

## Metabolomic NMR Fingerprinting to Identify and Predict Survival of Patients with Metastatic Colorectal Cancer

Ivano Bertini<sup>1,2</sup>, Stefano Cacciatore<sup>1,2</sup>, Benny V. Jensen<sup>3</sup>, Jakob V. Schou<sup>3</sup>, Julia S. Johansen<sup>3</sup>, Mogens Kruhøffer<sup>4</sup>, Claudio Luchinat<sup>1,2</sup>, Dorte L. Nielsen<sup>3</sup>, and Paola Turano<sup>1</sup>

### Abstract

Earlier detection of patients with metastatic colorectal cancer (mCRC) might improve their treatment and survival outcomes. In this study, we used proton nuclear magnetic resonance (<sup>1</sup>H-NMR) to profile the serum metabolome in patients with mCRC and determine whether a disease signature may exist that is strong enough to predict overall survival (OS). In 153 patients with mCRC and 139 healthy subjects from three Danish hospitals, we profiled two independent sets of serum samples in a prospective phase II study. In the training set, <sup>1</sup>H-NMR metabolomic profiling could discriminate patients with mCRC from healthy subjects with a cross-validated accuracy of 100%. In the validation set, 96.7% of subjects were correctly classified. Patients from the training set with maximally divergent OS were chosen to construct an OS predictor. After validation, patients predicted to have short OS had significantly reduced survival (HR, 3.4; 95% confidence interval, 2.06–5.50;  $P = 1.33 \times 10^{-6}$ ). A number of metabolites concurred with the <sup>1</sup>H-NMR fingerprint of mCRC, offering insights into mCRC metabolic pathways. Our findings establish that <sup>1</sup>H-NMR profiling of patient serum can provide a strong metabolomic signature of mCRC and that analysis of this signature may offer an independent tool to predict OS. *Cancer Res*; 72(1); 356–64. ©2011 AACR.

### Introduction

Pattern recognition technologies in the *omics* world have been used for the diagnosis and prognosis of several tumor types using a variety of experimental platforms (1–3). Metabolomics is a post-genomic research field concerned with the study of metabolic profiles in biologic samples such as isolated cells (4), tissues (5) or body fluids (6–10). In recent years, metabolomic studies have provided a significant contribution to the detailed understanding of the biochemical network and pathways in oncology (11), highlighting the potential of metabolomic analyses of blood samples for the detection of cancers, including epithelial ovarian (12), leukemia (13), oral (14), breast (15, 16), prostate (11), liver (17) as well as colorectal cancer (CRC; refs. 18–25). CRC is the second most common cause of cancer-related deaths in Europe (26) and the third most common cause of cancer-related deaths in the United States (27). The impact of colorectal carcinogenesis on the metabolic profile of tissues

and serum has been tested by mass spectrometry (MS; refs. 18–21), nuclear magnetic resonance (NMR; refs. 22, 23), or a combination of the two analytic tools (24).

Most of the symptoms associated with CRC do not manifest themselves until late in the progression of the disease and many patients have metastases at the time of initial diagnosis (27). The overall survival (OS) of a patient is reduced considerably when the disease is diagnosed at a stage where it has spread to distant sites. In the past decade, the OS of patients with metastatic CRC (mCRC) has improved because of new combinations of chemotherapy, including 5-fluorouracil (5-FU), irinotecan, and oxaliplatin (28–31). The introduction of the new targeted therapies directed against the epidermal growth factor receptor (EGFR) has further increased response rates (32) in some patients. Nevertheless, early selection of patients with mCRC for optimal treatment with these biologics is very difficult. For example, *KRAS* mutations occur in 35% to 45% of patients with mCRC and predict poor response to EGFR-targeted therapy with cetuximab and panitumumab (33, 34). The remaining patients with mCRC are *KRAS* wild-type and approximately 25% to 40% of them are nonresponders and do not benefit from this treatment (33, 34). OS HRs for mCRC have been evaluated from *KRAS* mutations (HR, 2.4; ref. 35), Eastern Cooperative Oncology Group (ECOG) performance status (PS; HR, 1.78–1.88; refs. 29, 36) and serum levels of alkaline phosphatase (HR, 1.71; ref. 36), bilirubin (HR, 1.89; ref. 29), and lactate dehydrogenase (HR, 2.12; ref. 29). Somewhat higher HR values for primary CRC are obtained from monocyte (HR, 2.86) and neutrophil (HR, 2.90) count (37). Novel noninvasive, sensitive, and inexpensive indicators for mCRC OS are therefore highly desirable.

**Authors' Affiliations:** <sup>1</sup>CERM and Department of Chemistry, University of Florence; <sup>2</sup>FiorGen Foundation, Florence, Italy; <sup>3</sup>Department of Oncology, Herlev Hospital, University of Copenhagen, Copenhagen; and <sup>4</sup>AROS Applied Biotechnology A/S, Aarhus N, Denmark

**Note:** Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org/>).

**Corresponding Author:** Ivano Bertini, CERM, University of Florence, Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Italy. Phone: 39-055-4574270; Fax: 39-055-4574271; E-mail: ivanobertini@cerm.unifi.it

doi: 10.1158/0008-5472.CAN-11-1543

©2011 American Association for Cancer Research.

In this study, conducted within the frame of the biomarker discovery activities of the FP7 project SPIDIA ([www.spidia.eu](http://www.spidia.eu)), we conducted  $^1\text{H-NMR}$  profiling of serum samples collected from 153 Danish patients with mCRC before third-line treatment with cetuximab and irinotecan (Table 1 and Supplementary Table S1) and of serum samples from 139 healthy subjects (Table 1). The statistical model generated from  $^1\text{H-NMR}$  profiles of serum samples of subjects from 2 of the 3 hospitals can robustly discriminate healthy subjects ( $n = 96$ ) from patients with mCRC ( $n = 45$ ) with 100% cross-validated accuracy. General applicability of the resultant classifier was successfully validated in an independent set of 43 healthy subjects and 108 patients with mCRC from the third hospital. The capability of the  $^1\text{H-NMR}$  profiles to predict OS after start of treatment with cetuximab and irinotecan using pretreatment serum samples was tested on a set of 20 subjects of the training set with maximally divergent OS. The classifier was validated on the independent set of 108 patients with mCRC. Furthermore, through multivariate analysis, a number of serum metabolites were identified whose levels were significantly different in patients with mCRC as compared to healthy subjects, as well

as between patients with short and long OS. These metabolites can provide hints both to define new biomarkers and to better understand the biochemistry involved in mCRC.

## Materials and Methods

### Patients with mCRC

Serum samples were collected from a cohort of 181 Danish patients with mCRC resistant to 5-FU, oxaliplatin, and irinotecan. Details on patient selection are provided in Supplementary Methods. The patients were included in a prospective phase II study evaluating cetuximab and irinotecan every second week. The patients were included at 3 Hospitals in Denmark during the period October 17, 2006, to October 3, 2008. The patients were treated until disease progression with third-line irinotecan (180 mg/m<sup>2</sup> of body surface area on day 1 of each 14-day period during the study) and cetuximab (500 mg/m<sup>2</sup> of body-surface area every second week) independent of their *KRAS* status. Serum samples used for metabolomics were collected before initiating the treatment. In the present study, we did not include any patient with hereditary CRC.

**Table 1.** Patient demographic and clinical characteristics in the training and the validation sets

Demographic or clinical characteristics	Training set ( $n = 141$ )		Validation set ( $n = 151$ )	
	Healthy subjects, $n$ (%)	mCRC patients, $n$ (%)	Healthy subjects, $n$ (%)	mCRC patients, $n$ (%)
Hospital				
Herlev			43 (100.0)	108 (100.0)
Aalborg	96 (100.0)	25 (55.6)		
Odense		20 (44.4)		
Sex				
Male	50 (52.1)	31 (68.9)	21 (48.8)	65 (60.2)
Female	46 (47.9)	14 (31.1)	22 (51.2)	43 (39.8)
Age, y				
Median	42	64	40	63
Range	30–66	49–78	26–69	46–87
Body mass index <sup>a</sup>				
Underweight		0 (0.0)		7 (6.6)
Normal weight		23 (51.1)		48 (45.3)
Overweight		17 (37.8)		35 (33.0)
Obese		5 (11.1)		16 (15.1)
ECOG-PS				
0		17 (37.8)		60 (55.6)
1		18 (40.0)		34 (31.5)
2		10 (22.2)		14 (13.0)
OS, mo				
Median		12.1		11.2
Range		1.5–31.9		0.5–45.4
<i>KRAS</i> <sup>b</sup>				
Wild-type		29 (67.4)		65 (63.7)
Mutated		14 (32.6)		37 (36.3)

Abbreviation:  $n$ , no. of subjects.

<sup>a</sup>Data not available for 2 patients of the validation set.

<sup>b</sup>Data not available for 2 patients of the training set and 6 patients of the validation set.

Patients were included in the "Phase II study of treatment with cetuximab and irinotecan administered biweekly to irinotecan-resistant patients with metastatic colorectal cancer—efficacy and biologic markers". Inclusion criteria were as follows: (i) patients with histologically verified adenocarcinoma of the colon or rectum with nonresectable or metastatic disease; (ii) patients with measurable disease as per Response Evaluation Criteria in Solid Tumors (RECIST; ref. 38); (iii) patients with progressive disease following either oxaliplatin- or irinotecan-based treatment; (iv) patients with irinotecan-resistant disease—defined as progressive disease after at least 6 weeks of treatment with irinotecan-based treatment or within 6 months following discontinuation of such treatment. Irinotecan-based treatment should not necessarily be given immediately before inclusion, but oxaliplatin-based treatment may be; (v) ECOG-PS less than 3; (vi) an expected survival time of at least 3 months; (vii) neutrophil count  $\geq 1.5 \times 10^9/L$  and thrombocytes  $\geq 100 \times 10^9/L$ ; (viii) normal liver function with bilirubin  $< 1.5 \times$  upper limit of normal and aspartate aminotransferase/alanine aminotransferase  $< 5 \times$  upper limit of normal; (ix) all should be exposed to oxaliplatin-based treatment; and (x) The signed informed consent as per requirements of the Scientific Ethical Committees.

The patients were followed until death or July 7, 2010 (39). The study (including biomarker analysis) was approved by the Regional Ethics Committee (VEK ref. KA-20060094). For 95% of the patients, mutations on the *KRAS* gene were determined, as reported in Supplementary Methods. Further clinical data about the study will be published elsewhere by B.V. Jensen. One hundred sixty-eight pretreatment serum samples were available for  $^1H$ -NMR profiling.

### Controls

Serum samples were collected from 96 Danish healthy blood donors and 43 healthy subjects (hospital staff; Table 1).

### Serum sample preparation

Blood was collected into standard blood collection tubes and allowed to clot at room temperature for 30 to 120 minutes before centrifugation ( $1,500 \times g$  for 10 minutes at  $4^\circ C$ ). Serum was aliquoted and stored at  $-80^\circ C$  in Greiner cryogenic vials. This procedure is compatible with the standard operating procedures (SOP) for metabolomic-grade serum samples recently defined (40).

### NMR sample preparation

Frozen serum samples were thawed at room temperature and shaken before use. A total of 300  $\mu L$  of buffer [70 mmol/L  $Na_2HPO_4$ ; 20% (v/v)  $H_2O$ ; 6.15 mmol/L  $NaN_3$ ; 6.64 mmol/L sodium trimethylsilyl [2,2,3,3- $^2H_4$ ] propionate; pH 7.4] was added to 300  $\mu L$  of each serum sample, and the mixture was homogenized by vortexing for 30 seconds. A total of 450  $\mu L$  of this mixture was transferred into a 4.25-mm NMR tube for analysis.

### NMR spectra

Carr-Purcell-Meiboom-Gill (CPMG)  $^1H$ -NMR spectra were acquired to observe metabolites signals and suppress reso-

nances arising from high-molecular weight molecules and J-resolved spectra projections (pJRES) to help identify biomarkers. Further details are provided in Supplementary Methods. The quality of the serum samples was assessed by visual inspection of the resulting NMR spectra according to the criteria defined in the study of Bernini and colleagues (40). Thirteen (8%) serum samples were not included in the statistical analysis because of the insufficient quality of the resulting NMR spectra that contained very broad features attributable to high-molecular weight impurities. Two patients with mCRC younger than 45 years were excluded as hereditary non-polyposis CRC is typically associated with increased inflammatory response and different prognosis than sporadic CRC. Demographic and clinical characteristics of the patient cohort studied by NMR are provided in Table 1 and Supplementary Table S1.

### Statistical analysis

Multivariate data analyses were conducted on processed data by combining established methods (6, 41). Projection to latent structure (PLS) was applied on processed data for dimension reduction using the SIMPLS algorithm as implemented in the R library "plsgenomics"; the support vector machines (SVM) method was used for data classification using the "libsvm" module of the R library "e1071." Canonical correlation analysis (CA) was conducted using the standard R function "cancor."

To assess the prediction ability of the model, a 10-fold cross-validation was conducted by generating splits with a ratio 1:9 for the data set, that is by removing 10% of samples prior to any step of the statistical analysis, including PLS component selection. Parameter selection (best number of components for PLS, kernel type, cost of constraints violation for SVM) was carried out by means of 5-fold cross-validation on the remaining 90%. The whole procedure was repeated 100 times. Data classification using SVM was conducted by applying the classifier on PLS scores. Accuracy, sensitivity, and specificity were calculated using standard definitions. To estimate the classification/feature selection performance of our predictors, the statistical significance of all observations was further assessed through permutation tests.

Score contribution weights were examined to assess which metabolites were most responsible for the observed discrimination between groups (42). The relative concentrations of each metabolite were calculated by integrating the signals in the spectra. Statistical significance was assessed using the nonparametric Kruskal-Wallis test. A  $P < 0.05$  was considered statistically significant. Groupwise comparisons of distributions of clinical and demographic data were conducted with the Fisher exact test for the categorical variables and Kruskal-Wallis test for the continuous variables.

Kaplan-Meier survival curves were created using the R library "survival." The Wald test was used to calculate the statistical significance ( $P$ ) of the differences between survival curves. Prognostic factors for OS were analyzed using the Cox proportional hazard regression.

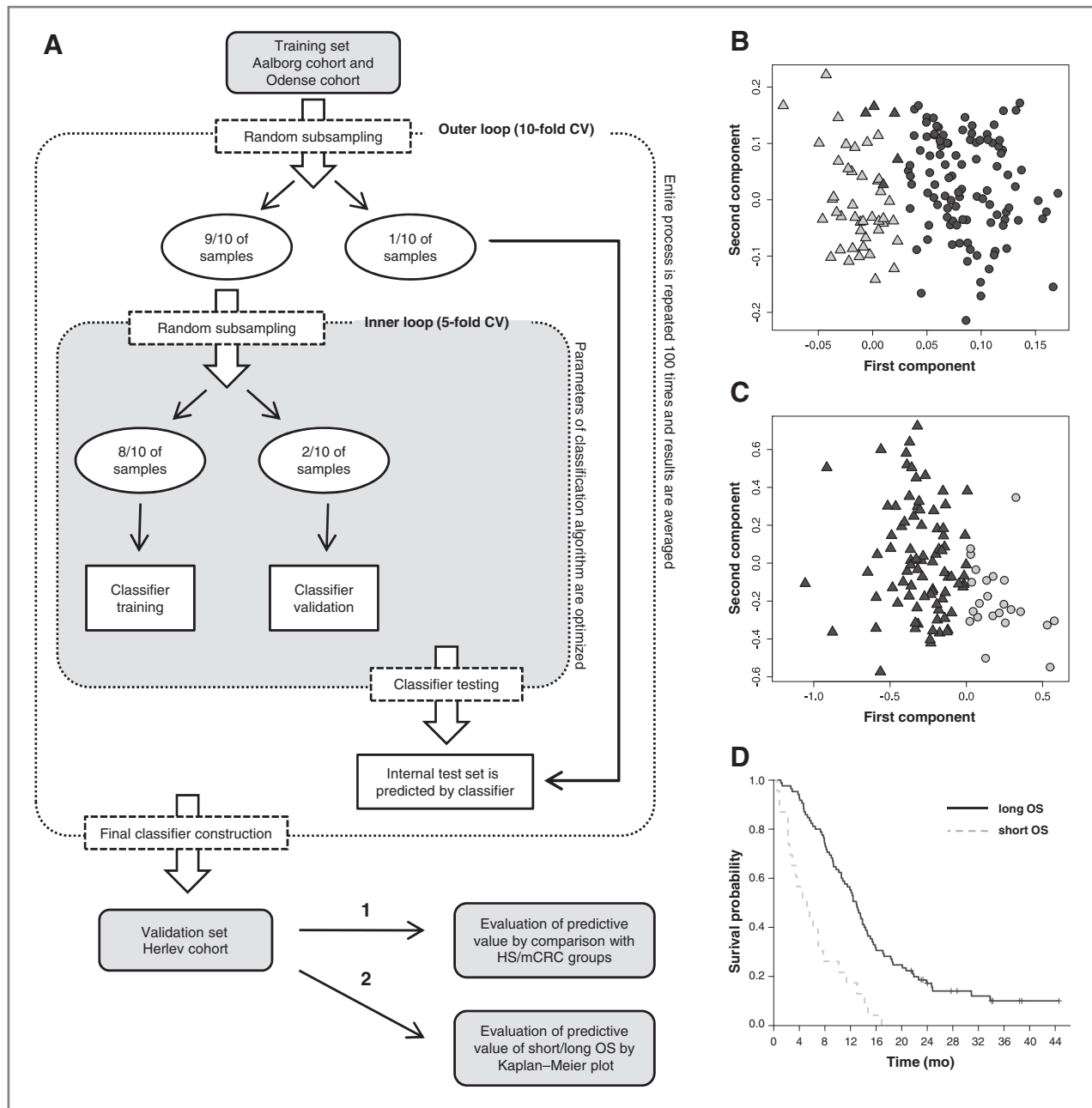
All calculations were made using R (41) scripts developed in-house.

## Results

### Fingerprinting of mCRC

A schematic diagram of the present strategy for constructing and validating a metabolic signature in patients with mCRC before third-line treatment with cetuximab and irinotecan is shown in Fig. 1A. Patients with mCRC and healthy subjects

(Table 1) were discriminated by using SVM on PLS scores (PLS-SVM) calculated on  $^1\text{H-NMR}$  serum spectra of samples from Aalborg and Odense hospitals ( $n = 141$ ). Accuracy, sensitivity, and specificity results for the training set were averaged on a total of 1,000 random subsamplings (100-times repeated 10-fold cross-validation). This procedure consists of two nested loops of cross-validation (Fig. 1A): an outer loop (10-fold



**Figure 1.** A, schematic overview of the procedure for the generation and validation of the  $^1\text{H-NMR}$  metabolomic profile classifier. Different approaches were used for the validation of the signature of mCRC (1) and for the prediction of OS (2). B, PLS-CA clustering for the healthy subjects (triangles) and patients with mCRC (circles) of the validation set based on the model built on the NMR spectra of the training set. Triangles are predicted as healthy subjects and circles as mCRC. C, PLS-CA clustering for the long OS group (triangles) and short OS group (circles) of the validation set based on the model built on the NMR spectra of the training set. D, Kaplan-Meier curves showing survival probability based on  $^1\text{H-NMR}$  profiling predictive model ( $P = 1.33 \times 10^{-6}$ ). CV, cross-validation; HS, healthy subjects.



cross-validation) for the estimation of the prediction accuracy and an inner loop (5-fold cross-validation) for the optimization of the parameters of the classifier training procedure. Classification performances on the validation set were calculated from the model built on the training set with a 10-fold parameter optimization procedure. Accuracy, sensitivity, and specificity of the PLS-SVM classifier are given in Table 2; excellent classification accuracies were found both for the training (100.0%) and the validation sets (96.7%). On a descriptive level, CA on PLS score (PLS-CA) was used to ascertain whether the spectral patterns of serum of patients with mCRC carry unique features with respect to those of healthy subjects. The resulting plot is shown in Fig. 1B: a very good discrimination is clearly apparent.

The different age distribution between the patients with mCRC and healthy subjects could be a possible source of inhomogeneity in the training set ( $P < 0.01$ ). Therefore, the same statistical analysis described above was repeated on a subgroup aged between 45 and 65 years ( $P = 0.23$ ), providing a classification accuracy of 100.0% in the training set and 92.1% in the validation set (Table 2). These results show that the difference in age distribution in the two groups does not affect significantly the classification accuracy values.

PLS-SVM cross-validation analysis was also repeated on a data set built with 96 healthy subjects and the 17 patients with mCRC and ECOG-PS of 0, obtaining a classification accuracy of 100.0% on the training set and 86.4% on the validation set (Table 2), which indicates that a clear metabolic signature of the disease exists even in the serum of patients with mCRC and ECOG-PS of 0.

### Prediction of survival

The strong metabolic signature of mCRC observed in the present study suggests that even within patients with mCRC, there may be statistically significant differences in the metabolic profile. This could be related to OS. The prognostic classification capabilities of  $^1\text{H-NMR}$  profiling were therefore tested. A model was built by selecting from the training set a subset of 20 patients with mCRC with maximally divergent OS: that is, the 10 patients with the shortest OS and the 10 patients with longest OS were studied. The PLS-SVM classifier, similar

to that used in the previous analysis, was used to assess whether patients with short or long OS could be identified by  $^1\text{H-NMR}$  profiling. The resulting accuracy was 78.5% (Table 2), suggesting that a signature for the progression state of the tumor is present in the  $^1\text{H-NMR}$  profiles. The statistical confidence of the prediction performance of the PLS-SVM classifier on the multimode data set with 10-fold cross-validation evaluation was investigated using a permutation test. Permutation test results (number of permutations = 1,000) showed that the accuracy difference between PLS-SVM and a random classifier was statistically significant ( $P = 0.026$ ). PLS-CA scores are shown in Fig. 1C.

The patients with mCRC from the validation set were examined by using the model built on the basis of the  $^1\text{H-NMR}$  profiles of the serum samples of the training set of patients with extreme values of OS, as described before. The model identified 85 (78.7%) patients as belonging to the long OS group and 23 (21.3%) patients to the short OS group, with features as in Table 3.

Fig. 1D shows the Kaplan–Meier plots of the survival data for the 108 patients with mCRC of the independent validation set, segregated according to the  $^1\text{H-NMR}$ -based clustering described above. Statistically significant differences ( $P = 1.33 \times 10^{-6}$ ) in the curve of OS were calculated on the basis of the Wald test. Cox hazard analysis shows that a significantly high HR exists between short and long OS groups [HR, 3.37; 95% confidence interval (CI), 2.06–5.50,  $P = 1.33 \times 10^{-6}$ ; Table 4].

ECOG-PS also provides a meaningful prediction (HR, 1.57; 95% CI, 1.20–2.06,  $P = 1.13 \times 10^{-5}$ ; Table 4), whereas the presence of *KRAS* mutations does not result as a meaningful predictor of OS ( $P = 2.11 \times 10^{-1}$ ; Table 4). Consistently, mutated and wild-type *KRAS* patients were found to be almost equally distributed between the long OS and short OS groups (Table 3). When a multivariate analysis on NMR and ECOG-PS is conducted, only NMR profiling remains meaningful (Table 4). The predictive power of NMR profiling with respect to that of classical serologic markers (YKL-40, C-reactive protein, carcinoembryonic antigen) was also tested (full analysis to be published elsewhere by B.V. Jensen); NMR profiling remains an independent and meaningful predictor, together with CRP levels.

**Table 2.** Classification parameters for the training set and validation set

Study and model	No. of Subjects	Sensitivity (%)	Specificity (%)	Accuracy (%)
<b>Training set</b>				
Healthy subjects vs. CRC	141 (96 + 45)	100.0	100.0	100.0
Healthy subjects vs. CRC (45–65 y)	61 (37 + 24)	100.0	100.0	100.0
Healthy subjects vs. CRC (ECOG-PS = 0)	113 (96 + 17)	100.0	100.0	100.0
Long vs. short OS	20 (10 + 10)	79.9	76.4	78.5
<b>Validation set</b>				
Healthy subjects vs. CRC	151 (43 + 108)	100.0	95.5	96.7
Healthy subjects vs. CRC (45–65 y)	76 (15 + 61)	80.0	95.1	92.1
Healthy subjects vs. CRC (ECOG-PS = 0)	103 (43 + 60)	95.3	80.0	86.4

**Table 3.** Patient demographic and clinical characteristics in the long and short OS classes predicted by the model built on the validation set

	Long OS, <sup>a</sup> n (%)	Short OS, <sup>a</sup> n (%)	<i>P</i> <sup>b</sup>
No. of patients	85 (78.7)	23 (21.3)	
OS, mo			
Median	12.9	5.2	<b>1.36 × 10<sup>-5</sup></b>
Range	1.1–45.4	0.5–16.9	
Age, y			
Median	64	60	5.14 × 10 <sup>-1</sup>
Range	46–87	55–76	
Sex			
Male	54 (63.5)	11 (47.8)	2.40 × 10 <sup>-1</sup>
Female	31 (36.5)	12 (52.2)	
Body mass index <sup>c</sup>			
Underweight	4 (4.7)	3 (13.0)	1.39 × 10 <sup>-1</sup>
Normal weight	35 (42.1)	13 (56.5)	
Overweight	29 (34.9)	6 (26.1)	
Obese	15 (18.1)	1 (4.4)	
ECOG-PS			
0	55 (64.7)	5 (21.6)	<b>3.36 × 10<sup>-5</sup></b>
1	25 (29.4)	9 (39.2)	
2	5 (5.9)	9 (39.2)	
<i>KRAS</i> <sup>d</sup>			
Wild-type	53 (64.3)	12 (60.0)	7.97 × 10 <sup>-1</sup>
Mutated	29 (35.7)	8 (40.0)	

NOTE: Bold indicates significant *P* values.Abbreviation: *n*, no. of subjects.<sup>a</sup>Predicted classes by the model built on the training set.<sup>b</sup>Groupwise comparisons of distributions of clinical and demographic data were conducted with the nonparametric Kruskal–Wallis rank-sum test for the continuous variables and the Fisher exact test for the categorical variables. Significant at *P* < 0.05.<sup>c</sup>Data not available for 2 patients of the validation set.<sup>d</sup>Data not available for 6 patients of the validation set.**Metabolites contributing to the mCRC fingerprint**

An analysis of the PLS loadings was conducted to identify the metabolites contributing to the mCRC fingerprint using the entire sets of patients with mCRC and healthy subjects. The values of the relative serum concentrations of metabolites were estimated through the integration of the signals in the NMR spectra. By comparing the spectra of the serum samples of patients with mCRC with those of healthy subjects, it appears that the patients with mCRC are characterized by lower serum levels (*P* < 0.05) of alanine, citrate, creatine, glutamine, peptide NHs, lactate, leucine, pyruvate, tyrosine, and valine and higher serum levels (*P* < 0.05) of 3-hydroxybutyrate, acetate, formate, glycerol, lipid (-CH<sub>2</sub>-OCOR), *N*-acetyl signal of glycoproteins, phenylalanine, and proline (Supplementary Table S2 and Fig. 2A).

To identify the metabolites that contribute to discriminate patients with short and long OS, we selected from the entire group of patients with mCRC those with extreme OS values. By setting OS thresholds of more than 24 months and less than 3 months, 18 and 17 patients were selected, respectively. Analysis of the PLS loadings shows that patients with short OS are characterized by lower serum levels (*P* < 0.05) of creatine, lipid

(-C=C-CH<sub>2</sub>-C=C-), lipid (-CH=CH-), and valine and higher serum levels (*P* < 0.05) of lipid (-CH<sub>2</sub>-OCOR) and *N*-acetyl signal of glycoproteins (Supplementary Table S3 and Fig. 2B). The worsening of the clinical conditions thus corresponds to a more marked difference in the concentration of several of the metabolites that are responsible of the disease fingerprint.

By comparing the spectra of the serum samples of the underweight (*N* = 7) and normal weight (*N* = 71) versus overweight (*N* = 52) and obese patients with mCRC (*N* = 21), it appears that the overweight and obese patients are characterized by lower levels of formate (*P* = 7.78 × 10<sup>-3</sup>) and low-density lipoprotein (LDL)/high-density lipoprotein (HDL; *P* = 1.42 × 10<sup>-4</sup>) and higher levels of valine (*P* = 1.98 × 10<sup>-3</sup>), *N*-acetyl signal of glycoproteins (*P* = 1.28 × 10<sup>-2</sup>), CH-CH<sub>2</sub>-CO and CH<sub>2</sub>-CO signals due to lipids (*P* = 1.24 × 10<sup>-2</sup>; *P* = 7.37 × 10<sup>-3</sup>) and very LDL (VLDL; *P* = 3.07 × 10<sup>-2</sup>).

**Discussion**

<sup>1</sup>H-NMR profiling was used to characterize the metabolomic signature of patients with mCRC before third-line treatment

**Table 4.** Univariate and multivariate Cox regression analysis of clinical factors associated with OS of the validation set ( $n = 108$ )

Variables	HR	95% CI	$P^d$
Univariate analysis <sup>a</sup>			
NMR profiling (long vs. short OS) <sup>b</sup>	3.37	2.06–5.50	<b><math>1.33 \times 10^{-6}</math></b>
ECOG-PS (grade 0 vs. 1 vs. 2)	1.57	1.20–2.06	<b><math>1.13 \times 10^{-3}</math></b>
BMI (underweight vs. normal weight vs. overweight vs. obese)	0.83	0.65–1.06	$1.32 \times 10^{-1}$
Sex (female vs. male)	0.92	0.61–1.38	$6.80 \times 10^{-1}$
Age (<60 vs. $\geq 60$ y)	0.93	0.62–1.40	$7.26 \times 10^{-1}$
KRAS (wild-type vs. mutant)	1.31	0.86–2.02	$2.11 \times 10^{-1}$
Multivariate analysis <sup>c</sup>			
NMR profiling (long vs. short OS)	2.79	1.61–4.84	<b><math>2.75 \times 10^{-4}</math></b>
ECOG-PS (grade 0 vs. 1 vs. 2)	1.27	0.94–1.71	$1.23 \times 10^{-1}$

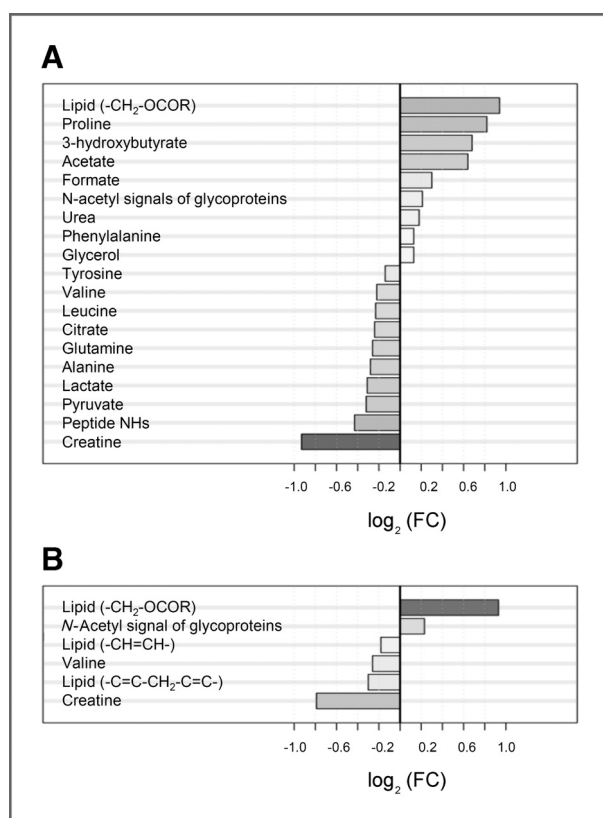
NOTE: Bold indicates significant  $P$  values.

<sup>a</sup>Univariate analysis, Cox proportional hazards regression.

<sup>b</sup>Predicted classes by the model built on the training set.

<sup>c</sup>Multivariate analysis, Cox proportional hazards regression adjusting for ECOG-PS.

<sup>d</sup> $P$  values were calculated according to the Wald method. Significant at  $P < 0.05$ .



**Figure 2.** Values of  $\log_2(\text{FC})$  resulting from the comparison between (A) healthy subjects and patients with mCRC and (B) patients with mCRC and long OS (>24 months) and short OS (<3 months). A,  $\log_2(\text{FC})$  with a positive value indicates a relatively higher serum concentration present in the patients with mCRC, whereas a negative value means a relatively lower serum concentration in the patients with mCRC. B,  $\log_2(\text{FC})$  with a positive value indicates a relatively higher serum concentration present in the patients with mCRC and short survival (<3 months), whereas a negative value means a relatively lower serum concentration in the patients with mCRC and short survival (<3 months). FC, fold change.

with cetuximab and irinotecan. The classification based on <sup>1</sup>H-NMR profiling was highly accurate (100.0%, as estimated by cross-validation) and confirmed on the validation set (96.7%) showing a clear separation between the patients with mCRC and healthy subjects. The validity of the approach was strengthened by the demonstrated ability to find a clear signature in serum also of patients with mCRC in good performance (i.e., ECOG-PS = 0).

The capability of <sup>1</sup>H-NMR profiling to predict OS is interesting. It probably reflects the fact that metabolites are the end products of the ensemble of processes occurring in living organisms and can be regarded as the ultimate response of the organisms to disease-induced metabolic alterations, inflammatory processes, and changes in lifestyle (diet, sedentary vs. active life, etc.) as a consequence of the pathologic status. The information content derived from classical approaches on the basis of single molecule markers is somehow embedded in the <sup>1</sup>H-NMR patterns. On the other hand, both NMR profiling and quantification of patients' general well-being like the definition of ECOG-PS are both complex multiparametric indicators. However, the former is based on an objective, albeit complex, evaluation of spectral data that contain molecular level information. On the contrary, the ECOG-PS is based on criteria that reflect how a patient's disease is progressing and how the disease affects the daily living abilities of the patient; it is therefore intrinsically associated with a certain degree of subjectivity (43).

It has been reported that indicators of an inflammatory response in serum NMR are represented by increased intensity for the signals of CH<sub>2</sub>-COOR of lipids and for the *N*-acetyl resonance of glycoproteins (44). We show here that the metabolomic fingerprint in serum of patients with mCRC is characterized by a higher intensity of these resonances; the effect is larger for patients with short OS after starting third-line treatment with cetuximab and irinotecan. This finding may reflect a nonspecific inflammatory response and is in

agreement with the proposal by Pages and colleagues (45) and Galon and colleagues (46) that the OS time is "governed in large part by the state of the local adaptive immune response."

Patients with short survival showed decreased serum concentrations of polyunsaturated lipids. Reduced levels of hydroxylated, polyunsaturated ultra long-chain fatty acids have been reported for patients with CRC on the basis of NMR/MS studies and proposed as a signature of this disease (21).

Moreover, in the serum of the patients with mCRC, we found a depletion of several potential precursors of glucose in gluconeogenesis (47, 48), such as lactate, pyruvate, alanine, glutamine, and other gluconeogenic amino acids, with no significant changes in glucose concentration. The observed trend may suggest an increased uptake of glucose precursors by the liver (49, 50) that could be consistent with an increased hepatic gluconeogenesis (51). Such an effect is found in patients with cachexia that occurs frequently in patients with metastatic cancer and is associated with more than 20% of cancer deaths (52). By comparing the spectra of the serum samples of the underweight and normal weight versus overweight and obese patients with mCRC, it appears that the overweight and obese patients are characterized by lower levels of formate and LDL/HDL and higher levels of valine, *N*-acetyl signal of glycoproteins, CH-CH<sub>2</sub>-CO and CH<sub>2</sub>-CO signals due to lipids and VLDL. Although information on nutritional status is not available, the *N*-acetyl signal of glycoproteins suggests that underweight and normal weight patients with mCRC may be characterized by a significantly higher immune response, highlighting a link between cachexia and inflammatory response.

The relative low serum levels of lactate and the nonstatistically significant changes in glucose levels are different from the observations in previous studies of patients with CRC (20, 22) and in other types of cancer where, commonly, serum

lactate levels are high (12, 53). However, relatively high serum levels of glucose and low levels of lactate, alanine, and gluconeogenic amino acids were observed in patients with oral cancer (14). This pattern was attributed to an altered energy metabolism (14). Gluconeogenesis in the present study of patients with mCRC may mask anaerobic dissimilation of glucose (48). From the comparison between patients with mCRC and healthy subjects it also emerges that high serum levels of 3-hydroxybutyrate are present in the patients with mCRC. This molecule is formed during fatty acid oxidation and its high levels may suggest a liver response to an increased energy demand needed for an increased hepatic gluconeogenesis (14). Depletion of polyunsaturated lipids observed in patients with mCRC also fits in this scheme.

In summary, we found that the serum of patients with mCRC before third-line therapy with cetuximab and irinotecan shows a signature of altered energy metabolism, which may reflect an increased gluconeogenesis and an accumulation of 3-hydroxybutyrate. Our findings also suggest that the overall <sup>1</sup>H-NMR profile of serum from patients with mCRC may reflect an inflammatory status, more serious in patients with short OS.

#### Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

#### Grant Support

The research was funded by the European Union Seventh Framework Programme [FP7/2007-2013] under grant agreement no. 222916.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received May 5, 2011; revised October 25, 2011; accepted November 1, 2011; published OnlineFirst November 11, 2011.

#### References

- Zhang XW, Li L, Wei D, Chen F. Moving cancer diagnostics from bench to bedside. *Trends Biotechnol* 2007;25:166-73.
- Spratlin JL, Serkova NJ, Eckhardt SG. Clinical applications of metabolomics in oncology: a review. *Clin Cancer Res* 2009;15:431-40.
- Claudino WM, Quattrone A, Biganzoli L, Pestrin M, Bertini I, Di Leo A. Metabolomics: available results, current research projects in breast cancer, and future applications. *J Clin Oncol* 2007;25:2840-6.
- Griffin JL, Shockcor JP. Metabolic profiles of cancer cells. *Nat Rev Cancer* 2004;4:551-61.
- Bathen TF, Sitter B, Sjøbakk TE, Tessem MB, Gribbestad IS. Magnetic resonance metabolomics of intact tissue: a biotechnological tool in cancer diagnostics and treatment evaluation. *Cancer Res* 2010;70:6692-6.
- Assfalg M, Bertini I, Colangioli D, Luchinat C, Schafer H, Schutz B, et al. Evidence of different metabolic phenotypes in humans. *Proc Natl Acad Sci U S A* 2008;105:1420-4.
- Bertini P, Bertini I, Nepi S, Saccenti E, Schaefer H, Schuetz B, et al. Individual human phenotypes in metabolic space and time. *J Proteome Res* 2009;8:4264-71.
- Bertini I, Calabro A, De carli V, Luchinat C, Nepi S, Porfirio B, et al. The metabolomic signature of celiac disease. *J Proteome Res* 2009;1:170-7.
- Pearson H. Meet the human metabolome. *Nature* 2007;446:8.
- Bertini P, Bertini I, Calabro A, La Marca G, Lami G, Luchinat C, et al. Are patients with potential celiac disease really potential? The answer of metabolomics. *J Proteome Res* 2011;10:714-21.
- Sreekumar A, Poisson LM, Rajendiran TM, Khan AP, Cao Q, Yu JD, et al. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature* 2009;457:910-4.
- Odunsi K, Wollman RM, Ambrosone CB, Hutson A, Mccann SE, Tammela J, et al. Detection of epithelial ovarian cancer using 1H-NMR-based metabolomics. *Int J Cancer* 2005;113:782-8.
- MacIntyre DA, Jimenez B, Lewintre EJ, Martin CR, Schafer H, Ballesteros CG, et al. Serum metabolome analysis by 1H-NMR reveals differences between chronic lymphocytic leukaemia molecular subgroups. *Leukemia* 2010;24:788-97.
- Tiziani S, Lopes V, Gunther UL. Early stage diagnosis of oral cancer using 1H NMR-based metabolomics. *Neoplasia* 2009;11:269-76.
- Oakman C, Tenori L, Biganzoli L, Santarpia L, Cappadona S, Luchinat C, et al. Uncovering the metabolomic fingerprint of breast cancer. *Int J Biochem Cell Biol* 2011;43:1010-20.
- Oakman C, Tenori L, Claudino WM, Cappadona S, Nepi S, Battaglia A, et al. Identification of a serum-detectable metabolomic fingerprint potentially correlated with the presence of micrometastatic disease in early breast cancer patients at varying risks of disease relapse by traditional prognostic methods. *Ann Oncol* 2011;22:1295-301.



17. Yang YX, Li CL, Nie X, Feng XS, Chen WX, Yue Y, et al. Metabonomic studies of human hepatocellular carcinoma using high-resolution magic-angle spinning 1H NMR spectroscopy in conjunction with multivariate data analysis. *J Proteome Res* 2007;6:2605–14.
18. Denkert C, Budczies J, Weichert W, Wohlgemuth G, Scholz M, Kind T, et al. Metabolite profiling of human colon carcinoma - deregulation of TCA cycle and amino acid turnover. *Mol Cancer* 2008;7:72–86.
19. Hirayama A, Kami K, Sugimoto M, Sugawara M, Toki N, Onozuka H, et al. Quantitative metabolome profiling of colon and stomach cancer microenvironment by capillary electrophoresis time-of-flight mass spectrometry. *Cancer Res* 2009;69:4918–25.
20. Qiu YP, Cai GX, Su MM, Chen TL, Zheng XJ, Xu Y, et al. Serum metabolite profiling of human colorectal cancer using GC-TOFMS and UPLC-QTOFMS. *J Proteome Res* 2009;8:4844–50.
21. Ritchie SA, Ahiaonu PWK, Jayasinghe D, Heath D, Liu J, Lu YS, et al. Reduced levels of hydroxylated, polyunsaturated ultra long-chain fatty acids in the serum of colorectal cancer patients: implications for early screening and detection. *BMC Med* 2010;8:13–32.
22. Ludwig C, Ward DG, Martin A, Viant MR, Ismail T, Johnson PJ, et al. Fast targeted multidimensional NMR metabolomics of colorectal cancer. *Magn Reson Chem* 2009;47:S68–73.
23. Piotto M, Moussalieh FM, Dillmann B, Imperiale A, Neuville A, Brigand C, et al. Metabolic characterization of primary human colorectal cancers using high resolution magic angle spinning 1H magnetic resonance spectroscopy. *Metabolomics* 2009;5:292–301.
24. Chan ECY, Koh PK, Mal M, Cheah PY, Eu KW, Backshall A, et al. Metabolic profiling of human colorectal cancer using high-resolution magic angle spinning nuclear magnetic resonance (HR-MAS NMR) spectroscopy and gas chromatography mass spectrometry (GC/MS). *J Proteome Res* 2009;8:352–61.
25. Bezabeh T, Somorjai R, Dolenko B, Bryskina N, Levin B, Bernstein CN, et al. Detecting colorectal cancer by 1H magnetic resonance spectroscopy of fecal extracts. *NMR Biomed* 2009;22:593–600.
26. Ferlay J, Parkin DM, Steliarova-Foucher E. Estimates of cancer incidence and mortality in Europe in 2008. *Eur J Cancer* 2010;46:765–81.
27. Jemal A, Siegel R, Xu JQ, Ward E. Cancer statistics, 2010. *CA Cancer J Clin* 2010;60:277–300.
28. Douillard JY, Cunningham D, Roth AD, Navarro M, James RD, Karasek P, et al. Irinotecan combined with fluorouracil compared with fluorouracil alone as first-line treatment for metastatic colorectal cancer: a multicentre randomised trial. *Lancet* 2000;355:1041–7.
29. Saltz LB, Cox JV, Blanke C, Rosen LS, Fehrenbacher L, Moore MJ, et al. Irinotecan plus fluorouracil and leucovorin for metastatic colorectal cancer. *N Engl J Med* 2000;343:905–14.
30. Giacchetti S, Perpoint B, Zidani R, Le Bail N, Faggiuolo R, Focan C, et al. Phase III multicenter randomized trial of oxaliplatin added to chronomodulated fluorouracil-leucovorin as first-line treatment of metastatic colorectal cancer. *J Clin Oncol* 2000;18:136–47.
31. de Gramont A, Figer A, Seymour M, Homerin M, Hmissi A, Cassidy J, et al. Leucovorin and fluorouracil with or without oxaliplatin as first-line treatment in advanced colorectal cancer. *J Clin Oncol* 2000;18:2938–47.
32. Cunningham D, Humblet Y, Siena S, Khayat D, Bleiberg H, Santoro A, et al. Cetuximab monotherapy and cetuximab plus irinotecan in irinotecan-refractory metastatic colorectal cancer. *N Engl J Med* 2004;351:337–45.
33. Siena S, Sartore-Bianchi A, Di Nicolantonio F, Balfour J, Bardelli A. Biomarkers predicting clinical outcome of epidermal growth factor receptor-targeted therapy in metastatic colorectal cancer. *J Natl Cancer Inst* 2009;101:1308–24.
34. Banck MS, Grothey A. Biomarkers of resistance to epidermal growth factor receptor monoclonal antibodies in patients with metastatic colorectal cancer. *Clin Cancer Res* 2009;15:7492–501.
35. Lievre A, Bachet JB, Boige V, Cayre A, Le Corre D, Buc E, et al. KRAS mutations as an independent prognostic factor in patients with advanced colorectal cancer treated with cetuximab. *J Clin Oncol* 2008;26:374–9.
36. Mityr E, Douillard JY, Van Cutsem E, Cunningham D, Magherini E, Mery-Mignard D, et al. Predictive factors of survival in patients with advanced colorectal cancer: an individual data analysis of 602 patients included in irinotecan phase III trials. *Ann Oncol* 2004;15:1013–7.
37. Leitch EF, Chakrabarti M, Crozier JEM, Mckee RF, Anderson JH, Horgan PG, et al. Comparison of the prognostic value of selected markers of the systemic inflammatory response in patients with colorectal cancer. *Br J Cancer* 2007;97:1266–70.
38. Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, et al. New guidelines to evaluate the response to treatment in solid tumors. *J Natl Cancer Inst* 2000;92:205–16.
39. Jensen BV, Schou JV, Johannesen HH, Christensen IJ, Nielsen D, Johansen JS, et al. Cetuximab every second week with irinotecan in patients with metastatic colorectal cancer refractory to 5-FU, oxaliplatin, and irinotecan: KRAS mutation status and efficacy. *J Clin Oncol* 2010;28(15):28(15) (suppl); abstr 3573.
40. Bernini P, Bertini I, Luchinat C, Nincheri P, Staderini S, Turano P. Standard operating procedures for pre-analytical handling of blood and urine for metabolomic studies and biobanks. *J Biomol NMR* 2011;49:231–43.
41. Ihaka R, Gentleman R. R: a language for data analysis and graphics. *J Comput Graph Stat* 1996;5:299–314.
42. Johansson E, Kettaneh-Wold N, Ericksson SWL. Multi- and megavariate data analysis principles and applications. Umea, Sweden: Umetrics Academy; 2005.
43. Ando M, Ando Y, Hasegawa Y, Shimokata K, Minami H, Wakai K, et al. Prognostic value of performance status assessed by patients themselves, nurses, and oncologists in advanced non-small cell lung cancer. *Br J Cancer* 2001;85:1634–9.
44. Torri GM, Torri J, Gulian JM, Vion-Dury J, Viout P, Cozzone PJ. Magnetic resonance spectroscopy of serum and acute-phase proteins revisited: a multiparametric statistical analysis of metabolite variations in inflammatory, infectious and miscellaneous diseases. *Clin Chim Acta* 1999;279:77–96.
45. Pages F, Kirilovsky A, Mlecnik B, Asslaber M, Tosolini M, Bindea G, et al. *In situ* cytotoxic and memory T cells predict outcome in patients with early-stage colorectal cancer. *J Clin Oncol* 2009;27:5944–51.
46. Galon J, Costes A, Sanchez-Cabo F, Kirilovsky A, Mlecnik B, Lagorce-Pages C, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* 2006;313:1960–4.
47. Lowry SF, Foster DM, Norton JA, Berman M, Brennan MF. Glucose disposal and gluconeogenesis from alanine in tumor-bearing Fischer 344 rats. *J Natl Cancer Inst* 1981;66:653–8.
48. Bongaerts GPA, van Hatteren HK, Verhagen CAM, Wagener DJT. Cancer cachexia demonstrates the energetic impact of gluconeogenesis in human metabolism. *Med Hypotheses* 2006;67:1213–22.
49. Leij-Halfwerk S, Dagnelie PC, van den Berg JWO, Wattimena JDL, Hordijk-Luijk CH, Wilson JHP. Weight loss and elevated gluconeogenesis from alanine in lung cancer patients. *Am J Clin Nutr* 2000;71:583–9.
50. Shearer JD, Buzby GP, Mullen JL, Miller E, Caldwell MD. Alteration in pyruvate metabolism in the liver of tumor-bearing rats. *Cancer Res* 2010;44:4443–6.
51. Holroyde CP, Gabuzda TG, Putnam RC, Paul P, Reichard GA. Altered glucose metabolism in metastatic carcinoma. *Cancer Res* 1975;35:3710–4.
52. Tisdale MJ. Cachexia in cancer patients. *Nat Rev Cancer* 2002;2:862–71.
53. Morvan D, Demidem A. Metabolomics by proton nuclear magnetic resonance spectroscopy of the response to chloroethylnitrosourea reveals drug efficacy and tumor adaptive metabolic pathways. *Cancer Res* 2007;67:2150–9.