Analysis of acute-phase proteins, AHSG, C3, CLI, HP and SAA, reveals distinctive expression patterns associated with breast, colorectal and lung cancer

Paul Dowling[1], Colin Clarke[1], Kim Hennessy[1], Beatriz Torralbo-Lopez[2], Jo Ballot[3], John Crown[3], Ingrid Kiernan[4], Kenneth J. O'Byrne[4], M. John Kennedy[5], Vincent Lynch[1] and Martin Clynes[1]

[1] National Institute for Cellular Biotechnology, Dublin City University, Dublin 9, Ireland
[2] The All-Ireland Cooperative Oncology Research Group (ICORG), Dublin 2, Ireland
[3] St Vincent's University Hospital, Dublin 4, Ireland/Molecular Therapeutics for Cancer Ireland (MTCI)
[4] Department of Clinical Medicine, Translational Cancer Research Group, Institute of Molecular Medicine, Trinity Centre for Health Sciences, St. James' Hospital, Dublin, Ireland
[5] Academic Unit of Clinical and Molecular Oncology St James Hospital and the University of Dublin, Trinity College, Ireland

Early detection, clinical management and disease recurrence monitoring are critical areas in cancer treatment in which specific biomarker panels are likely to be very important in each of these key areas. We have previously demonstrated that levels of alpha-2-heremans-schmid-glycoprotein (AHSG), complement component C3 (C3), clusterin (CLI), haptoglobin (HP) and serum amyloid A (SAA) are significantly altered in serum from patients with squamous cell carcinoma of the lung. Here, we report the abundance levels for these proteins in serum samples from patients with advanced breast cancer, colorectal cancer (CRC) and lung cancer compared to healthy controls (age and gender matched) using commercially available enzyme-linked immunosorbent assay kits. Logistic regression (LR) models were fitted to the resulting data, and the classification ability of the proteins was evaluated using receiver-operating characteristic curve and leave-one-out cross-validation (LOOCV). The most accurate individual candidate biomarkers were C3 for breast cancer [area under the curve (AUC) = 0.89, LOOCV = 73%], CLI for CRC (AUC = 0.98, LOOCV = 90%), HP for small cell lung carcinoma (AUC = 0.97, LOOCV = 88%), C3 for lung adenocarcinoma (AUC = 0.94, LOOCV = 89%) and HP for squamous cell carcinoma of the lung (AUC = 0.94, LOOCV = 87%). The best dual combination of biomarkers using LR analysis were found to be AHSG + C3 (AUC = 0.91, LOOCV = 83%) for breast cancer, CLI + HP (AUC = 0.98, LOOCV = 92%) for CRC, C3 + SAA (AUC = 0.97, LOOCV = 91%) for small cell lung carcinoma and HP + SAA for both adenocarcinoma (AUC = 0.98, LOOCV = 96%) and squamous cell carcinoma of the lung (AUC = 0.98, LOOCV = 84%). The high AUC values reported here indicated that these candidate biomarkers have the potential to discriminate accurately between control and cancer groups both individually and in combination with other proteins.

Despite the recent good news that cancer incidence and death rates for men and women continue to decline in developed countries, cancer is projected to become the leading cause of death worldwide this year.[1] For women, breast cancer is the most common, with lung cancer second and colorectal cancer (CRC) third.[2] For men, prostate cancer is the most prevalent form of cancer, followed by lung cancer and then CRC.[3]

The number of candidate biomarkers, which can be used for diagnosis, early detection or treatment monitoring, continues to grow rapidly; however, few of these reach the clinic mainly because of low levels of sensitivity and specificity.[4] Despite advances in new technologies, detection and routine measurement of low abundant tumor-specific proteins remain a significant challenge. The detection limits of modern mass spectrometers and conventional immunoassays are now, however approaching the relevant levels in biofluids such as serum, although it is likely that some early stage disease-specific biomarkers may still lie beyond current limits.[5] Coupled with the added difficulties of high-abundance resident proteins such as albumin masking the low abundant tumor-specific biomarkers and the fact that some biomarkers are rapidly degraded, and identification of low level proteins remains difficult.[6] Much of the focus on sample preparation for biomarker discovery experiments has focused on the removal of highly abundant proteins such as albumin and haptoglobin, a strategy that results in relative enrichment of

medium to low abundant potential biomarkers. Within the high abundant protein group reside the acute-phase proteins (APPs). APPs generally have a nonspecific rapid response to such processes as inflammation/infections, tissue damage, surgery, myocardial infarction or the presence of tumors. The relationship between the APPs and cancer has been well documented in the literature with numerous investigations reporting on altered levels of various APPs with different types of cancers and evidence that many APPs are actually produced directly by tumor tissue.[7,8]

It has been suggested that APPs were not likely to be specific for any type of cancer and would be expected to be elevated in all malignancies and in inflammatory diseases. In addition, high-abundance proteins such as the APPs were thought unlikely to be tumor-derived and thus to represent cancer epiphenomena rather than direct tumor-derived proteins. Recent progress in proteomics studies, however, which profiled the serum proteins of patients with cancer and those of normal individuals, indicated that the altered expression of APPs was different for distinct types, subtypes, and even stages of cancer.[9,10] It is likely that panels of biomarkers in the future will be comprised of biomarkers that reflect tumor-specific proteins together with proteins from the tumor microenvironment.

Our study focused on the measurement of five highly abundant serum proteins from the APP class, which we have previously found to be altered in serum from patients with squamous cell carcinoma of the lung,[11] in advanced breast, CRC and lung cancer serum. Specifically, the levels of alpha-2-heremans-schmid-glycoprotein (AHSG), complement component C3 (C3), clusterin (CLI), haptoglobin (HP) and serum amyloid A (SAA) were measured.

Material and Methods
Patients and sample collection
Samples were collected through a collaborative project involving participating sites (St Vincent's University Hospital, Dublin 4, Ireland and St. James's Hospital, Dublin 8, Ireland), coordinated through ICORG (the All Ireland Co-operative Oncology Research Group, www.icorg.ie). The samples were collected according to standard phlebotomy procedures from consented patients. Ethical consent was granted from each of the respective Hospital Ethics Committees. A total of 10 ml of blood was collected into additive free (serum) blood tubes and was allowed to clot for 30 min to 1 hr at room temperature. The serum was denuded by pipette from the clot and poured into a clean tube. The tubes were centrifuged at 400g for 30 min at 4°C. Serum was aliquoted in the cryovial tubes, labeled and stored at −80°C until time of analysis. The time from sample procurement to storage at −80°C was less than 3 hr. Each serum sample underwent not more than three freeze/thaw cycles prior to analysis.

Enzyme-linked immunosorbent assay
Enzyme-linked immunosorbent assay (ELISA)-based analysis was carried out using raw unfractionated serum samples.

Each sample was analyzed in triplicate using the following commercially available kits, for the measurement of serum haptoglobin (AssayPro), CLI (BioVendor R&D), complement C3 (AssayPro), SAA (Invitrogen) and alpha HS glycoprotein (AssayPro) kits were used. The ELISA assays were performed according to each manufacturer's protocol and guidelines. The haptoglobin and complement C3 assays use a quantitative and competitive enzyme immunoassay technique that takes 2–3 hr to perform. The assays for alpha HS glycoprotein, CLI and SAA use a quantitative sandwich enzyme immunoassay technique that takes 3–4 hr to perform. The optical density (OD) was measured using a microplate reader (Bio-Tek), and the concentration of each protein in the serum samples was determined by comparing the OD of the samples against the respective standard curve.

Statistical analysis
Box-and-whisker plots were generated through Microsoft Excel using the values from the ELISA data. The plots display a statistical summary including the median, quartiles and range. Student t-test was used to identify statistically significant changes in abundance levels for specific proteins between patient with cancer and control serum samples. Receiver-operating characteristic (ROC) curve analysis was performed as it is a useful tool in assessment of biomarker accuracy.[12,13] The ROC plots were obtained by plotting all sensitivity values (true positive fraction) on the y-axis against their equivalent (1-specificity) values (false positive fraction) for all available thresholds on the x-axis (MedCalc for Windows 8.1.1.0, Medcalc Software, Mariakerke, Belgium). The area under the curve (AUC) was calculated to provide a summary of overall classifier effectiveness. In our study, we consider AUC values ranging from 0.5→0.7 as poor, 0.7→0.8 as average, 0.8→0.9 as good and >0.9 as outstanding.

For multivariate analysis of biomarker combinations, logistic regression (LR) analysis of the serum biomarker levels in these patients groups was performed. As an additional measure of the potential of these biomarkers to distinguish between control and cancer samples, a commonly used internal validation technique known as leave-one-out cross-validation (LOOCV) was performed.[14,15] During the LOOCV procedure data from a single observation is removed from the dataset, and the remaining samples are then utilized to construct a LR model. The "test" sample is presented to the trained model and the performance assessed, LOOCV continues until each observation is designated as the "test." Thus, the LOOCV provides a more conservative estimate of model performance. LR, LR ROC curves and LOOCV evaluation were carried out in the open source R statistical computing environment (http://www.r-project.org/).

Results
Clinical data
Table 1 shows that the mean age of the female control group ($n = 15$) was 54 SD ± 8 years (range 35–81 years) and the

Table 1. Patient data table

	Control (m)	Control(f)	Breast	CRC	Small cell	Adeno	Squamous
n	15	15	33	32(17 male/ 15 female)	12(7 male/ 5 female)	25(14 male/ 11 female)	18(10 male/ 8 female)
	Age 59±6	54±8	57±13	64±10	62±11	65±10	65±7
			Stage IIIC/IV	IIIB/IIIC/IV	IIIB/IV	IIIB/IV	IIIB/IV

Patient data table including number of samples, average age plus/minus standard deviation and disease staging for controls, breast cancer, CRC and lung cancer (small cell lung carcinoma, lung adenocarcinoma and squamous cell carcinoma of the lung).
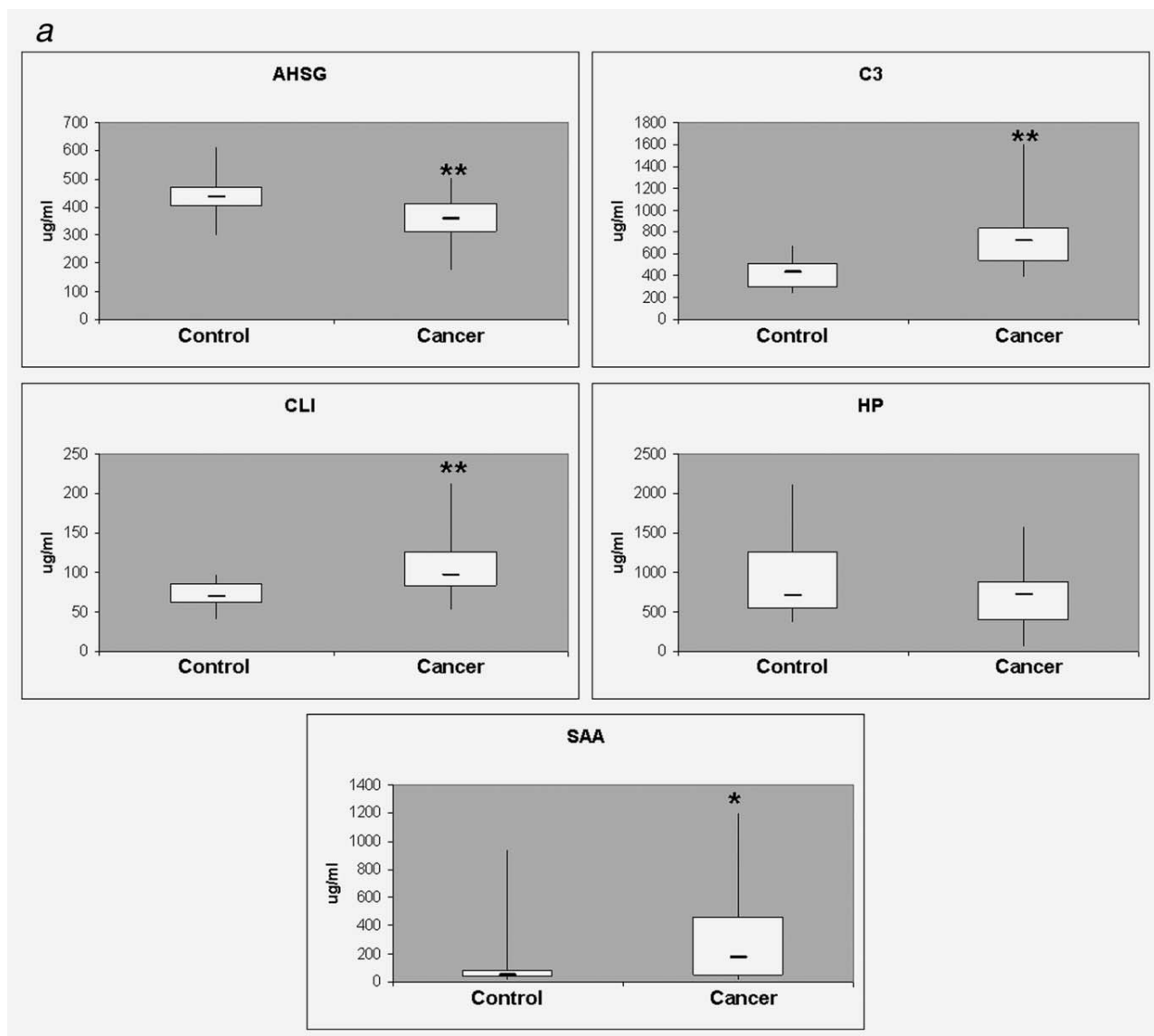

Figure 1. (*a,b*) Box-and-Whisker plots and Bar charts comparing AHSG, C3, CLI, HP and SAA levels in Breast cancer serum samples compared to a control group. Box-and-Whisker plots represent data with boxes ranging from the 25th to the 75th percentile of the observed distribution of values. The bold line of the box displays the median value for serum levels of AHSG, C3, CLI, HP and SAA in breast cancer compared to controls. Whiskers span minimum to maximum observed values. The bar charts display the mean and standard deviation for AHSG, C3, CLI, HP and SAA levels in breast cancer serum samples compared to controls. (*$p \leq 0.05$: significant; **$p \leq 0.01$: very significant).
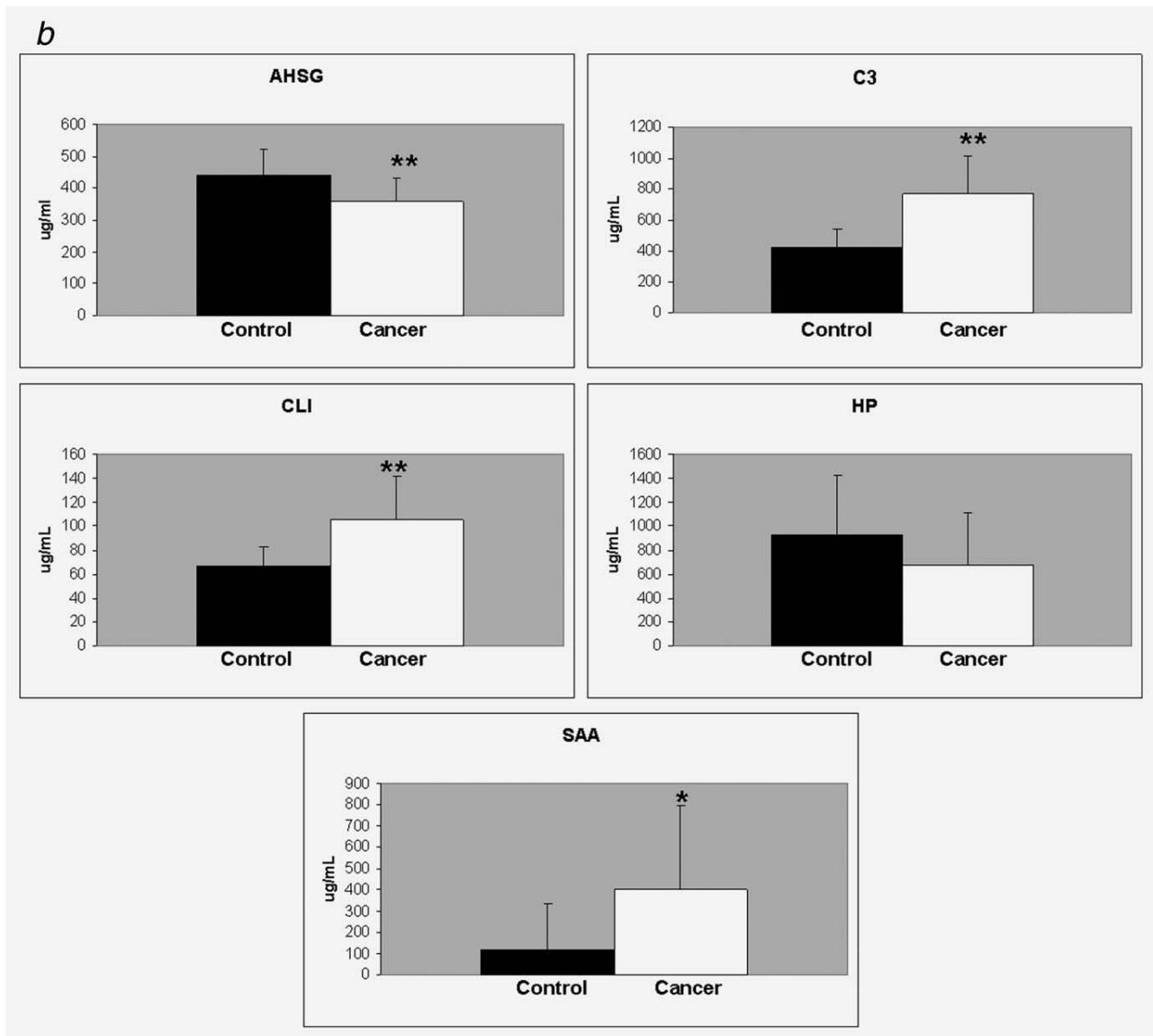
Early Detection and Diagnosis

b

Figure 1. Continued

mean age of the male control group ($n = 15$) was 59 SD $\pm$ 6 years (range 38–78 years). In the cancer groups, the mean age was 57 SD $\pm$ 13 years (range 31–79 years) for patients with breast cancer ($n = 33$), 64 SD $\pm$ 10 years (range 41–82 years) for patients with CRC ($n = 31$, 17 males/15 females), 62 SD $\pm$ 11 years (range 30–79 years) for patients with small cell lung cancer ($n = 12$, 7 males/5 females), 65 SD $\pm$ 10 years (range 33–84 years) for patients with lung adenocarcinoma ($n = 25$, 14 males/11 females) and 65 SD $\pm$ 7 years (range 35–81 years) for patients with lung squamous cell carcinoma ($n = 18$, 10 males/8 females). All the patients with breast cancer were stages IIIC or IV, the patients with CRC were stages IIIB, IIIC or IV and the patients with lung cancer were all stages IIIB or IV.

The data presented here on the control groups for CRC and the three subtypes of lung cancer are comprised of equal number of both male ($n = 15$) and female ($n = 15$) samples, with the breast cancer control group being comprised only of female samples ($n = 15$). Male and female control groups were also analyzed individually to identify if any of the candidate biomarkers are associated with one gender type over the other; however, no statistically significant changes were found in serum samples from male or female as individual groups in any of the comparisons. Therefore, all data presented here for control groups is a mixture of both male and female serum samples with the exception of the breast cancer group.

In addition, within the breast cancer group, human epidermal growth factor receptor 2 (HER2) and estrogen

Early Detection and Diagnosis
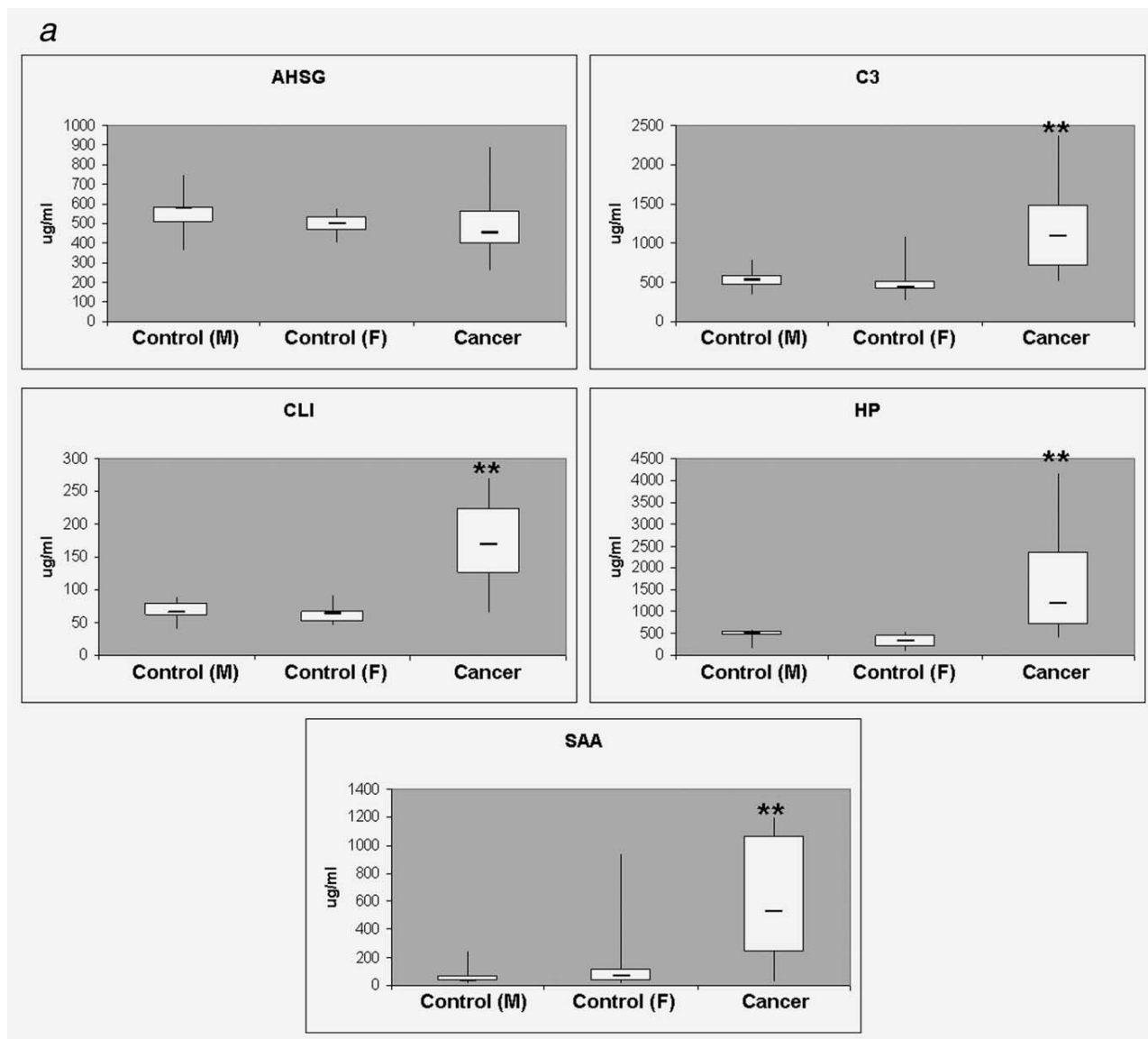
Figure 2. (*a,b*) Box-and-Whisker plots and Bar charts comparing AHSG, C3, CLI, HP and SAA levels in CRC serum samples compared to control groups. Box-and-Whisker plots represent data with boxes ranging from the 25th to the 75th percentile of the observed distribution of values. The bold line of the box displays the median value for serum levels of AHSG, C3, CLI, HP and SAA in CRC compared to controls. Whiskers span minimum to maximum observed values. The bar charts display the mean and standard deviation for AHSG, C3, CLI, HP and SAA levels in CRC serum samples compared to controls. (*$p \leq 0.05$: significant; **$p \leq 0.01$: very significant).

receptor (ER) status was available to us (from patient pathology reports) across the breast cancer cohort to identify if any specific biomarkers are associated with these molecular characteristics. Neither HER2 ($n = 19$: 4 HER2-positive/ 15 HER2-negative; AHSG $p = 0.12$, C3 $p = 0.57$, CLI $p = 0.73$, HP $p = 0.63$, SAA $p = 0.23$) nor ER ($n = 33$: 21 ER-positive/12 ER-negative; AHSG $p = 0.42$, C3 $p = 0.46$, CLI $p = 0.29$, HP $p = 0.46$, SAA $p = 0.77$) status were statistically significant in their association with any unique changes in abundance levels for the biomarkers discussed here.

Expression of biomarker panel in breast cancer
To assess the levels of AHSG, C3, CLI, HP and SAA in control and cancer samples, we assayed serum samples using ELISAs specific to each of these proteins for 33 patients with advanced stage breast cancer and compared the results to a control group. Significant increases in abundance levels for C3, CLI and SAA in breast cancer sera compared to a control group were observed. AHSG is shown to be significantly decreased in cancer samples (Fig. 1*a* & 1*b*). Box plot analysis (Fig. 1*a*) showed that the median level of AHSG was significantly increased in control *versus* cancer samples (436.0 µg/
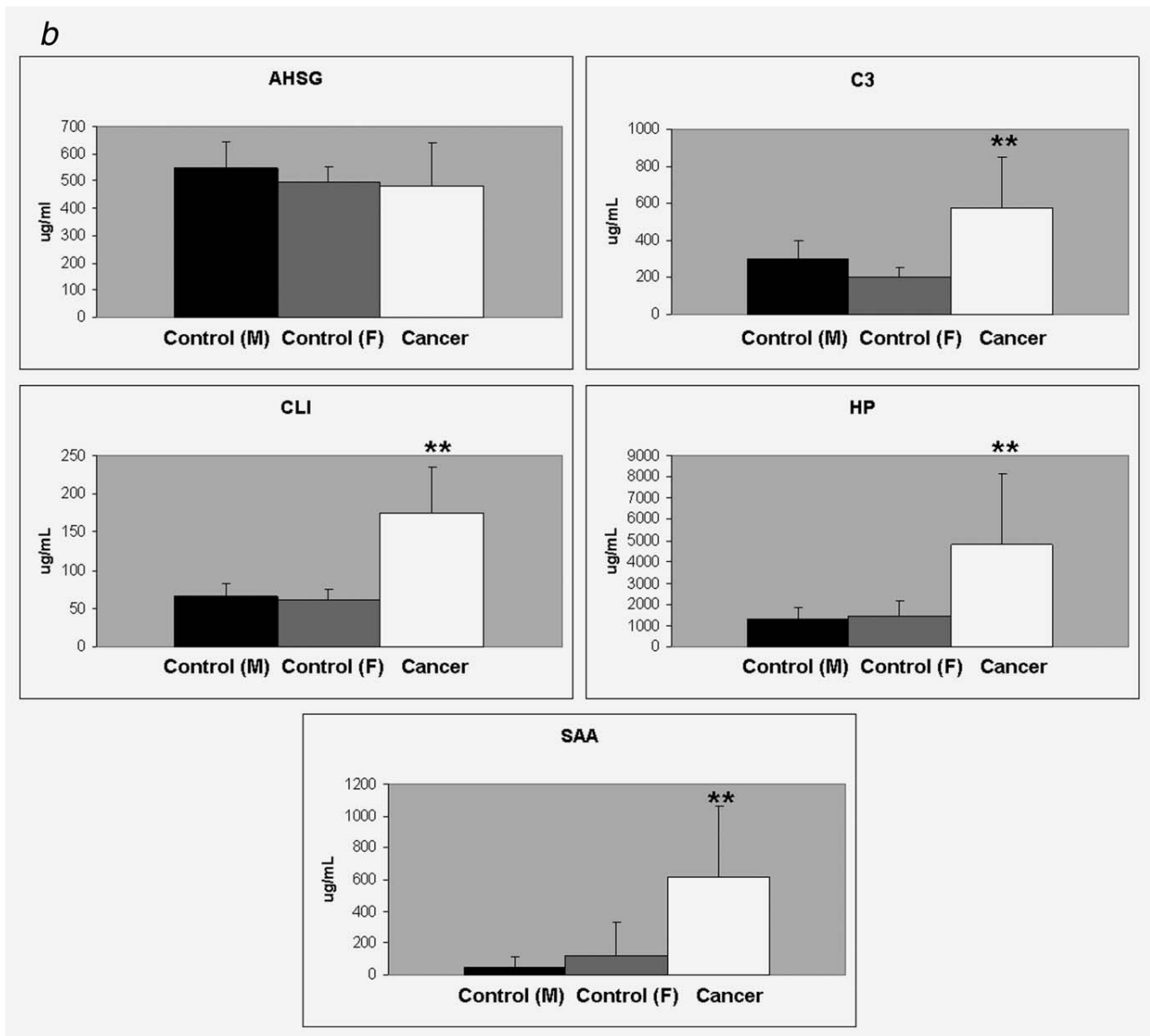
b

Figure 2. Continued

ml *vs.* 356.5 μg/ml, respectively; $p \leq 0.01$). The median serum levels in normal compared to cancer samples for C3 (431.8 μg/ml *vs.* 723.8 μg/ml, respectively; $p \leq 0.01$) and CLI (69.2 μg/ml *vs.* 96.8 μg/ml, respectively; $p \leq 0.01$) were also found to be very significantly different, with a clear increase in the cancer samples. Box-and-Whisker analysis of the ELISA data for SAA also demonstrated that the level of this protein was significantly elevated in serum from patients with cancer compared to controls (45.4 μg/ml *vs.* 172.6 μg/ml, respectively; $p \leq 0.05$).

Expression of biomarker panel in CRC

The CRC ELISA bar charts demonstrate significant increases in abundance levels for C3, CLI, HP and SAA in CRC sera compared to a control group. AHSG was found not to be significantly changed in CRC (Fig. 2*a* & 2*b*). Box-and-

Whiskers plots (Fig. 2*a*) were also generated for AHSG, C3, CLI, HP and SAA levels in control and cancer samples from 32 patients with CRC. Data from the control *versus* CRC sera analysis showed that median levels for C3 (511.0 μg/ml *vs.* 1092.6 μg/ml, respectively, $p \leq 0.01$), CLI (64.6 μg/ml *vs.* 170.3 μg/ml, respectively, $p \leq 0.01$), HP (447.4 μg/ml *vs.* 1200.2 μg/ml respectively, $p \leq 0.01$) and SAA (45.4 μg/ml *vs.* 528.6 μg/ml, respectively, $p \leq 0.01$) were all very significantly elevated in the cancer group.

Expression of biomarker panel in lung cancer

Box plot analysis of the five-candidate biomarkers in serum from patients with lung cancer was performed by comparing the median levels of the control group with that from small cell, adenocarcinoma and squamous cell carcinoma (Fig. 3*a*).

Early Detection and Diagnosis
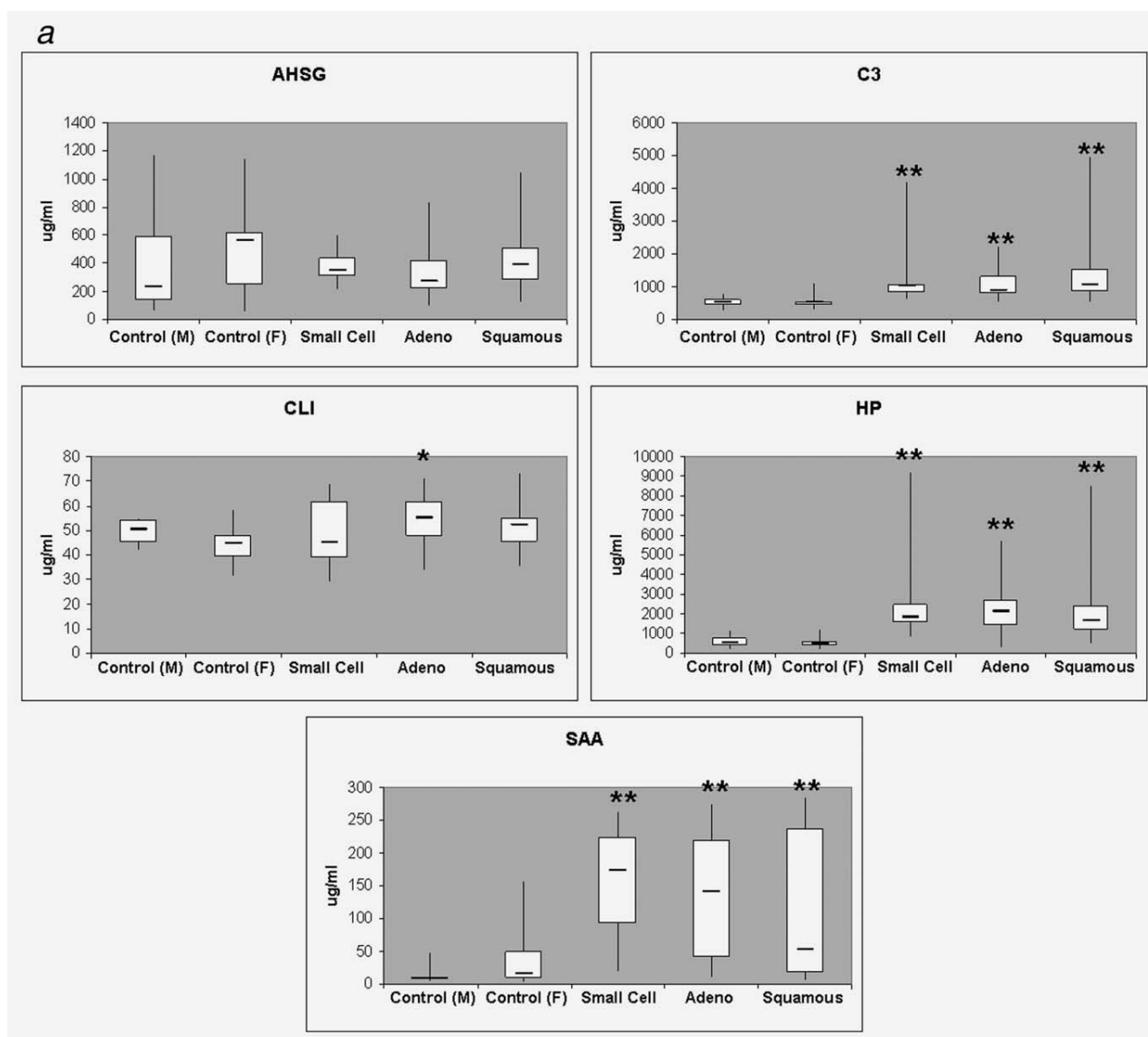
Figure 3. (*a,b*) Box-and-Whisker plots and Bar charts comparing AHSG, C3, CLI, HP and SAA levels in Lung cancer serum samples compared to control groups. Box-and-Whisker plots represent data with boxes ranging from the 25th to the 75th percentile of the observed distribution of values. The bold line of the box displays the median value for serum levels of AHSG, C3, CLI, HP and SAA in lung cancer (small cell lung carcinoma, lung adenocarcinoma and squamous cell carcinoma of the lung) compared to controls. Whiskers span minimum to maximum observed values. The bar charts display the mean and standard deviation for AHSG, C3, CLI, HP and SAA levels in Lung cancer serum samples compared to controls. (*$p \le 0.05$: significant; **$p \le 0.01$: very significant).

C3 (517.3 µg/ml *vs.* 1011.1 µg/ml, 873.8 µg/ml, 1043.4 µg/ml, respectively; $p \le 0.01$), HP (539.3 µg/ml *vs.* 1838.6 µg/ml, 2126.7 µg/ml, 1652.8 µg/ml, respectively; $p \le 0.01$) and SAA (9.8 µg/ml *vs.* 172.2 µg/ml, 141.6 µg/ml, 53.8 µg/ml, respectively; $p \le 0.01$) were all found to be very significantly elevated in all these lung cancer serum samples (Fig. 3). Median CLI levels very also found to be significantly increased in adenocarcinoma serum samples compared to controls (47.1 µg/ml *vs.* 54.8 µg/ml, respectively, $p \le 0.05$). In total, serum samples of 55 patients with lung cancer (12 small cell, 25 adenocarcinoma and 18 squamous cell carcinoma) were ana-

lyzed. Bar charts representing data collected for lung cancer samples show significant increases in abundance levels for C3, HP and SAA for small cell, adenocarcinoma and squamous serum samples compared to their respective control groups (Fig. 3*b*).

AUC and LOOCV analysis of individual biomarkers
The area under the receiver-operator characteristic curve (AUC ROC) value for these five candidate biomarkers was calculated (Table 2). The AUC for AHSG (0.764) was found to have good discriminatory power between control and
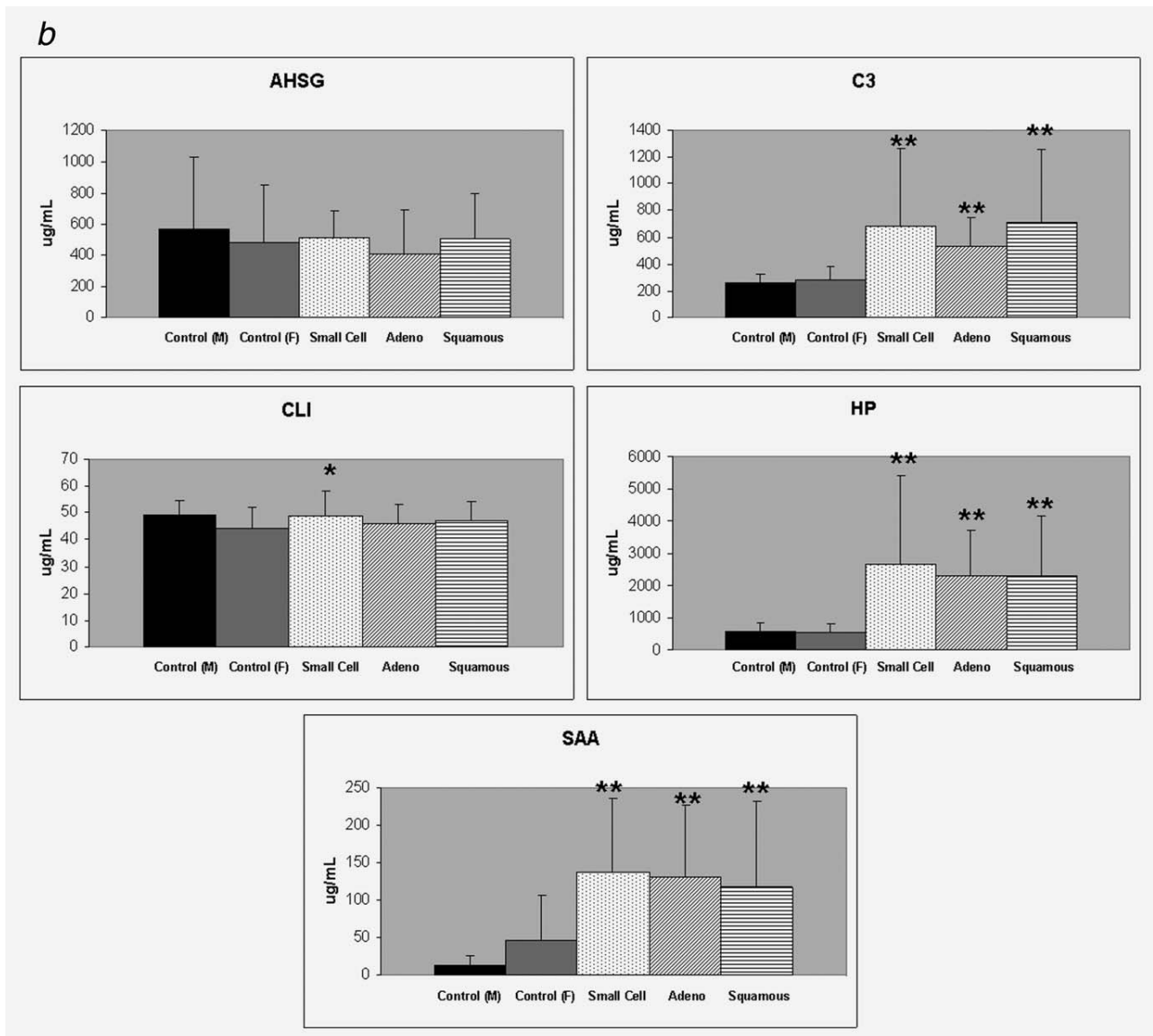
Figure 3. Continued

breast cancer serum samples according to guidelines published by Hosmer and Lemeshow.[16] In CRC and the three subtypes of lung cancer, AHSG was found not have significant AUC values when compared to the control groups.

C3 was found to have an AUC value of 0.888 in breast cancer, 0.901 in CRC, 0.933, 0.938 and 0.933 in small cell lung cancer (SCLC), lung adenocarcinoma and squamous cell lung cancer, respectively. These values represent excellent discriminatory power. CLI was calculated to have the highest AUC values in both breast cancer and CRC (0.83 and 0.98, respectively). AUC values for CLI in lung cancer were found not to achieve acceptable discrimination.

AUC values for HP were found to achieve outstanding discrimination in CRC, SCLC, lung adenocarcinoma and squamous cell lung cancer (0.958, 0.97, 0.92 and 0.944,

respectively). The AUC value for HP in breast cancer was low (0.614) and not found to be significant. AUC values for SAA were found to achieve excellent discrimination in CRC, SCLC, lung adenocarcinoma and squamous cell lung cancer (0.873, 0.929, 0.892 and 0.817, respectively). The AUC value for SAA in breast cancer, similar to HP was low (0.644).

Table 2 also shows LOOCV results for AHSG, C3, CLI, HP and SAA in breast, CRC, SCLC, lung adenocarcinoma and squamous cell lung cancer. LOOCV—where for each member of the training set, using a model built using the other $n - 1$ members, one tries to predict the class of the remaining member—was also performed. The results indicate that in terms of LOOCV the most accurate candidate biomarkers in their respective groups were CLI in breast cancer (76%), CLI in CRC (90%), HP in SCLC (88%), C3 in lung

Table 2. AUC, LOOCV, +LR and −LR values for individual and combined candidate biomarkers

	AHSG	C3	CLI	HP	SAA	Combination
Breast						AHSG+C3
AUC	0.76	0.89	0.83	0.61	0.64	0.91
LOOCV(%)	64.44	73.33	75.55	60.00	66.67	82.22
+LR	4.50	4.50	7.25	1.75	2.60	10.00
−LR	0.82	0.33	0.43	1.22	0.67	0.25
CRC						CLI+HP
AUC	0.65	0.90	0.98	0.96	0.87	0.98
LOOCV(%)	52.00	80.00	90.00	86.00	78.00	92.00
+LR	1.66	4.33	15.00	9.67	3.57	15.00
−LR	0.54	0.25	0.05	0.11	0.25	0.05
Small cell (Lung)						C3+SAA
AUC	0.53	0.93	0.53	0.97	0.93	0.97
LOOCV(%)	56.25	84.37	62.50	87.50	81.25	90.63
+LR	3.00	3.00	1.40	11.00	3.00	11.00
−LR	0.82	0.11	2.33	0.05	0.11	0.05
Adeno (Lung)						HP+SAA
AUC	0.56	0.94	0.71	0.92	0.89	0.98
LOOCV(%)	66.66	88.88	57.77	86.66	71.11	95.55
+LR	7.33	11.50	2.57	5.25	2.13	24.00
−LR	1.00	0.11	1.00	0.11	0.11	0.05
Squamous (Lung)						HP+SAA
AUC	0.52	0.93	0.65	0.94	0.82	0.98
LOOCV(%)	23.68	86.84	57.89	86.84	68.42	84.21
+LR	2.00	5.00	1.57	5.00	1.57	5.00
−LR	1.00	0.11	0.53	0.11	0.25	0.11

Area under the curve (AUC), leave-one-out cross-validation (LOOCV), positive likelihood ratio (+LR) and negative likelihood ratio (−LR) results for AHSG, C3, CLI, HP, SAA and the best performing two protein combination in breast, CRC, small cell carcinoma (lung), adenocarcinoma (lung) and squamous cell carcinoma (lung) are shown.

adenocarcinoma (89%) with both C3 and HP found to have identical LOOCV percentages in squamous cell carcinoma of the lung (87%). Positive likelihood ratios (+LR) and negative likelihood ratios (−LR) are also displayed in Table 2. The positive likelihood ratio is equal to the true positive rate/false positive rate (or sensitivity/(1 − specificity)) while the negative likelihood ratio is equal to the false negative rate/true negative rate (or (1 − sensitivity)/specificity).

Multivariate analysis of biomarker panel using LR

Table 2 also shows the best combination of two candidate biomarkers for distinguishing breast, CRC and lung cancer (small cell, adenocarcinoma and squamous) groups from their respective controls was calculated using LR analysis. LR analysis was performed for all possible combinations of two candidate biomarkers, with the top combination in terms of an AUC value presented here. A combination of AHSG + C3 was found to give an AUC value of 0.91 for breast cancer, a combination of CLI+HP gave an AUC value of 0.98 for CRC, while SCLC had a best combination of C3+SAA giving

an AUC value of 0.97. Interestingly, the best combination of two candidate biomarkers for both adenocarcinoma and squamous cell carcinoma of the lung was found to be HP+SAA, resulting in an AUC value of 0.98. Table 2 includes LOOCV results for the best combination of three proteins in breast cancer (AHSG + C3, 82%), CRC (CLI + HP, 92%), SCLC (C3 + SAA, 91%), adenocarcinoma (HP + SAA, 96%) and squamous cell carcinoma (HP + SAA, 84%). Positive likelihood ratios (+LR) and negative likelihood ratios (−LR) are also displayed in Table 2.

To summarize these findings, Table 3 shows the details of statistically significant abundance level trends for AHSG, C3, CLI, HP and SAA in serum samples from the patients with advanced breast, CRC and lung cancer compared to healthy individuals. Levels are indicated as increased (↑), decreased (↓) or unchanged (↔). From these data, it is clear that the levels of these five APPs are not uniformly changed across the different cancer types and that distinctive patterns present distinguish these cancers from each other. Trends for these five proteins from publications relating to their measurement

Early Detection and Diagnosis

Table 3. abundance level trends for AHSG, C3, CLI, HP and SAA

Protein	Breast	CRC	Small cell (lung)	Adeno (Lung)	Squamous (Lung)	Inflammation/Infection
AHSG	↓	↔	↔	↔	↔	↓ [46]
C3	↑	↑	↑	↑	↑	↑ [47]
CLI	↑	↑	↔	↑	↔	↑ [48]
HP	↔	↑	↑	↑	↑	↑ [49]
SAA	↑	↑	↑	↑	↑	↑ [50]

Table showing statistically significant abundance level trends for AHSG, C3, CLI, HP and SAA in serum samples from patients with advanced breast, CRC and lung cancer compared to healthy individuals. Levels are indicated as increased (↑), decreased (↓) or unchanged (↔). As a reference, levels for these five proteins are also included for general inflammatory conditions as indicated in the referenced literature.[17–21]

in infection/inflammatory conditions are also shown in the Table 3, with these data serving as a good control to compare the cancer-specific trends as well.

Discussion

APPs are characterized by a nonspecific rapid response to stimuli such as inflammation (infections, autoimmune diseases, etc.) or tissue damage (trauma, surgery, myocardial infarction or tumors). Some proteins (*e.g.*, haptoglobin, C3 and SAA) will increase in response to an acute-phase reaction and are known as positive APPs. Others (*e.g.*, AHSG) decrease and are known as negative APPs.

When APPs are found to be differentially expressed between disease and control populations in proteomics experiments, researchers often dismiss that these results as uninteresting because, by definition, these proteins have an increased or decreased abundance level in many conditions that cause inflammation. Because these changes are not specific for a condition or disease, they are often thought to have little potential as conventional biomarkers for diagnosis or prognosis. However, the potential of APPs as biomarkers for cancer may have been underestimated. All the APPs do not have an increased or decreased abundance level across all cancer types. A recent review by Pang *et al.* indicates the presence of putative APP fingerprints in patients with cancer.[8] Data was compiled from many international biomarker discovery studies and merged with data generated by Pang and colleges. From the overall pattern that was derived from the compiled data, the altered serum APP profiles of patients appear to be unique for each type of cancer. For example, pancreatic cancer is associated with increased levels of serum α1-antitrypsin, α1-antichymotrypsin, complement factor B and Leucine-rich glycoprotein, whereas breast cancer is associated with enhanced levels of α1-antichymotrypsin, complement factor B and CLI and decreased Kininogen.

In our study, based on earlier work analyzing serum from patients with squamous cell carcinoma of the lung,[11] we examined the abundance levels for AHSG, C3, CLI, HP and SAA in serum samples from patients with advanced breast, CRC and lung cancer compared to healthy individuals. AHSG, C3, HP and SAA are considered APPs while CLI is also called an APP but more often referred to as an abundant

secreted serum glycoprotein that may play a significant role in tumor growth and metastatic progression.[22]

SAA is a major acute-phase reactant and has been demonstrated to mediate proinflammatory cellular responses.[23] SAA is expressed mainly by hepatocytes in the liver. Bacterial lipopolysaccharides and several cytokines (not only IL1, IL6 and TNF but also LIF, CNTF, oncostatin M, IL11 and cardiotrophin-1) are involved in the induction of SAA synthesis.[24] In our investigation, we found that SAA levels were significantly changed between control and cancer groups, particularly in advanced CRC and subtypes of advanced lung cancer. Subsequent ROC curve analysis demonstrated that the AUC value was high, achieving excellent or outstanding discriminatory power. However, although the level of SAA in advanced breast cancer was also significant compared to the control group, the AUC value was low, indicating poorer discriminatory power.

AHSG and C3 are negative and positive APPs, respectively. AHSG is a 59 kDa glycoprotein predominantly synthesized in liver and has recently been shown to be involved in mediating growth signalling in breast tumor cells.[25] In our investigation, differential expression of AHSG was found to be significant only in breast cancer, with decreased abundance levels compared to the control group. Petrik and coworkers identified one peak using SELDI-TOF as the AHSG that was less prominent with increasing tumor grade in patients with glioblastoma. AHSG was therefore subsequently validated as a survival predictor in glioblastoma using ELISA in an independent group of 72 patients with glioblastoma.[26] C3 was found to be elevated in all cancer types investigated in our study with high AUC values. C3 is a major component of the complement system and elevated levels of this APP have previously been reported associated with cancer.[27] C3 has previously been found to be elevated in sera from patients diagnosed with pancreatic adenocarcinoma compared to sera from normal volunteers using gel electrophoresis.[28]

CLI is a pleiotropic protein with a broad range of functions including tumor growth and metastatic progression. The literature links expression levels of CLI (both nuclear and secreted forms) with CRC.[28–31] In CRC, an increase of secreted CLI expression occurs, whereas the nuclear proapoptotic form is decreased.[32] From the data presented here, it is also clear based on AUC and LOOCV values that CLI is the most accurate candidate biomarker for distinguishing CRC from its control group.

Several studies have shown decreased CLI levels in prostate cancer.[33,34] On the other hand, there are also reports on increased expression of CLI in prostate cancer, specifically after androgen ablation therapy.[35] These opposing results have been explained by the different isoforms of CLI, for example, proapoptotic nuclear CLI being decreased, while antiapoptotic, prosurvival secreted CLI could be increased.

An interesting observation from this data is that CLI levels are increased at a very significant level in both CRC and breast cancer. CLI was also found to be significantly elevated in adenocarcinoma of the lung but not in squamous cell or small cell carcinomas. As the majority of both CRC and breast cancers are adenocarcinomas, this data suggests a possible link between higher levels of secreted CLI and adenocarcinomas.

Haptoglobin is a tetrameric plasma glycoprotein produced primarily by hepatocytes but reports of extrahepatic site production have been reported.[36] Its synthesis is induced by various cytokines, including IL1, IL6 and CNTF.[37] Recently, some groups have examined the individual HP chains (α and β) and also various modified forms including glycosylated and deamidated variants to establish how significant such isoforms are for distinguishing between control and cancer groups,[38] with initial results indicate that specific HP isoforms may have a role to play in the detection of cancer.[39,40] Our data indicates that HP was significantly increased in CRC and the three subtypes of lung cancer when compared to their respective control groups. In all cases, high AUC values were achieved indicating excellent discriminatory power for this individual protein. However, no significant change was found when advanced breast cancer serum samples were compared to the control group.

Elevated levels of APPs in the serum of patients with cancer are generally believed to be of liver origin rather than a tumor cell product.[41] However, there is also evidence in the literature that APPs could also be directly produced by the disease tissue. Expression studies show local production of SAA proteins in histologically normal, inflammatory and tumor tissues.[42] For example, Gutfeld and coworkers showed, using nonradioactive *in situ* hybridization on paraffin tissue sections from 26 patients with CRC, detectable SAA mRNA expression in normal looking colonic epithelium. Expression was increased gradually as epithelial cells progressed through dysplasia to neoplasia, with deeply invading colon carcinoma cells showed the highest levels of SAA. Therefore, using APPs as part of a diagnostic panel may be more tumor specific that initially realized. Smeets and coworkers demonstrated that in human kidney tumors, both tumor and stromal cells produced haptoglobin mRNA.[43] Apparently, ectopic production of proteins is a frequent characteristic of tumors, with APPs proving to have functional roles in tumor development.[43,44]

Among the biomarkers most often used to monitor patients with advanced cancer or to determine recurrence are cancer antigen 15-3 (CA 15-3), cancer antigen 27.29 (CA 27.29), carcino-embryonic antigen (CEA) and neuron-specific enolase (NSE). CA 15-3 and CA27.29 are highly associated with breast cancer and are derived from the MUC1 gene. CA 15-3 is elevated in the serum/plasma of ~ 75% of women with metastasized breast cancer; however, these levels can also be raised due to the presence of other conditions or cancers. Pectasides and coworkers found that patients with postsurgical breast cancer prospectively monitored with simultaneous serum level estimations for CA 15-3 and CEA, resulted in sensitivity values of 68.2 and 34.1% respectively.[45] The markers most often elevated in advanced CRC are CEA and CA 19-9, but neither of these is useful as a screening test for CRC. More than 70% of patients with advanced CRC have elevated CEA levels.[46] Some of the tumor markers that may be elevated in advanced lung cancer are the carcinoembryonic antigen (CEA) in non-small cell lung cancer and the NSE in SCLC.

Molina and coworkers have recently published on serum levels of ProGRP (Pro-Gastrin-Releasing Peptide), CEA, squamous cell carcinoma antigen (SCC), CA 125 (cancer antigen 125), CYFRA 21-1 (cytokeratin 19 fragments) and NSE in patients with lung cancer and how these measurements correlated with histology.[47,48] It was concluded that these markers may be useful in the histological differentiation of non-small cell lung carcinoma (NSCLC) and SCLC. Patients with SCC serum levels >2 ng/ml were always NSCLC, while those with SCC <2 ng/ml and ProGRP >100 pg/ml and NSE >35 ng/ml were all patients with SCLC. Correct classification has important implications for clinical management and panels of biomarkers that can be used in combination with histology to accurately diagnose the specific type of lung cancer would have significant clinical utility.

From the literature, it is clear that while all these biomarkers are currently used in the clinic, these biomarkers still suffer from low levels of sensitivity and specificity, highlighting a need for additional biomarkers that will help to improve the overall accuracy for these tests. One difficulty in interpreting serologic data is the potential for spikes in protein abundance levels due to the presence of infection and signatures associated with its presence. Therefore, combination of biomarkers such as CA 15-3 and CEA, which are commonly used in the clinic together with some of the candidate biomarkers discussed in our study, may help improve the overall accuracy for monitoring advanced patients with cancer and help overcome the problems associated with infection-mediated APP signatures. Furthermore, the optimum pairs of biomarkers described here are different in each of the cancer types, and these cancer-related values should be persistent, whereas infection-related values would be expected to spike and to be less specific for any pair of APPs. Additionally, in terms of clinical management, other symptoms of infection would be apparent.

Prostate-specific antigen (PSA), an FDA-approved biomarker, has been found to be reasonably successful for prostate cancer diagnosis; however, false positives and negatives are common.[49] However, generally single biomarkers lack the sensitivity and specificity required for them to be considered worthwhile diagnostics. To overcome this problem, panels of biomarkers are now seen as providing that extra sensitivity

Early Detection and Diagnosis

and specificity to increase the overall accuracy of diagnostic tests.[50] We investigated various combinations of two candidate biomarkers using LR analysis, to examine if the AUC values for individual markers for each of the cancer types could be improved upon. Our data revealed that specific combinations of two candidate biomarkers outperformed (breast, lung adenocarcinoma and squamous cell carcinoma of the lung) or equalled (CRC and small cell lung carcinoma) their single counterpart. The combinations for breast cancer (AHSG + C3), CRC (CLI + HP) and SCLC (C3 + SAA) were all found to be different from each other. Interestingly, the best combination of two candidate biomarkers for both lung adenocarcinoma and squamous cell carcinoma of the lung was found to be HP + SAA. This data supports the possibility that patterns of altered serum APP in patients with different cancers exist as the best combinations were found to be different with the exception of lung adenocarcinoma and squamous cell carcinoma of the lung.

Counter-intuitively, combining the two candidate biomarkers with the best AUC values together using LR analysis did not always results in the best combination in terms of accuracy. For example, the best individual candidate biomarkers in the breast cancer comparison were C3 (AUC = 0.89, LOOCV = 73%) and CLI (AUC = 0.83, LOOCV = 76%). However, after analysis of all possible combination of two biomarkers, it was found that AHSG+C3 (AUC = 0.91, LOOCV = 82%) was found to be the best combination. The reasons for this antagonism are not entirely clear but this highlights the fact that members of a biomarker panel need not necessarily have the best individual score for distinguishing between groups with the real value of these biomarkers being recognized when used in combination.

In our study, we included both AUC and LOOCV values for all individual and combined biomarkers. ROC curves can sometimes present an overly optimistic result and the presence of LOOCV values helps to balance the results and paint a more realistic picture as regards the accuracy of candidate biomarkers. For example, in the dataset presented here, the AUC and LOOCV values for CLI in CRC were 0.98 and 90% respectively, but it is likely that the 90% value based on the LOOCV calculation better reflects the true accuracy for this candidate biomarker.

A variety of cytokines, chemokines and growth factors are produced by tumor cells and by different cells in the local tumor environment. This signature of active molecules is likely to differ between tumor types and disease stages resulting in the stimulation of APPs that reflect the type and stage of the malignancy. Although individual APPs may not satisfactory diagnosis biomarkers, in combination or with other serum biomarkers may enable more sensitive and specific tests. Therefore APPs may represent a worthwhile component in a panel of biomarkers that represent the tumor microenvironment. Our data indicate that specific patterns of altered serum APPs exist in patients with different cancers and further research is warranted to assess the value of theses patterns in applications such as diagnosis of recurrence and monitoring of therapy effectiveness.

Acknowledgement

References

1. Jemal A, Siegel R, Xu J, Ward E. Cancer Statistics, 2010. *CA Cancer J Clin* 2010;60: 277–300.

2. Higa GM. Breast cancer: beyond the cutting edge. *Exp Opin Pharmacother* 2009;10:2479–98.

3. Brooks DD, Wolf A, Smith RA, Dash C, Guessous I. Prostate cancer screening 2010: updated recommendations from the American Cancer Society. *J Natl Med Assoc* 2010;102:423–9.

4. Simpson RJ, Bernhard OK, Greening DW, Moritz RL. Proteomics-driven cancer biomarker discovery: looking to the future. *Curr Opin Chem Biol* 2008;12:72–7.

5. Boja E, Hiltke T, Rivers R, Kinsinger CR, Rahbar A, Mesri M, Rodriguez H. The evolution of clinical proteomics and its role in medicine. *J Proteome Res* 2011;10:66–84.

6. Tu C, Rudnick PA, Martinez MA, Cheek KL, Stein SE, Slebos RJ, Liebler DC. Depletion of abundant plasma proteins and limitations of plasma proteomics. *J Proteome Res* 2010;9: 4982–91.

7. Wood SL, Rogers M, Cairns DA, Paul A, Thompson D, Vasudev NS, Selby PJ, Banks RE. Association of serum amyloid A protein and peptide fragments with prognosis in renal cancer. *Br J Cancer* 2010;103:101–11.

8. Pang WW, Abdul-Rahman PS, Wan-Ibrahim WI, Hashim OH. Can the acute-phase reactant proteins be used as cancer biomarkers? *Int J Biol Markers* 2010;25:1–11.

9. Cho WC, Yip TT, Yip C, Yip V, Thulasiraman V, Ngan RK, Yip TT, Lau WH, Au JS, Law SC, Cheng WW, Ma VW, et al. Identification of serum amyloid a protein as a potentially useful biomarker to monitor relapse of nasopharyngeal cancer by serum proteomic profiling. *Clin Cancer Res* 2004;10:43–52.

10. Miyake H, Muramaki M, Furukawa J, Kurahashi T, Fujisawa M. Serum level of clusterin and its density in men with prostate cancer as novel biomarkers reflecting disease extension. *Urology* 2010;75:454–9.

11. Dowling P, O'Driscoll L, Meleady P, Henry M, Roy S, Ballot J, Moriarty M, Crown J, Clynes M. 2-D difference gel electrophoresis of the lung squamous cell carcinoma versus normal sera demonstrates consistent alterations in the levels of ten specific proteins. *Electrophoresis* 2007;28: 4302–10.

12. Petri AL, Simonsen AH, Hogdall E, Christensen IJ, Kjaer SK, Yip C, Risum S, Pedersen AT, Hartwell D, Fung ET, Hogdall C. Comparison of proteomic biomarker panels in urine and serum for ovarian cancer diagnosis. *Proteomics Clin Appl* 4:304–14.

13. Bitterlich N, Muley T, Schneider J. Centre-independent detection of non-small cell lung cancer (NSCLC) by means of classification with receiver operating characteristic (ROC)-based data transformation. *Anticancer Res* 30:1661–5.

14. Bloom GC, Eschrich S, Zhou JX, Coppola D, Yeatman TJ. Elucidation of a protein signature discriminating six common types of adenocarcinoma. *Int J Cancer* 2007;120:769–75.

15. Ragazzi E, Pucciarelli S, Seraglia R, Molin L, Agostini M, Lise M, Traldi P, Nitti D. Multivariate analysis approach to the plasma protein profile of patients with advanced colorectal cancer. *J Mass Spectrom* 2006;41: 1546–53.

16. Hosmer D, Lemeshow S. Applied logistic regression, Wiley Series in probability and statistics, 2nd edn., Wiley, 2000.

17. Kazama JJ. What does the circulating AHSG/fetuin-A level tell us? *Clin Exp Nephrol* 2007;11: 336–7.

18. Petersen NE, Elmgreen J, Teisner B, Svehag SE. Activation of classical pathway complement in chronic inflammation. Elevated levels of

circulating C3d and C4d split products in rheumatoid arthritis and Crohn's disease. *Acta Med Scand* 1988;223:557–60.

19. Nuutinen T, Suuronen T, Kauppinen A, Salminen A. Clusterin: a forgotten player in Alzheimer's disease. *Brain Res Rev* 2009;61: 89–104.

20. Quaye IK. Haptoglobin, inflammation and disease. *Trans R Soc Trop Med Hyg* 2008;102: 735–42.

21. dos Anjos BL, Grotto HZ. Evaluation of C-reactive protein and serum amyloid A in the detection of inflammatory and infectious diseases in children. *Clin Chem Lab Med* 48:493–9.

22. Hardardottir I, Kunitake ST, Moser AH, Doerrler WT, Rapp JH, Grunfeld C, Feingold KR. Endotoxin and cytokines increase hepatic messenger RNA levels and serum concentrations of apolipoprotein J (clusterin) in Syrian hamsters. *J Clin Invest* 1994;94:1304–9.

23. Malle E, Sodin-Semrl S, Kovacevic A. Serum amyloid A: an acute-phase protein involved in tumour pathogenesis. *Cell Mol Life Sci* 2009;66: 9–26.

24. Ganapathi MK, Rzewnicki D, Samols D, Jiang SL, Kushner I. Effect of combinations of cytokines and hormones on synthesis of serum amyloid A and C-reactive protein in Hep 3B cells. *J Immunol* 1991;147:1261–5.

25. Sakwe AM, Koumangoye R, Goodwin SJ, Ochieng J. Fetuin-A (ahsg) is a major serum adhesive protein that mediates growth signaling in breast tumor cells. *J Biol Chem*.

26. Petrik V, Saadoun S, Loosemore A, Hobbs J, Opstad KS, Sheldon J, Tarelli E, Howe FA, Bell BA, Papadopoulos MC. Serum alpha 2-HS glycoprotein predicts survival in patients with glioblastoma. *Clin Chem* 2008;54:713–22.

27. Qiu Y, Patwa TH, Xu L, Shedden K, Misek DE, Tuck M, Jin G, Ruffin MT, Turgeon DK, Synal S, Bresalier R, Marcon N, et al. Plasma glycoprotein profiling for colorectal cancer biomarker identification by lectin glycoarray and lectin blot. *J Proteome Res* 2008;7:1693–703.

28. Hanas JS, Hocker JR, Cheung JY, Larabee JL, Lerner MR, Lightfoot SA, Morgan DL, Denson KD, Prejeant KC, Gusev Y, Smith BJ, Hanas RJ, et al. Biomarker identification in human pancreatic cancer sera. *Pancreas* 2008;36:61–9.

29. Andersen CL, Schepeler T, Thorsen K, Birkenkamp-Demtroder K, Mansilla F, Aaltonen LA, Laurberg S, Orntoft TF. Clusterin expression in normal mucosa and colorectal cancer. *Mol Cell Proteomics* 2007;6:1039–48.

30. Pucci S, Bonanno E, Sesti F, Mazzarelli P, Mauriello A, Ricci F, Zoccai GB, Rulli F, Galata G, Spagnoli LG. Clusterin in stool: a new biomarker for colon cancer screening? *Am J Gastroenterol* 2009;104:2807–15.

31. Pucci S, Mazzarelli P, Nucci C, Ricci F, Spagnoli LG. CLU "in and out": looking for a link. *Adv Cancer Res* 2009;105:93–113.

32. Mazzarelli P, Pucci S, Spagnoli LG. CLU and colon cancer. The dual face of CLU: from normal to malignant phenotype. *Adv Cancer Res* 2009; 105:45–61.

33. Scaltriti M, Brausi M, Amorosi A, Caporali A, D'Arca D, Astancolle S, Corti A, Bettuzzi S. Clusterin (SGP-2, ApoJ) expression is downregulated in low- and high-grade human prostate cancer. *Int J Cancer* 2004;108:23–30.

34. Rauhala HE, Porkka KP, Saramaki OR, Tammela TL, Visakorpi T. Clusterin is epigenetically regulated in prostate cancer. *Int J Cancer* 2008; 123:1601–9.

35. July LV, Akbari M, Zellweger T, Jones EC, Goldenberg SL, Gleave ME. Clusterin expression is significantly enhanced in prostate cancer cells following androgen withdrawal therapy. *Prostate* 2002;50:179–88.

36. Smeets MB, Fontijn J, Kavelaars A, Pasterkamp G, De Kleijn DP. The acute phase protein haptoglobin is locally expressed in arthritic and oncological tissues. *Int J Exp Pathol* 2003;84: 69–74.

37. Yoshioka M, Watanabe A, Shimada N, Murata H, Yokomizo Y, Nakajima Y. Regulation of haptoglobin secretion by recombinant bovine cytokines in primary cultured bovine hepatocytes. *Domest Anim Endocrinol* 2002;23:425–33.

38. Shah A, Singh H, Sachdev V, Lee J, Yotsukura S, Salgia R, Bharti A. Differential serum level of specific haptoglobin isoforms in small cell lung cancer. *Curr Proteomics* 7:49–65.

39. Chen CB, Su YC, Huang TT, Ho HC, Chang YT, Tung YT, Lee WC. Differentially expressed serum haptoglobin alpha chain isoforms with potential application for diagnosis of head and neck cancer. *Clin Chim Acta* 2008;398:48–52.

40. Hamrita B, Chahed K, Trimeche M, Guillier CL, Hammann P, Chaieb A, Korbi S, Chouchane L. Proteomics-based identification of alpha1-

antitrypsin and haptoglobin precursors as novel serum markers in infiltrating ductal breast carcinomas. *Clin Chim Acta* 2009;404:111–8.

41. Kumar DM, Thota B, Shinde SV, Prasanna KV, Hegde AS, Arivazhagan A, Chandramouli BA, Santosh V, Somasundaram K. Proteomic identification of haptoglobin alpha2 as a glioblastoma serum biomarker: implications in cancer cell migration and tumor growth. *J Proteome Res* 9:5557–67.

42. Gutfeld O, Prus D, Ackerman Z, Dishon S, Linke RP, Levin M, Urieli-Shoval S. Expression of serum amyloid A, in normal, dysplastic, and neoplastic human colonic mucosa: implication for a role in colonic tumorigenesis. *J Histochem Cytochem* 2006;54:63–73.

43. Lau SH, Sham JS, Xie D, Tzang CH, Tang D, Ma N, Hu L, Wang Y, Wen JM, Xiao G, Zhang WM, Lau GK, et al. Clusterin plays an important role in hepatocellular carcinoma metastasis. *Oncogene* 2006; 25:1242–50.

44. Hsu SJ, Nagase H, Balmain A. Identification of Fetuin-B as a member of a cystatin-like gene family on mouse chromosome 16 with tumor suppressor activity. *Genome* 2004;47:931–46.

45. Pectasides D, Pavlidis N, Gogou L, Antoniou F, Nicolaides C, Tsikalakis D. Clinical value of CA 15-3, mucin-like carcinoma-associated antigen, tumor polypeptide antigen, and carcinoembryonic antigen in monitoring early breast cancer patients. *Am J Clin Oncol* 1996;19:459–64.

46. Moertel CG, O'Fallon JR, Go VL, O'Connell MJ, Thynne GS. The preoperative carcinoembryonic antigen test in the diagnosis, staging, and prognosis of colorectal cancer. *Cancer* 1986;58: 603–10.

47. Molina R, Holdenrieder S, Auge JM, Schalhorn A, Hatz R, Stieber P. Diagnostic relevance of circulating biomarkers in patients with lung cancer. *Cancer Biomark* 2010;6:163–78.

48. Molina R, Auge JM, Bosch X, Escudero JM, Vinolas N, Marrades R, Ramirez J, Carcereny E, Filella X. Usefulness of serum tumor markers, including progastrin-releasing peptide, in patients with lung cancer: correlation with histology. *Tumour Biol* 2009;30:121–9.

49. Schroder FH. PSA screening—a review of recent studies. *Eur J Cancer* 2009;45(Suppl 1):402–4.

50. Hanash SM, Pitteri SJ, Faca VM. Mining the plasma proteome for cancer biomarkers. *Nature* 2008;452:571–9.

Early Detection and Diagnosis