and hepatocyte growth factor can potentially serve as serum tumor biomarkers in small cell lung cancer. *Anticancer Res.* **24,** 1031–1038

39. Cheng, Z. Z., Corey, M. J., Pärepalo, M., Majno, S., Hellwage, J., Zipfel, P. F., Kinders, R. J., Raitanen, M., Meri, S., and Jokiranta, T. S. (2005) Complement factor H as a marker for detection of bladder cancer. *Clin. Chem.* **51,** 856–863

40. Junnikkala, S., Hakulinen, J., Jarva, H., Manuelian, T., Bjørge, L., Bützow, R., Zipfel, P. F., and Meri, S. (2002) Secretion of soluble complement inhibitors factor H and factor H-like protein (FHL-1) by ovarian tumour cells. *Br. J. Cancer* **87,** 1119–1127

41. Amornsiripanitch, N., Hong, S., Campa, M. J., Frank, M. M., Gottlin, E. B., and Patz, E. F., Jr. (2010) Complement factor H autoantibodies are associated with early stage NSCLC. *Clin. Cancer Res.* **16,** 3226–3231

42. Yu, K. H., Rustgi, A. K., and Blair, I. A. (2005) Characterization of proteins in human pancreatic cancer serum using differential gel electrophoresis and tandem mass spectrometry. *J. Proteome Res.* **4,** 1742–1751

43. Koomen, J. M., Shih, L. N., Coombes, K. R., Li, D., Xiao, L. C., Fidler, I. J., Abbruzzese, J. L., and Kobayashi, R. (2005) Plasma protein profiling for diagnosis of pancreatic cancer reveals the presence of host response proteins. *Clin. Cancer Res.* **11,** 1110–1118

44. Chen, Y., Lim, B. K., and Hashim, O. H. (2009) Different altered stage correlative expression of high abundance acute-phase proteins in sera of patients with epithelial ovarian carcinoma. *J. Hematol. Oncol.* **2,** 37–43

45. Andersen, J. D., Boylan, K. L., Jemmerson, R., Geller, M. A., Misemer, B., Harrington, K. M., Weivoda, S., Witthuhn, B. A., Argenta, P., Vogel, R. I., and Skubitz, A. P. (2010) Leucine-rich alpha-2-glycoprotein-1 is upregulated in sera and tumors of ovarian cancer patients. *J. Ovarian Res.* **3,** 21–34

46. Heo, S. H., Lee, S. J., Ryoo, H. M., Park, J. Y., and Cho, J. Y. (2007) Identification of putative serum glycoprotein biomarkers for human lung adenocarcinoma by multilectin affinity chromatography and LC-MS/MS. *Proteomics* **7,** 4292–4302

47. Gao, W. M., Kuick, R., Orchekowski, R. P., Misek, D. E., Qiu, J., Greenberg, A. K., Rom, W. N., Brenner, D. E., Omenn, G. S., Haab, B. B., and Hanash, S. M. (2005) Distinctive serum protein profiles involving abundant proteins in lung cancer patients based upon antibody microarray analysis. *BMC Cancer* **5,** 110–119

48. Siemes, C., Visser, L. E., Coebergh, J. W., Splinter, T. A., Witteman, J. C., Uiterlinden, A. G., Hofman, A., Pols, H. A., and Stricker, B. H. (2006) C-reactive protein levels, variation in the C-reactive protein gene, and cancer risk: the Rotterdam Study. *J. Clin. Oncol.* **24,** 5216–5222

49. Trichopoulos, D., Psaltopoulou, T., Orfanos, P., Trichopoulou, A., and Boffetta, P. (2006) Plasma C-reactive protein and risk of cancer: A prospective study from Greece. *Cancer Epidemiol. Biomarkers Prev.* **15,** 381–384

50. Proctor, M. J., Morrison, D. S., Talwar, D., Balmer, S. M., O'Reilly, D. S., Foulis, A. K., Horgan, P. G., and McMillan, D. C. (2011) An inflammation-based prognostic score (mGPS) predicts cancer survival independent of tumour site: A Glasgow Inflammation Outcome Study. *Br. J. Cancer* **104,**

726–734

51. Proctor, M. J., Talwar, D., Balmar, S. M., O'Reilly, D. S., Foulis, A. K., Horgan, P. G., Morrison, D. S., and McMillan, D. C. (2010) The relationship between the presence and site of cancer, an inflammation-based prognostic score and biochemical parameters: Initial results of the Glasgow Inflammation Outcome Study. *Br. J. Cancer* **103,** 870–876

52. Kiri, V. A., Fabbri, L. M., Davis, K. J., and Soriano, J. B. (2009) Inhaled corticosteroids and risk of lung cancer among COPD patients who quit smoking. *Respir. Med.* **103,** 85–90

53. Engels, E. A., Wu, X., Gu, J., Dong, Q., Liu, J., and Spitz, M. R. (2007) Systematic evaluation of genetic variants in the inflammation pathway and risk of lung cancer. *Cancer Res.* **67,** 6520–6527

54. Chen, C. B., Su, Y. C., Huang, T. T., Ho, H. C., Chang, Y. T., Tung, Y. T., and Lee, W. C. (2008) Differentially expressed serum haptoglobin alpha chain isoforms with potential application for diagnosis of head and neck cancer. *Clin. Chim. Acta* **398,** 48–52

55. Wieskopf, B., Demangeat, C., Purohit, A., Stenger, R., Gries, P., Kreisman, H., and Quoix, E. (1995) Cyfra 21–1 as a biologic marker of non-small cell lung cancer: Evaluation of sensitivity, specificity, and prognostic role. *Chest* **108,** 163–169

56. Dohmoto, K., Hojo, S., Fujita, J., Yang, Y., Ueda, Y., Bandoh, S., Yamaji, Y., Ohtsuki, Y., Dobashi, N., Ishida, T., and Takahara, J. (2001) The role of caspase 3 in producing cytokeratin 19 fragment (CYFRA21–1) in human lung cancer cell lines. *Int. J. Cancer* **91,** 468–473