

# BASIC AND TRANSLATIONAL—ALIMENTARY TRACT

## The Colorectal Cancer Lipidome: Identification of a Robust Tumor-Specific Lipid Species Signature



Josef Ecker,<sup>1,\*</sup> Elisa Benedetti,<sup>2,\*</sup> Alida S. D. Kindt,<sup>3,4,\*</sup> Marcus Höring,<sup>5</sup> Markus Perl,<sup>6</sup> Andrea Christel Machmüller,<sup>6,7</sup> Anna Sichler,<sup>6</sup> Johannes Plagge,<sup>1</sup> Yuting Wang,<sup>8</sup> Sebastian Zeissig,<sup>8</sup> Andrej Shevchenko,<sup>9</sup> Ralph Burkhardt,<sup>5</sup> Jan Krumsiek,<sup>2</sup> Gerhard Liebisch,<sup>5</sup> and Klaus-Peter Janssen<sup>6</sup>

<sup>1</sup>ZIEL–Institute for Food & Health, Research Group Lipid Metabolism, Technical University of Munich, Freising, Germany; <sup>2</sup>Institute of Computational Biomedicine, Department of Physiology and Biophysics, Weill Cornell Medical College, New York, New York; <sup>3</sup>Division of Analytical Biosciences, Leiden Academic Centre for Drug Research, Leiden University, Leiden, The Netherlands; <sup>4</sup>Institute of Computational Biology, Helmholtz Zentrum München, Neuherberg, Germany; <sup>5</sup>Institute of Clinical Chemistry and Laboratory Medicine, University Hospital Regensburg, Regensburg, Germany; <sup>6</sup>Technical University of Munich, School of Medicine, Klinikum rechts der Isar, Department of Surgery, Munich, Germany; <sup>7</sup>Institute for Diabetes and Obesity, Helmholtz Center Munich, Neuherberg, Germany; <sup>8</sup>Center for Regenerative Therapies Dresden, Technische Universität Dresden, Dresden, Germany; <sup>9</sup>Department of Medicine I, University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany; and <sup>9</sup>Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany

See Covering the Cover synopsis on page 740.

**OBJECTIVE:** Lipidomic changes were causally linked to metabolic diseases, but the scenario for colorectal cancer (CRC) is less clear. We investigated the CRC lipidome for putative tumor-specific alterations through analysis of 3 independent retrospective patient cohorts from 2 clinical centers, to derive a clinically useful signature. **DESIGN:** Quantitative comprehensive lipidomic analysis was performed using direct infusion electrospray ionization coupled with tandem mass spectrometry (ESI-MS/MS) and high-resolution mass spectrometry (HR-MS) on matched nondiseased mucosa and tumor tissue in a discovery cohort (n = 106). Results were validated in 2 independent cohorts (n = 28, and n = 20), associated with genomic and clinical data, and lipidomic data from a genetic mouse tumor model (Apc<sup>1638N</sup>). **RESULTS:** Significant differences were found between tumor and normal tissue for glycerol-, glycerophospho-, and sphingolipids in the discovery cohort. Comparison to the validation collectives unveiled that glycerophospholipids showed high interpatient variation and were strongly affected by preanalytical conditions, whereas glycerol- and sphingolipids appeared more robust. Signatures of sphingomyelin and triacylglycerol (TG) species significantly differentiated cancerous from nondiseased tissue in both validation studies. Moreover, lipogenic enzymes were significantly up-regulated in CRC, and FASN gene expression was prognostically detrimental. The TG profile was significantly associated with postoperative disease-free survival and lymphovascular invasion, and was essentially conserved in murine digestive cancer, but not associated with microsatellite status, KRAS or BRAF mutations, or T-cell infiltration. **CONCLUSION:** Analysis of the CRC lipidome revealed a robust TG-species signature with prognostic potential. A better understanding of the cancer-associated glycerolipid and sphingolipid metabolism may lead to novel therapeutic strategies.

**Keywords:** Biomarker; Mass Spectrometry; Signature; Sphingomyelin; Triacylglycerol.

Colorectal cancer (CRC) is among the main causes for morbidity and mortality worldwide and ranks as the second most common cause of cancer-related deaths.<sup>1</sup> CRC is currently increasing in young adults worldwide.<sup>2</sup> Despite recent advances in therapy, prevention, and early detection,<sup>3</sup> prognosis is dismal for patients with distant metastases, with a 5-year survival rate of only 10%–14%.<sup>2,4</sup> A deregulated lipid metabolism is a hallmark of cancer cells, suggesting a putative vulnerability that might be exploited therapeutically.<sup>5–7</sup> Highly proliferating cancer cells need lipids for the generation of cell membranes (phospholipids, cholesterol, and sphingolipids), but also as substrate of energy metabolism (triacylglycerol) or source for signaling molecules.<sup>8–10</sup> The lipidomic profile of cancer types like breast, kidney, lung, and liver is significantly altered compared with healthy tissue.<sup>11–13</sup> However, the evidence for an altered lipidome in CRC and its causal contribution to carcinogenesis is unclear. The expression of key enzymes for lipid synthesis is significantly induced in CRC cells,<sup>14</sup> and inhibitors of lipid synthesis have been proposed for CRC treatment.<sup>15,16</sup> Further, indirect evidence suggests that dietary lipids could influence CRC because obesity is a well-

\*Authors share co-first authorship.

**Abbreviations used in this paper:** AGC, automated gain control; AUC, area under the curve; CE, cholesteryl ester; Cer, ceramide; CRC, colorectal cancer; DG, diacylglycerol; EFS, event-free survival; ESI, electrospray ionization; FASN, human fatty acid synthase; Fasn, murine fatty acid synthase enzyme; FIA, flow injection analysis; FTMS, Fourier-Transform mass spectrometry; GL, glycerolipids; GPL, glycerophospholipids; HR, high-resolution; IT, injection time; LPC, lysophosphatidylcholine; MS, mass spectrometry; MSI, microsatellite instability; PC, phosphatidylcholine; PCR, polymerase chain reaction; PE, phosphatidylethanolamine; PI, phosphatidylinositol; PS, phosphatidylserine; PUFA, polyunsaturated fatty acids; RT, reverse transcription; SL, sphingolipids; SM, sphingomyelin; TG, triacylglycerol; TUM, Technical University of Munich.

Most current article

© 2021 by the AGA Institute  
0016-5085/\$36.00

<https://doi.org/10.1053/j.gastro.2021.05.009>

**WHAT YOU NEED TO KNOW****BACKGROUND AND CONTEXT**

Lipid synthesis and modification are thought to contribute causally to colorectal cancer, but preanalytical conditions in the clinical setting vary greatly and make it challenging to compare metabolite levels between individual cohorts.

**NEW FINDINGS**

Sphingolipids and glycerolipids are altered in 3 independent cohorts of colorectal cancer, accompanied by increased expression of lipogenic enzymes. A robust lipid signature, present in all genetic subgroups of colorectal cancer, separates tumor from normal tissue, has prognostic relevance, and is conserved in a genetic colorectal cancer mouse model.

**LIMITATIONS**

Mass spectrometry-based lipidomics requires analysis of fresh or frozen tissue samples, precluding widespread clinical application.

**IMPACT**

A robust lipid species signature, based on quantitative lipidomics, distinguishes normal from cancer tissue and is associated with prognosis, opening new ways for diagnosis, prognosis, and therapy monitoring.

known risk-factor.<sup>17</sup> High dietary fat content may promote carcinogenesis and the risk for disease recurrence.<sup>17–19</sup> In contrast, nutrition-derived n-3 polyunsaturated fatty acids or butyrate generated from dietary fiber by the gut microbiota have health-promoting and antineoplastic properties.<sup>17,20,21</sup> Notwithstanding, the existence, biochemical composition, and clinical relevance of a cancer-specific lipidome is under debate. Several studies have reported alterations of lipid species in CRC tissue, but the heterogeneity and variation between individual studies is considerable.<sup>22–25</sup> A multi-cohort study has, to the best of our knowledge, not yet been performed. Stability of analytes, strongly influenced by sample handling during endoscopic or surgical interventions, sample storage, and processing, is critical for cancer biomarker identification widely acknowledged for RNA- or protein-based markers, but not sufficiently addressed yet in lipidomics. Therefore, we quantified the comprehensive lipidome of colorectal tumors and matched nondiseased mucosa samples from patients of 3 independent cohorts sampled at 2 clinical centers. Further, we tested the putative association of the tumor lipidome with clinical parameters and prognosis, as well as with genomic markers, intratumoral T-cell densities, and expression of lipogenic enzymes. A data analysis strategy based on state-of-the-art bioinformatics substantiated with biological knowledge revealed significant alterations of sphingomyelin (SM) and triacylglycerol (TG) species profiles in the discovery cohort. These alterations were used to develop a robust tumor-related lipid signature that was successfully tested on 2 independent validation cohorts and was significantly associated with postoperative survival, indicating a putative clinical relevance. Moreover, we found a high degree of conservation regarding tumor-specific deregulated TG species in a genetic mouse model for digestive cancer.

**Materials and Methods***Analysis of Human Tissue Samples*

The discovery cohort (CI) consisted of  $n = 114$  patients with primary CRC, who underwent surgery at the Department of Surgery, Technical University of Munich (TUM), Munich, Germany, between 1990 and 2010. Patients with neoadjuvant treatment were excluded, and only cases with curative tumor resection (R0) were included, except for Union Internationale Contre Le Cancer stage IV patients who presented synchronous distant metastasis at diagnosis. Tissue samples were macro-dissected to separate carcinoma from adjacent nontumor tissue by experienced pathologists, and were immediately shock-frozen in liquid nitrogen, following established biobanking protocols and in accordance with local and national ethical and legal standards for data protection, after informed written consent (Ethics Committee of the Faculty of Medicine, TUM, #1926/07 and #5428/12). Histology-guided sample selection using H&E staining of consecutive tissue cryosections was performed as published earlier to ensure a tumor cell content of  $>70\%$ , leading to the exclusion of 6 cases.<sup>26,27</sup> Two cases were excluded due to missing values in lipidomics analysis ( $>20\%$ ), leaving  $n = 106$  patients for the discovery cohort CI. Matched tumor/normal samples were available from 41 patients. The paired differential analyses are thus based on 82 tissue samples from 41 individual patients. The validation cohort CII included 60 samples from 30 unique patients, with the same inclusion/exclusion criteria and sample processing protocol, recruited from years 2012–2018 at the Department of Surgery at TUM. Two samples were removed due to missing values ( $>20\%$ ), leaving 28 patients. Of these, 48 matched normal and tumor samples from 24 patients were available for analysis. A second independent validation cohort CIII ( $n = 20$ ) from the University Hospital of Kiel, Germany, from 2010–2014, was used with matched normal tissue for all 20 cases, as published elsewhere.<sup>23</sup> Clinical characteristics of all 3 cohorts are summarized in [Table 1](#).

*Lipidomics*

Tissue samples were homogenized in water/methanol = 1/1 by bead beating and subjected to lipid extraction according to the method of Bligh and Dyer<sup>28</sup> in the presence of not naturally occurring lipid species as internal standards. The following lipid species were added as internal standards: phosphatidylcholine (PC) 14:0/14:0, PC 22:0/22:0, phosphatidylethanolamine (PE) 14:0/14:0, PE 20:0/20:0 (di-phytanoyl), phosphatidylserine (PS) 14:0/14:0, PS 20:0/20:0 (di-phytanoyl), phosphatidylinositol (PI) 17:0/17:0, lysophosphatidylcholine (LPC) 13:0, LPC 19:0, lysophosphatidylethanolamine 13:0, ceramide (Cer) d18:1/14:0, Cer d18:1/17:0, D7-FC, cholesteryl ester (CE) 17:0, CE 22:0, TG 51:0, TG 57:0, diacylglycerol (DG) 28:0, and DG 40:0. Tissue homogenates representing a wet weight of 2 mg were extracted. The analysis of lipids was performed using direct flow injection analysis (FIA) either using a triple quadrupole mass spectrometer (FIA-MS/MS; QQQ triple quadrupole) or Fourier-Transform mass spectrometry (FIA-FTMS; high mass resolution). FIA-MS/MS (QQQ) was performed in positive ion mode using the analytical set-up and strategy described previously.<sup>29</sup> A fragment ion of  $m/z$  184 was used for PC, SM,<sup>29</sup> and LPC.<sup>30</sup> The following neutral losses were applied: PE 141, PS 185, phosphatidylglycerol 189, and PI 277.<sup>31</sup> PE-based plasmalogens (PE P) were

**Table 1.** Clinical Parameters for the Discovery and Validation Cohorts

Parameter	Variable	Discovery Cohort		Validation Cohort		Validation Cohort	
		CI n = 106	CI (%)	CII n = 28	CII (%)	CIII n = 20 <sup>a</sup>	CIII (%)
Sex	Male	62	58	19	67	10	50
	Female	44	42	9	33	10	50
Age	Median and range	66 y 31–86		68 y 30–84		64 y 51–72	
Histology (WHO)	Adeno-carcinoma	91	86	24	85	20	100
	Mucinous carcinoma	12	11	4	15	0	0
	Signet ring carcinoma	3	3	0	0	0	0
pTNM-stage (UICC/AJCC)	I	5	5	3	10	4	20
	II	27	26	10	33	5	25
	III	64	60	10	40	6	30
	IV	10	9	5	17	5	25
Anatomic localization	Colon	102	96	25	87	9	45
	Rectum	4	4	3	13	11	55
Alive status	Alive	48	45	9	37	/	/
	Tumor-related death	37	35	2	7	/	/
	Non-tumor-related death	21	20	17	66	/	/
Recurrence	No recurrence	69	65	7	65	/	/
	Disease recurrence	37	35	4	35	/	/
Nondiseased tissue	Matched control cases	41	46	24	63	20	100

pTNM, pathological tumor-node-metastasis staging; UICC/AJCC, Union Internationale Contre Le Cancer/American Joint Committee on Cancer; WHO, World Health Organization.

<sup>a</sup>Wang et al. 2020.

analyzed according to the principles described in detail elsewhere.<sup>32</sup> Sphingosine-based ceramides (Cer d18:1) and hexosylceramides were analyzed using a fragment ion of  $m/z$  264.<sup>33</sup> The Fourier Transform Mass Spectrometry (FIA-FTMS) setup has been previously described in detail.<sup>34</sup> TG, DG, and CE were recorded in positive ion mode FTMS in  $m/z$  range 500–1000 for 1 min with a maximum injection time (IT) of 200 ms, an automated gain control (AGC) of  $1 \times 10^6$ , 3 microscans, and a target resolution of 140,000 (at  $m/z$  200). Multiplexed acquisition was used for the  $[M+NH_4]^+$  of free cholesterol. Data processing details were described elsewhere in detail, using ALEX software,<sup>35</sup> which includes peak assignment and intensity picking. Extracted data were exported to Microsoft Excel 2010 and processed by generic macros. Lipid species were annotated according to the proposal for shorthand notation of lipid structures derived from MS.<sup>36</sup> For QQQ glycerophospholipid species annotation was based on the assumption of even-numbered carbon chains only. SM species annotation was based on the assumption that a sphingoid base with 2 hydroxyl groups is present.

### Data Analysis and Preprocessing

All analyses were performed in R (3.6.3). Each lipid species was represented as a percentage of the total lipid class abundance. For CI and CII all lipids detected were used, for CIII only

the SM, Cer, and TG species common with CI and CII were used (Cer d18:1: 16:0, 18:0, 18:1, 20:0, 22:0, 22:1, 23:0, 24:0, 24:1, 26:0, and 26:1; SM: 32:1, 33:1, 34:0, 34:1, 34:2, 35:0, 35:1, 36:0, 36:1, 36:2, 38:2, 40:0, 40:1, 40:2, 42:1, 42:2, and 42:3; TG: 48:2, 48:1, 50:1, 52:5, 52:3, 54:6, 54:5, 54:4, 54:3, 54:2, 55:4, 56:5, and 56:4). Lipids and samples with 20% or more missing values across the dataset were excluded from the analysis. This resulted in 149 lipid species left in CI, 172 lipids in CII, and 41 lipids in CIII; among the latter, 8 lipids were removed due to high missing value rate. Because CII was primarily used to validate the differential lipids detected in CI, the lipids in CII were filtered for those present in CI. This reduced the total number of lipids analyzed in CII to 148. For the prediction model, we first considered all SM, CER, and TG lipid species present in the discovery cohort CI, irrespective of their detection in the validation cohorts (Supplementary data). Next, we considered SM, Cer, and TG species only if they were detected in all 3 datasets. This filtering step, blind to the outcome of the analysis and, therefore, not leading to overfitting, reduced the set of analyzed lipids to 28 (see Prediction Model section for more details).

### Tumor vs Nondiseased Tissue: Differential Analysis

For paired analysis, only matched samples (tumor and matched normal mucosa from same patient and same surgical

intervention) were used (not available for all cases in CI and CII due to clinical constraints). In cohort CI,  $n = 41$  patients with matched samples were available,  $n = 24$  patients in cohort CII and  $n = 20$  in CIII. The paired differential analysis was performed with a Wilcoxon signed rank test using the `Wilcoxon.test` function in the stats R package. All results of the differential analysis were corrected for multiple testing using the Benjamini-Hochberg method to control the False Discovery Rate<sup>37</sup> at 0.05.

### Prediction Model

A prediction model was built using logistic regression with LASSO penalization using each of the 3 lipid classes separately (SM, Cer, and TG). Because the model was validated in cohorts CII and CIII, only lipids detected in all cohorts were considered. These included 7 lipids from the SM class, 4 lipids from the Cer class, and 6 lipids from the TG class (see also the Data Analysis and Preprocessing section). Additionally, a prediction model was built using all SM, CER, and TG lipids identified in CI. Lipid species that were not detected in CII or CIII were imputed with zeros in the respective cohorts (Supplementary Data). For each lipid class, a regularized linear model was trained from cohort CI using a 5-fold cross-validation approach. A final model was then built from cohort CI using the LASSO lambda parameter that gave the smallest cross-validation error in the previous step and validated on cohorts CII and CIII. Modeling was performed using the `cv.glmnet` function from the `glmnet` R package. The difference between prediction scores produced by each model on validation data was assessed using Wilcoxon rank test.

### Survival Modeling

In cohort CI,  $n = 32$  patients had an event (ie, disease relapse) during follow-up. As suggested,  $n/15$  variables should be used for survival modeling to avoid overfitting.<sup>38</sup> For this reason, a Ridge-regularized survival model was implemented based on Cox regression<sup>39</sup> with 2 degrees of freedom using the `coxph` and `ridge` function in the survival R package. The overall  $P$ -value of the model is reported in Figure 4. For visualization purposes, the survival curves were subdivided into risk strata. To verify that the observed predictive power of the lipid classes was not confounded by common covariates (age and gender), we developed a survival model using only the covariates and using covariates and lipid together. Analysis of variance (ANOVA) was performed to test the statistical significance of the difference between the full model and the models including only lipids or only covariates (Supplementary Data). The  $P$ -value of the ANOVA test indicates whether the included variables have a significant effect on the overall model. If the ANOVA test between survival model #1, including lipids and covariates, and survival model #2, performed using only lipids, is below the significance threshold, this would indicate that addition of the covariates substantially affected the model.

### Genomic, Immunologic, and Messenger RNA Expression Analysis of Tumor Tissue

Association of the tumor lipidome with genetic markers was assessed based on published data for discovery cohort CI.<sup>40</sup> DNA microsatellite instability (MSI) status was analyzed based on the type-it 170 Microsatellite PCR kit (Qiagen, Hilden, Germany) for the Bethesda panel on  $n = 61$  patients. Mutations

in KRAS exon 2 and BRAF exon 15 were assessed using high-resolution melting polymerase chain reaction (PCR) of genomic DNA on  $n = 75$  patients. Infiltrating T cells/ $\text{mm}^2$  of tumor tissue were assessed using quantitative immunohistochemistry, CD3- and CD8-positive cells ( $n = 62$  patients), and CD4-positive cells ( $n = 25$  cases). RNA isolation from  $n = 109$  cases from CI and CII and reverse transcription (RT) was performed as described, with the RNeasy Kit (Qiagen),<sup>41</sup> qRT-PCR was performed in triplicates with the LightCycler 480 II system (Roche, Penzberg, Germany). HPRT was used as reference transcript and a pool of human colon mucosa ( $n = 15$  cases) for normalization. Primers used: HPRT (UPL #22): 5'-GACCAGT-CAACAGGGGACAT-3' and 5'-GTGTCAATTATATCTTCCACAATC AAG-3'; FASN (UPL #11): 5'-CAGGCACACACGATGGAC-3' and 5'-CGGAGTGAATCTGGGTTGAT-3'; ELOVL5 (UPL #31): 5'-CCCTTCCATGCGTCCATA-3' and 5'-GATTGTCAGCA-CAAAGTGAAGC-3'; FADS1 (UPL #60): 5'-CCAGACATCAA-CATGCATCC-3' and 5'-TTTTCTTCTGTTTCCAAGCTC-3', and FADS2 (UPL #46), 5'-CTACCTCTCAGGCCAAGC-3' and 5'-GCGATGATTCCACCAGTTG-3'. To derive optimal cut-off values of gene expression, maximally selected log-rank statistics was performed using R Software version 2.13.0 (R Foundation for Statistical Computing, Vienna, Austria) (Supplementary Data) as published in detail earlier.<sup>40</sup> To consider multiple test issue within these analyses, the R-function 'maxstat.test' was used.

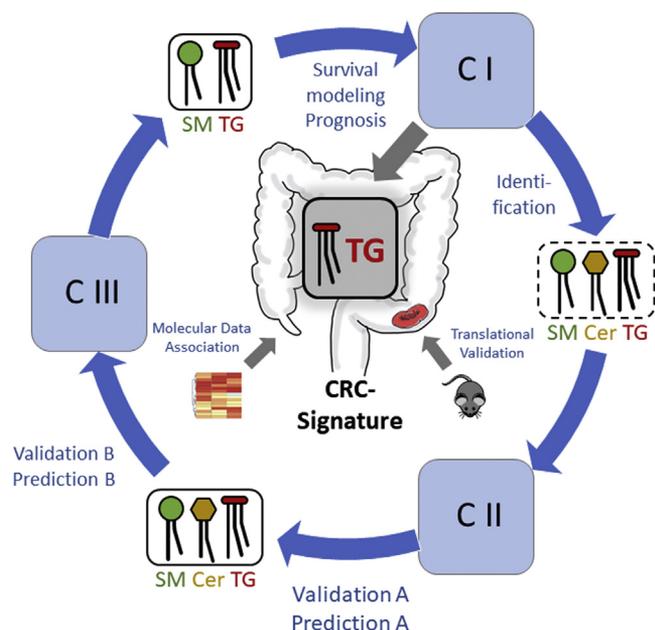
### Analysis of Genetic Tumor Mouse Model

Mice were kept at the animal facility of the Klinikum rechts der Isar (Munich, Germany) under specific pathogen-free conditions. Approval has been obtained by the local authorities (Approval ROB 55.2-1-54-2532-158-2015). Mice from the genetic model  $\text{Apc}^{1638\text{N}}$  were backcrossed to the C57BL/6N background for >10 generations with littermates serving as controls, as published.<sup>42</sup> Tissue samples were obtained at 12 months, frozen immediately in liquid nitrogen, and stored at  $-80^\circ\text{C}$ . Processing of murine tissue samples for lipidomics followed the protocol described for human samples. Isolation of RNA, complementary DNA synthesis, and qRT-PCR were also performed as described for human specimen. The primers used: beta-actin for normalization `Actb` (UPL #56): 5'-AAGCCAACCGTGAAAAGAT -3' and 5'-GTGGTACGACCA-GAGGCATAC-3'; Fatty acid synthase `Fasn` (UPL #58): 5'-GCTGCTGTTGGAAGTCAGC -3' and 5'-AGTGTTCGTTCTCG-GAGTG-3'.

## Results

### Study Design

The sample set was derived from 3 cohorts consisting of a total of 154 patients with primary CRC recruited at 2 different centers. Cohort CI (discovery cohort) included 106 patients (TUM, Munich, Germany) with median follow-up of 78 months ( $n = 41$  cases with both tumor and matched normal tissue); the validation cohort CII included 28 patients from the same clinical center ( $n = 24$  with matched tissues); and cohort CIII, an independent dataset derived from 20 patients (University of Kiel, Kiel, Germany), described in detail elsewhere, with matched normal tissue available for all cases.<sup>23</sup> Clinical data are indicated in Table 1; the study design is summarized in Figure 1 and



**Figure 1.** Study design and details on lipidome analysis. Discovery cohort I (CI;  $n = 106$ ) was used to identify and develop lipid species signatures specific for CRC, which were validated and used for prediction in validation cohorts CII ( $n = 28$ ) and CIII ( $n = 20$ ); its association with prognostic, clinical, and molecular data and evolutionary conservation in a tumor mouse model were further evaluated.

**Supplementary Table 1.** Quantitative comprehensive lipidomic analysis was performed using direct infusion electrospray ionization coupled to tandem mass spectrometry (ESI-MS/MS) and high-resolution mass spectrometry (HR-MS) (Supplementary Table 1), comprising: (1) glycerophospholipids (GPL, cell membrane components): PC, phosphatidylcholine-ether, LPC, PE, lysophosphatidylethanolamine, PE P, PI, PS; (2) sphingolipids (SL; cell membrane components): SM, Cer, and hexosylceramide; (3) sterols (cell membrane components and storage lipids): Free cholesterol and CE; and (4) glycerolipids (GL, storage lipids): DG and TG.

Samples of CI were measured randomized in 1 batch. In total, 173 lipid species were detected and quantified.

### Specific Lipid Species Patterns Are Associated With CRC

To assess whether CRC affects specific lipid profiles, tumor samples were compared with matched normal tissue from the same patients using a paired analysis (CI), based on a Wilcoxon signed-rank test. In particular, lipid species profiles of LPC (GPL), SM, and Cer (SL), but also TG (GL) differed significantly between the 2 groups (Figure 2 and Supplementary Figure 1). Tumor samples contained significantly higher fractions of mono- and polyunsaturated LPC, such as 16:1, 18:1, 20:4, and 22:6 (Figure 2A and B). Analysis of SL profiles revealed elevated proportions of SM species with 32–34 carbons in tumor samples, but lower levels with more than 34 carbons (Figure 2). In contrast, sphingosine-based Cer with longer chains (C24:0–C26:0) were enriched in tumor samples, whereas shorter <C22:0

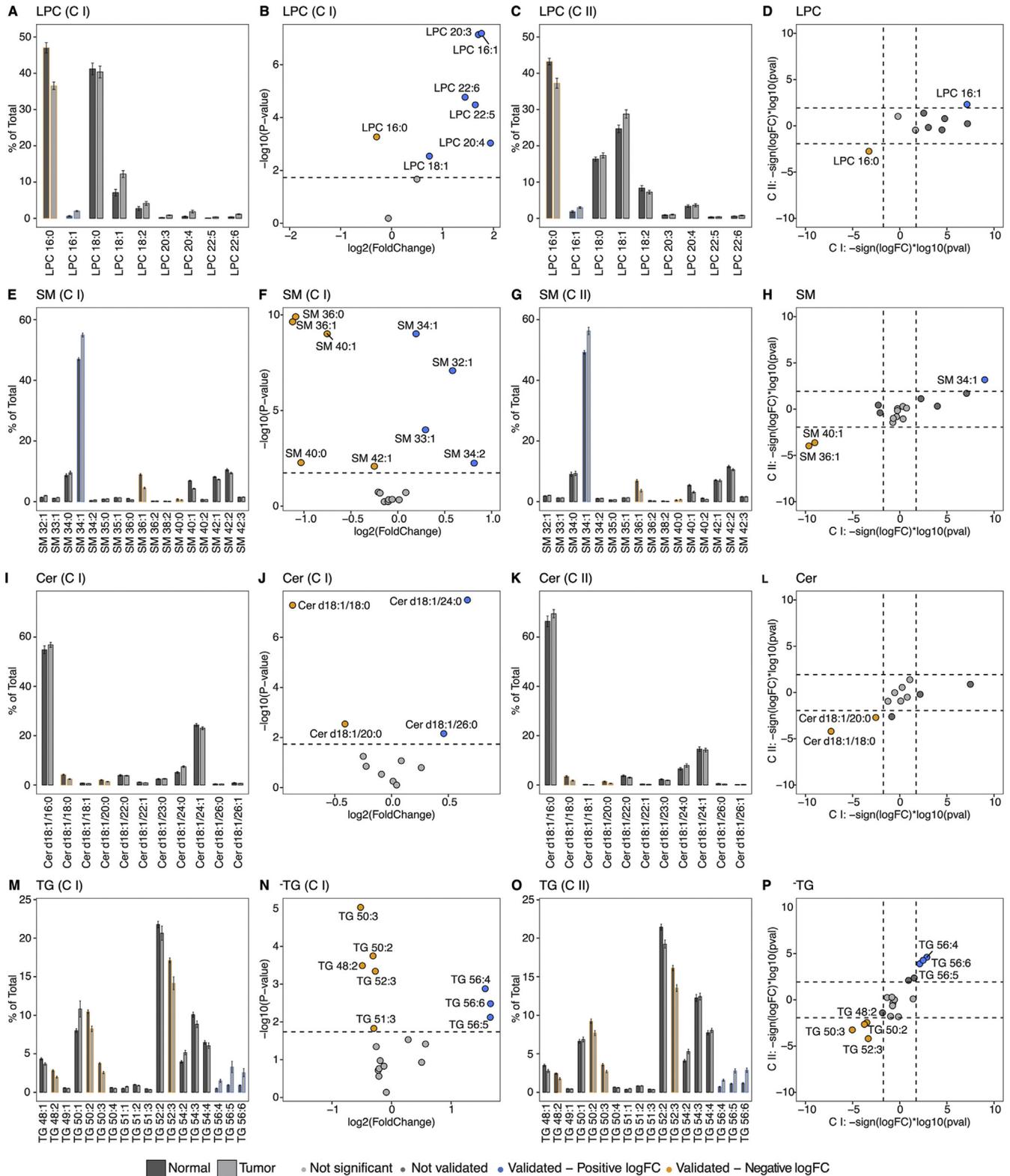
were decreased (Figure 2). For TG, species with <53 carbons were reduced while polyunsaturated with 56 carbons were enhanced in tumors (Figure 2). These results indicate that GPL, GL, and SM species pattern are significantly altered in tumor samples.

### Validation of the CRC-Specific Lipidome

To verify these results, an independent cohort (CII) was analyzed using identical methods as for CI. Surprisingly, dramatic differences occurred for GPL, despite essentially identical measurement conditions and highly similar clinicopathologic characteristics. The fraction of polyunsaturated GPL was up to 5-fold elevated in CII compared with CI. These included LPC 18:2, 20:4 (Figure 2C); PC 36:2, 36:3, 36:4, 38:4, 38:5; PE 36:2, 36:3, 38:4, 38:5; PE P-16:0/20:4, 18:0/20:4; and PI 38:4 (Supplementary Figure 1). In contrast, SL and GL profiles including SM, Cer, and TG were similar between both studies (Figure 2). In the paired comparison of matched normal and tumor samples, numerous lipid species showed similar changes in CI and CII. These included SM 34:1, 36:1, 40:1 (Figure 2); Cer d18:1/18:0, 20:0 (Figure 2); and TG 48:2, 50:2, 50:3, 52:3, 56:4, 56:5, 56:6 (Figure 2). These results indicate that SM, Cer, and TG profiles are significantly altered in human CRC, and that SL and GL species are the most suitable candidates to develop a CRC-specific lipid signature, whereas GPL species may be prone to degradation. To test whether pre-analytical conditions affect the lipidome, normal tissue and tumors from  $n = 3$  patients were split and either fresh-frozen immediately, stored for 1 hour at ambient temperature, or 24 hours at 4°C before lipidomic analysis (Supplementary Figure 2). Although SM, Cer, and TG profiles were stable even after 24 hours of sample storage, GPL profiles like LPC were strongly dependent on varying sample handling conditions (Supplementary Figure 2).

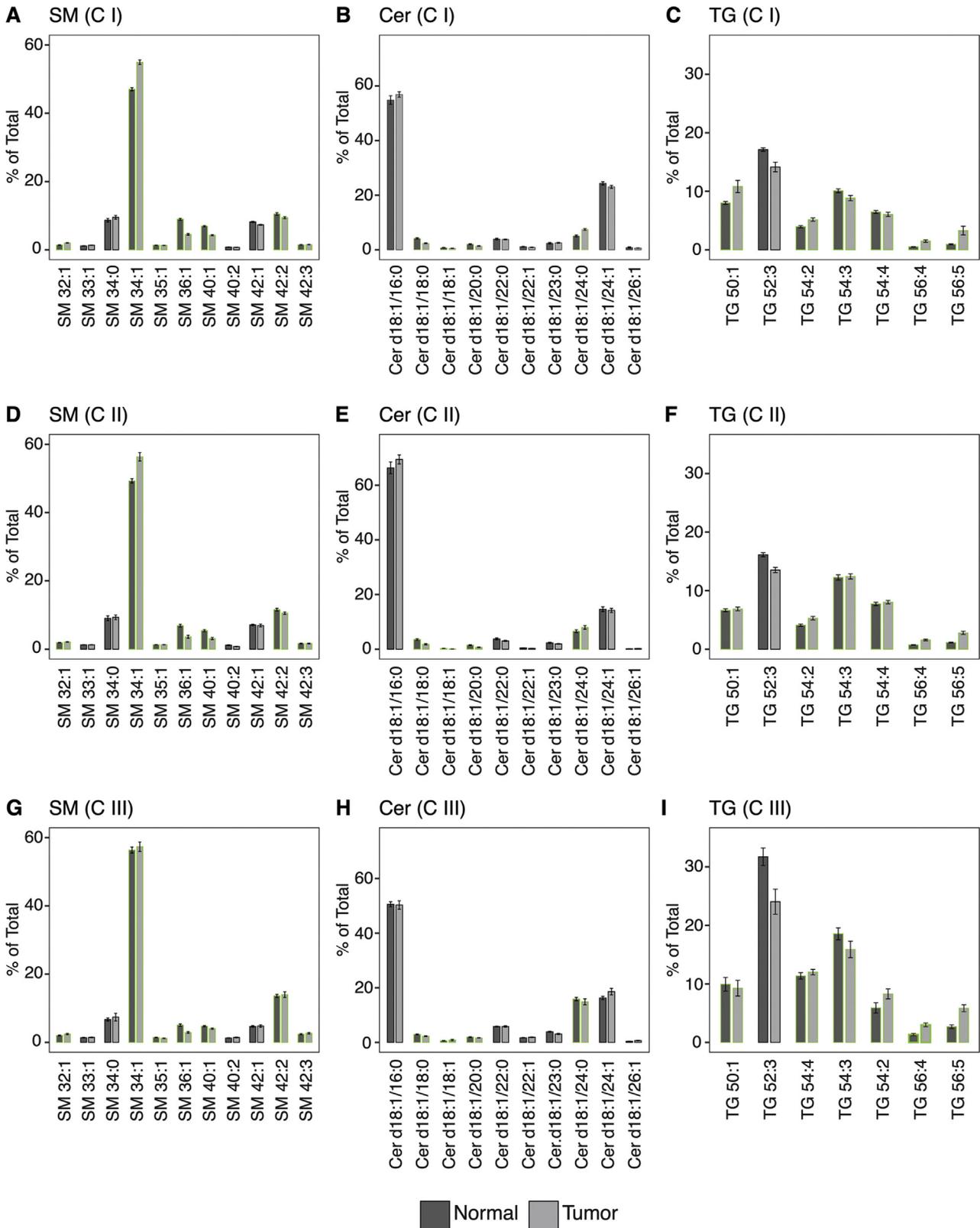
### Robust CRC-Specific Signatures Based on Glycerolipids and Sphingolipids

We investigated next whether SM, Cer, and TG species could be used to differentiate between tumor tissue and nondiseased mucosa. To this end, we trained a 5-fold cross-validated regression model based on LASSO-penalization on samples from CI and validated it on CII. For further and more stringent validation, an additional dataset was used (CIII), derived from a recently published study. This dataset was chosen deliberately as the most stringent control because it describes a clinically very similar set of patients with CRC, but it was derived in an entirely independent fashion, in a different clinical center from another city, using a different analysis method (HR-MS-based), and, furthermore, by a different team of clinicians and scientists. For this analysis, only the subset of lipid species common to all 3 cohorts was used. Hence, 7 of 11 for SM (Figure 3), 4 of 10 for Cer (Figure 3), and 6 of 7 for TG species (Figure 3) were selected by the model as predictors (Table 2). Classification scoring based on these TG, SM, and Cer signatures was effective in discriminating tumor from nondiseased tissue for in cohort CII with an area under the curve (AUC) >0.85



**Figure 2.** Differential analysis of lipid species in the discovery and validation cohorts: LPC, SM, Cer, and TG species in tumor vs nondiseased tissue samples. (A, E, I, M) Lipid species abundance in tumor and nondiseased tissue of C I. Bars indicate the mean abundance and standard error. (B, F, J, N) Volcano plots showing lipid species significantly different between tumor and nondiseased tissue. (C, G, K, O) Lipid species abundance in tumor and nondiseased tissue of C II. Right: (D, H, L, P) Comparison of the differential analysis results between Cohort C I and C II. Dark shaded bars, normal tissue; light shaded bars, tumor. Orange dots indicate validated species increased in tumors in both C I and C II; blue dots indicated validated lipid species decreased in tumors in C I and C II.

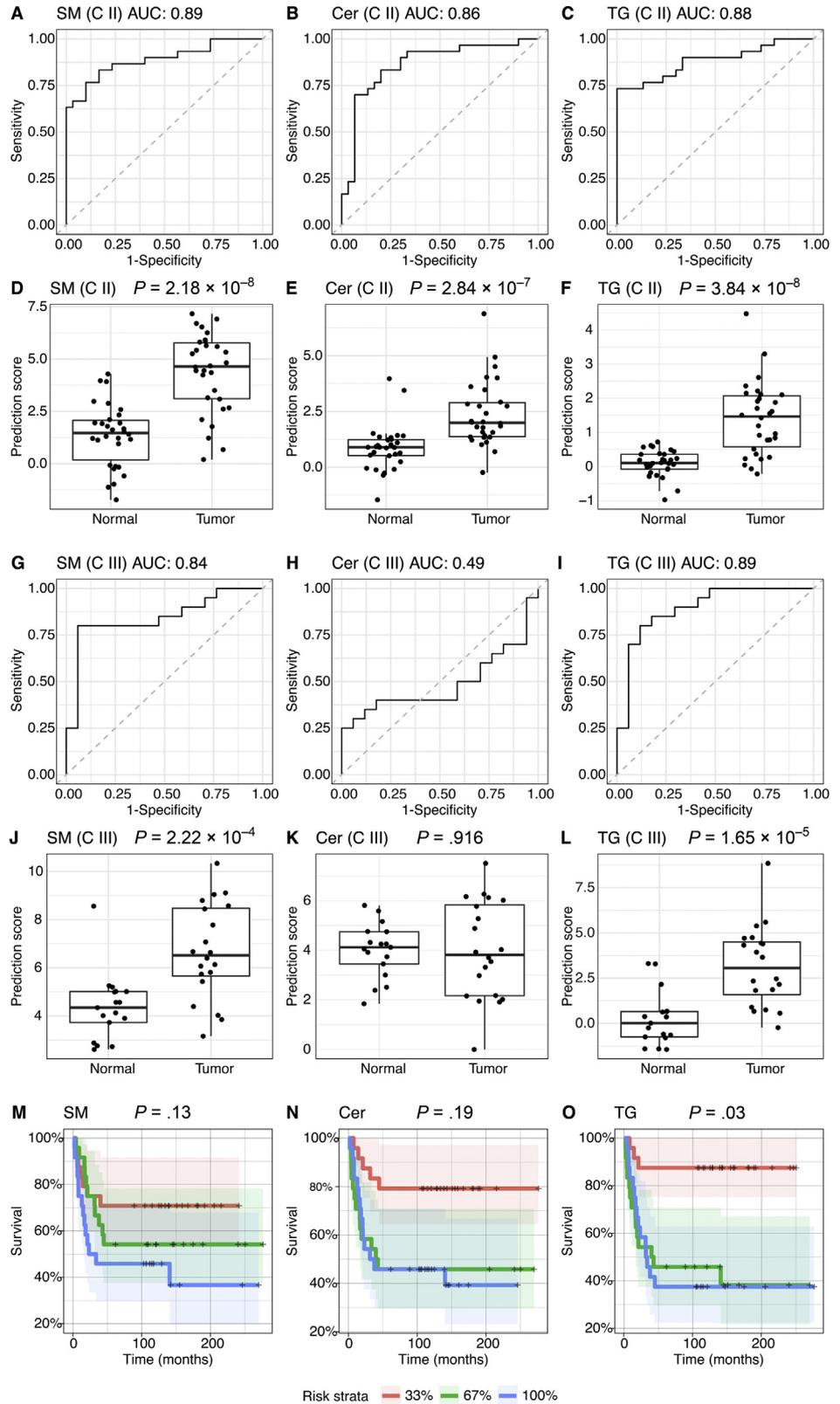
BASIC AND TRANSLATIONAL AT



**Figure 3.** Abundance of the subset of lipid species retained for the prediction model: SM, Cer, and TG species in tumor vs normal tissue (A, B, C) Discovery cohort CI. (D, E, F) Validation cohort CII. (G, H, I) Validation cohort CIII. *Dark shaded bars*, normal tissue; *light shaded bars*, tumor; *bars* designate mean abundance and standard error.

(Figure 4A-C). However, in cohort CIII-only signatures obtained from TG and SM, but not from Cer, allowed a stratification in tumor vs nontumor with an AUC >0.83.

(Figure 4G-I). This is also reflected in the prediction scores between the 2 groups, which showed strong significance in all cases except for Cer lipids in CIII (Figure 4D-F, J-L).



**Figure 4.** Prediction models for validation cohorts CII and CIII, and survival analysis based on lipid-based signatures. Stratification between tumor and nontumor tissue was based on lipid profiles and postoperative survival in CI. (A–C) AUC of the prediction model derived from CI tested on CII. (D–F) Prediction scores of the prediction model tested on CII. (G–I) AUC of the prediction model derived from cohort I tested on CIII. (J–L) Prediction scores of the prediction model tested on cohort CIII. (M–O) Recurrence-free postoperative survival for SM, Cer, and TG. Colored lines represent 3 risk strata, at 33.3% (red), 66.7% (green), and 100% (blue); colored areas represent confidence intervals.

Thus, validation on 2 independent clinical cohorts could confirm differences in the CRC lipidome for SM- and TG-based lipid species signatures. To exclude bias, an additional analysis was carried out including all SM, Cer, and TG

lipid species from CI, irrespective of whether they were detectable in the validation cohorts (Supplementary Figure 3 and Supplementary Figure 4). Of note, this approach significantly discriminated tumor from nontumor

**Table 2.** Lipid Species Used as Predictors for CRC-Related Signatures

SM	LASSO coefficients	Cer d18:1	LASSO coefficients	TG	LASSO coefficients
Intercept	1.09e+00	Intercept	5.44e-01	Intercept	4.76e-01
32:1	3.91e-01	18:0	-3.34e-01	50:1	1.70e-02
34:1	2.56e-02	18:1	-1.91e-01	54:2	1.27e-01
35:1	-6.06e-02	20:0	-5.05e-01	54:3	-1.31e01
36:1	-6.59e-01	24:0	3.49e-01	54:4	-5.51e-02
40:1	-3.78e-01			56:4	1.35e+00
42:2	2.82e-01			56:5	2.72e-02
42:3	1.01e+00				

tissue, with minor differences regarding the level of significance or the lipid species retained for the signature composition (Supplementary Figure 4, Supplementary Table 2).

### TG-Derived Signature is Prognostic For Postoperative Event-Free Survival

Finally, we tested whether TG, SM, and Cer species not only distinguish normal from cancer tissue, but also have prognostic potential. Based on postoperative follow-up data obtained in CI and CII, survival analysis was performed including clinical covariates. A ridge regularized model was used to avoid overfitting because the number of tumor samples ( $n = 101$ ) and events ( $n = 32$ ) was limiting. In fact, a significant event-free survival (EFS) prediction was achieved using a TG-based signature (Figure 4O). To ensure this presumed predictive power of TG lipids was not confounded by common covariates like age and gender, further survival models were considered including only the covariates, or the covariates and the lipids together (Supplementary Figure 5, ANOVA results in Supplementary Table 3). Our results indicate that these covariates are not confounding, confirming the observed predictive power of TG-derived signatures. The  $P$ -values for EFS based on SM and Cer failed to attain significance with values of .07 and .1, respectively (Figure 4M and N). Taken together, a TG species profile allowed an EFS prediction in patients with CRC. Elevated proportions of TG 56:4, 56:5, 56:6 (Figure 2M–P) were detected in tumor samples, containing polyunsaturated fatty acids (PUFA) (fatty acids 18:2, 20:2, 20:3, 20:4, 20:5, 22:4, 22:5, 22:6), which are solely of dietary origin (Supplementary Figure 6).

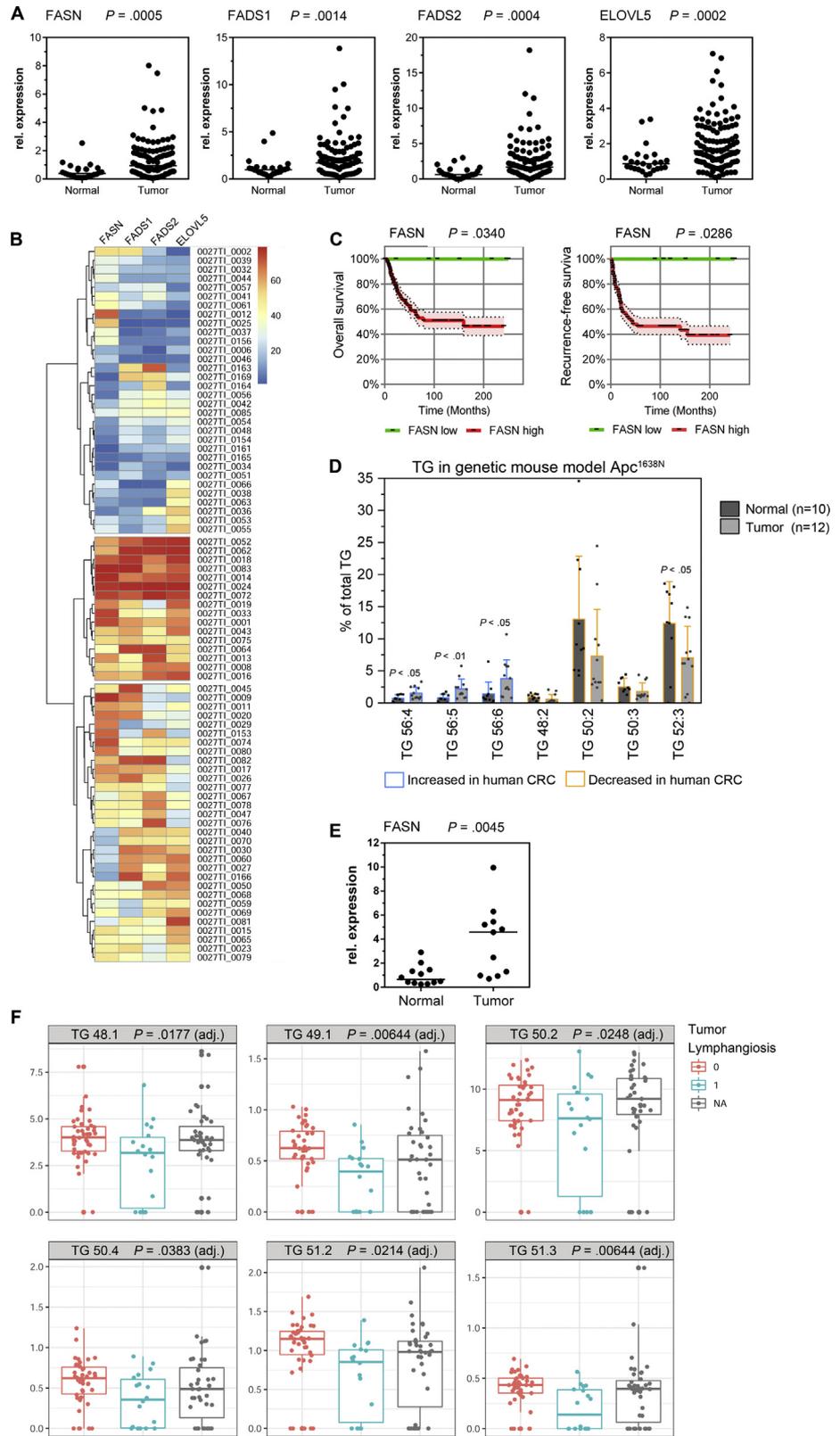
### Lipogenic and PUFA Metabolizing Enzymes Are Overexpressed in Tumors, and Deregulation of TG Species is Conserved in a Genetic Murine Tumor Model

To corroborate the biological basis of the findings, messenger RNA expression of lipogenic enzymes was analyzed using qPCR on 109 patients from CI and II

(Figure 5A). Expression of fatty acid synthase (FASN), the key enzyme of de novo lipogenesis, and of the PUFA metabolizing enzymes fatty acid elongase ELOVL5, fatty acid desaturases FADS1 and FADS2, were significantly increased in tumors ( $n = 109$ ) compared with matched colon mucosa ( $n = 25$ ). There was a remarkable degree of co-regulation among the 4 enzymes in individual tumors, shown in the form of a heatmap analysis (Figure 5B, Supplementary Table 4). There was no particular association between messenger RNA expression and specific lipid species (Supplementary Table 5, Supplementary Table 6), except for FASN expression and TG 52:3 (Supplementary Table 6). However, increased transcription of FASN was significantly associated with overall and recurrence-free survival (Figure 5C, Supplementary Figure 7). Next, in a translational approach, a well-established genetic mouse model for digestive cancer, the  $Apc^{1638N}$  line, was subjected to quantitative lipidomics. Tumors from small and large intestine and normal tissue were isolated from  $n = 10$  animals (aged 1 year), subjected to lipid extraction, and analyzed as described for the human specimen. Although the basic lipid species profile differs between humans and mice, TG containing poly-unsaturated acyl chains (TG 56:4, TG 56:5, TG 56:6) had a significantly higher abundance in tumors of mice, whereas shorter and less unsaturated TG like 52:3 were lower (Figure 5D), essentially identical to human tumors (Figure 2M–P), and the murine fatty acid synthase enzyme (Fasn) was significantly up-regulated in tumors (Figure 5E).

### TG Species Are Not Associated With Genetic Subtypes of CRC, But With Lymphatic Vessel Infiltration

To investigate putative associations of the cancer lipidome with genetic tumor subtypes, clinical and relevant molecular markers were tested in cohort CI. MSI was found in 14% (MSI-H), KRAS exon 2 mutations in 44%, and BRAF exon 15 mutations in 12% (mutually exclusive with KRAS; Supplementary Table 7, Supplementary Figure 8). Lipidomic alterations were detected throughout all subgroups, and none of the molecular features showed a specific association



**Figure 5.** Expression of lipogenic enzymes; TG profiles of tumors from murine cancer model; association of TG species with lymphangiosis. (A) Messenger RNA expression of lipogenic enzymes in colorectal cancer (n = 109, from CI and CII) vs normal colon (n = 25). (B) Co-regulation of FASN, FADS1/2, and ELOVL5 expression shown as heatmap after unsupervised clustering. (C) Kaplan-Meier analysis for overall and recurrence-free survival with respect to FASN expression (78 patients of CI and CII). *Colored lines* represent low and high expression levels, *colored areas* represent the corresponding confidence intervals. (D) TG profiles of tumors and normal tissue from *Apc*<sup>1638N</sup> mice (n = 10). (E) Messenger RNA expression of mouse Fasn in tumors compared with normal tissue. (F) TG species significantly associated with lymphatic vessel infiltration in human tumors (n = 61; CI) (0, no infiltration; 1, lymphatic vessel infiltration present; NA, not available).

with TG lipid species. Tumor grading or the density of tumor infiltrating lymphocytes (CD3, CD4, and CD8 positive cells) also failed to show an association (Supplementary Table 8,

Supplementary Figure 8). However, 6 TG species were significantly correlated with lymphatic vessel infiltration (Figure 5F, Supplementary Table 8).

## Discussion

There is an ongoing debate whether development and progression of CRC is related to a specific lipid environment, and it is well known that CRC development is influenced by diet, microbiome, and metabolic disorders.<sup>11–13,43</sup> However, no conclusive picture has yet emerged on lipidome of CRC and its clinical role.<sup>22–25</sup> Therefore, we conducted a comprehensive and quantitative analysis of clinical samples, revealing significantly and systematically altered SM and TG profiles in 3 independent CRC cohorts from 2 clinical centers, derived independently using different analysis methods.

No broad consensus has emerged on lipids as oncological biomarker so far, likely due to instability of many analytes. Hence, data analysis was applied here on species profiles of various lipid classes that were calculated from molar concentrations. This strategy is based on the following considerations: (1) lipid function in membranes is related to composition of its constituents that define their biochemical/biophysical properties and, therefore, needs to be accurately controlled and specific for individual cells and tissues;<sup>44–48</sup> (2) pathophysiology including oncological transformation may influence membrane homeostasis and change lipid profiles;<sup>49–51</sup> and (3) calculation of lipid profiles normalizes the data and factors out external references like wet weight or protein concentration. Because measurement of such reference parameters may convey inaccuracies, variations in lipid species profiles is commonly lower compared with species concentrations.

In our study, we noted a surprisingly high degree of variation between the cohorts regarding polyunsaturated GPL. Putative confounding variables, such as diet, prescribed drugs, and frequent comorbidities like obesity and diabetes, may inherently lead to inter-patient variability regarding lipid content and distribution of lipid species within the tissue. Further, any application of metabolite/lipid-based signatures as biomarkers requires biochemical stability of the contributing features, which is hard to implement and control during surgical or endoscopic interventions in daily clinical routine. However, in spite of preanalytical confounders, PUFA are inherently susceptible to lipid degradation due to their methylene bridges between the double bonds having reactive hydrogens that are easily abstracted by free radicals.<sup>52</sup> In contrast, sphingolipids like SM and Cer primarily contain saturated or mono-unsaturated acyl chains and, therefore, are not prone to radical-mediated degradation and are, hence, more suited as candidates for biomarker development. Sphingolipids are frequently discussed as disease biomarkers due to their bioactive properties.<sup>53,54</sup> Both Cer and SM are critical for cancer cell survival, growth, migration, and angiogenesis.<sup>55</sup> Cer promotes apoptosis through clustering of death receptors leading to induced cell cycle arrest and senescence, and is involved in therapy resistance against oxaliplatin in CRC.<sup>55–57</sup> Lipid species-specific effects are discussed for Cer, mediated by various Cer synthases, which are deregulated in cancer cells, leading to cancer-specific changes in Cer profiles that are also related to chemoresistance.<sup>58</sup>

Sphingomyelin synthase 2, which is an important regulator involved in Cer to SM conversion, was linked to colitis-associated colon cancer development.<sup>59</sup> In the present study, we were able to validate a SM-derived signature, but not a Cer-based signature, in the independent cohort CIII. This may be related to inter-patient or study center-related differences, or to analytical issues because CIII was obtained using a slightly different analytical method and may also be affected by the low absolute abundance of Cer.

Importantly, TG profiles proved to be robust, even after prolonged preanalytical storage, and were reproducibly detected in all tested cohorts. Thus, TGs may represent superior biomarker candidates for clinical samples. Biologically, this could be due to the localization of TGs in intracellular lipid droplets, which may contribute to their stability, and protects against radical damage. Lipid droplets are highly dynamic organelles, allowing cancer cells metabolic flexibility by providing fatty acids for generation of adenosine triphosphate after beta-oxidation or by storage of fatty acids to prevent lipotoxicity.<sup>60</sup> Our results demonstrate elevated proportions of TG 56:4, 56:5, and 56:6 containing PUFA (FA 18:2, 20:2, 20:3, 20:4, 20:5, 22:4, 22:5, 22:6) in tumor samples, which are solely of dietary origin. Fatty acids, as main lipid building blocks of TG, can either be synthesized *de novo* in cancer cells, reflected by increased intratumoral expression of lipogenic enzymes, and/or obtained exogenously. Exogenous sources from the tumor microenvironment include adipose tissue, the gut content containing dietary lipids that have not been absorbed in the small intestine, lipids transported via blood and lymph vessels, and short-chain fatty acids produced abundantly in the colon by intestinal microbiota that are further metabolized to more complex lipids.<sup>61</sup>

To assess the biological significance of the lipidome findings, they were complemented with molecular genetic evidence. Expression of major enzymes for lipogenesis, like FASN and FADS2,<sup>13</sup> was highly significantly up-regulated in cancer tissue, in accordance with earlier findings.<sup>14</sup> The enzymes were largely co-expressed and FASN expression was negatively associated with survival. Next, a putative association of the lipid signatures with molecular subgroups of CRC was assessed in cohort CI. Lipidomic signatures were not limited to one particular subgroup, but rather found in both mismatch repair-proficient and -deficient tumors, and were not significantly associated with oncogenic KRAS or BRAF status, nor with lymphocyte infiltration or tumor grading. Although the number of patients in our study is limited, it may be concluded that alteration of the lipidome is not restricted to a particular subset of CRC.

Because postoperative survival was significantly associated on TG species, they may be favored as biomarkers over SM. Of note, 6 TG species were significantly associated with lymphatic vessel infiltration. Lymph vessels play major roles for the transport of dietary lipids, especially chylomicrons and long-chain fatty acids.<sup>62</sup> In the tumor context, lymphatic vessels are mainly regarded as routes of metastasis, and whether tumor-associated lymph vessels shuttle dietary lipids to tumors is not well understood.<sup>63</sup> The present study

evaluated only lipid species but not molecular species, ie, acyl chain combinations. Analysis of such details is analytically more demanding, but may provide further insights into lipid biology and enhance the potential of the TG signature developed here for application in clinical routine, as a tool for risk stratification.<sup>64</sup> In the context of published studies, n-3 PUFAs were found increased and n-6 PUFAs decreased in CRC,<sup>65</sup> although opposing results were reported (decreased n-3 PUFAs and increased n-6 PUFAs).<sup>66</sup> Interestingly, 5 of 6 species from the TG-derived signature were also significantly up-regulated in hepatocellular cancer, indicating conserved patterns across tumor entities.<sup>67</sup> Last, analysis of an established genetic mouse model for human CRC, the *Apc*<sup>1638N</sup> line,<sup>42</sup> confirmed across species an up-regulation of Fasn expression, as well as a remarkable tumor-specific deregulation of TG lipids.

Taken together, we identified specific lipidomic changes in CRC compared with normal, nondiseased colorectal mucosa, using samples from 3 independent patient cohorts. Based on a 2-fold validation approach, we propose here a robust lipid signature based on TG that differentiates tumor from nontumor tissue and has prognostic significance.

## Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of Gastroenterology at [www.gastrojournal.org](http://www.gastrojournal.org), and at <https://doi.org/10.1053/j.gastro.2021.05.009>.

## References

- Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394–424.
- Siegel RL, Torre LA, Soerjomataram I, et al. Global patterns and trends in colorectal cancer incidence in young adults. *Gut* 2019;68:2179–2185.
- Guinney J, Dienstmann R, Wang X, et al. The consensus molecular subtypes of colorectal cancer. *Nat Med* 2015;21:1350–1356.
- Nitsche U, Maak M, Schuster T, et al. Prediction of prognosis is not improved by the seventh and latest edition of the TNM classification for colorectal cancer in a single-center collective. *Ann Surg* 2011;254:793–800; discussion 800–801.
- Shevchenko A, Simons K. Lipidomics: coming to grips with lipid diversity. *Nat Rev Mol Cell Biol* 2010;11:593–598.
- Snaebjornsson MT, Schulze A. Tumours use a metabolic twist to make lipids. *Nature* 2019;566:333–334.
- Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011;144:646–674.
- Santos CR, Schulze A. Lipid metabolism in cancer. *FEBS J* 2012;279:2610–2623.
- Mashima T, Sato S, Okabe S, et al. Acyl-CoA synthetase as a cancer survival factor: its inhibition enhances the efficacy of etoposide. *Cancer Sci* 2009;100:1556–1562.
- Ramirez de Molina A, Sarmentero-Estrada J, Belda-Iniesta C, et al. Expression of choline kinase alpha to predict outcome in patients with early-stage non-small-cell lung cancer: a retrospective study. *Lancet Oncol* 2007;8:889–897.
- Cifkova E, Holcapek M, Lisa M, et al. Determination of lipidomic differences between human breast cancer and surrounding normal tissues using HILIC-HPLC/ESI-MS and multivariate data analysis. *Anal Bioanal Chem* 2015;407:991–1002.
- Eggers LF, Muller J, Marella C, et al. Lipidomes of lung cancer and tumour-free lung tissues reveal distinct molecular signatures for cancer differentiation, age, inflammation, and pulmonary emphysema. *Sci Rep* 2017;7:11087.
- Vriens K, Christen S, Parik S, et al. Evidence for an alternative fatty acid desaturation pathway increasing cancer plasticity. *Nature* 2019;566:403–406.
- Zaytseva YY, Harris JW, Mitov MI, et al. Increased expression of fatty acid synthase provides a survival advantage to colorectal cancer cells via upregulation of cellular respiration. *Oncotarget* 2015;6:18891–18904.
- Peck B, Schug ZT, Zhang Q, et al. Inhibition of fatty acid desaturation is detrimental to cancer cell survival in metabolically compromised environments. *Cancer Metab* 2016;4:6.
- Zaytseva YY, Rychahou PG, Le AT, et al. Preclinical evaluation of novel fatty acid synthase inhibitors in primary colorectal cancer cells and a patient-derived xenograft model of colorectal cancer. *Oncotarget* 2018;9:24787–24800.
- O’Keefe SJ. Diet, microorganisms and their metabolites, and colon cancer. *Nat Rev Gastroenterol Hepatol* 2016;13:691–706.
- Goday A, Barneto I, Garcia-Almeida JM, et al. Obesity as a risk factor in cancer: a national consensus of the Spanish Society for the Study of Obesity and the Spanish Society of Medical Oncology. *Clin Transl Oncol* 2015;17:763–771.
- Mehta RS, Song M, Nishihara R, et al. Dietary patterns and risk of colorectal cancer: analysis by tumor location and molecular subtypes. *Gastroenterology* 2017;152:1944–1953.e1.
- Hofmanova J, Vaculova A, Lojek A, et al. Interaction of polyunsaturated fatty acids and sodium butyrate during apoptosis in HT-29 human colon adenocarcinoma cells. *Eur J Nutr* 2005;44:40–51.
- Kolar SS, Barhoumi R, Callaway ES, et al. Synergy between docosahexaenoic acid and butyrate elicits p53-independent apoptosis via mitochondrial Ca(2+) accumulation in colonocytes. *Am J Physiol Gastrointest Liver Physiol* 2007;293:G935–G943.
- Pakiet A, Kobiela J, Stepnowski P, et al. Changes in lipids composition and metabolism in colorectal cancer: a review. *Lipids Health Dis* 2019;18:29.
- Wang Y, Hinz S, Uckermann O, et al. Shotgun lipidomics-based characterization of the landscape of lipid metabolism in colorectal cancer. *Biochim Biophys Acta Mol Cell Biol Lipids* 2020;1865:158579.

24. Yan G, Li L, Zhu B, et al. Lipidome in colorectal cancer. *Oncotarget* 2016;7:33429–33439.
25. Bien T, Perl M, Machmuller AC, et al. MALDI-2 mass spectrometry and immunohistochemistry imaging of Gb3Cer, Gb4Cer, and further glycosphingolipids in human colorectal cancer tissue. *Anal Chem* 2020;92:7096–7105.
26. Kistner L, Doll D, Holtorf A, et al. Interferon-inducible CXCL chemokines are crucial immune modulators and survival predictors in colorectal cancer. *Oncotarget* 2017;8:89998–90012.
27. Nitsche U, Rosenberg R, Balmert A, et al. Integrative marker analysis allows risk assessment for metastasis in stage II colon cancer. *Ann Surg* 2012;256:763–771; discussion 771.
28. Bligh EG, Dyer WJ. A rapid method of total lipid extraction and purification. *Can J Biochem Physiol* 1959;37:911–917.
29. Liebisch G, Lieser B, Rathenberg J, et al. High-throughput quantification of phosphatidylcholine and sphingomyelin by electrospray ionization tandem mass spectrometry coupled with isotope correction algorithm. *Biochim Biophys Acta* 2004;1686:108–117.
30. Liebisch G, Drobnik W, Lieser B, et al. High-throughput quantification of lysophosphatidylcholine by electrospray ionization tandem mass spectrometry. *Clin Chem* 2002;48:2217–2224.
31. Matyash V, Liebisch G, Kurzchalia TV, et al. Lipid extraction by methyl-tert-butyl ether for high-throughput lipidomics. *J Lipid Res* 2008;49:1137–1146.
32. Zemski Berry KA, Murphy RC. Electrospray ionization tandem mass spectrometry of glycerophosphoethanolamine plasmalogen phospholipids. *J Am Soc Mass Spectrom* 2004;15:1499–1508.
33. Liebisch G, Drobnik W, Reil M, et al. Quantitative measurement of different ceramide species from crude cellular extracts by electrospray ionization tandem mass spectrometry (ESI-MS/MS). *J Lipid Res* 1999;40:1539–1546.
34. Horing M, Ejsing CS, Hermansson M, et al. Quantification of cholesterol and cholesteryl ester by direct flow injection high-resolution fourier transform mass spectrometry utilizing species-specific response factors. *Anal Chem* 2019;91:3459–3466.
35. Husen P, Tarasov K, Katafiasz M, et al. Analysis of lipid experiments (ALEX): a software framework for analysis of high-resolution shotgun lipidomics data. *PLoS One* 2013;8:e79736.
36. Liebisch G, Vizcaino JA, Kofeler H, et al. Shorthand notation for lipid structures derived from mass spectrometry. *J Lipid Res* 2013;54:1523–1530.
37. Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. *Stat Med* 1990;9:811–818.
38. Harrell FE Jr, Lee KL, Califf RM, et al. Regression modelling strategies for improved prognostic prediction. *Stat Med* 1984;3:143–152.
39. Simon N, Friedman J, Hastie T, et al. Regularization paths for Cox's proportional hazards model via coordinate descent. *J Stat Softw* 2011;39:1–13.
40. Sperlich A, Balmert A, Doll D, et al. Genetic and immunological biomarkers predict metastatic disease recurrence in stage III colon cancer. *BMC Cancer* 2018;18:998.
41. Franke FC, Slusarenko BO, Engleitner T, et al. Novel role for CRK adaptor proteins as essential components of SRC/FAK signaling for epithelial-mesenchymal transition and colorectal cancer aggressiveness. *Int J Cancer* 2020;147:1715–1731.
42. Janssen KP, Alberici P, Fsihi H, et al. APC and oncogenic KRAS are synergistic in enhancing Wnt signaling in intestinal tumor formation and progression. *Gastroenterology* 2006;131:1096–1109.
43. Cifkova E, Holcapek M, Lisa M, et al. Lipidomic differentiation between human kidney tumors and surrounding normal tissues using HILIC-HPLC/ESI-MS and multivariate data analysis. *J Chromatogr B Analyt Technol Biomed Life Sci* 2015;1000:14–21.
44. van Meer G, de Kroon AI. Lipid map of the mammalian cell. *J Cell Sci* 2011;124:5–8.
45. Ecker J, Liebisch G, Englmaier M, et al. Induction of fatty acid synthesis is a key requirement for phagocytic differentiation of human monocytes. *Proc Natl Acad Sci U S A* 2010;107:7817–7822.
46. Hermansson M, Hokynar K, Somerharju P. Mechanisms of glycerophospholipid homeostasis in mammalian cells. *Prog Lipid Res* 2011;50:240–257.
47. Levental KR, Malmberg E, Symons JL, et al. Lipidomic and biophysical homeostasis of mammalian membranes counteracts dietary lipid perturbations to maintain cellular fitness. *Nat Commun* 2020;11:1339.
48. Schweizer S, Liebisch G, Oeckl J, et al. The lipidome of primary murine white, brite, and brown adipocytes: Impact of beta-adrenergic stimulation. *PLoS Biol* 2019;17:e3000412.
49. Snaebjornsson MT, Janaki-Raman S, Schulze A. Greasing the wheels of the cancer machine: the role of lipid metabolism in cancer. *Cell Metab* 2020;31:62–76.
50. van der Veen JN, Kennelly JP, Wan S, et al. The critical role of phosphatidylcholine and phosphatidylethanolamine metabolism in health and disease. *Biochim Biophys Acta Biomembr* 2017;1859:1558–1572.
51. Wang B, Tontonoz P. Phospholipid remodeling in physiology and disease. *Annu Rev Physiol* 2019;81:165–188.
52. Rudy MD, Kainz MJ, Graeve M, et al. Handling and storage procedures have variable effects on fatty acid content in fishes with different lipid quantities. *PLoS One* 2016;11:e0160497.
53. Ekroos K, Lavrynenko O, Titz B, et al. Lipid-based biomarkers for CVD, COPD, and aging - a translational perspective. *Prog Lipid Res* 2020;78:101030.
54. Hannun YA, Obeid LM. Sphingolipids and their metabolism in physiology and disease. *Nat Rev Mol Cell Biol* 2018;19:175–191.
55. Morad SA, Messner MC, Levin JC, et al. Potential role of acid ceramidase in conversion of cytostatic to cytotoxic end-point in pancreatic cancer cells. *Cancer Chemother Pharmacol* 2013;71:635–645.

56. Madigan JP, Robey RW, Poprawski JE, et al. A role for ceramide glycosylation in resistance to oxaliplatin in colorectal cancer. *Exp Cell Res* 2020;388:111860.
57. Morad SA, Cabot MC. Ceramide-orchestrated signalling in cancer cells. *Nat Rev Cancer* 2013;13:51–65.
58. Brachtendorf S, El-Hindi K, Grosch S. Ceramide synthases in cancer therapy and chemoresistance. *Prog Lipid Res* 2019;74:160–185.
59. Ohnishi T, Hashizume C, Taniguchi M, et al. Sphingomyelin synthase 2 deficiency inhibits the induction of murine colitis-associated colon cancer. *FASEB J* 2017;31:3816–3830.
60. Corbet C, Bastien E, Santiago de Jesus JP, et al. TGFbeta2-induced formation of lipid droplets supports acidosis-driven EMT and the metastatic spreading of cancer cells. *Nat Commun* 2020;11:454.
61. Kindt A, Liebisch G, Clavel T, et al. The gut microbiota promotes hepatic fatty acid desaturation and elongation in mice. *Nat Commun* 2018;9:3760.
62. Dixon JB. Lymphatic lipid transport: sewer or subway? *Trends Endocrinol Metab* 2010;21:480–487.
63. Garnier L, Gkoutidi AO, Hugues S. Tumor-associated lymphatic vessel features and immunomodulatory functions. *Front Immunol* 2019;10:720.
64. Ellis SR, Paine MRL, Eijkel GB, et al. Automated, parallel mass spectrometry imaging and structural identification of lipids. *Nat Methods* 2018;15:515–518.
65. Yang K, Li H, Dong J, et al. Expression profile of polyunsaturated fatty acids in colorectal cancer. *World J Gastroenterol* 2015;21:2405–2412.
66. Zhang J, Zhang L, Ye X, et al. Characteristics of fatty acid distribution is associated with colorectal cancer prognosis. *Prostaglandins Leukot Essent Fatty Acids* 2013;88:355–360.
67. Li Z, Guan M, Lin Y, et al. Aberrant lipid metabolism in hepatocellular carcinoma revealed by liver lipidomics. *Int J Mol Sci* 2017;18:2550–2565.

Munich, Gregor-Mendel-Str. 2, 85354 Freising, Germany. e-mail: [josef.ecker@tum.de](mailto:josef.ecker@tum.de); Gerhard Liebisch, PhD, Institute of Clinical Chemistry and Laboratory Medicine, University Hospital Regensburg, Franz-Josef-Strauß-Allee 11, 93053 Regensburg, Germany. e-mail: [gerhard.liebisch@ukr.de](mailto:gerhard.liebisch@ukr.de); or Klaus-Peter Janssen, PhD, Department of Surgery, Klinikum rechts der Isar, TUM, Ismaninger Str. 22, 81675 Munich, Germany. e-mail: [Klaus-Peter.Janssen@tum.de](mailto:Klaus-Peter.Janssen@tum.de).

#### Acknowledgments

The authors thank Anja Conrad, Widya Johannes, and Doreen Müller for expert technical assistance and Matthias Maak, MD for scientific illustrations.

#### CRedit Authorship Contributions

Josef Ecker, PhD (Data curation: Lead; Formal analysis: Lead; Funding acquisition: Lead; Investigation: Lead; Methodology: Lead; Project administration: Lead; Supervision: Lead; Writing – original draft: Lead; Writing – review & editing: Lead)

Elisa Benedetti, PhD (Data curation: Equal; Investigation: Equal; Software: Equal; Validation: Equal; Visualization: Equal)

Alida S.D. Kindt, PhD (Data curation: Equal; Formal analysis: Equal; Software: Equal; Validation: Equal; Visualization: Equal)

Marcus Hoering, MSc (Data curation: Supporting; Formal analysis: Supporting; Investigation: Supporting; Validation: Supporting)

Markus Perl, MD (Data curation: Supporting; Formal analysis: Supporting; Investigation: Supporting; Methodology: Supporting; Writing – original draft: Supporting)

Andrea C. Machmueller, MSc (Data curation: Supporting; Formal analysis: Supporting; Investigation: Supporting; Methodology: Supporting)

Anna Sichler, MSc (Data curation: Supporting; Formal analysis: Supporting; Investigation: Supporting)

Johannes Plagge, MSc (Data curation: Supporting; Formal analysis: Supporting; Investigation: Supporting; Methodology: Supporting)

Yuting Wang, PhD (Data curation: Equal; Formal analysis: Equal; Validation: Equal; Writing – original draft: Supporting)

Sebastian Zeissig, MD (Conceptualization: Supporting; Funding acquisition: Supporting; Resources: Equal; Writing – original draft: Supporting)

Andrej Shevchenko, PhD (Conceptualization: Supporting; Methodology: Supporting; Supervision: Supporting; Validation: Supporting)

Ralph Burkhardt, MD (Funding acquisition: Supporting; Project administration: Supporting; Supervision: Supporting; Writing – original draft: Supporting)

Jan Krumsiek, PhD (Investigation: Supporting; Methodology: Equal; Software: Equal; Supervision: Equal; Validation: Equal; Visualization: Equal)

Gerhard Liebisch, PhD (Conceptualization: Equal; Data curation: Lead; Formal analysis: Lead; Funding acquisition: Lead; Investigation: Lead; Methodology: Equal; Project administration: Equal; Resources: Equal; Supervision: Equal; Validation: Lead; Visualization: Equal; Writing – original draft: Equal)

Klaus-Peter Janssen, PhD (Conceptualization: Lead; Formal analysis: Equal; Funding acquisition: Lead; Investigation: Equal; Project administration: Lead; Resources: Lead; Supervision: Lead; Validation: Supporting; Writing – original draft: Lead).

#### Conflicts of interest

The authors disclose no conflicts.

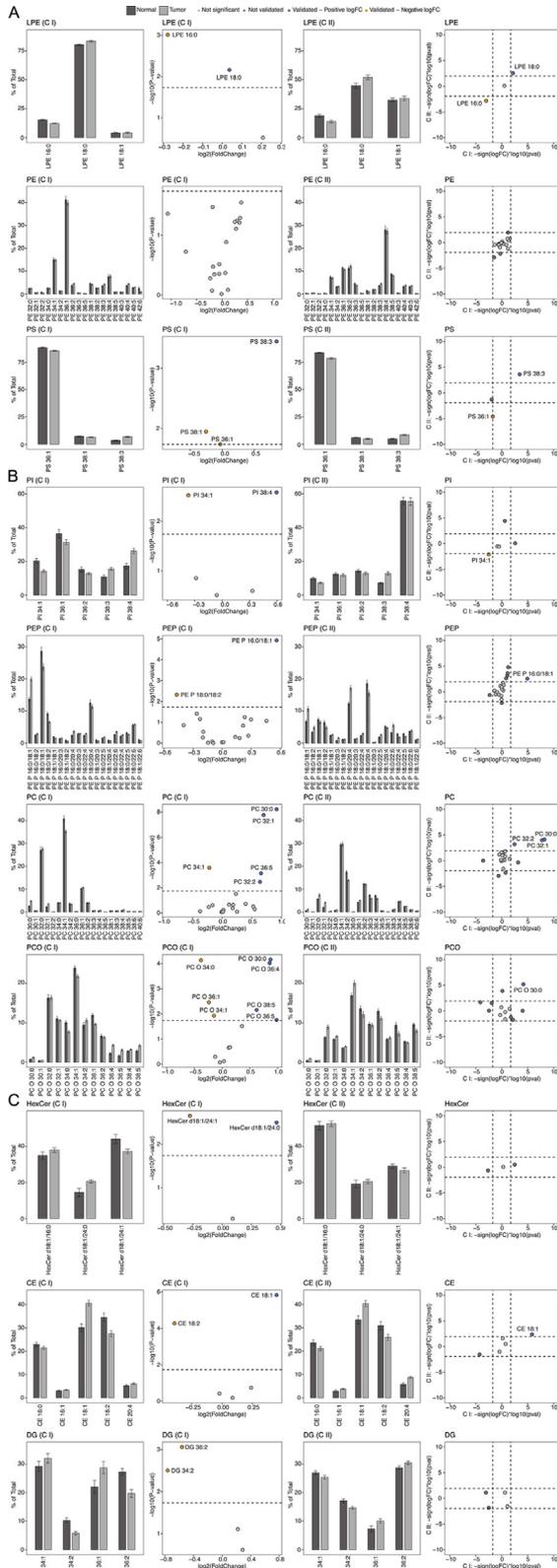
#### Funding

This work was funded by Deutsche Forschungsgemeinschaft grant 395357507 - SFB 1371 to J.E., S.Z., and K.P.J., and grant 209933838 - SFB1052 to R.B., and LI 923/4-1 - SPP 1656 to G.L.

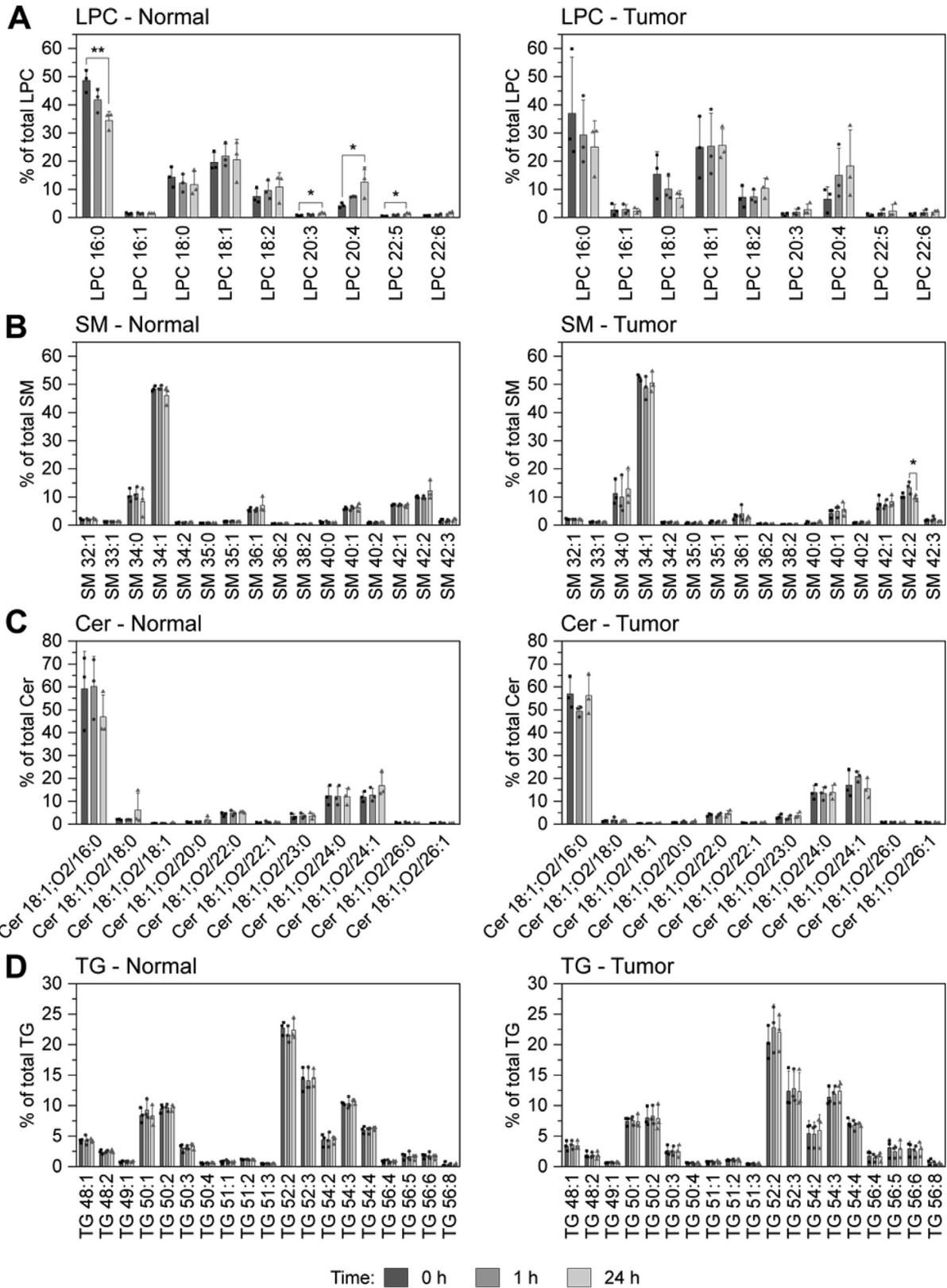
Received August 21, 2020. Accepted May 6, 2021.

#### Reprint requests

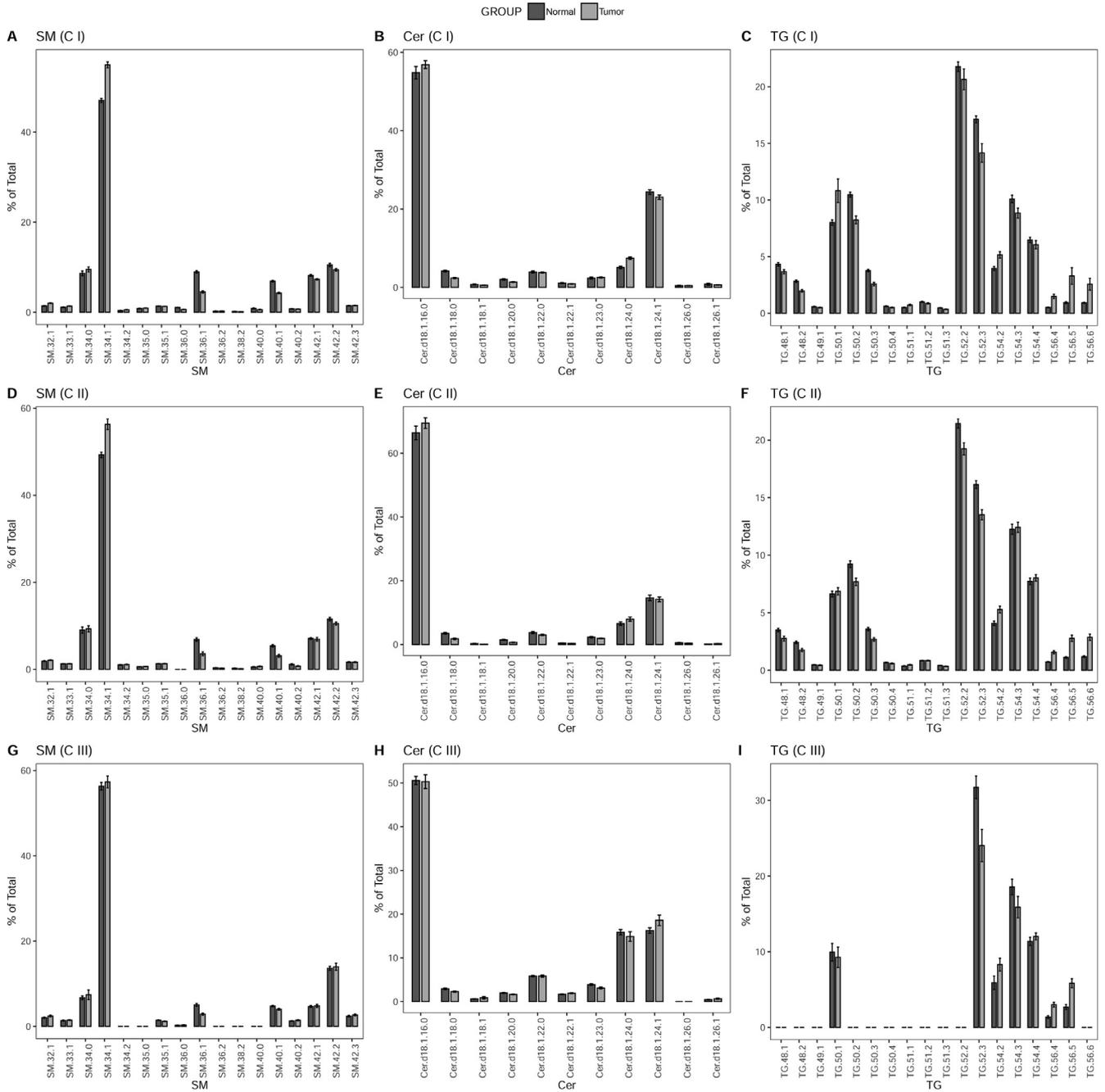
Address requests for reprints to: Josef Ecker, PhD, Research Group Lipid Metabolism, ZIEL – Institute for Food & Health, Technical University of



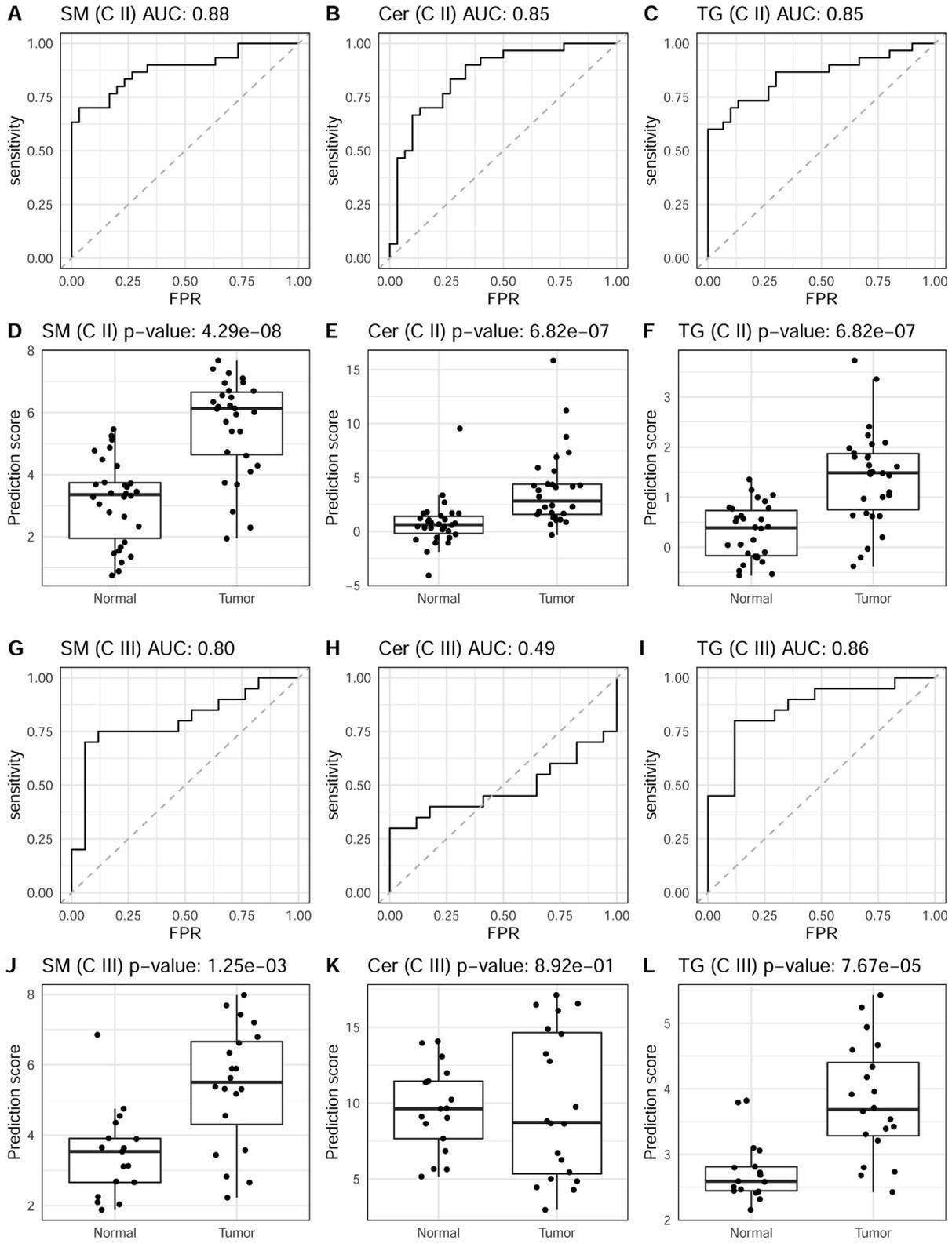
**Supplementary Figure 1.** Differential analysis of lipid species abundance in tumor vs nondiseased tissue samples of the discovery cohort CI, and the validation cohort CII, respectively. Bars indicate the mean abundance and standard error in either CI (*left column*) or CII. Volcano plots showing lipid species significantly differing between tumor and nondiseased tissue. Lipid species abundance in tumor and nondiseased tissue of Cohort II. *Right column*: comparison of the differential analysis results between cohorts CI and CII. (A) LPE, PE, and PS. (B) PI, PEP, PC, and PCO. (C) HexCer, CE, and DG.



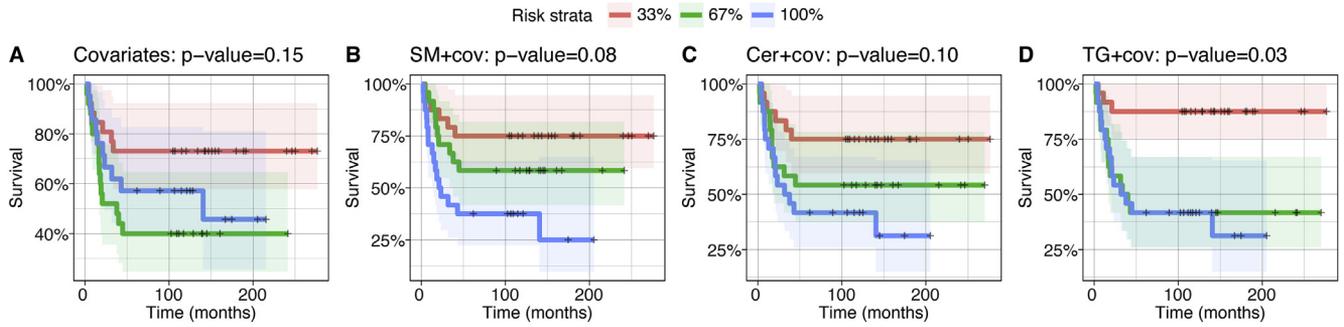
**Supplementary Figure 2.** Stability of lipid species profiles of normal and tumor tissue of  $n = 3$  patients after storage for 1 h at room temperature or 24 h at 4°C for LPC (A), SM (B), Cer (C), and TG (D). Bars designate mean abundance and standard error, \* $P < .05$ , \*\* $P < .01$  using a Student  $t$  test.



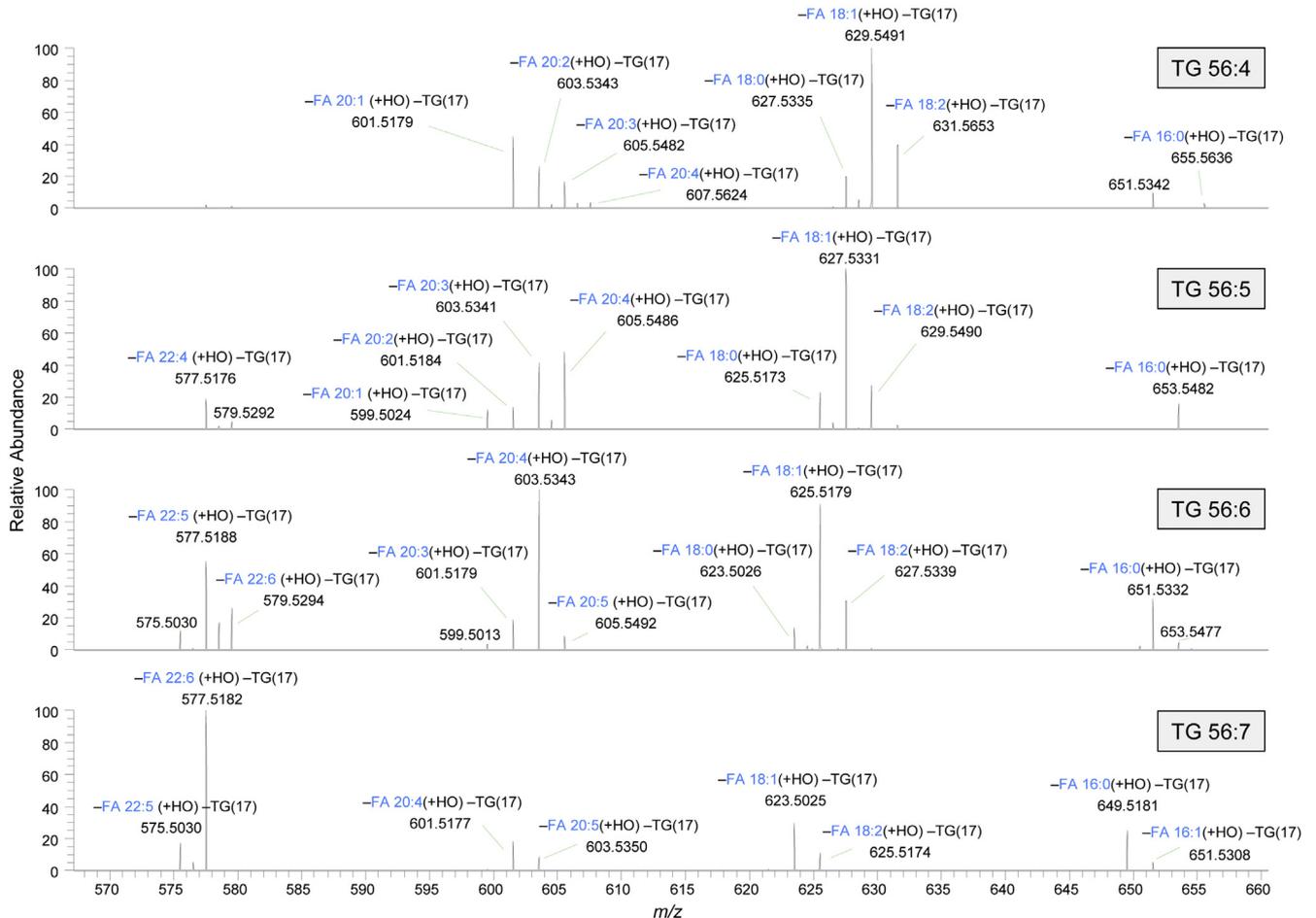
**Supplementary Figure 3.** Abundance of lipid species used for the prediction model containing all SM, CER, and TG lipid species detected in tumor vs normal tissue for the cohort C I. (A–C) Results for discovery cohort C I. (D–F) Validation cohort C II. (G–I) Validation cohort C III. *Dark shaded bars*, normal tissue; *light shaded bars*, tumor; *bars* designate mean abundance and standard error.



**Supplementary Figure 4.** Prediction models for validation cohorts CII and CIII using all lipid species detected in CI on lipid-based signatures. Stratification between tumor and nontumor tissue was based on lipid profiles. (A–C) AUC of the prediction model derived from CI tested on CII. (D–F) Prediction scores of the prediction model tested on CII. (G–I) AUC of the prediction model derived from cohort I tested on CIII. (J–L) Prediction scores of the prediction model tested on cohort CIII.

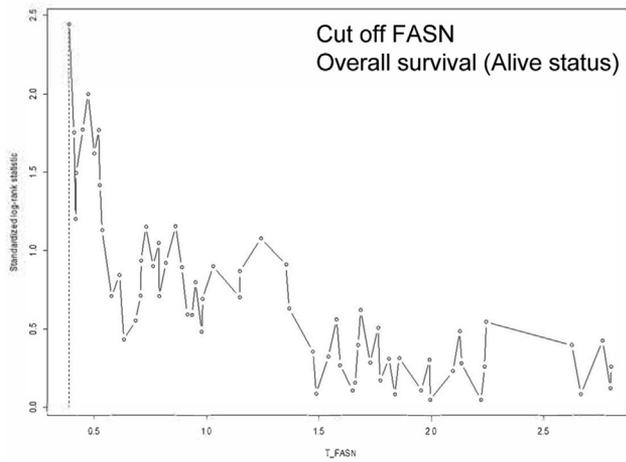


**Supplementary Figure 5.** Survival analyses including only the covariates (age and gender) (A), as well as the lipid signatures and covariates for SM (B), Cer (C), and TG (D). The colored lines represent 3 risk strata, at 33.3% (red), 66.7% (green), and 100% (blue). The colored areas represent the corresponding confidence intervals.

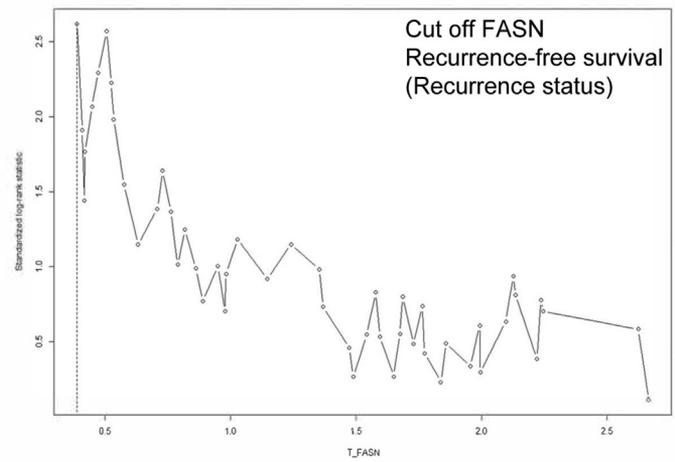


**Supplementary Figure 6.** Product ion spectra of TG 56:4 ( $m/z$  928.8328), TG 56:5 ( $m/z$  926.8171), TG 56:6 ( $m/z$  924.8015), and TG 56:7 ( $m/z$  922.7858) of a human CRC sample. Spectra were recorded in positive ion mode at mass resolution setting 35,000 (at  $m/z$  200) using an isolation window of 1 d, normalized collision energy of 18%, maximal IT of 64 ms, and AGC of  $1 \times 10^5$ .

Threshold derivation by maximally selected logRank statistics for FASN mRNA expression in tumor samples, for stratified Kaplan-Meier survival analysis

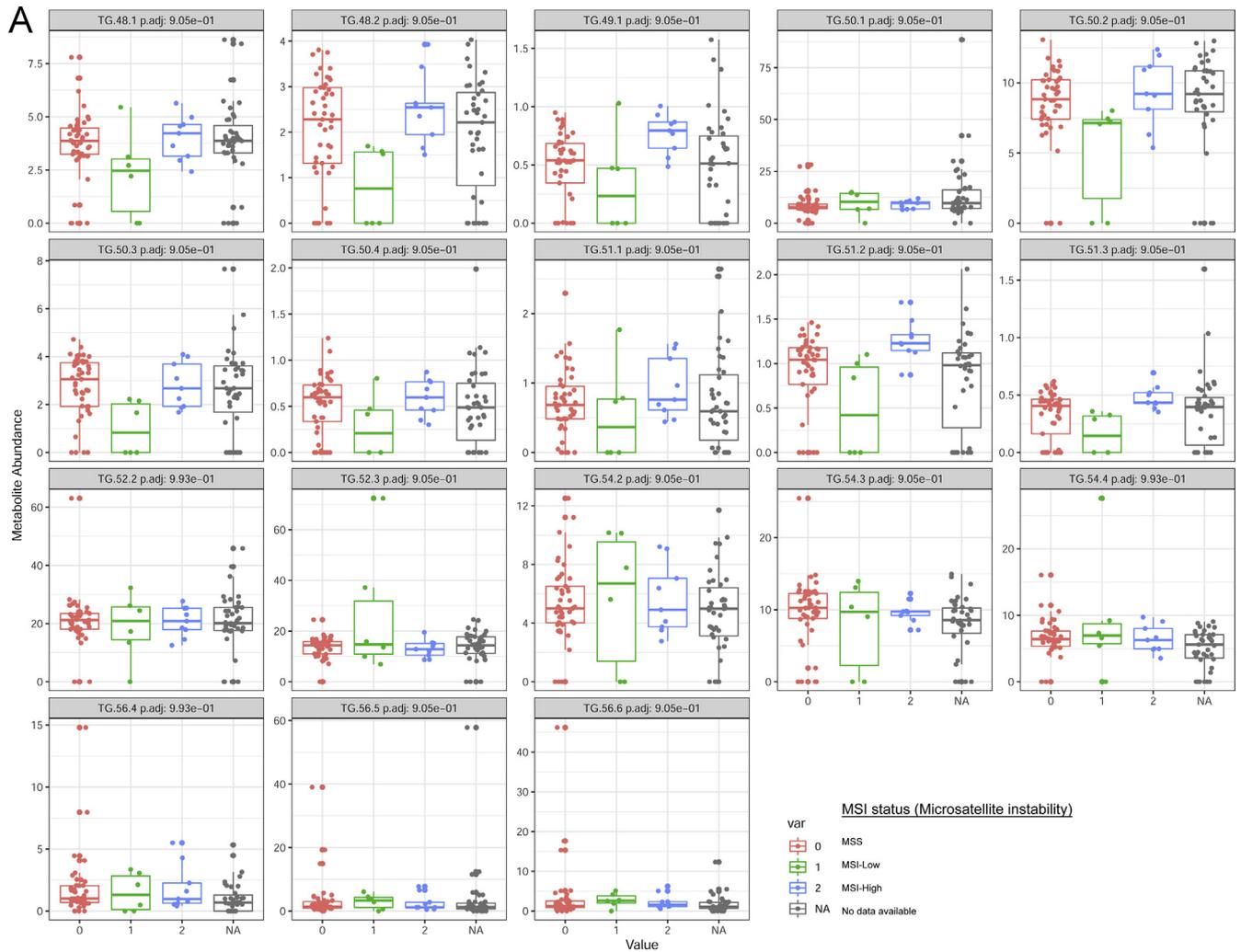


$M = 2.4439$ ,  $p\text{-value} = 0.1592$   
estimated cutpoint: 0.3879 (rel. expression)

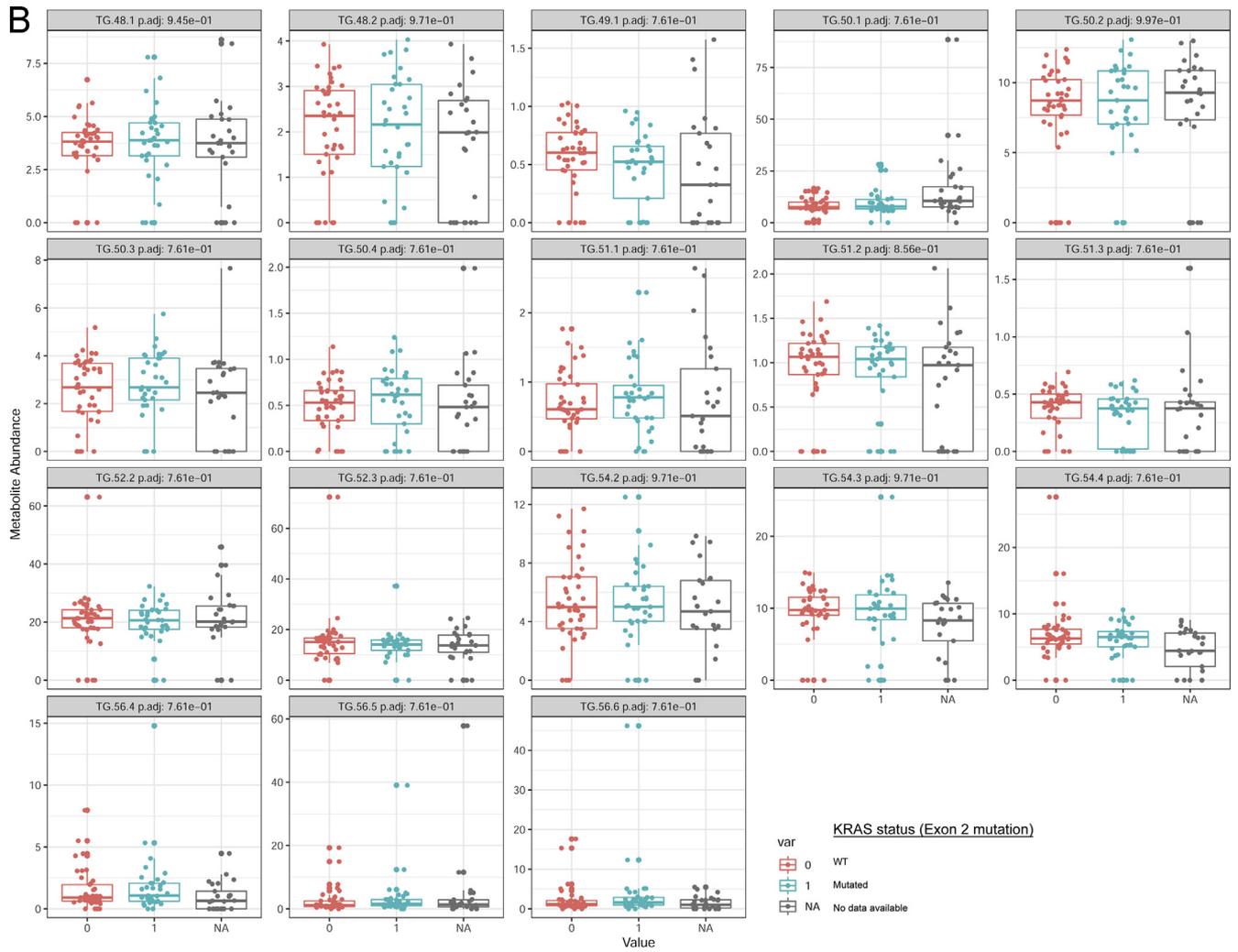


$M = 2.6179$ ,  $p\text{-value} = 0.1111$   
estimated cutpoint: 0.3879 (rel. expression)

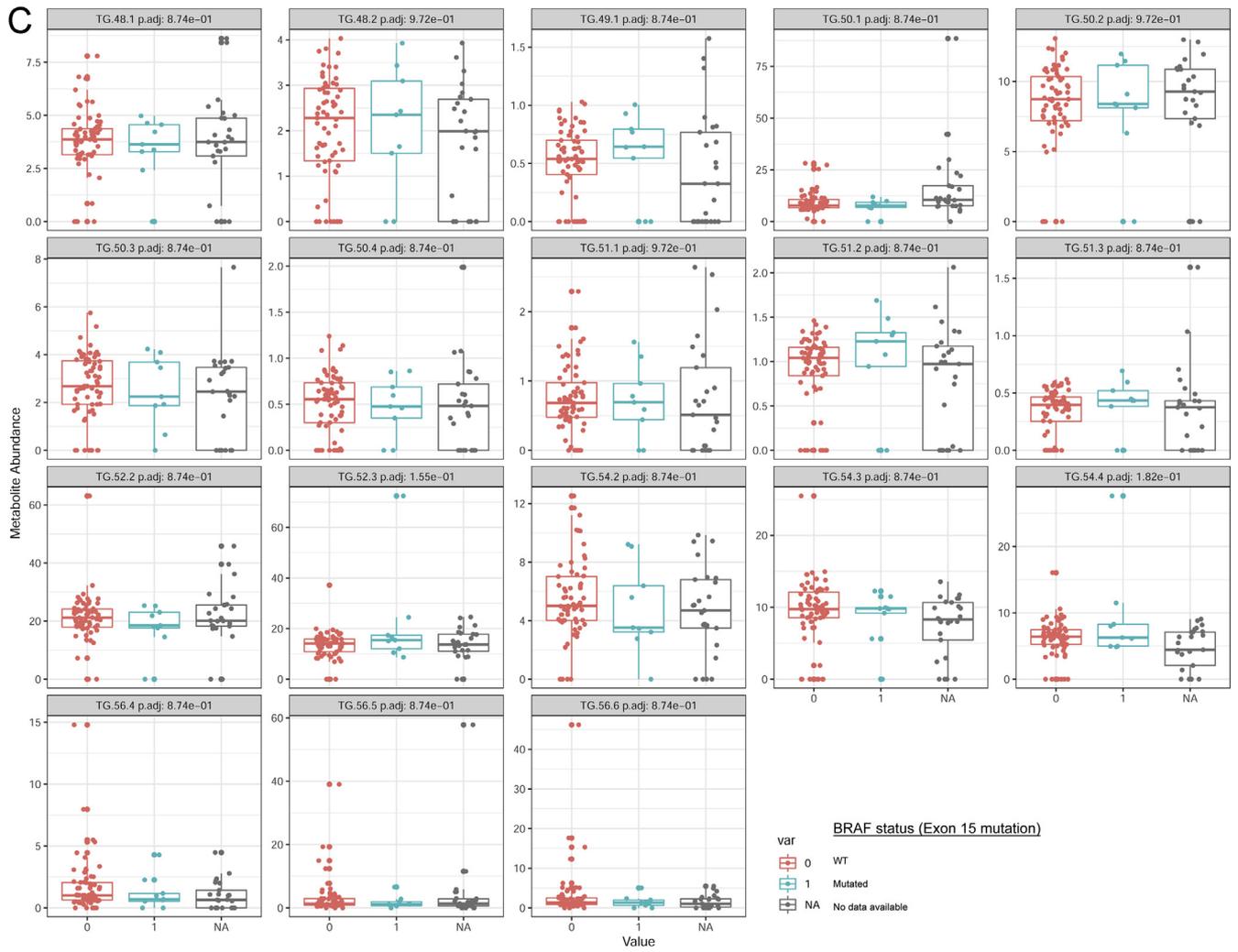
**Supplementary Figure 7.** Threshold derivation by maximally selected log-rank statistics for FASN messenger RNA expression in tumor samples, to derive optimal cut-offs for Kaplan-Meier survival analysis (postoperative overall survival, and postoperative tumor recurrence-free survival, respectively). The method for threshold derivation has been described in detail elsewhere (Nitsche et al<sup>27</sup>).



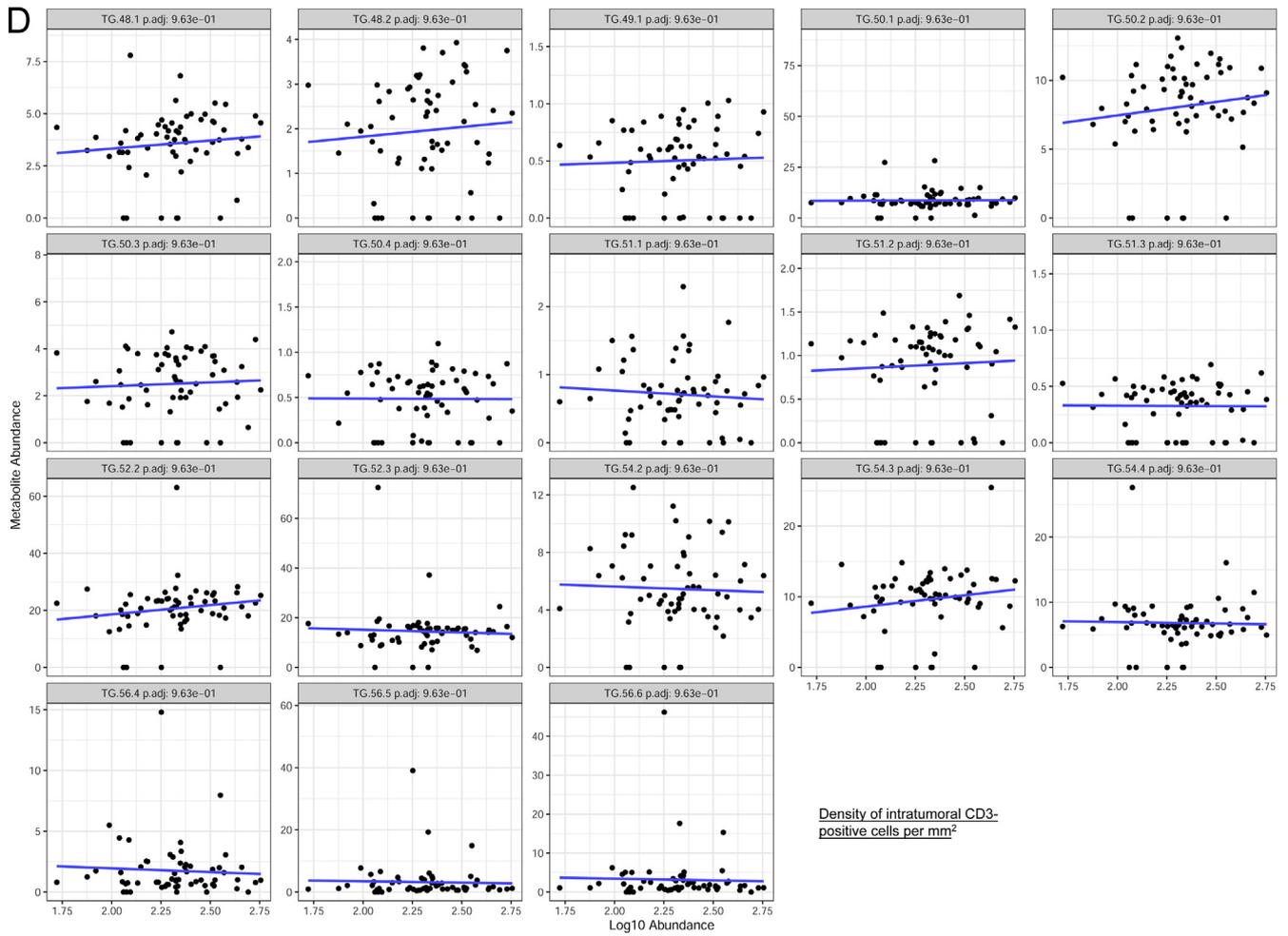
**Supplementary Figure 8.** Association of TG species with molecular genomic and clinical parameters determined for patients from the discovery cohort CI. Association was tested for the following parameters, each depicted on separate pages: (A) MSI (DNA MSI), (B) KRAS exon 2 mutations; (C) BRAF exon 15 mutations; (D) density of intratumoral CD3-positive cells/mm<sup>2</sup> of tumor tissue; (E) density of intratumoral CD4-positive cells/mm<sup>2</sup> of tumor tissue; (F) density of intratumoral CD8-positive cells/mm<sup>2</sup> of tumor tissue; (G) with histopathologic tumor grading; and (H) with intratumoral lymphatic vessel infiltration (lymphangiogenesis).



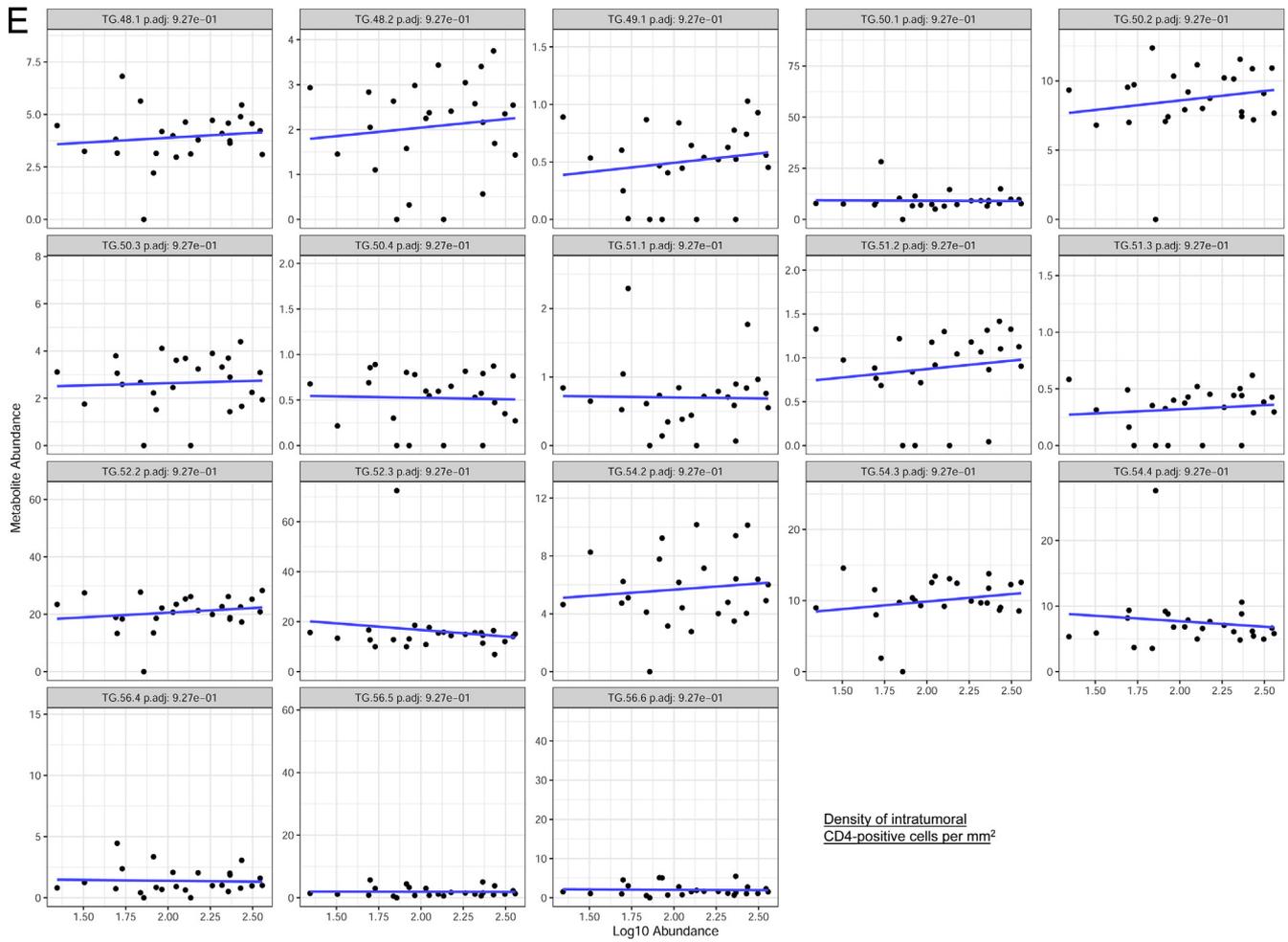
Supplementary Figure 8. Continued.



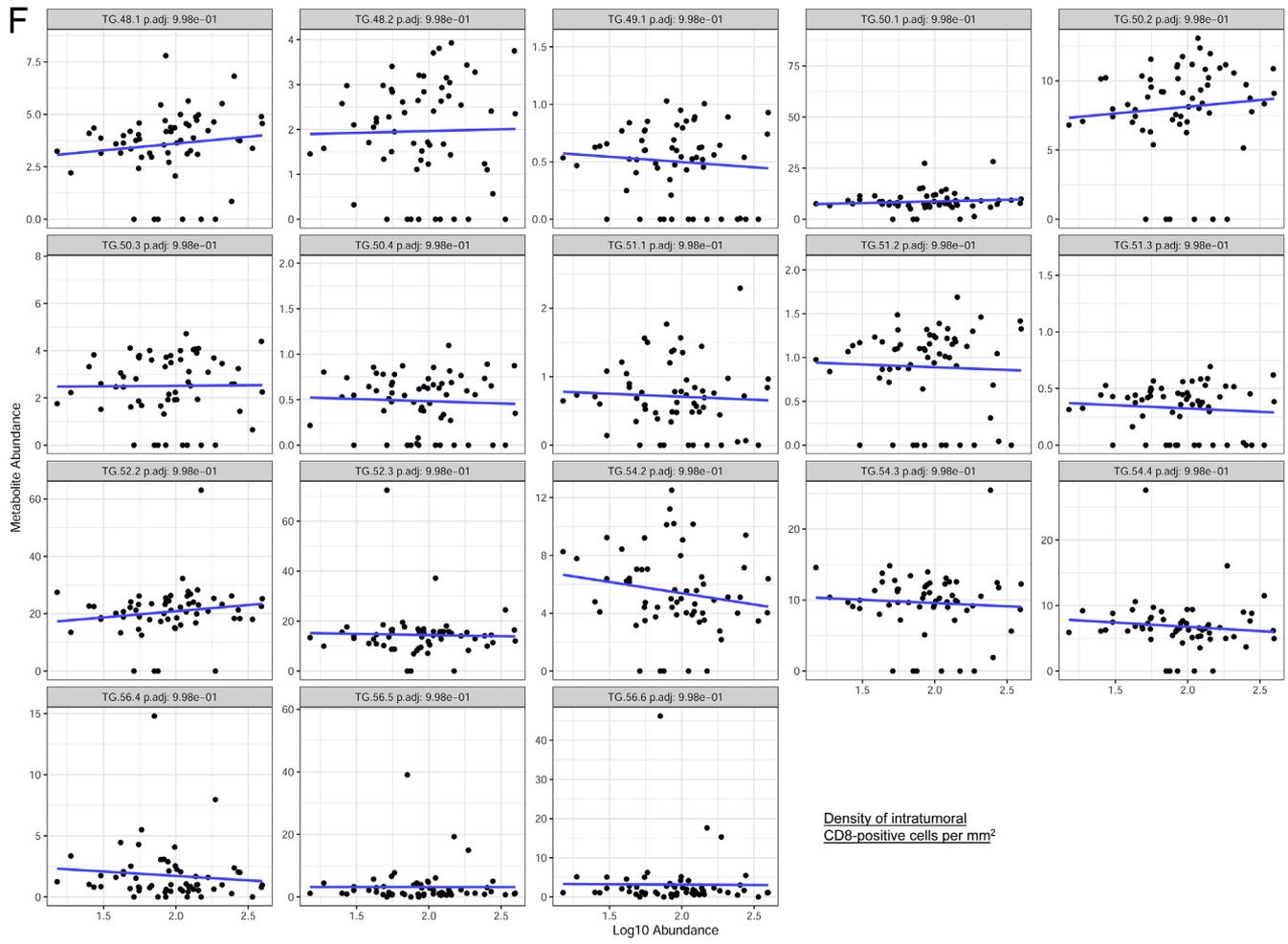
Supplementary Figure 8. Continued.



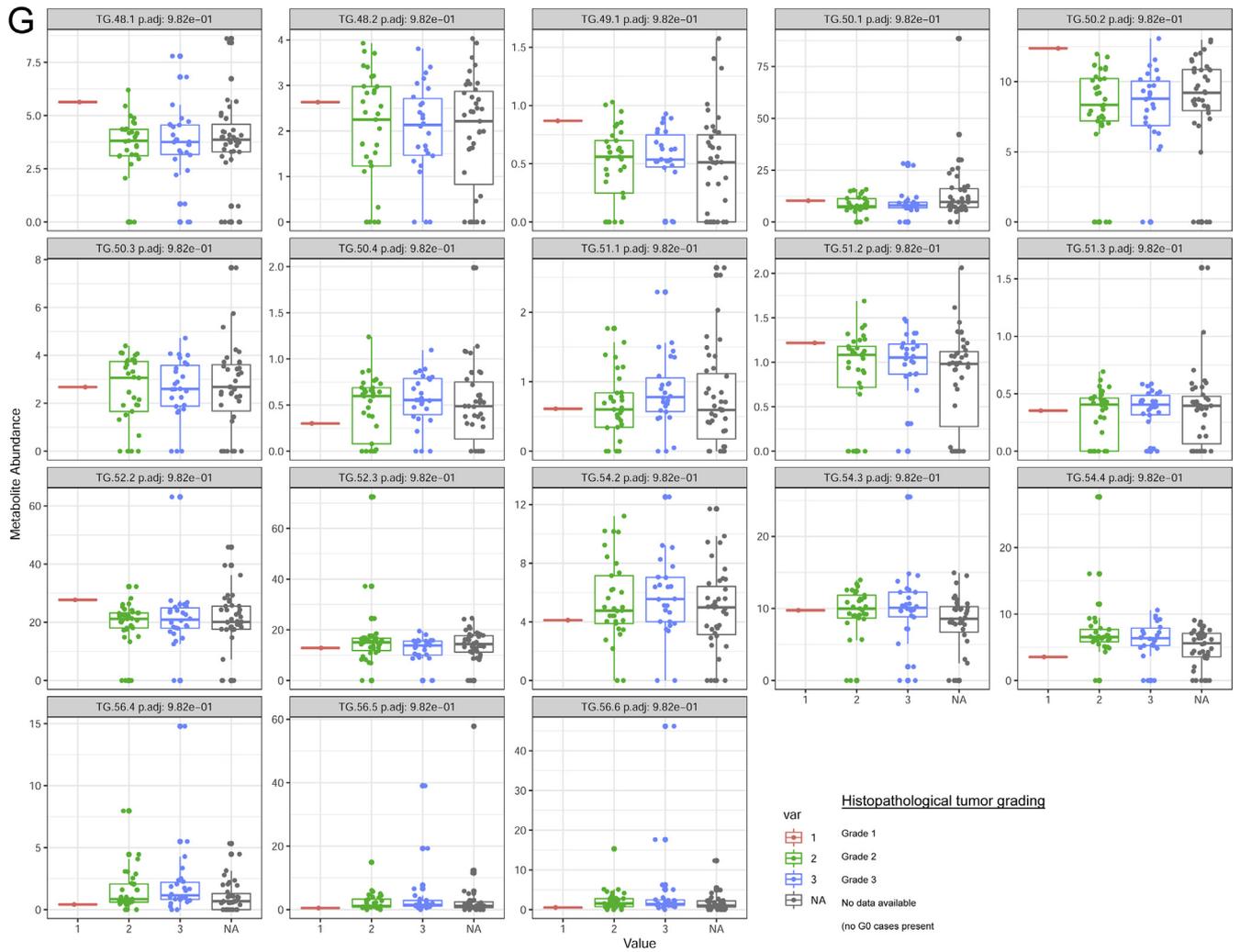
Supplementary Figure 8. Continued.



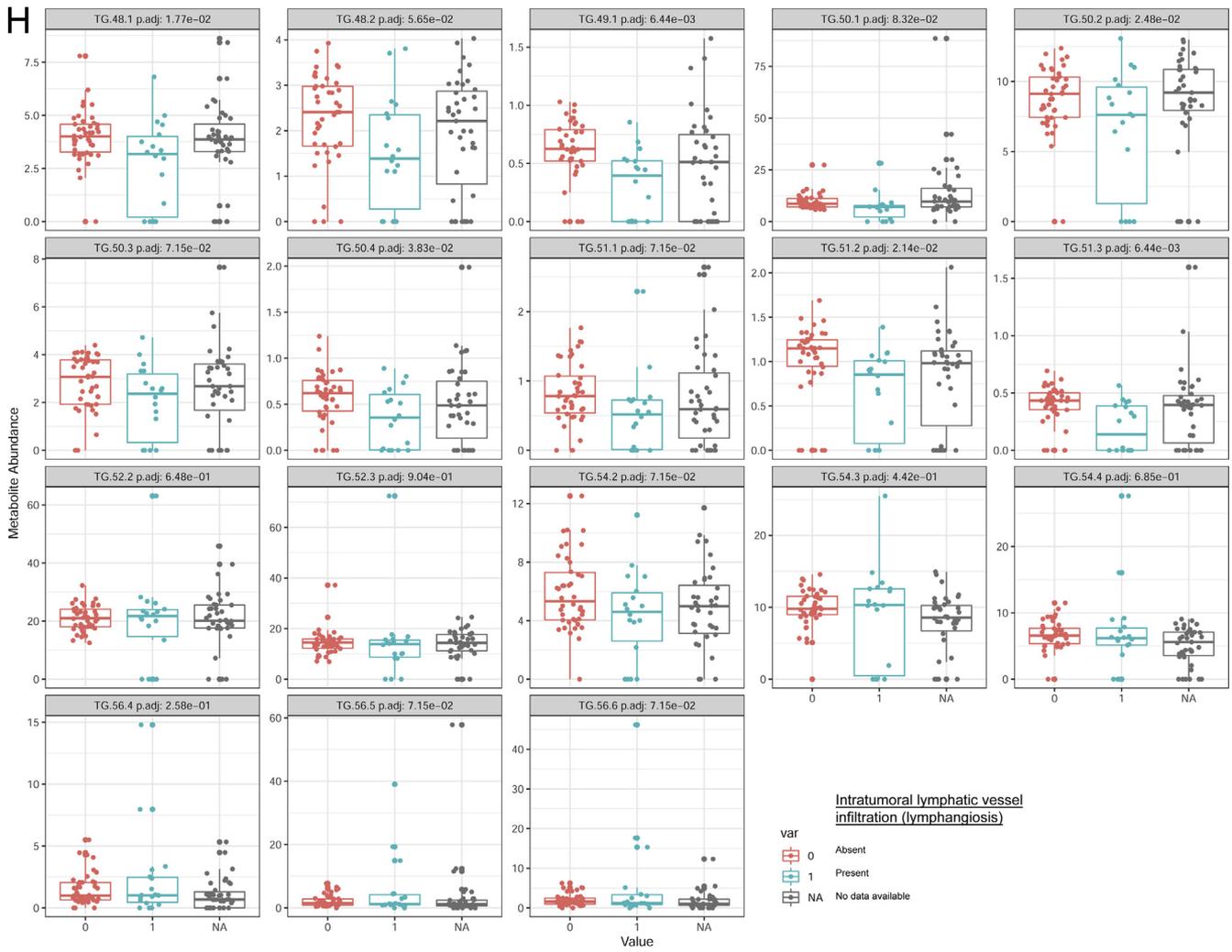
Supplementary Figure 8. Continued.



Supplementary Figure 8. Continued.



Supplementary Figure 8. Continued.



Supplementary Figure 8. Continued.

**Supplementary Table 1.** Methods Used for Quantitative Lipidomic Analysis of GPL, SL, GL, and Sterols

Category	Lipid class	Analysis
GPL	PC, PC O, PE, PS, PI, PE P, LPC, LPE	FIA-MS/MS
SL	SM, Cer, Hex-Cer	FIA-MS/MS
GL	DG, TG	FIA-FTMS
Sterols	FC CE	FIA-MSX FIA-FTMS

Hex-Cer, hexosylceramide; MSX, multiplexed high-resolution spectrometry.

**Supplementary Table 2.** Lipid Species Used as Predictors for CRC-Related Signatures, Including All Lipid Species Detected CI Without A Priori Changes (Lipid Species Not Detected in CII or CIII Were Imputed by Zero Values)

SM	LASSO coefficients	CER	LASSO coefficients	TG	LASSO coefficients
Intercept	7.18e+00	Intercept	-1.28e+00	Intercept	2.01e+00
34:2	5.60e-01	18:0	-7.22e-01	50:2	-5.07e-02
35:1	-1.12e+00	18:1	-2.67e-02	50:3	-4.99e-01
36:0	-2.02e+00	23:0	-9.20e-01	51:1	4.45e-01
36:1	-5.79e-01	24:0	1.00e+00	54:3	-6.86e-03
36:2	8.25e-01	24:1	4.64e-02	56:4	6.32e-01
40:1	-2.91e-01	26:0	-9.95e-01		
42:2	7.09e-02	26:1	-1.20e-01		
42:3	7.60e-01				

**Supplementary Table 3.** Results of ANOVA Test Between the Different Survival Models

	ANOVA <i>P</i> - value (only covariates vs covariates and lipids)	ANOVA <i>P</i> value (only lipids vs covariates and lipids)
SM	9.42 E-2	1.07 E-1
CER	1.34 E-1	9.99 E-2
TG	2.85 E-2	1.44 E-1

**Supplementary Table 4.** Unsupervised Cluster Analysis for Messenger RNA Expression of Lipogenic Enzymes (qPCR Relative Expression Data Normalized to Housekeeping Transcript HPRT, Heatmap)

SAMPLE_NAME	FASN	FADS1	FADS2	ELOVL5	RankSum	Cluster
0027TI_0082	57	72	74	22	225	3
0027TI_0017	64	60	59	35	218	3
0027TI_0060	24	63	60	67	214	3
0027TI_0030	23	68	64	57	212	3
0027TI_0074	72	38	49	49	208	3
0027TI_0026	60	69	50	27	206	3
0027TI_0166	13	73	51	69	206	3
0027TI_0050	37	32	68	64	201	3
0027TI_0020	68	65	34	30.5	197.5	3
0027TI_0068	39	41	56	59	195	3
0027TI_0078	44	45	65	40	194	3
0027TI_0045	48	74	29	42	193	3
0027TI_0009	77	67	25	20	189	3
0027TI_0011	62	59	27	41	189	3
0027TI_0015	51	37	40	60	188	3
0027TI_0040	20	58	55	53.5	186.5	3
0027TI_0027	16	64	41	63	184	3
0027TI_0081	27	47	35	74	183	3
0027TI_0067	32	49	63	37	181	3
0027TI_0080	67	34	44	36	181	3
0027TI_0023	47	53	28	50	178	3
0027TI_0076	45	36	72	24	177	3
0027TI_0069	41	20	52	62	175	3
0027TI_0153	69	24	61	21	175	3
0027TI_0065	43	40	32	58	173	3
0027TI_0047	46	42	58	26	172	3
0027TI_0079	50	39	31	45	165	3
0027TI_0077	49	48	43	23	163	3
0027TI_0070	17	50	48	46	161	3
0027TI_0059	34	21	46	44	145	3
0027TI_0029	66	13	23	33	135	3
0027TI_0024	76	78	79	79	312	2
0027TI_0072	75	71	78	75	299	2
0027TI_0083	79	79	62	71	291	2
0027TI_0062	58	77	77	78	290	2
0027TI_0018	71	76	66	76	289	2
0027TI_0014	78	66	70	73	287	2
0027TI_0052	61	70	76	77	284	2
0027TI_0001	73	62	57	65	257	2

Supplementary Table 4. Continued

SAMPLE_NAME	FASN	FADS1	FADS2	ELOVL5	RankSum	Cluster
0027TI_0016	59	56	69	72	256	2
0027TI_0064	40	75	75	53.5	243.5	2
0027TI_0008	54	46	67	66	233	2
0027TI_0013	63	43	73	52	231	2
0027TI_0043	55	61	45	68	229	2
0027TI_0033	74	44	38	61	217	2
0027TI_0019	65	54	26	70	215	2
0027TI_0075	53	57	47	48	205	2
0027TI_0163	10	52	71	10	143	1
0027TI_0169	2	55	53	29	139	1
0027TI_0085	29	35	36	38	138	1
0027TI_0042	18	33	42	34	127	1
0027TI_0002	52	51	19	2	124	1
0027TI_0055	28	11	22	56	117	1
0027TI_0036	7	14	39	55	115	1
0027TI_0041	38	22	33	17	110	1
0027TI_0053	25	15	21	47	108	1
0027TI_0056	5	31	37	28	101	1
0027TI_0054	11	30	24	32	97	1
0027TI_0061	42	29	18	7	96	1
0027TI_0038	19	9	9	51	88	1
0027TI_0164	8	17	54	9	88	1
0027TI_0057	22	28	30	6	86	1
0027TI_0012	70	5	8	1	84	1
0027TI_0039	31	26	14	12	83	1
0027TI_0032	30	18	17	15	80	1
0027TI_0154	6	27	16	30.5	79.5	1
0027TI_0048	9	25	15	25	74	1
0027TI_0044	33	12	12	14	71	1
0027TI_0066	26	2	3	39	70	1
0027TI_0063	12	8	5	43	68	1
0027TI_0051	14	23	10	16	63	1
0027TI_0025	56	1	2	3	62	1
0027TI_0156	36	7	7	8	58	1
0027TI_0037	35	3	6	13	57	1
0027TI_0161	1	16	20	19	56	1
0027TI_0006	15	10	1	18	44	1
0027TI_0034	3	19	11	4	37	1
0027TI_0046	21	4	4	5	34	1
0027TI_0165	4	6	13	11	34	1

**Supplementary Table 5.** Association of Lipid Signature With Expression of Lipogenic Enzymes in Tumor Tissue

TG species	FASN	FADS1	FADS2	ELOVL5
TG.48.1	0.753	0.873	0.960	0.707
TG.48.2	0.753	0.873	0.692	0.707
TG.49.1	0.795	0.873	0.269	0.707
TG.50.1	0.753	0.278	0.269	0.707
TG.50.2	0.753	0.873	0.960	0.707
TG.50.3	0.753	0.888	0.834	0.707
TG.50.4	0.753	0.873	0.834	0.744
TG.51.1	0.753	0.903	0.960	0.707
TG.51.2	0.753	0.873	0.546	0.708
TG.51.3	0.753	0.848	0.269	0.707
TG.52.2	0.753	0.997	0.939	0.707
TG.52.3	0.868	0.848	0.879	0.707
TG.54.2	0.753	0.531	0.960	0.707
TG.54.3	0.876	0.997	0.692	0.707
TG.54.4	0.876	0.848	0.692	0.707
TG.56.4	0.753	0.278	0.960	0.707
TG.56.5	0.753	0.278	0.960	0.707
TG.56.6	0.753	0.280	0.960	0.707

**Supplementary Table 6.** Association of Lipid Signature With Expression of Lipogenic Enzymes in Normal Tissue

TG species	FASN	FADS1	FADS2	ELOVL5
TG.48.1	0.484	0.208	0.727	0.262
TG.48.2	0.639	0.066	0.727	0.170
TG.49.1	0.484	0.304	0.727	0.371
TG.50.1	0.578	0.634	0.716	0.813
TG.50.2	0.971	0.066	0.727	0.170
TG.50.3	0.495	0.502	0.727	0.747
TG.50.4	0.848	0.066	0.727	0.170
TG.51.1	0.495	0.066	0.365	0.170
TG.51.2	0.578	0.066	0.716	0.170
TG.51.3	0.578	0.900	0.727	0.976
TG.52.2	0.848	0.304	0.727	0.648
TG.52.3	0.012 <sup>a</sup>	0.370	0.727	0.603
TG.54.2	0.307	0.807	0.727	0.976
TG.54.3	0.578	0.918	0.916	0.976
TG.54.4	0.639	0.477	0.727	0.747
TG.56.4	0.848	0.370	0.727	0.648
TG.56.5	0.495	0.304	0.727	0.262
TG.56.6	0.358	0.304	0.727	0.416

<sup>a</sup>Significance  $P < .05$ .

**Supplementary Table 7.** Genomic Data (Cohort CI): Relative Frequencies

Parameter/no. of cases tested	Observed frequency: tumor/normal tissue	Analysis method
MSI-status (DNA MSI) n = 61 tumors, n = 15 normal	Tumor MSI-high n = 9 (14%)/MSI-low n = 6 (9%) MSS n = 46 (75%) Normal MSI-high n = 0 (0%)/MSS: n = 15 (100%)	Multiplex PCR (Bethesda panel)
KRAS (exon 2 mutations) n = 75 tumors, n = 15 normal	Tumor Mutated: n = 33 (44%)/WT: n = 42 (56%) Normal Mutated: n = 0 (0%)/WT: n = 15 (100%)	HR-melting PCR
BRAF (exon 15 mutations n = 75 tumors, n = 15 normal)	Tumor Mutated: n = 9 (12%)/WT: n = 66 (88%) Normal Mutated: n = 0 (0%)/WT: n = 15 (100%)	HR-melting PCR
Double-positive cases/oncogenic mutations	Tumor KRAS mutated and BRAF mutated: n = 0 (0%)	HR-melting PCR

MSS, microsatellite stable.

**Supplementary Table 8.** Association of TG Species With Genomic and Clinical Data

TG species	KRAS mut	BRAF mut	MSI-status	CD3 <sup>+</sup> cells/mm <sup>2</sup>	CD4 <sup>+</sup> cells/mm <sup>2</sup>	CD8 <sup>+</sup> cells/mm <sup>2</sup>	Grading	Lymphangiosis
TG.48.1	0.945	0.874	0.905	0.963	0.927	0.998	0.982	0.018 <sup>a</sup>
TG.48.2	0.971	0.972	0.905	0.963	0.927	0.998	0.982	0.056
TG.49.1	0.761	0.874	0.905	0.963	0.927	0.998	0.982	0.006 <sup>a</sup>
TG.50.1	0.761	0.874	0.905	0.963	0.927	0.998	0.982	0.083
TG.50.2	0.997	0.972	0.905	0.963	0.927	0.998	0.982	0.025 <sup>a</sup>
TG.50.3	0.761	0.874	0.905	0.963	0.927	0.998	0.982	0.072
TG.50.4	0.761	0.874	0.905	0.963	0.927	0.998	0.982	0.038 <sup>a</sup>
TG.51.1	0.761	0.972	0.905	0.963	0.927	0.998	0.982	0.072
TG.51.2	0.856	0.874	0.905	0.963	0.927	0.998	0.982	0.021 <sup>a</sup>
TG.51.3	0.761	0.874	0.905	0.963	0.927	0.998	0.982	0.006 <sup>a</sup>
TG.52.2	0.761	0.874	0.993	0.963	0.927	0.998	0.982	0.648
TG.52.3	0.761	0.155	0.905	0.963	0.927	0.998	0.982	0.904
TG.54.2	0.971	0.874	0.905	0.963	0.927	0.998	0.982	0.072
TG.54.3	0.971	0.874	0.905	0.963	0.927	0.998	0.982	0.442
TG.54.4	0.761	0.182	0.993	0.963	0.927	0.998	0.982	0.685
TG.56.4	0.761	0.874	0.993	0.963	0.927	0.998	0.982	0.258
TG.56.5	0.761	0.874	0.905	0.963	0.927	0.998	0.982	0.072
TG.56.6	0.761	0.874	0.905	0.963	0.927	0.998	0.982	0.072

<sup>a</sup>Significant P < .05 (Adjusted P-values indicated).