Metabolomics (2013) 9:202–212
DOI 10.1007/s11306-012-0439-z

ORIGINAL ARTICLE

Plasma choline-containing phospholipids: potential biomarkers for colorectal cancer progression

Song Li · Bin Guo · Jianwen Song · Xiaoli Deng · Yusheng Cong · Pengfei Li · Ke Zhao · Lihong Liu · Gang Xiao · Feng Xu · Yingjiang Ye · Zhenwen Zhao · Menggang Yu · Yan Xu · Jianli Sang · Junjie Zhang

Received: 9 March 2012 / Accepted: 18 May 2012 / Published online: 29 June 2012

Abstract Colorectal cancer (CRC) is believed to progress through the adenoma–carcinoma sequence. The adenoma–carcinoma transition is an important window for early detection and intervention of CRC. In the present study, plasma samples from patients with CRC ($n = 120$), patients with adenomatous polyps (AP) ($n = 120$), and healthy controls ($n = 120$) were collected. Plasma phospholipid levels were analyzed with liquid chromatography–tandem mass spectrometry. It was found that the plasma levels of major lysophosphatidylcholine (LPC) species were gradationally decreased from healthy controls, AP to CRC subjects. A formula including total saturated LPCs, 18:2 LPC and sphingosylphosphorylcholine (SPC) yielded a sensitivity and specificity of 88.3 and 80 % for separating CRC from healthy controls. An optimized model with total saturated LPCs, 20:4 LPC and sphingomyelins (SM) as markers yielded a sensitivity and specificity of 89 and 80 % for separating AP from the healthy controls. Moreover, with SM, SPC and saturated LPCs as markers, a model was made to separate CRC from AP with the sensitivity and specificity of 90 and 92.5 %, respectively. These data indicate that the plasma choline-containing phospholipid levels represent potential biomarkers to distinguish between healthy controls, AP and CRC cases, implying their clinical usage in CRC and/or AP-CRC progression detection.

Keywords Colorectal cancer · Adenomatous polyps · Phospholipids · Mass spectrometry · Biomarkers

Abbreviations

CRC	Colorectal cancer
AP	Adenomatous polyps
LPC	Lysophosphatidylcholine
SPC	Sphingosylphosphorylcholine
SM	Sphingomyelin
LPA	Lysophosphatidic acid
S1P	Sphingosine-1-phosphate

1 Introduction

Colorectal cancer (CRC) is the second leading cause of cancer-related deaths in the world with a worldwide incidence

Electronic supplementary material The online version of this article (doi:10.1007/s11306-012-0439-z) contains supplementary material, which is available to authorized users.

S. Li · B. Guo · J. Song · X. Deng · Y. Cong · J. Sang (✉) · J. Zhang (✉)
Key Laboratory of Cell Proliferation and Regulation Biology, Ministry of Education, Institute of Cell Biology, College of Life Sciences, Beijing Normal University, Beijing 100875, China
e-mail: jlsang@bnu.edu.cn

J. Zhang
e-mail: jjzhang@bnu.edu.cn

P. Li · K. Zhao · L. Liu
General Hospital of Second Artillery Force of the Chinese PLA, Beijing 100088, China

G. Xiao
Beijing Hospital, Beijing 100730, China

F. Xu · Y. Ye
People's Hospital of Peking University, Beijing 100044, China

Z. Zhao · Y. Xu
Department of Obstetrics and Gynecology, Indiana University Cancer Center, Indiana University School of Medicine, 975 West Walnut Street IB355A, Indianapolis, IN 46202, USA

M. Yu
Division of Biostatistics, Indiana University School of Medicine, 410 West 10th Street, Suite 3000, Indianapolis, IN 46202, USA

more than a million cases annually (Center 2009; Jemal 2010, 2011). Epidemiological studies suggest that CRC is highly sensitive to environmental changes, such as lifestyle including diet, while genetic differences also partially account for the differences in the degree of susceptibility to CRC (Moshkowitz and Arber 2005). In recent decades, many Asian countries, including China, Japan, South Korea, and Singapore, have experienced an increase of 2–4 times in the incidence of CRC (Sung et al. 2005), and CRC has become one of the leading cancers in China. In the United States, around 149,000 Americans were estimated to be diagnosed with CRC and approximately 50,000 died from the disease in 2008 (Jemal et al. 2008, 2009). When CRC is detected early, the 5-year relative survival rate is approximately 90 %, while the 5-year survival rate for CRC subjects presenting with distant metastases is only 8–11 % (Guittet et al. 2010; Lieberman 2009). Although colonoscopy has the sensitivity and specificity values exceeding 95 % in CRC detection, the degree of patient compliance is low due to the cost, perceived inconvenience and discomfort associated with this test. In developing countries, such as China, these factors as well as inaccessibility to the colonoscopy service in rural areas further reduce the rate of this test in the general population. As a result, many patients still present with the late-stage of CRC (Guittet et al. 2010; Sarfaty 2007; Umar and Greenwald 2009). A number of biomarkers have been identified for the detection and/or prognosis of CRC in tissues, feces, and serum (Srivastava et al. 2001; Wild et al. 2010). However, none of these markers provide a level of accuracy of detection comparable to colonoscopy (Berger et al. 2003a, b; Calistri et al. 2003; Dong et al. 2001; Doolittle et al. 2001; Huang et al. 2010; Kim et al. 2009; Muller 2003). For example, the test based on detecting colon cancer-specific methylation in fecal DNA gives sensitivity less than 50 % when the specificity is 90 % (Chen et al. 2005). In addition, stool-based tests are less convenient than blood-based tests (Loitsch et al. 2008). In resent years, several blood biomarkers have been developed, including serum colon cancer-specific antigen (CCSA)-3 and CCSA-4 and autoantibodies against 6 phage-expressed antigens derived from colon cancer tissues (Alexandrov et al. 2009; Han et al. 2008; Leman et al. 2007; Liu et al. 2009; Ran et al. 2008). Moreover, some serum metabolites have been recognized as biomarkers with high sensitivity for CRC detection with the metabolomic approach (Ikeda et al. 2011). These potential CRC markers remains to be further validated in clinical investigations.

Most of the colorectal malignant lesions are believed to progress through the adenoma-to-carcinoma sequence (Kountouras et al. 2000; Tierney et al. 1990). The transition from polyps to CRC is reflected not only in the overt histological transition to carcinoma, but also in the disruption of molecular and biochemical control mechanisms of cellular functions (Markowitz and Bertagnolli 2009). The progression of adenomatous polyps (AP) to carcinoma, i.e., the 'polyps dwelling window', is estimated at 5–10 years. Thus, the adenoma–carcinoma transition represents an important window for early detection and intervention of CRC. Identification and validation of highly sensitive and specific serologic markers in this window are therefore critically important. However, the blood-based markers which reliably detect the CRC progression are still lacking.

Phospholipids have long since been recognized as signaling mediators that have the capacity to trigger profound physiological responses (Wymann and Schneiter 2008). Among them, there is a group of choline-containing phospholipids, such as lysophosphatidylcholine (LPC), sphingosylphosphorylcholine (SPC), and sphingomyelin (SM). These choline-containing phospholipids and their metabolic products can affect fundamental cellular functions (Rivera and Chun 2008) and, therefore, are involved in various biological processes such as angiogenesis, wound healing, immunity, atherosclerosis and carcinogenesis (Fuchs and Schiller 2009). The metabolism of choline-containing phospholipids plays a critical role in their functional regulation. LPC can be converted into LPA by lysoPLD (Autotaxin) in plasma, and then functions through LPA receptors and contributes to signaling events that influence cell behaviors (Noguchi et al. 2009). S1P can be derived from both SPC and SM through different pathways to induce the S1P receptor-mediated signaling (Graler 2010). As the bioactive phospholipids, LPA and S1P are the simplest lysophospholipids with growth factor-like activities. Lots of reports indicate that LPA and S1P markedly affect cell proliferation, survival and motility, which are hallmarks of cancer promoting factors (Peyruchaud 2009).

We have identified several LPC isoforms as potential markers to separate the healthy controls form CRC subjects in a previous study, which was conducted in the United States (the majority of the subjects were Caucasians) (Zhao et al. 2007). In the present study, we recruited a total of 360 subjects to determine whether these LPCs and/or additional phospholipids are useful markers to distinguish healthy versus CRC cases, healthy versus AP cases, and/or AP versus CRC cases in a Chinese population. Sets of choline-containing phospholipids, including LPC, SPC and SM, were identified as potential markers for CRC and/or AP-CRC progression detection, and the optimized formula/models were developed to distinguish between healthy controls, AP, and CRC cases.

2 Materials and methods

2.1 Study population

The 120 patients with CRC, 120 patients with AP, and 120 healthy controls were enrolled between May 2007 and

2010 at the General Hospital of Second Artillery Force of Chinese PLA, Beijing Hospital, and People's Hospital of Peking University. The AP and CRC subjects went through colonoscopy with pathologically verified adenoma polyps and CRC, respectively. Healthy controls in this study were recruited form the asymptomatic individuals who visited the hospitals for regular examination and volunteered to take the colonoscopy screening. All controls were confirmed with no colorectal malignancy or polyps by colonoscopy. They were also free of other gastrointestinal diseases and other types of cancer. The vein blood samples were collected early in the morning from the participants under limosis condition. The project was approved by the Institutional Review Board, and written informed consent forms were signed by participants.

2.2 Blood processing, lipid extraction and MS analyses of phospholipids

Blood samples were collected in EDTA-containing tubes and centrifuged at 1,750 g for 15 min at room temperature. Plasma samples were aliquoted into siliconized Eppendorf tubes and frozen at −80 °C until use. The lipid extraction and MS analyses were performed with the methods as same as described previously (Zhao et al. 2007) (detailed in Supplementary Materials), except that a different mass spectrometer (Applied Biosystems Sciex 3200Qtrap™ mass spectrometer, Applied Biosystems Sciex, Ontario, Canada) was used.

2.3 Statistical analysis

Absolute and relative levels (%) of individual phospholipid forms within the total plasma phospholipid group and within the saturated or unsaturated fraction were evaluated. Chi-square test was used to test the association between disease status and gender. Pearson's correlations were used to assess the associations between individual lysophospholipids and age. Student *t* test was used to compare phospholipid levels in univariate analyses whereas logistic regression was used for multivariate analyses. All factors (include age and gender) were initially considered as potential candidates. Changes in the Akaike information criteria (AIC) served as the basis for model selection. We choose the top 3 or 4 factors from the stepwise selection as our diagnostic markers because choosing more factors added little improvement. A final logistic regression model was then fitted to determine the optimum linear combination of the factors for classification. Cut-off points for determination of disease status were chosen empirically from the associated ROC curves using an arbitrarily predefined sensitivity. For evaluating LPC markers identified in the previous study (Zhao et al. 2007), we directly constructed the ROC curve using both the established formula (Zhao et al. 2007) and a new formula with the same markers. All our estimated sensitivity and specificity are internally validated using a bootstrapping method, which is a preferred method versus two other possible methods, the data splitting and cross-validation methods. All tests of statistical significance were two-sided and there were no adjustments for multiple comparisons. *P* values <0.05 were considered to be statistically significant. All analyses were performed using TIBCO Spotfire Splus.

3 Results and discussion

3.1 Results

3.1.1 Study population

The demographical data of subjects (age and gender) are summarized in Table 1. All of the subjects were Chinese, and overall 63 % (228/360) of all subjects were male (68, 63, and 59 % in the healthy control, AP, and CRC groups, respectively). The median age was 55 years (55.7 ± 7.5, 54.5 ± 14.2, and 55.7 ± 11.8 years for the healthy control, AP and CRC groups, respectively). The ages of subjects in these three groups were not statistically different (Table 1). Table 2 summarizes CRC tumor characteristics. Overall, 53 % (64/120) of CRC cases had colon tumors. Higher than three quarters of the tumors were T_3/T_4 (93/120, 77 %) and almost half of all tumors were N_0 (58/120, 48 %). Tumor T stages were not identified for 12 patients and N stages were not identified for 13 patients.

3.1.2 The effects of age and gender on plasma phospholipid levels

While 100 % of subjects were Chinese, we analyzed the effects of age and gender on plasma phospholipid levels. It is well known that CRC incidence is strongly associated with age (Brenner et al. 2008). However, if age alone was used as the factor, a poor sensitivity and specificity (79.8 and 23.9 %)

Table 1 Descriptive characteristics of study subjects

Group	*N*	Age	Gender	
		Mean ± SD	Male *N* (%)	Female *N* (%)
Controls	120	55.7 ± 7.5	81 (68)	39 (32)
Adenoma	120	54.5 ± 14.2	76 (63)	44 (37)
CRC	120	55.7 ± 11.8	71 (59)	49 (41)
Total	360	55. 3 ± 11.5	228 (63)	132 (37)

CRC colorectal cancer, *SD* standard deviation

Table 2 Clinical characteristics of the colorectal cancer cases

Factors	Overall *N* (%)
Primary site	
Colon	64 (53)
Rectum	51 (43)
Other	5 (4)
T-Stage	
T1, T2	15 (13)
T3, T4	93 (77)
Unknown	12 (10)
N-Stage	
N0	58 (48)
N1, N2, N3	49 (41)
Unknown	13 (11)
Differentiation	
Well, Moderate	82 (69)
Poor	24 (20)
Unknown	14 (11)

were obtained for the healthy vs. CRC group. In addition, we did not observe any significant association between the age and the plasma phospholipid levels (Pearson's correlation coefficients range from −0.16 to 0.11), indicating that phospholipids are independent markers. Males are over-represented (63 %) in the subjects collected and analyzed in this study. Although gender leads to significantly different distributions among certain lipids (Supplementary Table 1), it has limited classification power for disease diagnosis.

3.1.3 Plasma phospholipid levels in healthy controls, AP and CRC patients

The phospholipids listed in Table 3, as well as S1P and LPAs were initially analyzed for their plasma levels in a subset of samples. Among them, choline-containing phospholipids, including SPC, SM and different forms of LPCs, were found to have significantly different plasma levels among the subject groups. Thus, the plasma levels of these choline-containing phospholipids were measured in all of the subjects.

Table 3 Comparison of plasma choline-containing phospholipid levels among CRC, AP, and control groups

	Control	AP	CRC	P^a		
	Mean ± SD[b]	Mean ± SD[b]	Mean ± SD[b]	Control vs. AP	Control vs. CRC	AP vs. CRC
SPC	0.078 ± 0.051	0.073 ± 0.043	0.038 ± 0.018	0.2099	<0.0001	<0.0001
Lyso-PAF§	1.613 ± 0.719	1.121 ± 0.434	0.719 ± 0.254	<0.0001	<0.0001	<0.0001
14:0 LPC	1.915 ± 0.993	1.256 ± 0.594	1.015 ± 0.664	<0.0001	<0.0001	0.0017
16:0 LPC*§	240.326 ± 97.203	157.448 ± 46.259	107.426 ± 39.074	<0.0001	<0.0001*	<0.0001
18:2 LPC*§	78.828 ± 30.140	62.490 ± 20.801	44.630 ± 23.338	<0.0001	<0.0001*	<0.0001
18:1 LPC*§	43.401 ± 15.575	30.936 ± 9.517	23.482 ± 9.381	<0.0001	<0.0001*	<0.0001
18:0 LPC*§	109.458 ± 46.779	67.457 ± 19.875	48.835 ± 19.728	<0.0001	<0.0001*	<0.0001
Total 18 LPC*§	231.690 ± 83.879	160.883 ± 40.740	116.946 ± 49.469	<0.0001	<0.0001*	<0.0001
20:4 LPC	0.290 ± 0.124	0.327 ± 0.115	0.177 ± 0.079	0.0096	<0.0001#	<0.0001
20:0 LPC*	1.087 ± 0.973	1.143 ± 0.721	0.627 ± 0.312	0.3060	<0.0001*	<0.0001
22:6 LPC§	5.956 ± 2.349	5.203 ± 1.979	3.889 ± 1.989	0.0039	<0.0001#	<0.0001
22:0 LPC	0.836 ± 0.935	1.059 ± 1.478	0.472 ± 0.299	0.0814	<0.0001	<0.0001
Sat LPC*§	353.621 ± 143.841	228.363 ± 63.234	158.376 ± 58.002	<0.0001	<0.0001*	<0.0001
Unsat LPC*§	128.478 ± 45.869	98.956 ± 27.119	72.178 ± 33.647	<0.0001	<0.0001*	<0.0001
Sat LPC/unsat LPC ratio	2.792 ± 0.785	2.371 ± 0.618	2.350 ± 0.585	<0.0001	<0.0001	0.4048
Total LPC*§	482.099 ± 179.935	327.319 ± 81.043	230.553 ± 88.257	<0.0001	<0.0001*	<0.0001
16:0 SM	176.776 ± 51.363	126.264 ± 38.786	159.489 ± 41.842	<0.0001	0.0023	<0.0001
18:0 SM	34.423 ± 9.941	25.127 ± 7.727	31.609 ± 10.340	<0.0001	0.0175	<0.0001
Sum of 16:0 and 18:0 SM	211.205 ± 59.450	151.391 ± 44.705	191.098 ± 50.501	<0.0001	0.0026	<0.0001

All lipid concentrations are in μM. *Sat* saturated; *Unsat* unsaturated

[a] *P* value from the Wilcoxon rank sum test; [b] *SD* standard deviation

* The same trend of results obtained when compared to the USA study (Zhao et al. 2007), where $P < 0.001$ were reported, and $P < 0.0001$ are reported in present study

No statistical differences were observed in these two lipids in the USA study (Zhao et al. 2007)

§ Lipids with their levels decreased from healthy to AP and from AP to CRC with significance in all three comparisons

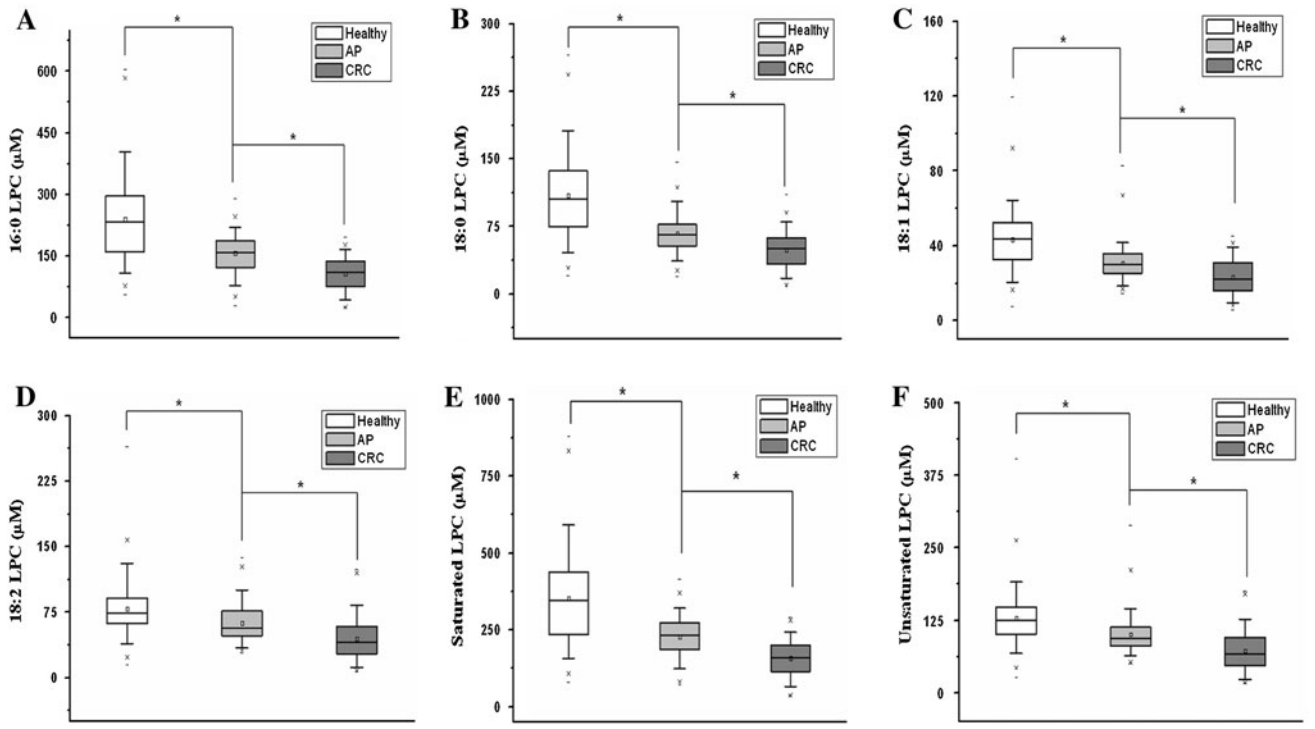

Fig. 1 The plasma levels of major LPC species are gradually decreased from healthy controls, AP to CRC subjects. The plasma levels of 16:0 LPC (**a**), 18:0 LPC (**b**), 18:1 LPC (**c**), 18:2 LPC (**d**), saturated LPC (**e**), and unsaturated LPC (**f**) in healthy control, AP and CRC groups were detected as described in Materials and Methods. The collected data were analyzed by origin 7.5 software (OriginLab). Statistical analysis was performed using Student's *t* tests. Significance was assumed at a *P* value <0.0001

The summary of plasma choline-containing phospholipid levels in the healthy controls, AP and CRC subjects is shown in Table 3. The similar plasma SPC levels were detected in healthy controls and AP patients, while the plasma SPC levels were decreased obviously in CRC patients (Table 3; Supplementary Fig. 1a). Compared with healthy control and CRC groups, AP subjects presented significant lower levels of SM in plasma (Table 3; Supplementary Fig. 1b–d), suggesting a special relationship between AP and plasma SM levels. Plasma levels of the major species of LPCs, including 16:0 LPC, 18:0 LPC, 18:1 LPC, 18:2 LPC, saturated LPC, unsaturated LPC, and total LPC were reduced significantly in the CRC group, compared with in the control group (Table 3, lipids flagged by a *), which is highly consistent with the trend observed previously in the US population (Zhao et al. 2007). While $P < 0.001$ were reported in the previous report (Zhao et al. 2007), it was observed in the present study that the plasma levels of these LPC species were significantly reduced in the CRC group versus the control group with *P* values < 0.0001 (Table 3). These data indicate that the reduction of the plasma levels of major LPC species in CRC versus healthy controls is validated in a completely independent population, and that these potential markers are likely to be race-independent. We also observed a gradual decrease in the plasma levels of major LPC species from healthy controls, AP to CRC subjects, which demonstrated that LPCs were potential markers for CRC progression detection (Fig. 1). When AP and CRC subjects were pooled together as a diseased group, the plasma levels of SPC, SM and the major species LPCs were significantly lower in the diseased group than in the healthy controls. However, it concealed the relationship between the phospholipid levels and each specific stage in CRC progress (Supplementary Table 2). It is worth to emphasize that all of healthy control subjects in the present study were colonoscopically confirmed to be negative for any colorectal abnormality, which was not done in the previous US study (Zhao et al. 2007). The CRC group was composed of patients with different tumor characteristics as shown in Table 2. However, there were no statistically significant differences in the plasma levels of the LPCs with respect to tumor location, tumor stage, nodal status, or differentiation (data not shown).

3.1.4 Evaluation of the LPCs as markers for distinguishing CRC from healthy controls

Previously, we have established a formula using several LPC levels in plasma for discrimination of CRC and healthy controls (Zhao et al. 2007). Interestingly, we found that, highly consistent with the reported trends, the plasma levels

of major LPC isoforms were significantly decreased in CRC cases in the Chinese population analyzed in the current work. When the exactly same formula [18:2-LPC%–18:1-LPC% + 0.053 × total saturated LPC = 16.8] published previously in the USA population was used in the current set of data, a sensitivity and specificity of 80 and 58 %, were obtained respectively to separate CRC from healthy controls (see Supplementary Fig. 2). Although the sensitivity and specificity are relatively poorer compared to the data in the previous study (with a sensitivity and specificity of 82 and 93 %, respectively) (Zhao et al. 2007), implying a potential population-based difference, the consistent trends indicate that plasma LPCs are common CRC biomarkers in the two totally different race population.

Based on statistical analysis of the data obtained in the present study, a new formula was established with the LPC markers identified from the previous study in US (Zhao et al. 2007). When the adjusted formula (11.406 × 18:2 LPC% + 7.43 × 18:1 LPC% + 0.024 × Saturated LPC = 14.58) was applied, a sensitivity and specificity of 83 and 86 %, respectively, were achieved (Fig. 2a). Overall, 100 of 120 (83 %) CRC cases were correctly classified, including 5 of 6 (83 %) T_{0-1} stage, 5 of 9 (55 %) T_2 stage, 50 of 58 (86 %) T_3 stage, and 29 of 35 (83 %) T_4 stage CRC. Among the 20 misclassified cancer cases, 1 patient was at stage T_{0-1}, 4 at T_2, 8 at T_3, and 6 at T_4. Meanwhile, 17 of 120 (14.2 %) healthy controls were falsely classified as CRC. These results indicate that LPCs could be the biomarkers for the CRC detection, but not for stage determination of CRC.

Importantly, in the present study, plasma levels of additional phospholipids besides LPCs were analyzed. Taken all of these analyzed phospholipids into consideration, a formula including saturated LPC, 18:2 LPC and SPC (0.023 × Saturated LPC + 5.96 × 18:2 LPC% + 17.06 × SPC) yielded a sensitivity and specificity 88.3 and 80 %, respectively (Fig. 2b), where the cutoff value for the formula is −9.66. Overall, 14 of 120 (11.7 %) CRC cases were misclassified, and 24 of 120 (20 %) healthy controls were falsely classified as CRC. These results suggest that, compared with the LPCs alone, one set of choline-containing phospholipids with LPCs and SPC can be potentially useful markers with higher sensitivity for CRC.

To validate these sensitivity and specificity estimates, we use an internal validation method based on bootstrapping (Harrell et al. 1996). In particular, the classification rule is obtained from bootstrap samples using the above markers (saturated LPC, 18:2 LPC and SPC). The rule is then applied to the original data to estimate sensitivity and specificity. In 95 % of the 500 bootstrap samples we created from the original data, the predicted sensitivity and specificity on the original data are all above 82.5 and 75 %, indicating that these markers may be valid classifier for external validation data.

3.1.5 Identification of potential phospholipid markers for distinguishing AP from healthy controls and CRC from AP cases

It is widely accepted that the adenoma-to-carcinoma sequence represents the process by which most, if not all, CRCs arise. The malignancy potential of adenomatous polyp (AP) is dependent on both genetic as well as environmental risk factors, such as diet and lifestyle behavior (Kountouras et al. 2000; Tierney et al. 1990). Therefore, the unique 'polyps dwelling window' represents an important time period for early detection and intervention of CRC. Currently, there is no reliable and convenient method (such as a blood test) to detect AP and/or AP–CRC transition. Therefore, we have included the AP group in our study to determine whether some phospholipids can be good markers for CRC progression.

A formula with four kinds of phospholipids (saturated LPC, 20:4 LPC, 16:0 SM, and 18:0 SM) as markers [0.0093 × Sat LPC − 1225 × 20:4 LPC% + 0.02 × sum of 16:0 and 18:0 SM] yield a sensitivity and specificity of 89 and 80 % for separating AP from the healthy controls (Fig. 3), where the comparison value for the formula is −3.3. If the sensitivity was set on 85 %, the specificity was 81.7 %. Using this formula, only 18 of 120 AP cases were missed with 22 of 120 healthy controls wrongly grouped (Fig. 3). These results suggest that these markers potentially be very useful for distinguishing AP from the healthy controls. Bootstrap validated sensitivity and specificity are 83.2 and 74.1 % respectively for 95 % of 500 bootstrap samples.

We also identified phospholipids that could be used to distinguish between the AP and CRC groups. These markers could potentially be very valuable for the detection of the transition from AP to CRC. Using a formula with SPC, 16:0 SM, 18:0 SM, and saturated LPC as markers [0.044 × sum of 16:0 and 18:0 SM − 60.1 × SPC − 0.027 × Total Saturated LPC], a sensitivity and specificity of 90 and 92.5 %, respectively, were achieved to separate CRC from AP subjects (Fig. 4), where the comparison value for the formula is 1.06. With this formula, only 9 of 120 AP cases were missed with 12 of 120 CRC wrongly grouped (Fig. 4). Bootstrap validated sensitivity and specificity are 85 and 88.3 % respectively for 95 % of 500 bootstrap samples.

Interestingly, we have observed that the plasma levels of major LPC species, including 14:0 LPC, 16:0 LPC, 18:0 LPC, 18:1 LPC, 18:2 LPC, total saturated LPC, total unsaturated LPC, and total LPC, were gradationally decreased from healthy controls, AP to CRC, with AP cases presenting the middle levels of these lipids (Fig. 1; Table 3, lipids flagged by §), suggesting that these phospholipids are pertinent to CRC progression.

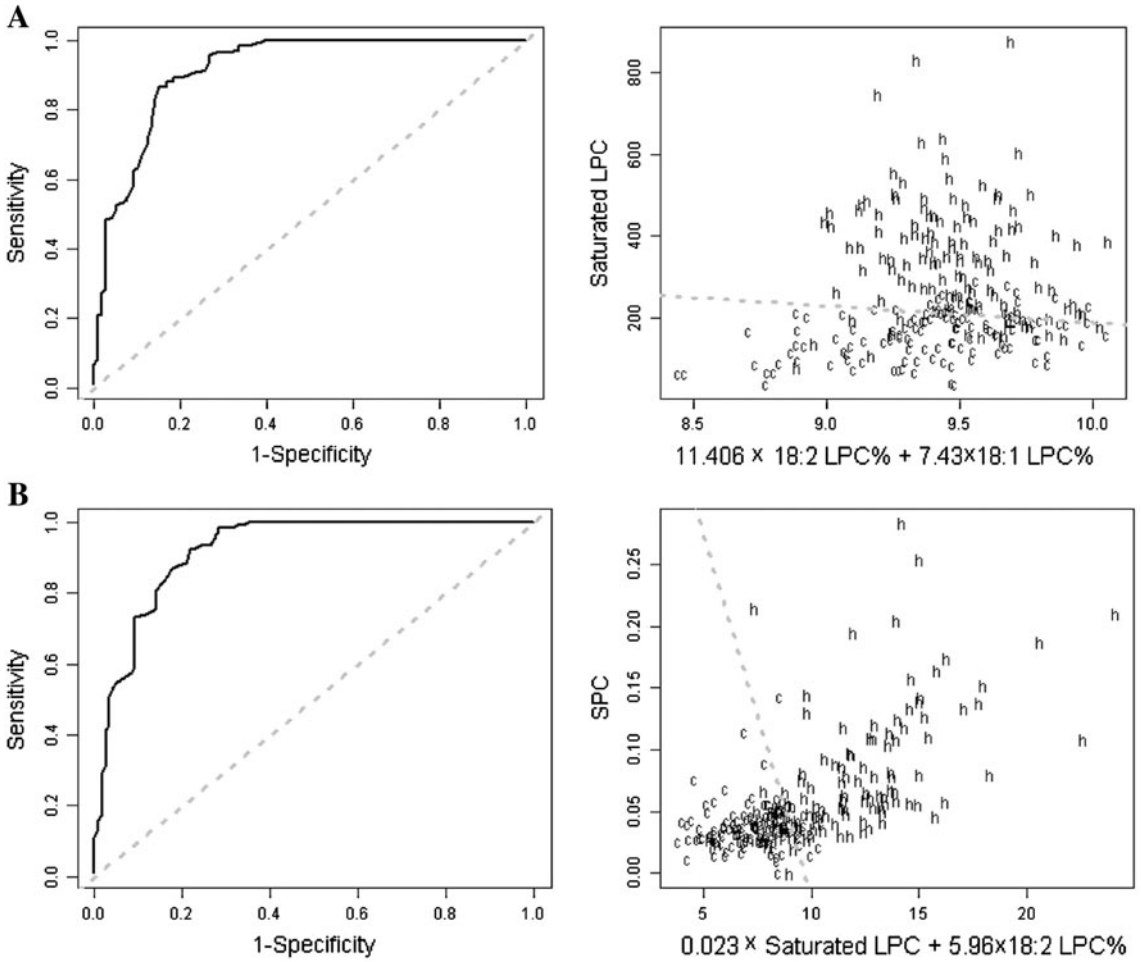

Fig. 2 Multiple phospholipid forms as markers for distinguishing CRC from healthy controls. **a** LPCs as markers for distinguishing CRC from healthy controls. Plasma levels of (11.406 × 18:2 LPC% + 7.43 × 18:1 LPC%) from CRC cases (c) and healthy controls (h) are plotted against total saturated LPC levels in each sample. The equation [11.406 × 18:2 LPC% + 7.43 × 18:1 LPC% + 0.024 × Saturated LPC = 14.58] is also plotted (*the line in the figure*). Subjects whose values fall to the down of this line are classified as diseased, while subjects whose data are to the up of the line would be classified as unaffected (healthy) subjects. A total of 120 CRC and 120 healthy controls are involved. The 18:2 LPC% and 18:1 LPC% in the formula are calculated as the percentages (with range from 0 to 100 %) of 18:2 LPC and 18:1 LPC with respect to unsaturated LPC. The ROC curve for healthy versus CRC cases is shown (*left*). **b** Multiple phospholipid forms as markers for distinguishing CRC from healthy controls. Plasma levels of (0.023 × saturated LPC + 5.96 × 18:2 LPC%) from CRC cases (*c*) and healthy controls (*h*) are plotted against SPC levels in each sample. The equation (0.023 × Saturated LPC + 5.96 × 18:2 LPC% + 17.06 × SPC = 9.66) is plotted (*the line in the figure*). Subjects whose values fall to the left of this line are classified as diseased, while subjects whose data are to the right of this line would be classified as unaffected (healthy) subjects. A total of 120 CRC and 120 healthy controls are involved. The 18:2 LPC% in the formula is calculated as the percentage (with range from 0 to 100 %) of 18:2 LPC with respect to unsaturated LPC. The ROC curve for healthy versus CRC cases is shown (*left*)

3.2 Discussion

Colorectal cancer is one of the most diagnosed cancers in the world, which is not only common in developed countries, but also has an increasing incidence in developing countries. Currently, colonoscopy is the gold standard for CRC screening, but there is no satisfactory and convenient method for the early detection of CRC (Sarfaty 2007). Identification of CRC biomarkers in plasma will be helpful to establish the convenient blood-based diagnostic techniques. In previous study, we have identified several LPC isoforms as potential markers to separate the healthy controls and CRC subjects in an American population (the majority of the subjects were Caucasians). In the present study, we have demonstrated that the plasma levels of major LPC species were significantly reduced in CRC subjects compared with in healthy controls in the Chinese population, which is highly consistent with the results obtained in the American population, indicating that LPCs are the common CRC biomarkers in plasma in the two totally different race populations. Nevertheless, when the exactly same formula, which was established with total saturated LPCs, 18:2 LPC and 18:1 LPC as markers in the previous

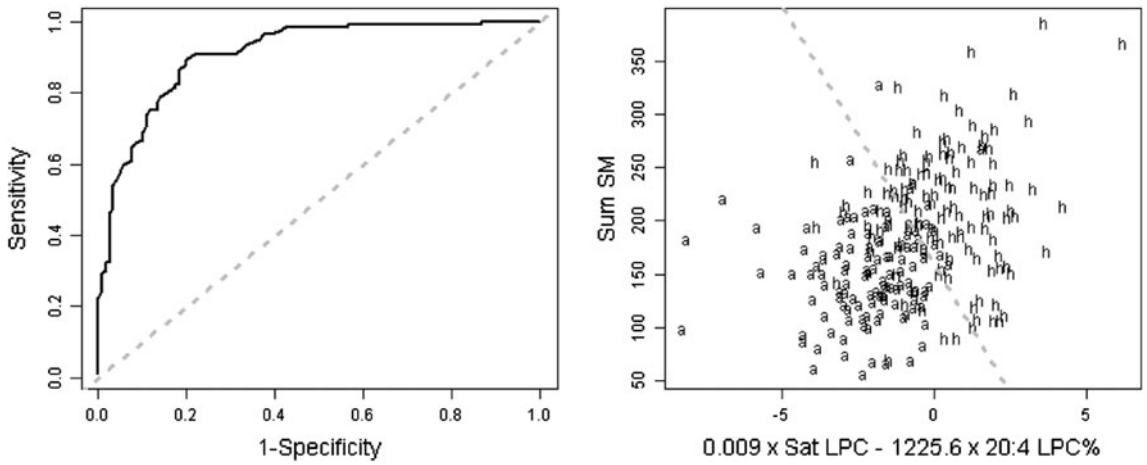

Fig. 3 Multiple phospholipid forms as markers for distinguishing AP from healthy controls. Plasma levels of (0.0093 × saturated LPC − 1225 × 20:4 LPC%) from AP cases (*a*) and healthy controls (*h*) are plotted against sum of 16:0 and 18:0 SM levels in each sample. The equation (0.0093 × saturated LPC − 1225 × 20:4 LPC% + 0.02 × sum of 16:0 and 18:0 SM = 3.3) is plotted (*the line in the figure*). Subjects whose values fall to the left of this line are classified as AP, while subjects whose data are to the right of this line would be classified as unaffected (healthy) subjects. A total of 120 AP and 120 healthy controls are involved. The 20:4 LPC% in the formula is calculated as the percentage (with range from 0 to 100 %) of 20:4 LPC with respect to unsaturated LPC. The ROC curve for healthy versus AP cases is shown (*left*)

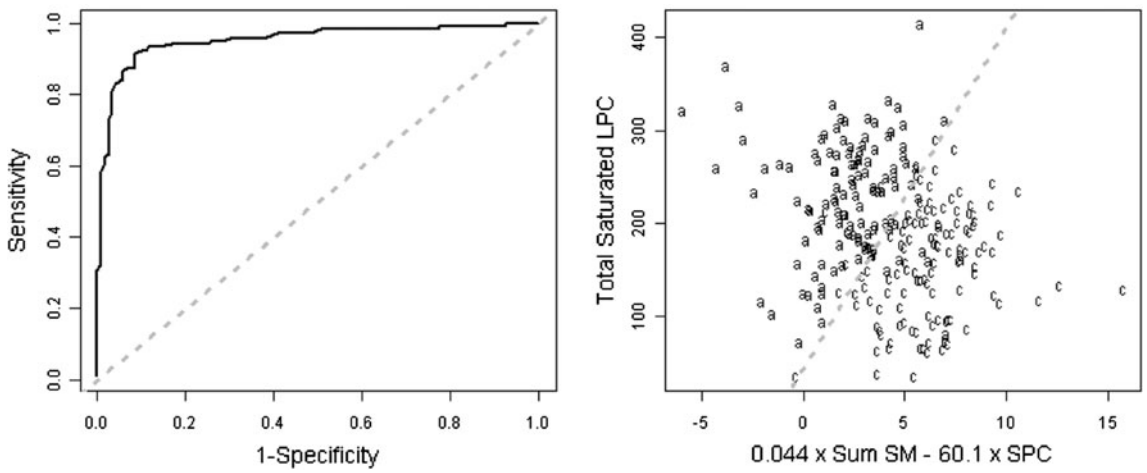

Fig. 4 Multiple phospholipid forms as markers for distinguishing CRC from AP cases. Plasma levels of (0.044 × sum of 16:0 and 18:0 SM − 60.1 × SPC) from AP cases (*a*) and CRC (*c*) are plotted against total saturated LPC levels in each sample. The equation (0.044 × sum of 16:0 and 18:0 SM − 60.1 × SPC − 0.027 total saturated LPC = −1.06) is also plotted (*the line in the figure*). Subjects whose values fall to the left of this line are classified as AP, while subjects whose data are to the right of this line would be classified as CRC subjects. A total of 120 AP and 120 CRC cases are involved. The ROC curve for AP versus CRC cases is shown (*left*)

USA study (Zhao et al. 2007), was applied for CRC detection in the Chinese population, the specificity was lower than the reported value. However, an adjusted formula with the same LPC makers used in the USA study was developed with 83 and 86 % of sensitivity and specificity, respectively, for separating CRC form healthy controls in the Chinese population. The reasons for such a discrepancy may include: (1) with the same trends decreasing from healthy control to CRC group, the plasma LPC levels are quantitatively different between the two ethnic populations (100 % Chinese vs. 83 % Caucasian, 6 % black, 7 % Asia and 4 % others) tested in these two studies and (2) colonoscopic confirmed healthy controls were used in this, but not in the previous study (Zhao et al. 2007). Such a derivation among populations may reflect metabolic differences in different races and countries and suggests that it is important to evaluate metabolomic biomarkers in different populations. It is well known that the incidence of CRC is dependent on both genetic as well as environmental factors, such as diet and lifestyle behavior. Therefore, it will be interesting to test whether the same formula can be used in the populations with similar genetic and environmental background in the future studies.

In the present study, additional phospholipids were analyzed, and an optimized formula was developed with saturated LPC, 18:2 LPC and SPC as makers, resulting a sensitivity and specificity of 88.3 and 80 %, respectively, to separate CRC form healthy controls. These performances are in fact better than most recently developed tests (Alexandrov et al. 2009; Han et al. 2008; Imperiale et al. 2004; Leman et al. 2007; Liu et al. 2009; Ran et al. 2008), and thus has a potential usage as the first line screening tool for CRC.

Most, if not all, of CRCs are believed to progress through the adenoma-to-carcinoma sequence. The adenoma–carcinoma transition represents an important window for early detection and intervention of CRC. Mortality from CRC could be largely preventable if the disease is detected early, thus the identification of such biomarkers for adenoma–carcinoma transition is critical for early CRC detection and intervention. In the present study, a group of choline-containing phospholipids, LPCs, SPC and SMs, were identified as plasma markers for CRC progression. One set of choline-containing phospholipid markers (saturated LPCs, 20:4 LPC, 16:0 SM, and 18:0 SM) was identified to distinguish AP from healthy controls with a sensitivity and specificity 89 and 80 % respectively, and another set of markers (SPC, 16:0 SM, 18:0 SM, and saturated LPCs) to further separate CRC from AP cases with a sensitivity and specificity of 90 and 92.5 %. These data indicate that the choline-containing phospholipids makers are clinical useful to detect precancerous AP and AP-CRC transition.

In the present study, a very few samples fell on the classification boundary. These samples were classified as the diseased when distinguishing AP or CRC subjects from the healthy controls, and as CRC when separating CRC from AP subjects. Although male subjects were over-represented (63 %) in this study, the sensitivity and specificity of the established models were not biased towards male subjects, except the specificity for separating AP from CRC was better in male subjects (95.5 % for males and 86.5 % for females) (Supplementary Table 3).

All of the CRC progression markers identified in this study are choline-containing phospholipids. The choline-containing phospholipids and their metabolites have been recognized as important cell signaling molecules in cancer progression. Here, we have observed a gradual decrease in the plasma levels of major LPC species from healthy controls, AP to CRC subjects (Table 3; Fig. 1). LPC is a major plasma lipid component and transports fatty acids and choline to tissues. LPC regulates various cell activities and plays an important role in atherosclerosis and inflammatory diseases (Xu 2002). LPC can be converted into LPA, a bioactive phospholipid mediator that evokes a vast variety of physiological and pathological actions, including cell proliferation and differentiation, cell to cell interactions, cytoskeletal rearrangement and tumorigenesis (Mills and Moolenaar 2003). LPA has been identified as a biomarker of ovarian cancer (Xu et al. 1998), and LPA can enhance CRC cell proliferation and invasion (Lee and Yun 2010; Yang et al. 2005). It has been reported recently that the ATX-positive mast cells are detected in early CRC tissue by immunohistochemical staining (Kazama et al. 2010), suggesting that local production of LPA from LPC may contribute to the CRC progression. In the present study, we did not observe the increase of LPA plasma levels in CRC patients compared with in the healthy controls, which may due to the effective degradation of LPA by lysophosphatidic acid phosphatases (LPPs) in plasma after its generation (Noguchi et al. 2009). The SPC levels are decreased obviously in CRC patients' plasma (Table 3; Supplementary Fig. 1a), suggesting that SPC is another biomarker for CRC. SPC is a potentially important lipid mediator with cell type-specific functions in major tissues, such as heart, blood vessels, skin, brain and immune system (Nixon et al. 2008). SPC can be converted into S1P, an oncogenic lipid that promotes tumor growth, migration and metastasis. The S1P receptor inhibitors are tested to be anti-cancer drugs (Nixon et al. 2008; Pyne and Pyne 2010). The plasma SM levels were down-regulated in AP subjects compared with in healthy control and CRC groups (Table 3; Supplementary Fig. 1b–d), indicating that SM is a valuable maker for AP during the CRC progression. SM, a prominent phospholipid component of cell membranes, has been involved in diverse functions beyond its role in membrane structural organization. SM can be metabolized to ceramide, a bioactive lipid in its own right, but also a precursor molecule to other signaling lipids and a central hub of the sphingolipid network (Milhas et al. 2010). The molecular mechanism of biomarkers is very important for the metabolomics-based cancer diagnosis. The functions of these choline-containing phospholipids and their metabolic mechanisms during the CRC development remain to be further studied.

In the present study, we have identified sets of plasma choline-containing phospholipids, including LPC, SPC and SM, as biomarkers for CRC and/or AP-CRC progression detection. So far, the sensitivity and specificity of plasma biomarkers, including the choline-containing phospholipids in our present study and other markers reported previously, are still lower than those of colonoscopy in CRC detection. However, the degree of patient compliance to colonoscopy is low due to the inconvenience and discomfort. Thus, the convenient and minimal invasive blood marker test will be very useful in the population-based CRC screening.

4 Conclusions

The results of this study demonstrate that the plasma levels of a group of choline-containing phospholipids can be used to distinguish healthy, AP, and CRC cases, and thus be clinically useful as the CRC early detection and/or CRC progression markers. These markers may also be functionally involved in the CRC development.

Acknowledgments This work was supported in part by the National Nature Science Foundation of China (No. 31070714), the National High Technology Research and Development Program, China (No. 2006AA02Z4A6), the Science and Technology Program of Beijing Municipality (No. Z0006329000191), the Fundamental Research Funds for the Central Universities (105566GK), the Beijing NOVA Program (No. 2005B47) to Junjie Zhang, and an Overseas Research Grant from Indiana University to Yan Xu.

Conflict of interest There are no conflicts of interest.

References

Alexandrov, J. D., Mertens, B., Deelder, A. M., Rob, A. E. M., Tollenaar, P. M., & Thiele, H. (2009). Biomarker discovery in MALDI-TOF serum protein profiles using discrete wavelet transformation. *Bioinformatics, 25*, 643–649.

Berger, B. M., Robison, L., & Glickman, J. (2003a). Colon cancer-associated DNA mutations: Marker selection for the detection of proximal colon cancer. *Diagnostic Molecular Pathology, 12*, 187–192.

Berger, B. M., Vucson, B. M., & Ditelberg, J. S. (2003b). Gene mutations in advanced colonic polyps: Potential marker selection for stool-based mutated human DNA assays for colon cancer screening. *Clinical Colorectal Cancer, 3*, 180–185.

Brenner, H., Hoffmeister, M., & Haug, U. (2008). Family history and age at initiation of colorectal cancer screening. *American Journal of Gastroenterology, 103*, 2326–2331.

Calistri, D., Rengucci, C., Bocchini, R., Saragoni, L., Zoli, W., & Amadori, D. (2003). Fecal multiple molecular tests to detect colorectal cancer in stool. *Clinical Gastroenterology and Hepatology, 1*, 377–383.

Center, M. M., Jemal, A., Smith, R. A., & Ward, E. (2009). Worldwide variations in colorectal cancer. *CA Cancer Journal for Clinicians, 59*(6), 366–378.

Chen, W. D. Z., Han, J., Skoletsky, J., Shuber, A., Markowitz, S. D., et al. (2005). Detection in fecal DNA of colon cancer-specific methylation of the nonexpressed vimentin gene. *Journal of the National Cancer Institute, 97*, 1124–1132.

Dong, S. M., Traverso, G., Johnson, C., Vogelstein, B., Jen, J., et al. (2001). Detecting colorectal cancer in stool with the use of multiple genetic targets. *Journal of the National Cancer Institute, 93*, 858–865.

Doolittle, B. R., Emanuel, J., Tuttle, C., & Costa, J. (2001). Detection of the mutated K-Ras biomarker in colorectal carcinoma. *Experimental and Molecular Pathology, 70*, 289–301.

Fuchs, B., & Schiller, J. (2009). Lysophospholipids: Their generation, physiological role and detection. Are they important disease markers? *Mini Reviews in Medicinal Chemistry, 9*, 368–378.

Graler, M. H. (2010). Targeting sphingosine 1-phosphate (S1P) levels and S1P receptor functions for therapeutic immune interventions. *Cellular Physiology and Biochemistry, 26*, 79–86.

Guittet, L., Launoy, G., Mohammed, F., & Lieberman, D. A. (2010). Screening for colorectal cancer. *New England Journal of Medicine, 362*, 84–85.

Han, M., et al. (2008). Novel blood-based, five-gene biomarker set for the detection of colorectal cancer. *Clinical Cancer Research, 14*, 455–460.

Harrell, F. E., Jr, Lee, K. L., & Mark, D. B. (1996). Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine, 15*, 361–387.

Huang, Z., Huang, D., Ni, S., Peng, Z., Sheng, W., & Du, X. (2010). Plasma microRNAs are promising novel biomarkers for early detection of colorectal cancer. *International Journal of Cancer, 127*, 118–126.

Ikeda, A., Nishiumi, S., Shinohara, M., Yoshida, M., et al. (2011). Serum metabolomics as a novel diagnostic approach for gastrointestinal cancer. *Biomedical Chromatography, 26*(5), 548–558.

Imperiale, T. F., Ransohoff, D. F., Itzkowitz, S. H., Turnbull, B. A., & Ross, M. E. (2004). Fecal DNA versus fecal occult blood for colorectal-cancer screening in an average-risk population. *New England Journal of Medicine, 351*, 2704–2714.

Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., & Forman, D. (2011). Global cancer statistics. *CA Cancer Journal for Clinicians, 61*(2), 69–90.

Jemal, A., Center, M. M., Ward, E., & Thun, M. J. (2009). Cancer occurrence. *Methods in Molecular Biology, 471*, 3–29.

Jemal, A., Siegel, R., Ward, E., Hao, Y., et al. (2008). Cancer statistics, 2008. *CA Cancer Journal for Clinicians, 58*, 71–96.

Jemal, A., Siegel, R., Xu, J., & Ward, E. (2010). Cancer statistics, 2010. *CA Cancer Journal for Clinicians, 60*, 277–300.

Kazama, S., Kitayama, J., Aoki, J., Mori, K., & Nagawa, H. (2010). Immunohistochemical detection of autotaxin (ATX)/lysophospholipase D (lysoPLD) in submucosal invasive colorectal cancer. *Journal of Gastrointestinal Cancer, 42*(4), 204–211.

Kim, H. J., Kang, H. J., Lee, S. T., et al. (2009). Identification of S100A8 and S100A9 as serological markers for colorectal cancer. *Journal of Proteome Research, 8*, 1368–1379.

Kountouras, J., Boura, P., & Lygidakis, N. J. (2000). New concepts of molecular biology for colon carcinogenesis. *Hepato-Gastroenterology, 47*, 1291–1297.

Lee, S. J., & Yun, C. C. (2010). Colorectal cancer cells: Proliferation, survival and invasion by lysophosphatidic acid. *International Journal of Biochemistry & Cell Biology, 42*, 1907–1910.

Leman, E. S., et al. (2007). Initial analyses of colon cancer-specific antigen (CCSA)-3 and CCSA-4 as colorectal cancer-associated serum markers. *Cancer Research, 67*, 5600–5605.

Lieberman, D. A. (2009). Screening for colorectal cancer. *New England Journal of Medicine, 361*, 1179–1187.

Liu, W., Wang, P., Li, Z., Xu, W., et al. (2009). Evaluation of tumour-associated antigen (TAA) miniarray in immunodiagnosis of colon cancer. *Scandinavian Journal of Immunology, 69*, 57–63.

Loitsch, S. M., Shastri, Y., & Stein, J. (2008). Stool test for colorectal cancer screening: It's time to move! *J Clinical Laboratory, 54*, 473–484.

Markowitz, S. D., & Bertagnolli, M. M. (2009). Molecular basis of colorectal cancer. *New England Journal of Medicine, 361*, 2449–2460.

Milhas, D., Clarke, C. J., & Hannun, Y. A. (2010). Sphingomyelin metabolism at the plasma membrane: Implications for bioactive sphingolipids. *FEBS Letters, 584*, 1887–1894.

Mills, G. B., & Moolenaar, W. H. (2003). The emerging role of lysophosphatidic acid in cancer. *Nature Reviews Cancer, 3*, 582–591.

Moshkowitz, M., & Arber, N. (2005). Differences in incidence and distribution of colorecetal cancer among races and ethnic societies: Lifestyle, genes or both? *Digestion, 72*, 219–222.

Muller, O. (2003). Identification of colon cancer patients by molecular diagnosis. *Digestive Diseases, 21*, 315–319.

Nixon, G. F., Mathieson, F. A., & Hunter, I. (2008). The multifunctional role of sphingosylphosphorylcholine. *Progress in Lipid Research, 47*, 62–75.

Noguchi, K., Herr, D., Mutoh, T., & Chun, J. (2009). Lysophosphatidic acid (LPA) and its receptors. *Current Opinion in Pharmacology, 9*, 15–23.

Peyruchaud, O. (2009). Novel implications for lysophospholipids, lysophosphatidic acid and sphingosine 1-phosphate, as drug targets in cancer. *Anti-Cancer Agents in Medicinal Chemistry, 9*, 381–391.

Pyne, N. J., & Pyne, S. (2010). Sphingosine 1-phosphate and cancer. *Nature Reviews Cancer, 10*, 489–503.

Ran, Y., Hu, H., Zhou, Z., Yu, L., Sun, L., Pan, J., et al. (2008). Profiling tumor-associated autoantibodies for the detection of colon cancer. *Clinical Cancer Research, 14*, 2696–2700.

Rivera, R., & Chun, J. (2008). Biological effects of lysophospholipids. *Reviews of Physiology Biochemistry and Pharmacology, 160*, 25–46.

Sarfaty, M., & Wender, R. (2007). How to increase colorectal cancer screening rates in practice. *CA Cancer Journal for Clinicians, 57*, 354–366.

Srivastava, S., Verma, M., & Henson, D. E. (2001). Biomarkers for early detection of colon cancer. *Clinical Cancer Research, 7*(5), 1118–1126.

Sung, J. J., Lau, J. Y., Goh, K. L., & Leung, W. K. (2005). Increasing incidence of colorectal cancer in Asia: Implications for screening. *Lancet Oncol, 6*, 871–876.

Tierney, R. P., Ballantyne, G. H., & Modlin, I. M. (1990). The adenoma to carcinoma sequence. *Surgery, Gynecology & Obstetrics, 171*, 81–94.

Umar, A., & Greenwald, P. (2009). Alarming colorectal cancer incidence trends: A case for early detection and prevention. *Cancer Epidemiology, Biomarkers and Prevention, 18*, 1672.

Wild, N., Andres, H., Rollinger, W., Krause, F., et al. (2010). A combination of serum markers for the early detection of colorectal cancer. *Clinical Cancer Research, 16*, 6111.

Wymann, M. P., & Schneiter, R. (2008). Lipid signalling in disease. *Nature Reviews Molecular Cell Biology, 9*, 162–176.

Xu, Y. (2002). Sphingosylphosphorylcholine and lysophosphatidylcholine: G protein-coupled receptors and receptor-mediated signal transduction. *Biochimica et Biophysica Acta, 1582*, 81–88.

Xu, Y., Shen, Z., Wiper, D. W., Wu, M., et al. (1998). Lysophosphatidic acid as a potential biomarker for ovarian and other gynecologic cancers. *JAMA, 280*, 719–723.

Yang, M., Zhong, W. W., Srivastava, N., Slavin, A., et al. (2005). G protein-coupled lysophosphatidic acid receptors stimulate proliferation of colon cancer cells through the β-catenin pathway. *Proceedings of the National Academy of Sciences of the USA, 102*, 6027–6032.

Zhao, Z., Xiao, Y., Elson, P., Tan, H., et al. (2007). Plasma lysophosphatidylcholine levels: Potential biomarkers for colorectal cancer. *Journal of Clinical Oncology, 25*, 2696–2701.